



HAL
open science

Contrôle traductionnel du SARS-CoV-2

Lionel Condé

► **To cite this version:**

Lionel Condé. Contrôle traductionnel du SARS-CoV-2. Virologie. Ecole normale supérieure de Lyon - ENS LYON, 2024. Français. NNT : 2024ENSL0010 . tel-04694234

HAL Id: tel-04694234

<https://theses.hal.science/tel-04694234v1>

Submitted on 11 Sep 2024

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.



THESE

en vue de l'obtention du grade de Docteur, délivré par
l'ÉCOLE NORMALE SUPÉRIEURE DE LYON

Ecole Doctorale N° 340
Biologie Moléculaire, Intégrative et Cellulaire (BMIC)

Discipline : Sciences de la vie et de la santé

Soutenue publiquement le 06/06/2024, par :

Lionel CONDÉ

Contrôle traductionnel du SARS-CoV-2

Devant le jury composé de :

Dr. FIORINI TREGOUET Francesca, Chargée de recherche – HDR, CNRS

Dr. PAILLART Jean-Christophe, Directeur de recherche, CNRS

Dr. DE BREYNE Sylvain, Chargé de recherche, École normale supérieure de Lyon

Dr. ETIENNE Lucie, Directrice de recherche, École normale supérieure de Lyon

Dr. SARGUEIL Bruno, Directeur de recherche, Université Paris Cité

Dr. OHLMANN Théophile, Directeur de recherche, École normale supérieure de Lyon

Rapporteur

Rapporteur

Examineur

Examinatrice

Examineur

Directeur de thèse



THESE

en vue de l'obtention du grade de Docteur, délivré par
l'ÉCOLE NORMALE SUPÉRIEURE DE LYON

Ecole Doctorale N° 340
Biologie Moléculaire, Intégrative et Cellulaire (BMIC)
Discipline : Sciences de la vie et de la santé

Soutenue publiquement le 06/06/2024, par :

Lionel CONDÉ

Contrôle traductionnel du SARS-CoV-2

Devant le jury composé de :

Dr. FIORINI TREGOUET Francesca , Chargée de recherche – HDR, CNRS	Rapporteur
Dr. PAILLART Jean-Christophe , Directeur de recherche, CNRS	Rapporteur
Dr. DE BREYNE Sylvain , Chargé de recherche, École normale supérieure de Lyon	Examineur
Dr. ETIENNE Lucie , Directrice de recherche, École normale supérieure de Lyon	Examinatrice
Dr. SARGUEIL Bruno , Directeur de recherche, Université Paris Cité	Examineur
Dr. OHLMANN Théophile , Directeur de recherche, École normale supérieure de Lyon	Directeur de thèse

Résumé

Pendant l'infection virale, la régulation de l'expression des gènes est au cœur des interactions complexes entre l'hôte et le pathogène. Les virus exploitent la machinerie cellulaire de l'hôte pour assurer la synthèse de leurs protéines nécessaires pour la réplication et la propagation de l'infection. C'est notamment le cas lors de l'infection par le SARS-CoV-2, qui induit rapidement une inhibition globale de la traduction cellulaire grâce à l'action de facteurs viraux tels que la protéine Nsp1. Pour produire efficacement ses protéines, le virus doit alors mettre en place des stratégies pour contourner cette inhibition. Le génome du SARS-CoV-2 s'exprime à partir de 10 ARN, l'ARN génomique (ARNg) et 9 ARN sous-génomiques qui possèdent une région leader commune mais des régions 5'UTR uniques pour chacun des transcrits. Mon travail s'est concentré sur les éléments structuraux qui régulent la traduction des différents ARN du SARS-CoV-2.

À travers un ensemble d'expériences *in vitro* (lysats de réticulocytes) et en cellules, nous avons découvert que l'efficacité de traduction variait significativement entre les différents ARN viraux. En particulier, l'ARN génomique, malgré sa structure complexe, se distingue par une efficacité de traduction particulièrement élevée. Nous avons aussi déterminé que la structure tige-boucle SL1, présente dans l'ensemble des transcrits viraux, était un déterminant majeur pour l'expression des ARN et qu'elle jouait également un rôle crucial pour contrer l'inhibition induite par la protéine virale Nsp1. Nous avons établi que l'initiation de la traduction se déroulait par un mécanisme dépendant de la coiffe et nécessitait le complexe eIF4F. Enfin notre étude a également permis de caractériser le rôle de deux courtes phases de lecture ouvertes (uORF) retrouvées dans certaines régions 5'UTR des ARN du SARS-CoV-2; ces uORF ont des impacts variables selon leur position.

Mots-clés : SARS-CoV-2, contrôle traductionnel, régulation de l'expression des gènes, uORF, codon d'initiation non canonique

Abstract

During viral infection, the regulation of gene expression is central to the complex interactions between the host and the pathogen. Viruses exploit the host's cellular machinery to ensure the synthesis of their proteins, which are necessary for replication and the spread of the infection. This is particularly the case with SARS-CoV-2 infection, which rapidly induces a global inhibition of cellular translation through the action of viral factors such as the Nsp1 protein. To efficiently produce its proteins, the virus must implement strategies to bypass this inhibition. The SARS-CoV-2 genome is expressed from 10 RNAs, the genomic RNA (gRNA) and 9 subgenomic RNAs that possess a common leader region but unique 5'UTR regions for each of the transcripts. My work focused on the structural elements that regulate the translation of the different SARS-CoV-2 RNAs.

Through a series of *in vitro* (reticulocyte lysate) and in-cell experiments, we discovered that the translation efficiency varied significantly among the different viral RNAs. In particular, the genomic RNA, despite its complex structure, distinguishes itself by its remarkably high translation efficiency. We also determined that the SL1 stem-loop structure, present in all viral transcripts, was a major determinant for RNA expression and also played a crucial role in countering the inhibition induced by the Nsp1 viral protein. We established that translation initiation occurred through a cap-dependent mechanism and required the eIF4F complex. Finally, our study also characterized the role of two short upstream open reading frames (uORF) found in certain 5'UTR regions of SARS-CoV-2 RNAs; these uORF have variable impacts depending on their position.

Keywords : SARS-CoV-2, translation control, gene regulation, uORF, non-canonical initiation codon.

Table des matières

Résumé	5
Abstract.....	6
Table des matières.....	7
Liste des tableaux	11
Liste des figures d'Introduction	13
Liste des sigles et abréviations	15
Remerciements	21
Introduction.....	25
Chapitre 1– La traduction chez les eucaryotes.....	25
1 Aperçu général	25
2 L'ARNm	27
2.1 Généralités.....	27
2.2 La coiffe	27
2.3 La région 5'UTR de l'ARNm.....	28
a) Longueur de la 5'UTR	28
b) Structure de la région 5'UTR.....	29
c) Les uORF.....	30
3 Les facteurs d'initiation de la traduction	32
3.1 eIF4E.....	32
3.2 eIF4G	33
3.3 eIF4A.....	34
3.4 eIF3	36
3.5 eIF2	37
3.6 eIF4B/eIF4H	38

3.7 PAPB.....	38
4 Le mécanisme canonique de traduction.....	40
4.1 La formation du complexe ternaire: eIF2•GTP•Met-tRNAi	42
4.2 Liaison du complexe ternaire au ribosome et formation du 43 S.....	42
4.3 Liaison du complexe 43S sur l'ARNm: formation du complexe 48S.....	44
4.4 Le balayage de la région 5'UTR.....	47
a) Mécanisme moléculaire du balayage de la région 5'UTR par le complexe 48S	48
b) Dynamique d'association du complexe eIF4F au cours du balayage	49
4.5 Reconnaissance du codon AUG et assemblage du ribosome 80S.....	51
5 Mécanismes alternatifs de traduction: quelques exemples.....	52
5.1 La traduction des uORF.....	53
5.2 Initiation dépendante des codons non AUG.....	56
a) Présentation générale.....	56
b) Mécanisme d'initiation de la traduction à partir des codons non canoniques	57
5.3 Initiation à partir de bases modifiées.....	57
Chapitre 2 – La traduction virale	61
1 CARACTERISTIQUE STRUCTURALES.....	62
1.1 L'expression de polyprotéines.....	62
1.2 Le ribosomal <i>Frameshifting</i>	63
1.3 Initiation alternative par <i>leaky scanning</i>	64
1.4 Les IRES	65
a) Classification des IRES	65
b) Les IRES de classe I et II.....	66
c) Les IRES de type III.....	68
d) Les IRES de type IV.....	70

1.5 Le shunt ribosomal	72
2 Le contrôle de la traduction par les virus	73
2.1 Aperçu général.....	73
2.2 Le contrôle des facteurs d'initiation.....	74
a) Clivage protéolytique des facteurs d'initiation	75
b) Relocalisation des facteurs d'initiation	76
c) Contrôle de l'expression des facteurs d'initiation	76
2.3 Le contrôle de l'interaction 4E-BP/4E	77
2.4 Le contrôle de eIF2 α et PKR.....	80
a) Présentation générale.....	80
b) Résistance à l'inactivation d'eIF2.....	80
c) Stratégie d'échappement à PKR.....	81
d) Formation des granules de stress et stratégies virales d'échappement	82
Chapitre 3 – Le contrôle traductionnel du SARS-CoV-2	83
1 Le virus du SARS-CoV-2.....	83
1.1 Présentation générale du SARS-CoV-2.....	83
1.2 Résumé du cycle de réplication virale	85
2 Organisation du génome viral	89
2.1 Les caractéristiques des transcrits du SARS-CoV-2.....	89
2.2 Les éléments fonctionnels des régions 5 'UTR	92
3 Contrôle de la traduction pendant l'infection par le SARS-CoV-2.....	96
3.1 Nsp1 : protéine centrale de l'inhibition traductionnelle	96
3.2 Comment les ARN viraux échappent à l'inhibition induite par Nsp1.....	99
3.3 Implication des autres protéines virales	100
Projet de thèse	101
Résultats : partie 1	103

Résultats : partie 2	127
1 Caractérisation de la fonction de la tige boucle SL1	127
1.1 La tige boucle SL1 module l'efficacité de traduction	127
1.2 Plusieurs éléments de la tige-boucle SL1 sont impliqués dans l'efficacité traductionnelle médiée par la 5'UTR du SARS-CoV-2	128
1.3 Effet de la protéine Nsp1 sur la traduction des ARNm du SARS-CoV-2: effet de la tige boucle SL1	130
2 Des uORF régulent l'expression des ARN viraux.....	132
2.1 Impact sur la traduction de l'ARN génomique (ORF 1a/b)	133
2.2 Impact sur l'ORF 7a et 8.....	134
2.3 Impact sur l'ORF 6.....	135
2.4 Impact sur l'ORF S.....	135
Matériel et Méthodes	155
Discussion et Perspectives	159
Nouvelles perspectives et Projets futurs	167
Références bibliographiques	173
Annexes	207
1 Annexe I	207
Revue contrôle traductionnel des coronavirus.....	207
2 Annexe II.....	231
3 Annexe III.....	235
4 Annexe IV	236

Liste des tableaux

Tableau 1. –	Présentation des facteurs d’initiation de la traduction27
Tableau 2. –	Tableau récapitulatif des protéines structurales du SARS-CoV-2 ...88
Tableau 3. –	Tableau récapitulatif des protéines Nsp du SARS-CoV-289
Tableau 4. –	Tableau récapitulatif des protéines accessoires du SARS-CoV-2 ...89

Liste des figures d'Introduction

Figure 1. – Traduction eucaryote	26
Figure 2. – Les différents types d'uORF.....	31
Figure 3. – Structure du facteur d'initiation de la traduction eIF4E.....	33
Figure 4. – Représentation schématique des sites de liaison d'eIF4G	34
Figure 5. – Structure du facteur d'initiation de la traduction eIF3.....	37
Figure 6. – Structure et fonction de PABP	39
Figure 7. – Schéma bilan de l'initiation canonique de la traduction.....	41
Figure 8. – Formation du complexe ternaire	42
Figure 9. – Assemblage du complexe de préinitiation 43S.....	43
Figure 10. – Activation de l'ARNm.....	44
Figure 11. – Recrutement du complexe 43S sur l'ARNm.....	45
Figure 12. – Modèles de recrutement du complexe 43S sur l'ARNm	46
Figure 13. – Association du complexe eIF4F au cours du balayage.....	50
Figure 14. – Structure d'eIF4A en complexe avec la sous unité 40S.....	51
Figure 15. – Mécanisme régulateur des uORF	54
Figure 16. – Initiation de la traduction par l'intermédiaire des bases modifiées m6A	58
Figure 17. – Ribosomal <i>frameshifting</i>	64
Figure 18. – Les IRES de type I et de type II	67
Figure 19. – IRES de type III	69
Figure 20. – IRES de type IV	71
Figure 21. – Shunt ribosomal.....	72
Figure 22. – Contrôle traductionnel médié par 4EBP	78
Figure 23. – Arbre phylogénétique.....	84
Figure 24. – Organisation du génome viral du SARS-CoV-2.....	85
Figure 25. – Cycle de réplication virale du SARS-CoV-2	86
Figure 26. – Mécanisme de fusion du SARS-CoV-2	87
Figure 27. – Représentation des polyprotéines du SARS-CoV-2.....	87
Figure 28. – Modèle de la transcription discontinue.....	90

Figure 29. –	Représentation des ARN sous-génomiques du SARS-CoV-2	91
Figure 30. –	Organisation des ARN génomiques et sous génomiques du SARS-CoV-2	92
Figure 31. –	Structure de la région 5'UTR du SARS-CoV-2	94
Figure 32. –	Position du codon non canonique CUG-59 dans les différentes 5'UTR du SARS-CoV-2	95
Figure 33. –	Structure de la protéine virale Nsp1	97
Figure 34. –	Mécanisme d'action de la protéine Nsp1	98

Liste des sigles et abréviations

3'UTR	Région 3' non traduite
40S	Sous-unité ribosomique 40S
43S	Complexe de préinitiation 43S
4E-BP	4E-binding protein
5'UTR	Région 5' non traduite
60S	Sous-unité ribosomique 60S
aa	Acide aminé
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
ARNm	ARN messenger
ARNg	ARN génomique
ARNsg	ARN sous-génomique
ARNt	ARN de transfert
ATP	Adénosine triphosphate
C-terminale	Extrémité C-terminale
CaMV	Virus de la mosaïque du chou-fleur
CBV3	Coxsackievirus B3
CMV	Cytomégalovirus
CrPV	Virus de la Paralysie du Cricket
Cryo-EM	Cryo-microscopie électronique
CSFV	Virus de la peste porcine classique
DDX	ARN hélicase à boîte DEAD
DEAD	Acide aspartique-Acide glutamique-Alanine-Acide aspartique
dsRNA	ARN double brin
EBER	Epstein–Barr virus–encoded small RNAs
eIF	Facteur d'initiation de la traduction
EMCV	Virus de l'encéphalomyocardite
FMDV	Virus de la fièvre Aphteuse
G3BP1	Ras GTPase-activating protein-binding protein 1

GCN2	General Control Non-derepressible 2
GDP	Guanosine diphosphate
GTP	Guanosine triphosphate
HCV	Virus de l'hépatite C
HRI	Heme Regulated Inhibitor
IRE	Iron Responsive Element
IRES	Site d'entrée interne du ribosome
IRP	Iron Reponsive Protein
KDa	Kilo Dalton
m5C	5-Méthylcytosine
m6A	N6-Méthyladénosine
SL	Tige-boucle
MAPK	MAP kinases
MERS-CoV	Coronavirus du syndrome respiratoire du Moyen-Orient
Met-tRNAi	ARNt initiateur
METTL3	N6-adenosine-methyltransferase 70 kDa subunit
miR	Micro ARN
Mnk1	MAP kinase-interacting serine/threonine-protein kinase 1
mTOR	mammalian Target Of Rapamycin
N-Terminale	Extrémité N-terminale
NMD	Nonsense Mediated mTNA Decay
Nsp	Protéine non structurale
nT	Nucléotide
ORF	Cadre ouvert de lecture
PABP	Poly (A)-binding protein
PACT	Protein activator of PKR
Paip	PABP-Binding protein
PCBP	poly (rC)-binding protein 1
PERK	protein kinase R (PKR)-like endoplasmic reticulum kinase
PI3K	Phosphoinositide 3-kinase
PK	Pseudo-nœuds
PKR	Protein Kinase R

poly(A)	Queue Poly(A)
PTEN	Phosphatase and TENsin homolog
PV	Virus de la poliomyélite
RBD	Domaine de liaison à l'ARN
RBP	Protéine de liaison à l'ARN
RRM	Motif de reconnaissance de l'ARN
rRNA	ARN ribosomique
SARS-CoV	Coronavirus du syndrome respiratoire aigu sévère
SARS-CoV-2	Coronavirus 2 du syndrome respiratoire aigu sévère
uORF	Upstream open reading frame
VIH	Virus de l'immunodéficience humaine
VPg	Viral protein genome-linked
YTHDF3	YTH domain family protein 3

Je dédie cette thèse à toutes celles et ceux qui, comme moi, sont convaincus que la recherche et l'enseignement, unis dans un effort commun, œuvrent pour un monde où le savoir sert l'Humain dans son épanouissement collectif

« Science is a way of thinking much more than it is a body of knowledge »

Carl Sagan

« Sans la curiosité de l'esprit, que serions-nous ? Telle est bien la beauté et la noblesse de la science : un désir sans fin de repousser les frontières du savoir, de traquer les secrets de la matière et de la vie sans idée préconçue des conséquences éventuelles »

Marie Curie

« Le progrès, ce n'est pas l'acquisition de biens. C'est l'élévation de l'individu, son émancipation, sa compréhension du monde. Et pour ça il faut du temps pour lire, s'instruire, se consacrer aux autres »

Christiane Taubira

Remerciements

Le temps passé dans le bâtiment M5 de l'ENS, et plus particulièrement dans l'équipe REVE, a été très enrichissant et agréable. Au cours de ces années, j'ai eu la chance de m'épanouir dans un environnement propice à la poursuite de ma thèse. Si toutes ces journées sont passées si vite, c'est sans aucun doute grâce aux personnes que j'y ai côtoyées.

Pour ne pas déroger à la coutume, je souhaite commencer par remercier mon directeur de thèse, Théophile Ohlmann, pour m'avoir accueilli dans son équipe et m'avoir guidé tout au long de cette thèse, en compagnie de Sylvain de Breyne. Leur aide a été précieuse et toujours apportée avec efficacité et bienveillance. Théophile et Sylvain sont deux scientifiques passionnés par les mécanismes moléculaires de la traduction eucaryote. Ils n'ont jamais manqué une occasion de partager leur enthousiasme, leur expertise et leur savoir au cours de nos nombreuses discussions. Lors des années passées au laboratoire, j'ai toujours été encouragé et soutenu pour mener à bien un projet stimulant, et j'en suis tout particulièrement reconnaissant. Je les remercie également pour l'autonomie qu'ils m'ont toujours accordée et pour avoir maintenu en permanence un environnement où il était facile et agréable de faire de la recherche. Vous m'avez permis de continuer à apprécier la recherche fondamentale en laquelle je crois énormément. Vous m'avez permis de continuer à aimer la Science, en étant animé par la curiosité de la découverte, l'attrait du savoir et la compréhension du monde qui nous entoure. Je suis persuadé que vous avez eu un impact sur la poursuite de ma carrière scientifique, notamment en me poussant à développer un solide esprit critique. Et bien sûr, puisque c'est ce dont il est question ici, je souhaite vous remercier pour la correction de ce manuscrit de thèse dans un effort collectif.

Je remercie également ma rapportrice Francesca Fiorini Tregouet et mon rapporteur Jean-Christophe Paillart pour la relecture attentive de ce manuscrit. Vos commentaires et suggestions ont été précieux et utiles pour améliorer ce travail. J'adresse également des remerciements à mes examinateurs, Lucie Etienne et Bruno Sargueil.

Je tiens ensuite à remercier avec humilité tous les membres de l'équipe REVE. Vous m'avez souvent conseillé, parfois écouté, et surtout toujours soutenu, ce qui, je le reconnais, n'est pas une tâche aisée! Donc, merci Caroline, Evelyne, Fabrice, Henri, Laurent, Olivia, Philippe, Sylvain, Théophile, Thibaut, ainsi que toutes les autres personnes rencontrées dans l'équipe. Au fil des années, plusieurs d'entre vous sont devenus des amis.

Olivia, merci pour ton accueil au laboratoire et pour les innombrables discussions que nous avons partagées. Ta présence a rendu mon intégration dans cette équipe très facile. Merci également de m'avoir permis de rencontrer nos amis du LBMC. Tes conseils ont été précieux dès le début de cette thèse. Je sais que nous n'avons pas toujours eu le temps de nous voir autant que nous l'aurions souhaité, mais nous aurons tout le temps de rattraper cela bientôt.

Caroline, merci pour ta bienveillance et ton oreille attentive. J'ai toujours apprécié nos échanges, qu'ils portent sur la science ou sur tout autre sujet du quotidien.

Thibaut, on dit souvent que parfois peu de mots valent plus que de longs discours, et je pense que cela s'applique très bien ici. Cette thèse n'aurait pas été la même sans ta présence, au laboratoire et en dehors. Merci d'être mon voisin de bureau presque depuis le premier jour. Désolé d'ailleurs pour le bazar permanent qui y règne. Merci pour tous nos échanges scientifiques et nos discussions (souvent politiques) au détour d'une manip, autour d'un café, tard le soir au laboratoire ou sur le chemin du retour d'une virée quelconque. Merci pour toutes les fois où l'un de nous a dit « j'ai une idée » (bonne ou mauvaise en y repensant) et que l'autre s'y est engouffré sans hésiter. J'espère que notre amitié durera longtemps. Je te souhaite le meilleur pour la suite, je ne m'inquiète pas une seule seconde pour toi. Voici le mot de la fin : « Celui qui sait briser la table basse sait briser les barrières du prime editing ». Je te laisse méditer là-dessus.

Philippe, merci d'avoir été mon voisin de paillasse depuis le premier jour. J'ai énormément apprécié ton accueil au laboratoire et la facilité avec laquelle tu m'as aidé à m'intégrer bien au-delà des frontières de l'équipe et du CIRI. Quelle chance d'occuper un bout de paillasse avec quelqu'un avec qui je partage une vision similaire de la science, des projets communs, et une vision de la société identique. Merci pour tous les moments de brainstorming où tu as partagé avec moi ta créativité et ton ingéniosité. Je pense que tu apportes énormément à notre équipe. Pour reprendre tes mots, nous sommes certes des collaborateurs scientifiques mais aussi et avant tout des amis. J'ai une très grande gratitude envers toi. Fabrice, je te remercie pour toute l'aide que tu m'as apportée. J'ai beaucoup apprécié travailler avec toi ces derniers mois. Merci pour tout le travail commun que tu fais pour l'équipe, ça fait de toi un pilier de notre groupe. Merci pour tous tes précieux conseils et tes protocoles si bien rédigés. Merci pour les discussions tardives devant la porte de ton bureau.

Laurent, merci pour ta bonne humeur permanente et ton attitude joviale de tous les instants.

Sylvain, je te remercie particulièrement pour m'avoir transmis la rigueur scientifique. Je suis arrivé au laboratoire avec un fort intérêt pour la régulation de l'expression des gènes. J'ai maintenant un très fort intérêt pour la traduction à part entière ; tu as donc su me transmettre ta passion et m'encourager à persévérer dans l'exploration de ce domaine. Merci également pour tes nombreux conseils tant techniques que scientifiques que tu as partagés avec moi et qui m'ont beaucoup fait réfléchir sur mon projet.

Théophile, je te remercie de m'avoir accueilli dans une équipe en parfaite adéquation avec mes thématiques d'intérêt. Ton oreille attentive, tes encouragements constants et ta méthode d'encadrement ont été très appréciés. Je pense que tu as su trouver un juste équilibre entre conseil, liberté et soutien, toujours avec bienveillance. Je te suis particulièrement reconnaissant de m'avoir permis de m'engager dans de nombreux projets, tant au laboratoire qu'en dehors. Ta capacité d'écoute et tes précieux conseils ont été déterminants pour le bon déroulement de cette thèse.

Merci aux différents stagiaires que nous avons accueillis dans l'équipe. Pour celles et ceux que j'ai encadrés, j'espère l'avoir fait avec toute la bienveillance nécessaire. Tanaïs, Tristan, Tom, Hélène, Clara, Swann, je veux vous remercier pour tout ce que vous avez apporté au projet. Votre motivation et votre énergie m'ont toujours encouragé à développer de nouveaux axes dont vous vous êtes ensuite emparés. Merci pour toutes les questions que vous m'avez posées, pour tous les moments où vous

m'avez challengé et où vous avez chamboulé certaines de mes certitudes. Sans le savoir, vous m'avez énormément appris. Je suis certain que vous vous épanouirez dans la suite de vos parcours. Gardez à l'esprit que tout est possible avec de la motivation et une bonne dose de travail collectif.

Je n'oublie pas non plus les membres des équipes LR et OR de l'autre côté du magnifique aquarium qui nous sépare (je ne peux d'ailleurs que remercier celles et ceux qui ont contribué à sa décoration avec goût et conviction). Cela a toujours été un plaisir de discuter, de rigoler et de travailler avec vous. Vous êtes trop nombreux pour que je vous nomme toutes et tous ici, mais vous savez que je vous apprécie énormément.

Je voudrais remercier en particulier Xuan-Nhi. Tu es définitivement un pilier du bâtiment M5. Merci pour tous les moments que nous avons partagés et pour ton énergie permanente. C'est sans aucun doute le temps que tu accordes aux autres qui est ton meilleur trait de caractère. Je suis sûr que tu resteras longtemps le ciment social du M5.

Je voudrais également prendre quelques instants pour remercier Stéphanie. Je me souviens très clairement d'une des premières fois où nous avons discuté au coin de ta paillasse ; nous sommes depuis devenus des amis. Ta motivation et ta détermination m'impressionnent. Cela m'a encouragé tout au long de cette thèse, faisant de toi une vraie source d'inspiration. Je suis ravi de tout ce qui t'est arrivé récemment et j'espère que nous continuerons à rester en contact encore longtemps.

Je tiens également à remercier Andrea pour ne pas nous avoir systématiquement chassés de son équipe lorsque deux acolytes de l'équipe REVE y semaient la pagaille. Merci pour nos échanges scientifiques et pour être un si bon public de nos fourberies (j'ai longtemps cherché le mot adéquat). Merci à mes amis du LBMC ; merci Thibault pour toutes nos discussions autour de la traduction, merci Paul, Julien et Matéo pour vos inventions plus folles les unes que les autres. Merci également pour tous les moments festifs que nous avons passés ensemble ; vous avez participé à la quasi-totalité des événements qui auront marqué cette thèse.

Merci aussi à tout le personnel du CIRI et de l'ENS de Lyon qui œuvre pour faciliter le travail. Je pense, entre autres, à vous Rahnia et Danielle, votre bonne humeur fait la différence. Merci également à la SFR Biosciences pour l'ensemble des plateaux techniques. Je n'oublie pas non plus de remercier notre secrétariat pour m'avoir encouragé à cultiver la polyvalence et la persévérance.

Je souhaite garder quelques lignes pour remercier toutes les personnes qui m'ont permis d'arriver là où j'en suis aujourd'hui. Je remercie l'école doctorale BMIC pour m'avoir fait confiance et m'avoir attribué une bourse ministérielle ainsi que les membres de mon CST pour leurs conseils avisés. Je remercie également mon alma mater, l'Université de Montréal, qui m'a donné une solide formation. J'y ai rencontré Luc DesGroseillers, qui m'a accueilli au sein de son laboratoire pour compléter mon stage de L2 puis mon master. Luc est un scientifique passionné qui m'a transmis avec facilité son enthousiasme pour la recherche. Sa rencontre est sans aucun doute un des événements qui a marqué positivement mon parcours universitaire, et je l'en remercie chaleureusement.

Je pense également à tous mes professeurs qui, sur les bancs de l'école de la République, m'ont transmis leurs connaissances pendant toutes ces années. Aujourd'hui, je me rappelle particulièrement de ce jour où, au détour d'une discussion, mon proviseur Jacques Puel m'a dit: « Lionel, trouve un

métier qui te plaît et tu ne travailleras plus une seule journée de ta vie. » M. Puel, je crois que j'ai trouvé : c'est la recherche fondamentale et l'enseignement.

Ceux qui me connaissent bien savent combien il est essentiel pour moi de m'investir dans de nombreux projets pour trouver un certain équilibre. Ceux auxquels j'ai participé ces dernières années m'ont sans aucun doute aidé à traverser cette thèse avec joie et sérénité. Pour commencer, je voudrais remercier l'UCBL et l'ENS de Lyon pour m'avoir permis d'enseigner. Je suis à titre personnel convaincu que la recherche est difficilement dissociable de la notion de transmission des connaissances. C'était donc pour moi, au-delà d'une chance, un vrai plaisir. Merci Dominique de m'avoir donné ma chance comme ACE à l'UCBL puis Nathalie à l'ENS. Je remercie également ceux qui m'ont ensuite encouragé à poursuivre vers un ATER pour rejoindre une magnifique équipe pédagogique. Je pense en particulier à vous, Marie et Chloé. Vos encouragements m'ont conforté dans mon choix et ont été extrêmement appréciés. Merci Auriane pour toutes nos discussions et tes précieux conseils. Je suis heureux de te compter comme co-ATER et comme amie. Ensuite, je voudrais remercier mes collègues représentants des doctorants et doctorantes avec qui nous avons pu mener de jolis projets qui, je l'espère, dureront dans le temps. Merci Gregory, Blaise et Claudia.

Enfin, encore et toujours, je tiens à remercier mes amis qui continuent à me soutenir avec une patience bien plus grande que de raison, même lorsque nous ne pouvons pas nous voir pendant de longues périodes. Je pense particulièrement à celles et ceux qui sont de l'autre côté de l'Atlantique. Votre amitié est précieuse. Encore une fois, je tiens à remercier tout particulièrement mes parents, mes grands-parents, mes frères, mon oncle et ma tante. Ils m'ont sans cesse encouragé et soutenu dans la poursuite de mes rêves, me donnant l'énergie nécessaire pour me dépasser et atteindre mes objectifs. Merci, Maman, pour tous les moments où tu m'as aidé pendant mes études. Bien sûr, je ne parle pas uniquement des moments où tu as corrigé mes fautes de français (y compris dans cette thèse), mais aussi de toutes les occasions où tu m'as encouragé à faire ce qui me plaisait avant tout. Je suis très fier des valeurs que tu m'as transmises ; elles enrichissent sans aucun doute mon quotidien et dictent ma façon d'interagir avec les gens qui m'entourent. Ton dévouement total au service du public m'impressionne toujours autant et m'incite à te prendre comme modèle les yeux fermés. Merci, Papa, de m'avoir montré que le monde est riche de différentes cultures et de m'avoir appris à toujours travailler avec rigueur et enthousiasme ; mais aussi tout en me poussant à me donner les moyens de mes ambitions. J'ai une chance incommensurable de tous vous avoir autour de moi.

Séverine, j'aurais clairement pu te citer dans l'ensemble de ce texte, mais je préfère te dédier les derniers remerciements de cette thèse. Tu m'as apporté un équilibre inestimable tout au long de ce parcours. Merci pour nos discussions, qu'elles soient scientifiques ou non, que ce soit sur la trottinette en venant au labo, pendant les pauses café, ou quelque part à l'autre bout du monde. Je te remercie pour tes encouragements constants dans cette thèse et pour toutes les fois où tu me remets sur les rails quand je suis déterminé à faire la révolution. Merci pour toute la patience et la bienveillance dont tu fais preuve au quotidien ainsi que pour tous tes conseils. Je sais que tu t'amuses dans ta thèse autant que moi dans la mienne, et j'ai hâte de voir la suite. Et surtout, merci que tout soit toujours aussi simple avec toi !

Introduction

Chapitre 1– La traduction chez les eucaryotes

1 Aperçu général

La traduction est un processus extrêmement régulé qui permet le décodage de l'information génétique portée par les ARNm. Au cœur de ce mécanisme complexe se trouvent les ribosomes formant un assemblage ribonucléoprotéique responsable de la synthèse protéique. Les ribosomes opèrent en synergie avec une multitude de facteurs cellulaires, notamment les facteurs d'initiation (eIF, eukaryotic initiation factor), essentiels pour garantir l'efficacité de la traduction d'un ARNm donné (tableau 1). D'autres protéines tels que les facteurs d'élongation et de terminaison sont également impliquées. L'efficacité avec laquelle l'ARNm est traduit détermine la quantité de protéine finale produite et ce processus se déroule en quatre étapes clés (Figure 1): (i) l'initiation, au cours de laquelle le ribosome va se fixer sur l'ARNm et rejoindre le codon d'initiation; (ii) l'élongation, caractérisée par la synthèse de la chaîne polypeptidique; (iii) la terminaison, marquée par l'arrêt de la synthèse protéique lorsque le ribosome rencontre le codon stop; et (iv) le recyclage des composants de la machinerie traductionnelle pour préparer un nouveau cycle de traduction. L'étape d'initiation est particulièrement régulée grâce à la nature et à la structure de la région 5' non traduite de l'ARNm (5'UTR, untranslated region) et par le jeu d'interactions avec de nombreuses protéines qui assurent la fixation du ribosome.

La synthèse des protéines débute par le recrutement de la machinerie traductionnelle constituée de la petite sous-unité du ribosome (40S) et sa migration jusqu'au codon d'initiation. L'attachement du ribosome sur l'ARNm est un processus qui est catalysé par de nombreuses protéines cellulaires dont les facteurs d'initiation. Dans le cadre d'une initiation classique, la machinerie traductionnelle est recrutée à l'extrémité 5' des ARNm par une interaction avec la coiffe et elle migre de façon linéaire en balayant la séquence d'ARN jusqu'à la rencontre du codon d'initiation AUG. Lorsque celui-ci est identifié, la sous-unité ribosomique 40S s'arrête et la sous-unité 60S est recrutée pour

former le ribosome 80 S qui va synthétiser la protéine. L'ensemble de ce processus est finement contrôlé par le jeu d'interactions protéines-protéines ou protéines-ARN et une consommation d'énergie sous forme d'hydrolyse d'ATP et de GTP; c'est un processus actif et relativement énergivore.

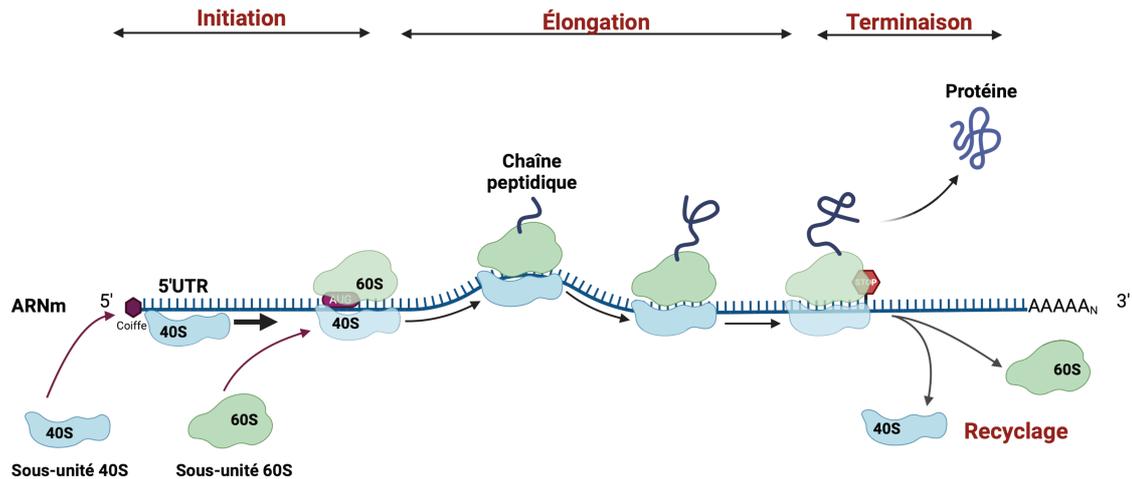


Figure 1. – Traduction eucaryote

La traduction des ARNm se déroule en trois étapes principales : l'initiation, l'élongation et la terminaison. Durant l'initiation, la sous-unité ribosomique 40S est recrutée à l'extrémité 5' de l'ARNm et balaye la région 5'UTR à la recherche du codon d'initiation. Lorsque ce codon est reconnu, la sous-unité 60S est recrutée pour former le ribosome 80S complet. L'élongation est l'étape où le ribosome 80S synthétise le polypeptide, correspondant à l'information encodée par l'ARNm. La terminaison survient lorsque le ribosome rencontre un codon stop et relâche la protéine néosynthétisée. S'ensuit une étape de recyclage, qui prépare les sous-unités ribosomiques pour un nouveau cycle de traduction. Dessinée sur BioRender.

Tableau 1. – Présentation des facteurs d'initiation de la traduction

Facteur d'initiation de la traduction	Fonction moléculaire	Nombre de sous-unités	Poids moléculaire (total*) (en kDa)
eIF1	Possède un rôle clé dans la sélection du codon d'initiation en empêchant la libération de eIF2 du complexe de pré-initiation 43S Favorise la liaison complexe ternaire à la sous-unité ribosomique 40S.	1	13
eIF1A	Coopère avec eIF1 pour contrôler la fidélité de la sélection du codon d'initiation. Participe au recyclage du complexe ribosomique post-termination. Favorise la liaison complexe ternaire à la sous-unité ribosomique 40S.	1	19
eIF3	Participe au recrutement du complexe de pré-initiation 43S au niveau l'ARNm en interagissant avec eIF4G. Favorise la liaison du complexe ternaire à la sous-unité ribosomique 40S.	13	800*
eIF2	Etablissement du complexe ternaire.	3	120*
eIF4A	Déroule les structures secondaires de l'ARN grâce à son activité hélicase avant le recrutement de complexe 43S et pendant l'étape de balayage de la région 5'UTR	1	46
eIF4G	Protéine d'échafaudage permettant la formation du complexe eIF4F (avec eIF4E et eIF4A) Participe au recrutement du complexe de préinitiation à l'extrémité 5' des ARN par une interaction avec eIF3	1	188
eIF4E	Protéine de liaison de la coiffe des ARNm.	1	25
eIF4B	Stimulateur de l'activité hélicase de eIF4A	1	69
eIF4H	Stimulateur de l'activité hélicase de eIF4A.	1	27
eIF5	Induction de la libération de eIF2 lors de la reconnaissance du codon d'initiation en hydrolysant le GTP lié à eIF2.	1	17
eIF5B	Permet l'association des sous-unités ribosomiques au codon d'initiation.	1	139
eIF6	Évite l'assemblage prématurée des sous-unités ribosomales 40S et 60S	1	27
PABP	Protéine de liaison à la queue poly(A) des l'ARNm et à eIF4G permettant la circularisation des l'ARNm	1	71

2 L'ARNm

2.1 Généralités

L'ARNm est transcrit dans le noyau à partir de l'ADN (Cramer 2019) sous forme d'un pré-ARNm qui est épissé puis exporté vers le cytoplasme où il est traduit (Hocine et al. 2010; Hershey et al. 2012). L'ARNm est constitué d'une séquence linéaire d'acides ribonucléiques simple brin contenant une région codante, flanquée de deux régions non traduites (5' et 3' UTR) qui interviennent à différents niveaux de sa régulation.

2.2 La coiffe

Chez les eucaryotes, le premier nucléotide de l'extrémité 5' de l'ARNm s'appelle la coiffe et c'est une guanosine méthylée en position 7 qui établit un pont phospho-diester 5'-5' avec le second nucléotide. La coiffe est ajoutée par l'action séquentielle de plusieurs enzymes. L'extrémité de l'ARN est d'abord hydrolysée par une triphosphatase pour passer d'une configuration triphosphate à une configuration

diphosphate. Parallèlement, une molécule de GTP est liée de façon covalente à une guanylyltransférase, qui transfère le GMP sur l'ARNm. Finalement, la position N7 de la guanosine 5' est méthylée par une méthyltransférase, formant ainsi la coiffe dite cap0 (Decroly et al. 2012). La coiffe est une structure nécessaire et joue un rôle dans l'épissage, l'export, la stabilité et, surtout, la traduction des ARNm en permettant le recrutement de la petite sous unité du ribosome.

2.3 La région 5'UTR de l'ARNm

Au cours de l'évolution, les régions 5' et 3' UTR se sont complexifiées et elles sont au centre des mécanismes de la traduction (Hernández et al. 2010; Chen et al. 2012). La 5'UTR est principalement impliquée dans le contrôle de la traduction alors que la région 3'UTR a un rôle dans le contrôle de la stabilité des ARNm bien que cette dernière puisse également, dans une certaine mesure, moduler la traduction (Hinnebusch et al. 2016; Mayr 2019). La 5'UTR peut renfermer différents éléments qui régulent la traduction tels que de courts cadres ouverts de lecture (uORF : upstream open reading frame), des motifs de liaison à des protéines ou des ARN non codants, ou encore des structures secondaires et tertiaires de l'ARN (Pesole et al. 2001). De fait, la longueur, la composition et la structure de la 5'UTR modulent l'efficacité de l'étape d'initiation et donc la quantité de protéines produites (Castello et al. 2016).

a) Longueur de la 5'UTR

Chez les mammifères la longueur moyenne de la 5'UTR est de 150 à 200 nucléotides (nt) et on observe de grandes variations d'un gène à l'autre (Mignone et al. 2002). Chez les humains, la 5'UTR est légèrement plus longue que chez les autres mammifères (Leppek et al. 2018). Une région trop longue et peu structurée peut recruter des ribosomes de manière efficace, à l'opposé une 5'UTR trop courte, en dessous de 12 nucléotides, peut entraîner une non-reconnaissance du codon d'initiation (Pestova and Kolupaeva 2002). Cela s'explique par le fait que le ribosome couvre une région de 28 nucléotides sur l'ARNm avec le site de 'reconnaissance' du codon AUG (site A) positionné à partir du 14^e nt ce qui explique pourquoi le codon

AUG peut ne pas être reconnu dans des 5'UTR très courtes. De fait, la taille de la 5'UTR est un déterminant important de l'efficacité traductionnelle.

b) Structure de la région 5'UTR

La présence de structures secondaires et tertiaires au sein de la région 5'UTR joue un rôle prépondérant dans le contrôle de l'expression de l'ARNm. En règle générale, la présence de structures secondaires dans la région 5'UTR réduit l'efficacité de traduction en interférant avec le déplacement du ribosome (Pelletier and Sonenberg 1985a). Cela est d'autant plus vrai pour des structures secondaires stables telles que les G-quadruplexes (Halder et al. 2009; Beaudoin and Perreault 2010). Cependant, la position de ces structures sur l'ARNm est également un facteur déterminant; ainsi, une structure secondaire, même de faible énergie, aura un effet d'autant plus important qu'elle se trouve à une position proche de la coiffe (autour de 10 nt) car elle empêche la formation du complexe qui la reconnaît (le complexe eIF4F) et donc le recrutement de la petite sous-unité du ribosome (Pelletier and Sonenberg 1985b).

Dans certains cas des protéines spécifiques peuvent s'associer aux structures secondaires pour les stabiliser. C'est par exemple le cas de l'ARNm codant pour la ferritine qui possède une structure secondaire appelée IRE (*Iron Responsive Element*) au tout début de sa région non traduite. En l'absence de fer, la protéine IRP (*Iron Responsive Protein*) se lie sur cette structure ce qui inhibe fortement la synthèse de ferritine. À l'inverse lorsque le niveau de fer augmente, la protéine IRP se retire et l'ARNm est traduit (Hentze et al. 1987; Gray and Hentze 1994).

Mais il existe aussi des exemples dans lesquels la présence d'une structure secondaire de faible énergie immédiatement en aval du codon AUG favorise l'initiation en diminuant le phénomène de leaky-scanning (voir le chapitre 2 section 1.3), (Wolin and Walter 1988; Kozak 1990). En effet, la structure d'ARN ralentit le ribosome ce qui augmente le temps de contact entre le codon AUG et le Met-tRNA_i dans le site du ribosome permettant une meilleure reconnaissance de ce dernier.

Nous verrons plus tard que les structures secondaires sont déroulées par le facteur eIF4A au sein du complexe eIF4F. Cela permet de faciliter le balayage de la petite sous-unité du ribosome et de traduire des ARNm contenant des 5'UTR longues et structurées (Dmitriev et al. 2007). La présence de structures secondaires complexes

telles que des G-quadruplexes dans certains ARNm entraîne aussi une augmentation du recrutement de eIF4A (Kozak 1990; Wolfe et al. 2014). Des protéines cellulaires comme la protéine DDX3, qui est une hélicase à ARN, assistent aussi le déroulement des structures proches de la coiffe pour faciliter l'attachement du ribosome (Soto-Rifo et al. 2012).

Enfin, les structures secondaires des ARN peuvent réguler l'efficacité de traduction de façon indirecte. Par exemple, la 5'UTR de l'interféron gamma contiennent une région d'ARN très structurée sur laquelle se lie la kinase PKR ce qui induit la phosphorylation de eIF2 α et l'arrêt de la traduction cellulaire (Ben-Asouli et al. 2002; Cohen-Chalamish et al. 2009).

c) Les uORF

Les uORF sont de courtes phases de lecture contenues dans la 5'UTR des ARNm et qui peuvent réguler la traduction de l'ORF principale en contrôlant le flux de ribosomes (Figure 2). Des études à large échelle ont montré que les ARNm qui possèdent des uORF sont moins traduits avec une baisse de 30 à 60% de leur expression (Calvo et al. 2009; Johnstone et al. 2016). Cela s'explique par le fait que le ribosome reconnaît et initie sur l'uORF ce qui l'empêche d'atteindre la cadre de lecture principal. Les uORF sont très abondantes (Ingolia et al. 2011; McGillivray et al. 2018; Dever et al. 2020) et il est estimé qu'entre 30 et 50 % des ARNm humains contiendraient au moins une uORF dans leur 5'UTR (Ji et al. 2015; Mackowiak et al. 2015; Chen et al. 2017; Cheng et al. 2018; Zhang et al. 2018). On les retrouve également dans des ARN viraux (Dever et al. 2020).

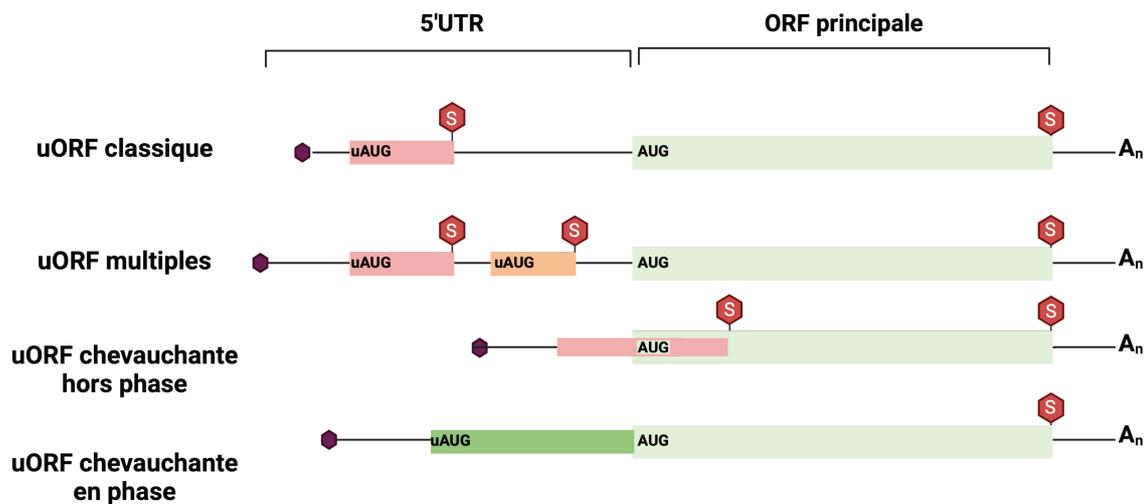


Figure 2. – Les différents types d'uORF

Les uORF ont des localisations et des configurations différentes. L'initiation de la traduction chez les eucaryotes peut être significativement modulée par la présence d'uORF situées dans la 5'UTR des ARNm. Les uORF peuvent se terminer dans la 5'UTR (uORF classiques) ou dans la séquence codante. Lorsqu'elles terminent dans la séquence codante, on parle d'uORF chevauchantes. Les uORF chevauchantes peuvent être en phase ou hors phase avec l'ORF principale, médiant des régulations distinctes. Les uORF en phase peuvent générer des isoformes de protéines allongées en N-terminal.

En termes de structure génétique, il existe différents types d'uORF qui sont classés selon leur position par rapport à la phase de lecture principale (Figure 2). On dénombre les uORF non chevauchantes qui se trouvent entièrement en amont du codon AUG principal ; les uORF chevauchantes hors phase dont le codon-stop se trouve en aval du codon AUG de l'ORF principale ; les uORF chevauchantes en phase, qui sont un prolongement 5'-terminal de l'ORF principale (Johnstone et al. 2016). En général, les uORF, hors phase ou chevauchantes inhibent l'expression de la phase de lecture principale; c'est aussi le cas lorsque le codon AUG de l'uORF est situé bien en aval de la coiffe car il a plus de chances d'être reconnu par le ribosome (Young and Wek 2016). Il existe également des ARNm qui possèdent plusieurs uORF au sein de leur région 5'UTR comme c'est le cas de l'ARNm codant pour GCN4 chez la levure qui comporte 4 uORF (Hinnebusch 2005) ; ou encore l'ARNm codant pour ATF4 chez l'humain (Young and Wek 2016). La régulation de ces ARNm est particulièrement complexe car les différentes uORF agissent de façon synergique ou antagoniste (Hinnebusch 2005; Young and Wek 2016).

Enfin, il est à noter que l'initiation des uORF peut se faire sur un codon non-AUG comme c'est le cas de ceux retrouvés au sein des ARNm codant pour AZIN1 et BIP qui débutent toutes deux par le codon d'initiation non canonique AUU (Ivanov et al. 2008; Starck et al. 2016).

Le mécanisme de traduction des uORF sera décrit en détail à la fin de ce chapitre.

3 Les facteurs d'initiation de la traduction

Dans le modèle de la traduction dépendante de la coiffe, la première étape est l'association du complexe hétérotrimérique eIF4F au niveau de la coiffe sur la partie 5' terminale de l'ARNm. Celui-ci est composé de trois facteurs d'initiation : eIF4E, qui est la protéine de liaison à la coiffe ; eIF4G, qui est une protéine d'échafaudage ; et eIF4A, qui grâce à son activité hélicase déroule les structures secondaires présentes dans la région 5'UTR. Le recrutement du complexe eIF4F sur l'ARNm est accompagné de protéines annexes telle que eIF4B et eIF4H qui renforcent l'activité de eIF4A. La formation du complexe eIF4F s'accompagne également de la pseudo-circularisation du transcrit grâce à la protéine PABP fixée au niveau de la queue poly (A) et qui va interagir avec la 5'UTR. Nous allons brièvement décrire les caractéristiques de chacun de ces facteurs d'initiation dans la partie ci-dessous.

3.1 eIF4E

C'est le premier acteur de l'initiation de la traduction qui s'attache sur l'extrémité 5' coiffée de l'ARNm (Sonenberg et al. 1978). Au niveau moléculaire, eIF4E possède une poche formée par les chaînes latérales de deux tryptophanes (W56 and W102) dans laquelle vient se loger la guanosine méthylée de la coiffe (Figure 3) (Marcotrigiano et al. 1997; Matsuo et al. 1997). Les résidus impliqués dans l'interaction sont très conservés au cours de l'évolution depuis la levure jusqu'à l'humain (Altmann et al. 1987; Matsuo et al. 1997), soulignant la fonction importante de cette interaction et son caractère essentiel pour la viabilité cellulaire puisque son KO est létal (Truitt et al. 2015). L'interaction coiffe-eIF4E est fortement réduite par la présence de structures secondaires d'ARNm localisées immédiatement en aval de la coiffe (entre 1 et 10 nucléotides) (Pelletier and Sonenberg 1985b; Parkin et al. 1988).

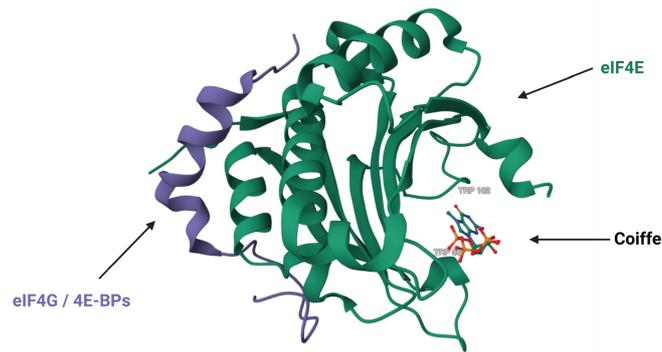


Figure 3. – Structure du facteur d’initiation de la traduction eIF4E

Le facteur d’initiation eIF4E se lie à la coiffe notamment grâce aux tryptophanes (56 et 102) localisées dans sa cavité. eIF4E possède un site mutuel de liaison à eIF4G et 4EBP sur sa face opposée.

La protéine eIF4E interagit avec la protéine d’échafaudage eIF4G (Grüner et al. 2016) et cela repose sur une interaction de sa face dorsale avec un domaine localisé dans la partie amino-terminale d’eIF4G (Mader et al. 1995; Marcotrigiano et al. 1999; Grüner et al. 2016). L’association avec eIF4G renforce considérablement l’affinité d’eIF4E pour la coiffe qui est intrinsèquement modeste (von Der Haar et al. 2000).

3.2 eIF4G

La protéine eIF4G joue le rôle de plateforme d’assemblage pour les différents facteurs d’initiation du complexe eIF4F sur la sous-unité ribosomique 40S (Gingras et al. 1999). Il existe chez les mammifères trois isoformes de eIF4G (Figure 4). Les isoformes eIF4G1 et eIF4G3 partagent 46% d’identité de séquence et ont toutes deux des sites de liaison à eIF4E dans leur région N-terminale, deux sites de liaison à eIF4A, l’un dans leur région centrale et l’autre dans leur région C-terminale (Imataka and Sonenberg 1997), un domaine de liaison aux sous-unités c-d-e d’eIF3 dans leur domaine central (Imataka and Sonenberg 1997; Villa et al. 2013) et un domaine de liaison à PABP dans leur région N-terminale (Gradi et al. 1998a). À l’inverse, l’isoforme eIFG2, ne possède pas le site de liaison à eIF4E et PABP (Imataka et al. 1997; Levy-Strumpf et al. 1997) et elle n’est pas impliquée dans la traduction dépendante de la coiffe (Lieberman et al. 2015); cette isoforme ne sera pas décrite davantage dans ce chapitre.

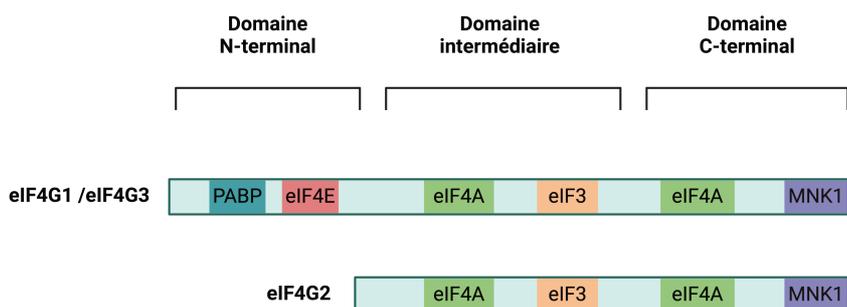


Figure 4. – Représentation schématique des sites de liaison d'eIF4G

Le facteur d'initiation de la traduction, eIF4G est un composant clé du complexe d'initiation de la traduction eIF4F auquel il sert de plateforme d'assemblage. Chez les eucaryotes, eIF4G se décline en 3 isoformes. eIF4G1 et eIF4G3 qui sont les isoformes les plus courantes et possèdent des sites de liaison aux autres facteurs d'initiation de la traduction (PABP, eIF4E, eIF3, eIF4A) ainsi qu'à la kinase Mnk1. eIF4G 2 quant à elle ne possède pas les sites de liaison à PABP et eIF4E. Dessinée sur BioRender.

En plus de son activité d'échafaudage dans l'assemblage du complexe, eIF4G joue également un rôle dans la modulation de l'activité du complexe eIF4F. D'une part, l'interaction eIF4G/eIF4A renforce l'activité ATPase et hélicase de cette dernière (Rogers et al. 2001; Korneeva et al. 2005) et d'autre part, eIF4G sert de plateforme de liaison aux kinases Mnk1 et Mnk2 impliquées dans la phosphorylation de eIF4E (Pyronnet et al. 1999; Waskiewicz et al. 1999). Le mécanisme de cette régulation sera détaillé dans la suite du manuscrit.

3.3 eIF4A

eIF4A appartient à la grande famille des hélicases à ARN, ATP-dépendantes, qui contiennent le motif DEAD (pour D : acide aspartique, E : acide glutamique A : alanine et D: acide aspartique) (Linder et al. 1989; Linder and Jankowsky 2011; Weis and Hondele 2022). Cette famille de protéines riche de plus de 500 membres participe à de nombreux aspects du métabolisme des ARN (Rocak and Linder 2004). eIF4A possède une activité ATPase (Grifo et al. 1984) et hélicase bidirectionnelle (5' vers 3' et 3' vers 5') *in vitro* (Ray et al. 1985; Rozen et al. 1990). Cependant, l'association de eIF4A au sein du complexe eIF4F restreint son activité enzymatique avec une directionnalité 5' vers 3' *in vivo* ce qui lui permet d'accompagner le ribosome dans l'étape de balayage (Rozen et al. 1990) en déroulant les structures secondaires pour

faciliter sa progression vers le codon AUG. eIF4A est une enzyme qui reste faiblement processive seule (García-García et al. 2015) mais dont l'activité est fortement augmentée lors de son association avec les facteurs d'initiation de la traduction eIF4B/eIF4H et eIF4G (Rogers et al. 2001; Marintchev et al. 2009).

Chez les mammifères, il existe deux isoformes principales d'eIF4A : eIF4AI et eIF4AII, qui partageant environ 90% d'identité de séquence chez l'humain. Elles présentent une activité similaire *in vitro* et sont interchangeable au sein du complexe eIF4F (Yoder-Hill et al. 1993). Cependant, il a été montré que eIF4AI est indispensable pour la survie alors que le KO de eIF4AII n'est pas létal (Galicia-Vázquez et al. 2012). De plus, la déplétion de eIF4AI s'accompagne d'une augmentation du niveau d'expression de eIF4AII, mais sans que cela ne compense totalement la perte de eIF4AI au niveau fonctionnel (Galicia-Vázquez et al. 2012). Enfin, eIF4AI est davantage exprimé dans les cellules en croissance alors qu'on retrouve eIF4AII principalement dans les cellules quiescentes (Marintchev et al. 2009). Ces différents éléments suggèrent que chacune des isoformes pourrait avoir des fonctions uniques et non redondantes bien qu'elles n'aient pas encore été clairement identifiées.

Le facteur eIF4A est le plus abondant de tous les facteurs d'initiation en termes de quantité et disponibilité cellulaire (Duncan and Hershey 1983; Kulak et al. 2014). Cela s'explique par le fait que eIF4A s'associe et se dissocie constamment du complexe eIF4F à chaque fois qu'il déroule les structures secondaires en consommant une molécule d'ATP. De fait, plus un ARNm est structuré, plus il aura besoin du facteur eIF4A pour être bien traduit. (Marintchev et al. 2009). Ainsi, la traduction d'ARNm avec de courtes séquences 5'UTR non structurées comme les TISU (Translation Initiation of Short 5'UTR) est indépendante d'eIF4A (Elfakess et al. 2011). À l'opposé l'insertion de motifs ARN stables dans la région 5'UTR augmente la dépendance à eIF4A (Pestova and Kolupaeva 2002).

Enfin, il est à noter qu'une troisième isoforme eIF4AIII a été décrite (65% d'homologie de séquence avec eIF4AI); celle-ci n'a pas de rôle au niveau de l'initiation de la traduction mais elle exerce une fonction essentielle dans la voie du *Non Sense Mediated Decay* (NMD) (Li et al. 1999; Shibuya et al. 2004).

3.4 eIF3

Le facteur d'initiation eIF3 joue un rôle très important tout au long du processus d'initiation de la traduction (Trachsel and Staehelin 1979; Hinnebusch 2006). Il est de loin le facteur d'initiation le plus complexe avec une organisation multimérique composée de 13 sous-unités différentes (Figure 5), atteignant un poids moléculaire total de plus de 800 kDa. Sa structure moléculaire repose sur un cœur protéique très structuré formé par l'assemblage des sous-unités eIF3a, eIF3c, eIF3e, eIF3k, eIF3l, eIF3m, eIF3f, eIF3h (Querol-Audi et al. 2013; des Georges et al. 2015; Smith et al. 2016). Les autres sous-unités (eIF3b, eIF3d, eIF3g, eIF3i, eIF3j) se situent à la périphérie de ce cœur et présentent une organisation plus dynamique et versatile (des Georges et al. 2015). Chez la levure, on ne retrouve que 6 sous-unités au total mais qui sont toutes essentielles (Asano et al. 1998; Phan et al. 1998). Plusieurs sous-unités d'eIF3 interagissent avec le ribosome ainsi qu'avec les facteurs d'initiation de la traduction. Les sous-unités eIF3a, eIF3c et eIF3h se lient directement sur la sous-unité ribosomique 40S (Valásek et al. 2003; Erzberger et al. 2014; des Georges et al. 2015; Brito Querido et al. 2024a). La cohésion du complexe est renforcée par une interaction entre les sous-unités eIF3a-d et l'ARN ribosomal 18S (Nygård and Nika 1982; Valásek et al. 2003). En outre, eIF3 interagit avec le facteur eIF4G par l'intermédiaire des sous-unités eIF3e, eIF3g, eIF3d, et eIF3e (Korneeva et al. 2000; LeFebvre et al. 2006; Villa et al. 2013). Enfin, les sous-unités eIF3a, eIF3b et eIF3c interagissent avec le complexe ternaire par association avec eIF2 (Phan et al. 2001; Valásek et al. 2002). Nous verrons dans la suite de ce chapitre qu'eIF3 intervient aussi au cours de l'étape de balayage de la région 5'UTR (Pestova and Kolupaeva 2002; Valásek et al. 2004; Damoc et al. 2007).

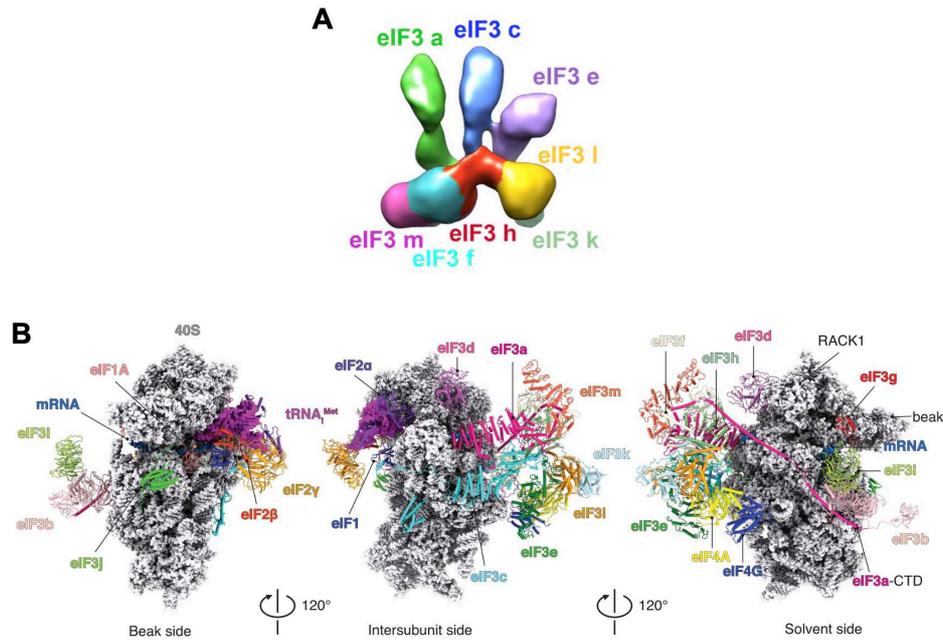


Figure 5. – Structure du facteur d’initiation de la traduction eIF3

eIF3 facilite la liaison de la petite sous-unité ribosomique 40S à l'ARNm et régule l'assemblage du complexe d'initiation, jouant un rôle critique dans la sélection du site d'initiation de la traduction. (A) Structure du cœur protéique de eIF3 libre. Structure Cryo-EM du facteur eIF3 associé à la sous unité ribosomique 40S.

3.5 eIF2

Le facteur eIF2 a une structure relativement flexible (Stolboushkina et al. 2008) et tient une place centrale dans la formation du complexe ternaire. Dans son ensemble, il promeut l’interaction entre l’ARNt initiateur et la petite sous-unité ribosomique (Hashem and Frank 2018). L’interaction entre eIF2 et l’ARNt initiateur est contrôlée par la liaison à une molécule de GTP qui augmente considérablement l’interaction entre les deux (Kapp and Lorsch 2004). D’un point de vue structural, eIF2 est composé de trois sous-unités eIF2α, eIF2β, et eIF2γ, c’est cette dernière qui est responsable de la liaison entre le GTP et l’ARNt initiateur et elle possède une activité GTPase (Erickson and Hannig 1996). Cependant, des données plus récentes indiquent que la sous-unité α interagit également avec l’ARNt initiateur (Schmitt et al. 2012) et avec la sous-unité 40S une fois le complexe de préinitiation formé (Hashem et al. 2013a). En outre, sous sa forme phosphorylée, eIF2α participe à la régulation du recyclage d’eIF2γ-GDP, en modulant son association avec le facteur d’échange nucléotidique eIF2B (Rowlands et al. 1988; Pavitt et al. 1998). La sous-unité eIF2β est moins

caractérisée et semble surtout avoir un rôle de cohésion au sein du complexe ternaire notamment en facilitant l'interaction entre eIF2 α et eIF2 γ . Plusieurs études rapportent que eIF2 β interagit avec eIF5 et eIF1 pendant le processus de balayage de la région 5'UTR et participe à la reconnaissance et la sélection du codon d'initiation (Asano et al. 1999; Jennings et al. 2016; Thakur et al. 2019).

3.6 eIF4B/eIF4H

Les facteurs d'initiation eIF4B et eIF4H interagissent fonctionnellement avec le complexe eIF4F pour stimuler les activités ATPase et hélicase portées par eIF4A (Rozen et al. 1990; Altmann et al. 1993; Rozovsky et al. 2008). C'est également le cas avec eIF4G, ce qui permet aussi d'augmenter l'activité enzymatique de eIF4A de plus de 100 fois (Rogers et al. 2001; Özeş et al. 2011).

Le facteur d'initiation eIF4B n'est pas essentiel à l'initiation de la traduction et n'est pas fortement conservé chez les eucaryotes. Il joue cependant un rôle important lors de la traduction d'ARNm spécifiques qui présentent des régions 5'UTR longues et structurées chez les mammifères (Sen et al. 2016). Il possède un site de liaison à eIF3 (Methot et al. 1996) ainsi que deux sites de liaison à l'ARN 18S (Méthot et al. 1994; Naranda et al. 1994). Ces différents sites lui permettent de favoriser le recrutement du complexe de préinitiation 43S (Pestova et al. 1996a). eIF4B possède un homologue fonctionnel appelé eIF4H avec lequel il partage environ 40% d'homologie de séquence. Cependant, eIF4H ne possède qu'un des deux domaines de liaison à l'ARN suggérant qu'il ne remplit que partiellement les fonctions de eIF4B. Néanmoins, tout comme eIF4B, eIF4H se lie à eIF4A1 et stimule son activité enzymatique (Richter-Cook et al. 1998).

3.7 PABP

Les *poly-A-binding-proteins* (PABP) (Figure 6) sont une famille de protéines qui se lie sur la queue poly(A) des ARNm eucaryotes. Elles interviennent dans la régulation post-transcriptionnelle de l'expression des ARNm, notamment au cours du développement embryonnaire (Bernstein et al. 1989). Chez les mammifères, on retrouve six isoformes de PABP (Mangus et al. 2003), certaines ayant une localisation

cytoplasmique telle que PABPC1 alors que d'autres ont une localisation nucléaire telle que PABPN1. La protéine PABPC1 est exprimée de façon ubiquitaire (Görlach et al. 1994) et possède 4 domaines de liaison à l'ARN dans sa région C-terminale ce qui lui permet d'interagir avec la queue poly(A) (Burd et al. 1991). Plusieurs molécules de PABP se fixent sur le même ARNm, chacune recouvrant de 25 à 27 nucléotides sur le transcrit (Baer and Kornberg 1983). Les protéines PABP et leurs rôles dans le métabolisme de l'ARNm sont décrites de façon exhaustive dans les revues suivantes (Mangus et al. 2003; Gorgoni and Gray 2004).

Cependant, depuis 1996 on sait que la protéine PABP joue aussi un rôle crucial au niveau de l'initiation de la traduction en interagissant avec le complexe eIF4F pour induire la pseudo-circularisation du messager. Ainsi l'interaction entre PABPC1 avec le facteur eIF4F augmente l'affinité de liaison de eIF4E pour la coiffe, renforçant ainsi l'efficacité de traduction (Haghighat and Sonenberg 1997; Borman et al. 2000).

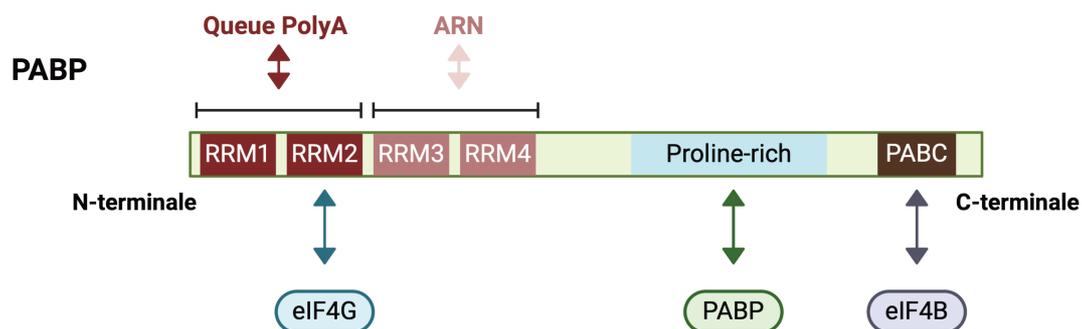


Figure 6. – Structure et fonction de PABP

PABP possède quatre domaines de liaison à l'ARN (RRM). Les domaines RRM1 et RRM2 interagissent exclusivement avec la queue poly(A), tandis que les RRM3 et RRM4 interagissent de façon aspécifique avec les ARN. PABP possède dans son domaine N-terminal un site de liaison pour eIF4G et dans son domaine C-terminal, un site de liaison à eIF4B. Enfin, PABP dispose également d'un domaine proline-riche dans sa partie centrale, qui lui permet de s'homodimériser. Dessinée sur BioRender.

Grâce à son interaction simultanée avec la queue poly(A) des ARNm et le facteur d'initiation eIF4G, PABP favorise le rapprochement des extrémités 5' et 3' des ARNm, qui adoptent ainsi une conformation circulaire (Imataka et al. 1998; Wells et al. 1998; Amrani et al. 2008; Kopeina et al. 2008). Celle-ci est favorable au processus de traduction sans que le mécanisme moléculaire n'ait été complètement élucidé et il

existe actuellement deux hypothèses principales qui peuvent être complémentaires. La première suggère que la circularisation stabilise la machinerie de traduction sur l'ARNm augmentant ainsi son efficacité. La seconde hypothèse propose que le rapprochement entre la partie terminale de l'ARNm et la coiffe pourraient faciliter le recyclage des ribosomes pour les rendre disponibles lors du prochain cycle de traduction (Mangus et al. 2003). Il a également été évoqué que la forme circulaire de l'ARN le rendait moins vulnérable aux *exo-ribonucléases* 5' et 3' (Imataka et al. 1998). Ces hypothèses sont supportées par des expériences qui mettent en avant une synergie fonctionnelle entre la coiffe 5' et la queue poly(A), mécanisme qui avait été mis en évidence initialement chez les plantes (Munroe and Jacobson 1990; Gallie 1991). Il faut toutefois noter que, chez les mammifères, la circularisation des ARNm est difficilement observable par fluorescence *in situ* à l'exception de certains contextes particuliers tels que la localisation des ARNm dans des granules de stress (Adivarahan et al. 2018; Khong and Parker 2018). Cela montre qu'il reste des zones d'ombre sur ce mécanisme et que l'activité de PABP pourrait être multifactorielle comme le démontre l'ajout de séquences poly(A) libres, *en trans*, qui favorise l'association de eIF4E avec la coiffe *in vitro* sans pour autant induire la circularisation de l'ARNm (Borman et al. 2002). Néanmoins, il ne fait pas de doute que PABP est une protéine indispensable pour assurer une traduction efficace et qu'elle est, à ce titre, considérée comme un facteur d'initiation à part entière (Kahvejian et al. 2005).

4 Le mécanisme canonique de traduction

Chez les eucaryotes, la sous-unité 40 S seule ne peut pas spontanément se fixer sur l'ARNm pour initier la traduction de manière autonome. Elle doit au préalable, s'associer avec différents facteurs eIF et le complexe ternaire incluant l'ARNt initiateur pour former le complexe de préinitiation 43S. C'est celui-ci qui s'attache sur l'extrémité 5' de l'ARNm pour former le complexe 48S qui balaye ensuite la région 5'UTR à la recherche du codon initiateur. Lors de la formation du complexe 43S, la petite sous-unité du ribosome va adopter une conformation ouverte propice pour accepter l'ARNt initiateur et permettre le déplacement du ribosome sur la 5'UTR de l'ARNm (un schéma bilan est proposé à la figure 7).

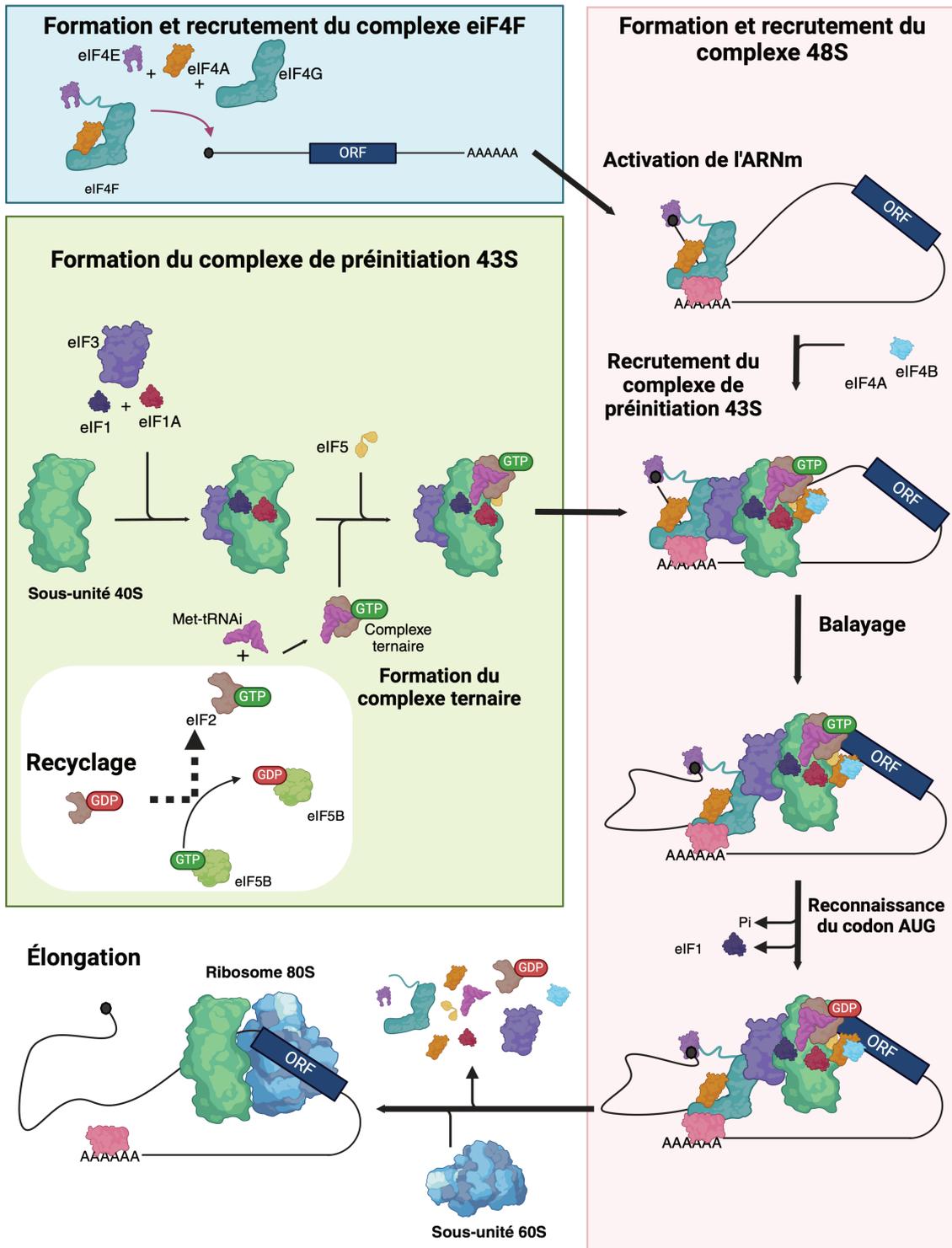


Figure 7. – Schéma bilan de l'initiation canonique de la traduction

L'initiation de la traduction commence avec l'assemblage simultané du complexe eIF4F et du complexe de préinitiation 43S. Le complexe eIF4F est recruté à l'extrémité 5' des ARNm par une interaction entre eIF4E et la coiffe. Cela permet l'activation de l'ARNm. Ensuite, le complexe eIF4F attire le complexe de préinitiation 43S grâce à une interaction entre eIF4G et eIF3. Le complexe 43S balaye ensuite la région 5'UTR jusqu'à ce qu'il trouve un codon AUG dans un contexte optimal. Cette reconnaissance du codon AUG entraîne la libération des facteurs d'initiation et le recrutement de la sous-unité 60S, permettant ainsi le début de l'élongation. Dessinée sur BioRender.

4.1 La formation du complexe ternaire: eIF2•GTP•Met-tRNA_i

Pour se fixer sur le ribosome, l'ARNt initiateur s'associe avec le facteur d'initiation eIF2, qui est lui-même lié à une molécule de GTP (Figure 8), formant ainsi le complexe ternaire qui va permettre de localiser et reconnaître le codon AUG de départ.

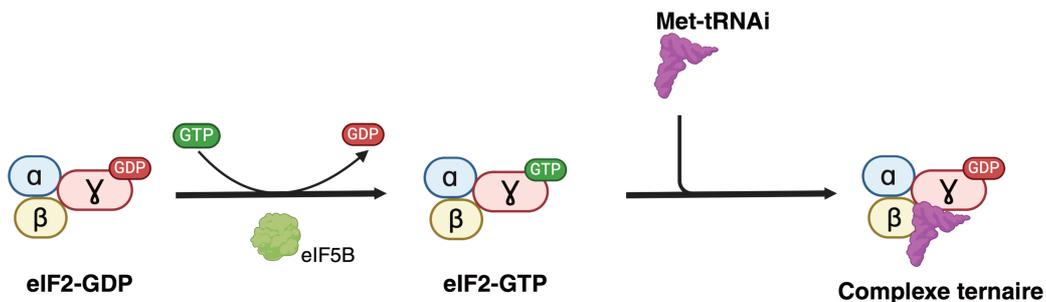


Figure 8. – Formation du complexe ternaire

Le complexe ternaire est formé par l'association de l'ARN de transfert (ARNt initiateur) avec le facteur d'initiation eIF2 (composé de 3 sous-unités α , β , γ) et une molécule de GTP. La protéine eIF5B catalyse cette réaction. Dessinée sur BioRender.

L'ARNt initiateur reconnaît spécifiquement le codon AUG de démarrage de la traduction car il diffère de l'ARNt méthionine d'élongation classique. La particularité de l'ARNt méthionine initiateur est de pouvoir entrer directement dans le site P du ribosome au lieu du site A (Kolitz and Lorsch 2010) et il contient plusieurs bases modifiées dans les boucles T et D qui forment une structure spécifique qui empêche la liaison au facteur d'élongation eEF1A, réservée aux ARNt d'élongation (Basavappa and Sigler 1991). Il comporte aussi un appariement de base spécifique (A1 : U72) au sommet de la tige acceptrice qui permet de le différencier des ARNt élongation (Drabkin et al. 1993) et de s'associer avec eIF2 (Farruggio et al. 1996).

4.2 Liaison du complexe ternaire au ribosome et formation du 43 S

Le complexe ternaire eIF2-GTP-Met-tRNA_i se lie à la sous-unité 40S du ribosome avec l'aide de 3 facteurs: eIF3, eIF1 et eIF1A ce qui forme le complexe 43 S (Figure 9). La formation et la cohésion de ce complexe repose sur eIF3 qui fait le pont entre la sous-unité 40S, le complexe ternaire et le facteur eIF4F ce qui permet de positionner le complexe de préinitiation 43S sur l'extrémité de l'ARNm (Korneeva et al. 2000; Villa et

al. 2013). En agissant de concert avec eIF1 et eIF1A, eIF3 favorise le recrutement du complexe ternaire en stabilisant l'interaction d'eIF2 α avec la sous-unité 40S (Chaudhuri et al. 1999; des Georges et al. 2015) (Chaudhuri et al. 1999).

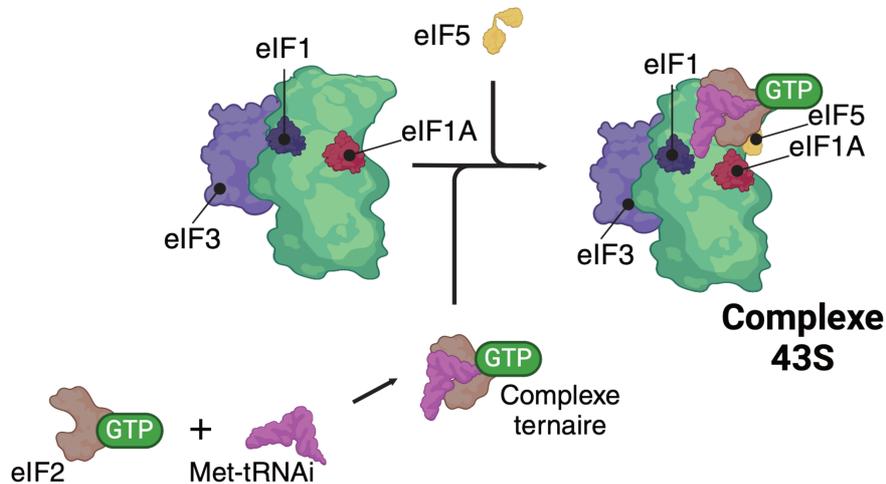


Figure 9. – Assemblage du complexe de préinitiation 43S

Le ribosome associé à eIF1, eIF1A et eIF3 recrute successivement le complexe ternaire et eIF5 pour former le complexe de préinitiation 43S. Celui-ci pourra par la suite être recruté sur l'ARNm. Dessinée sur BioRender.

Les facteurs eIF1 et eIF1A interviennent dans le recrutement du complexe ternaire sur la petite sous-unité ribosomique 40S et des expériences chez la levure montrent que la délétion du domaine C-terminal ou la mutation ponctuelle G1017K d'eIF1 suffisait pour diminuer ce recrutement (Fekete et al. 2005; Nanda et al. 2009). Des expériences *in vitro* montrent que la présence d'eIF3 était requise pour former un complexe de préinitiation stable (Chaudhuri et al. 1999). Ces différentes observations sont renforcées par des données de biologie structurale plus récentes, présentant la structure de la sous-unité 40S associée à eIF1 et eIF1A. Ainsi, en présence d'eIF3 les protéines eIF1 et eIF1A sont associées à proximité du site P (Rabl et al. 2011; Weisser et al. 2013) ce qui induit un changement conformationnel des régions proches du site A, et à proximité immédiate du site de liaison du complexe ternaire (Passmore et al. 2007; Weisser et al. 2013). L'ensemble de ces résultats suggère que eIF1 et eIF1A agissent de concert pour assurer le changement conformationnel de la sous-unité 40S ce qui permet le recrutement et la stabilisation du complexe ternaire en présence d'eIF3.

4.3 Liaison du complexe 43S sur l'ARNm: formation du complexe 48S

Le mécanisme moléculaire précis de recrutement du complexe de préinitiation 43S sur l'ARNm comporte encore des zones d'ombre. Il est toutefois bien admis que c'est l'interaction entre eIF4E et la coiffe qui permet d'associer le complexe à l'extrémité 5' de l'ARNm (Figure 10).

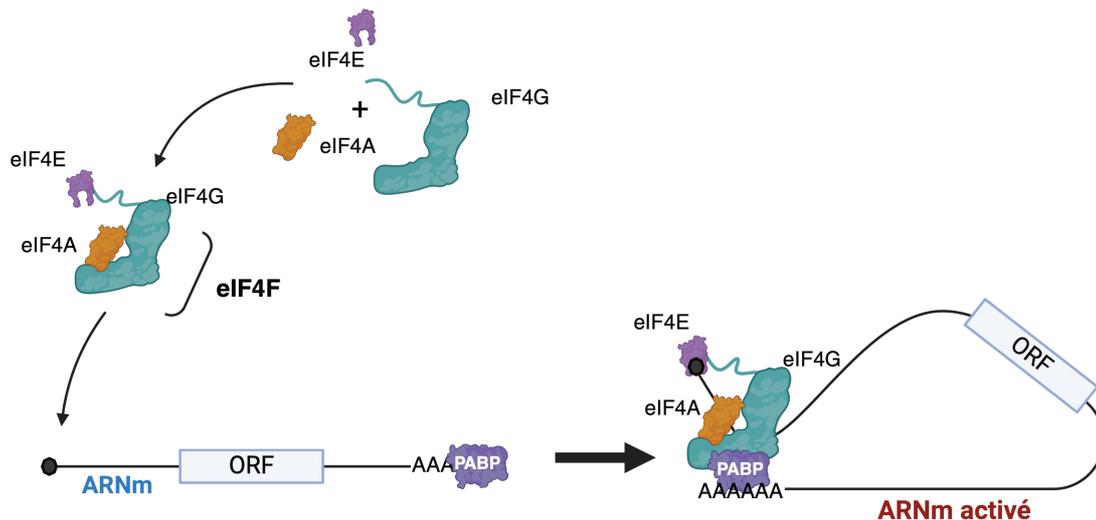


Figure 10. – Activation de l'ARNm

Les facteurs d'initiation eIF4E, eIF4G et eIF4A s'assemblent pour former le complexe eIF4F. Celui-ci est ensuite recruté à l'extrémité 5' de l'ARNm grâce à l'activité de liaison de la coiffe par eIF4E. PABP, en interagissant à la fois avec la queue poly(A) et avec eIF4G, permet la circularisation des ARNm, événement qui est proposé pour stabiliser le complexe et stimuler la traduction. Dessinée sur BioRender.

Le modèle actuel basé sur des observations biophysiques propose que eIF4F est déjà préformé avant de se fixer sur la coiffe (Niedzwiecka et al. 2002; Slepnev et al. 2008; O'Leary et al. 2013; Izidoro et al. 2022). Au sein de eIF4F, eIF4G interagit avec PABP afin de circulariser l'ARNm et eIF4A déroule les structures secondaires immédiatement en aval de la coiffe pour libérer l'espace nécessaire à la fixation de la petite sous-unité ribosomique 40S. C'est ensuite l'interaction entre eIF4G et eIF3 qui ancre le complexe de préinitiation 43S sur l'ARNm (Brito Querido et al. 2024a) (Figure 11)

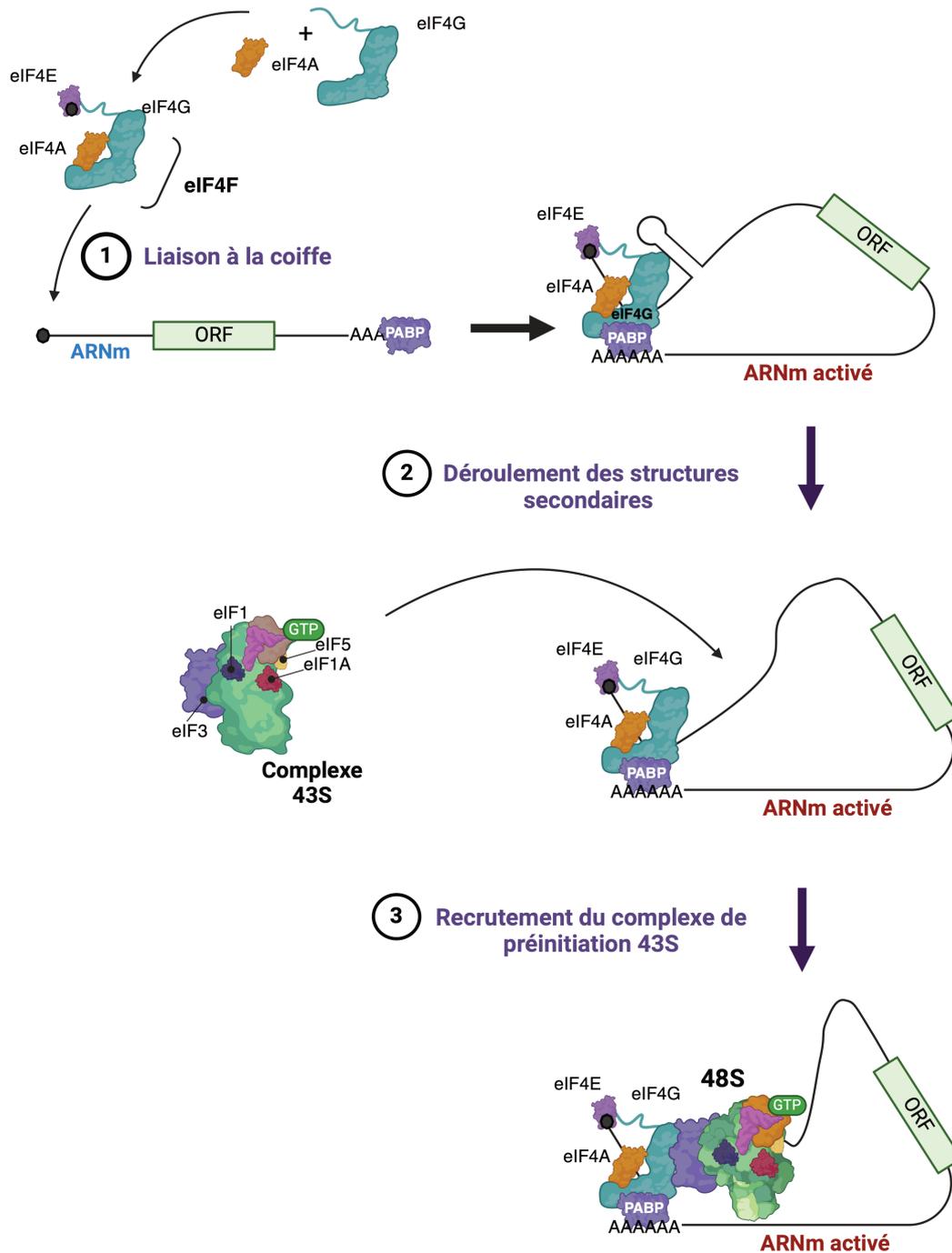


Figure 11. – Recrutement du complexe 43S sur l'ARNm

Une fois formé, le complexe eIF4F se lie à l'ARNm au niveau de la coiffe. L'activité hélicase de eIF4A permet de dénouer les structures secondaires afin de libérer un espace d'atterrissage pour le complexe de préinitiation 43S. Celui-ci est recruté grâce à l'interaction entre eIF3 et eIF4G. Dessinée sur BioRender.

Il existe deux modèles principaux qui résument la liaison du complexe 43S à l'ARNm pour former le complexe 48S (Figure 12). Le premier dit du « *slotting* » propose que l'ARNm est directement inséré dans le site de décodage de la sous-unité ribosomique

de manière latérale. Dans ce modèle, le complexe eIF4F est en amont du complexe 43S au niveau de la sortie du canal de l'ARNm de la sous-unité 40S. Dans cette configuration, l'encombrement stérique généré par la taille du complexe implique que le processus de balayage ne débute pas immédiatement à la coiffe mais seulement quelques dizaines de nucléotides en aval de celle-ci (Brito Querido et al. 2020).

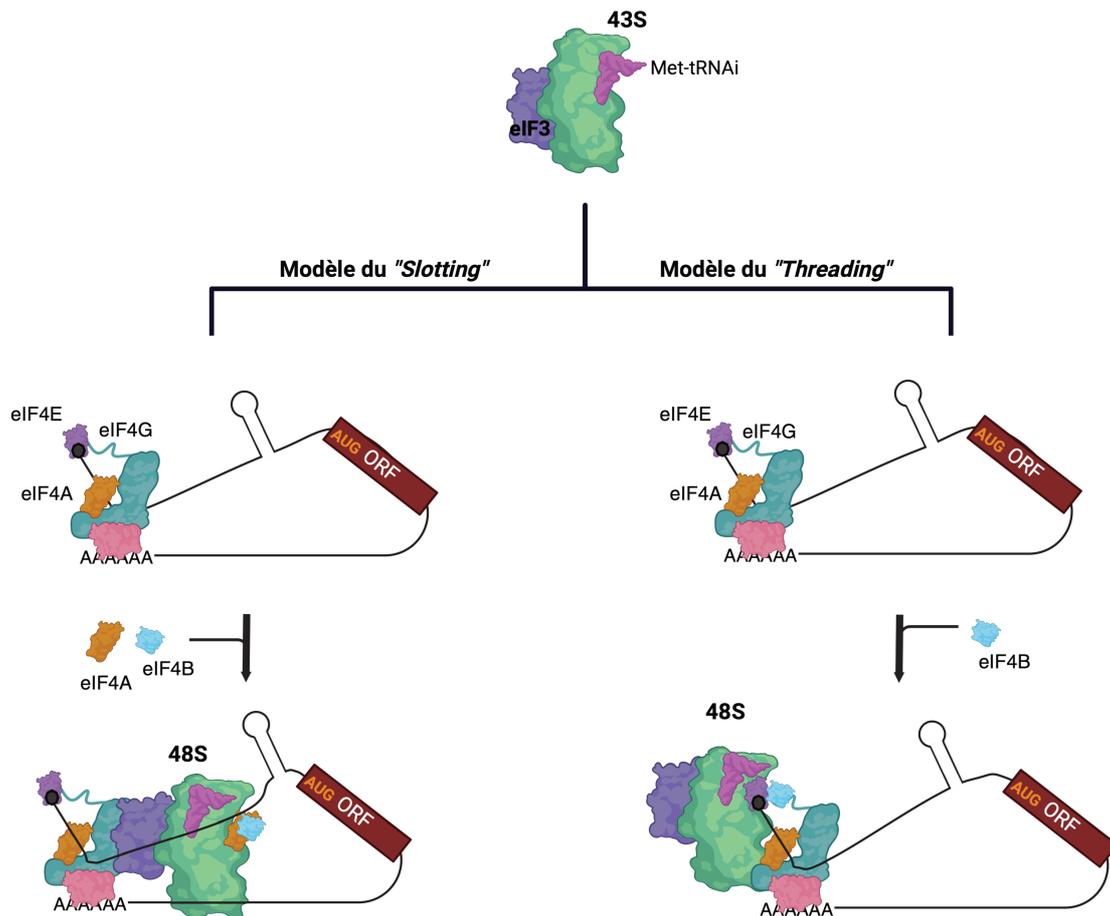


Figure 12. – Modèles de recrutement du complexe 43S sur l'ARNm

Deux modèles de recrutement de la sous unité 43 S sur l'ARN ont été proposés. Dans le modèle du « *slotting* » l'ARNm est directement inséré dans le site de décodage de la sous-unité ribosomique de manière latérale. Dans cette configuration le complexe eIF4F se trouve en amont du complexe 43S et une seconde protéine eIF4A s'associe à l'opposé de celui-ci. Dans le modèle du « *threading* » l'extrémité de l'ARNm est insérée dans le canal du ribosome, cela suggère une dissociation du complexe eIF4F de l'extrémité de l'ARN. Dans ce modèle ce complexe eIF4F est localisé en aval du complexe 43S. Dessinée sur BioRender.

Le second modèle dit du « *threading* » propose un mécanisme de fixation d'eIF4F en aval du ribosome avec une inspection de l'ARNm dès les premiers nucléotides par le site de décodage du ribosome (Kumar et al. 2016). Ce second modèle laisse penser toutefois un détachement, *a minima* temporaire, du complexe eIF4F de l'extrémité 5'

de l'ARNm. En revanche, cela est compatible avec l'idée d'un processus de balayage débutant dès le tout premier nucléotide en aval de la coiffe (Brito Querido et al. 2024a). Le modèle du *threading* est séduisant puisqu'il permet d'expliquer la traduction des ARNm médiée par des 5'UTR très courtes telles que les séquences TISU (Translation Initiator of Short 5' UTR) dont la taille peut être de 5 nt seulement (Elfakess and Dikstein 2008) (Elfakess et al. 2011). Cela suppose un balayage très court de la 5'UTR et la reconnaissance d'un codon d'initiation se trouvant à proximité de la coiffe ce qui paraît peu compatible avec le modèle du « *slotting* ». (Elfakess and Dikstein 2008; Elfakess et al. 2011). Les données structurales récentes du complexe 48S, par Cryo-EM avec une résolution de $\sim 3\text{\AA}$, suggèrent un positionnement du facteur eIF4F en amont du complexe de préinitiation 48S (Brito Querido et al. 2020; Brito Querido et al. 2024a) . Cependant, les travaux précurseurs du Dr Marilyn Kozak sur l'étape de balayage des 5'UTR, suggéraient déjà que les codons AUG proximaux étaient peu reconnus s'ils étaient trop proches de la coiffe (Kozak 1989b), renforçant l'hypothèse du modèle de « *slotting* » où les premiers nucléotides ne seraient pas inspectés par le complexe de préinitiation 43S. Le modèle du « *slotting* » est également compatible avec le mécanisme d'initiation de la traduction dépendant des sites d'entrée interne du ribosome, où le recrutement du ribosome se fait à distance de l'extrémité 5' des ARN (voir chapitre 2 section 1.4).

4.4 Le balayage de la région 5'UTR

Marilyn Kozak a défini le modèle du balayage comme un déplacement linéaire le long de l'ARNm dans la direction 5'-3'. Cela reposait sur le fait que l'insertion d'un codon AUG dans un bon contexte nucléotidique en amont du codon AUG principal induisait une diminution de l'initiation de la traduction de l'ORF principale tout comme l'insertion de structures secondaires stables au sein de la 5'UTR qui réduisait considérablement l'efficacité de traduction (Kozak 1986; Kozak 1987).

40 ans après, ce mécanisme est toujours valide mais des données récentes permettent d'apporter de nouvelles perspectives quant aux rôles précis des facteurs d'initiation et de suggérer des alternatives à ce modèle.

Les facteurs d'initiation requis pour l'étape de balayage et de reconnaissance du codon AUG sont maintenant bien identifiés et leurs fonctions caractérisées; il s'agit des

facteurs d'initiation eIF1, eIF1A, eIF3, eIF4F, eIF4B/eIF4H et eIF2 (Pestova and Kolupaeva 2002). Après attachement du ribosome au niveau de la coiffe, celui-ci va migrer en direction 5'-3' et il reste associé à eIF4G et eIF4A qui vont assurer, avec eIF4B/eIF4H le déroulement processif des structures secondaires rencontrées dans la 5'UTR. Le binôme eIF1-eIF1A est requis à la fois pour le balayage et pour la reconnaissance du codon AUG (Pestova and Kolupaeva 2002). *In vitro*, la déplétion d'eIF1 ou d'eIF1A provoque le blocage du complexe 48S à l'extrémité 5' de l'ARNm (Pestova et al. 1998a) et, chez la levure, l'introduction d'une mutation ponctuelle à la position G107 de eIF1 réduit son affinité pour le complexe 48S et inhibe le processus de balayage (Nanda et al. 2009). À l'inverse, une surexpression de eIF1 renforce l'interaction de ce dernier avec le complexe 48S, stimule l'étape de balayage et augmente la fidélité d'initiation au codon AUG (Martin-Marcos et al. 2014). De manière fonctionnelle, le rôle de eIF1 serait d'empêcher la reconnaissance d'un codon AUG cryptique en inhibant l'hydrolyse du GTP d'eIF2 en GDP, étape qui marque la fin du processus d'initiation et le recrutement de la sous unité 60S (Algire et al. 2005; Martin-Marcos et al. 2014; Llácer et al. 2018). Ainsi, eIF1 agirait comme un facteur de contrôle pour reconnaître le site d'initiation; ce rôle est co-catalysé par le facteur eIF1A qui agit de concert dans cette fonction.

Enfin, il est à noter qu'au-delà des facteurs canoniques décrits ci-dessus, certains ARNm spécifiques très structurés (viraux et cellulaires) requièrent l'intervention de protéines à activité hélicase, notamment DHX29 et DDX3, pour assurer le trajet du ribosome (Pisareva et al. 2008; Soto-Rifo and Ohlmann 2013; Calviello et al. 2021).

a) Mécanisme moléculaire du balayage de la région 5'UTR par le complexe 48S

Il persiste deux questions essentielles concernant le mécanisme de balayage: sa dépendance en une source d'énergie (ATP) et sa directionnalité dans le sens 5'-3'.

En utilisant un ARNm rapporteur dépourvu de structure secondaire et uniquement composé d'une succession de triplets CAA, il a été montré que l'initiation de la traduction pouvait avoir lieu de façon totalement indépendante d'eIF4F et sans consommation d'ATP (Pestova and Kolupaeva 2002). Une expérience complémentaire utilisant un ARNm rapporteur dont la 5'UTR ne contient que des

Adénines a montré des résultats similaires (Shirokikh and Spirin 2008) avec le complexe 48S qui se déplace par diffusion passive et qui peut localiser le codon AUG en absence de eIF4F et eIF3. A l'inverse, un ARNm très structuré nécessite une concentration plus forte d'ATP et donc une forte dépendance dans l'activité d'eIF4A (Svitkin et al. 2001).

L'ensemble de ces données a permis d'établir le modèle du cliquet brownien (*Brownian ratchet*) dans lequel le complexe 48S peut se déplacer de façon passive le long de l'ARNm mais ce déplacement est bi-directionnel (5'-3' et 3'-5'). Dans ce modèle, c'est le complexe eIF4F qui assure la direction du mouvement en empêchant le complexe 48S de reculer et progresser de 5' vers 3' jusqu'au codon AUG (Spirin 2009; Vassilenko et al. 2011; Li et al. 2022). Cependant, des mouvements de recul du complexe 48S (3'-5') persistent malgré la présence du complexe eIF4F (Spirin 2009; Gu et al. 2021; Wang et al. 2022) ce qui provoque des phénomènes d'oscillation qui pourraient jouer un rôle important au niveau de la reconnaissance de codons non canoniques (Wang et al. 2022).

b) Dynamique d'association du complexe eIF4F au cours du balayage

Du fait de la difficulté technique à suivre son déplacement, il n'existe que peu de données factuelles pour établir un modèle précis de la dynamique d'association/dissociation du complexe eIF4F et deux hypothèses s'opposent (Figure 13). Dans le premier cas de figure, le complexe 48S reste lié à la coiffe par l'intermédiaire de la protéine eIF4E et ce serait l'ARNm qui s'engagerait dans le ribosome par un mécanisme de 'threading'. Cette hypothèse n'est pas supportée par les données récentes de ribosome profiling ni par des expériences de cinétique en molécule unique (Wang et al. 2022) qui montrent la présence simultanée de plusieurs complexes 48S sur l'ARNm (Archer et al. 2016).

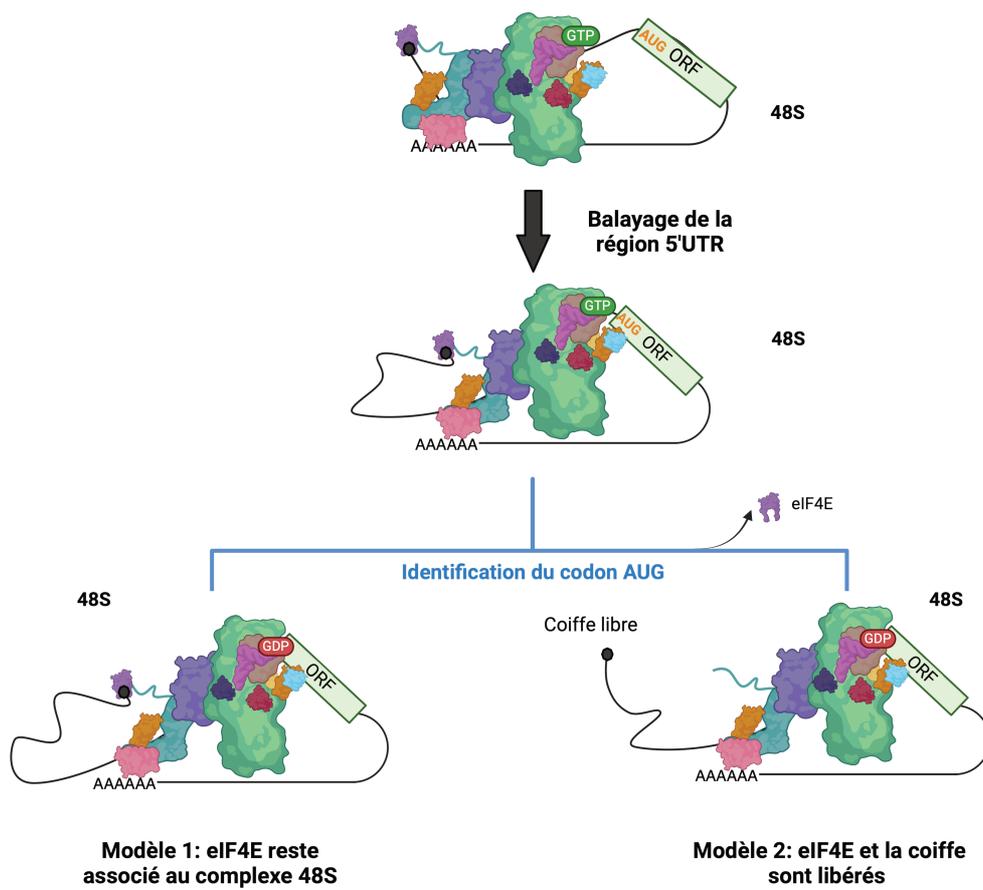


Figure 13. – Association du complexe eIF4F au cours du balayage

Après son recrutement, le complexe de préinitiation parcourt la 5' UTR jusqu'à tomber sur un codon d'initiation. Concernant le devenir du complexe eIF4F, deux modèles sont envisagés. Dans le premier, eIF4E reste lié à la coiffe durant le balayage, impliquant qu'un seul complexe de préinitiation peut scanner la 5'UTR à la fois. Le second modèle suggère que eIF4E se détache de la coiffe dès le début du balayage, permettant potentiellement à plusieurs complexes de préinitiation de scanner simultanément la 5'UTR. Dessinée sur BioRender.

De plus, les travaux de Marilyn Kozak montraient déjà qu'en présence de l'antibiotique edeine qui déstabilise la formation du complexe ternaire, plusieurs sous-unités 40S s'accumulaient sur l'ARNm du virus de la mosaïque du tabac et sur celui de la globine (Kozak and Shatkin 1978). Récemment, la résolution par Cryo-EM de la structure du complexe 48S montre que eIF4G et eIF4A restent au contact du ribosome pendant le balayage ce qui n'est pas le cas de eIF4E suggérant la dissociation partielle du complexe eIF4F mais le maintien des autres facteurs (eIF4G et eIF4A) qui accompagnent la progression du ribosome jusqu'au codon d'initiation.

La résolution du complexe 48S par Cryo-EM nous éclaire aussi sur le fonctionnement de eIF4A au sein du complexe (Figure 14). En effet, deux molécules d'eIF4A sont associées au complexe 48 S, la première se trouve au sein de eIF4F au niveau de l'extrémité 5' de l'ARNm (le modèle du « *slotting* »). La seconde molécule d'eIF4A serait libre et à proximité du site d'entrée de l'ARNm du ribosome en fonctionnant indépendamment du complexe eIF4F (Brito Querido et al. 2024a). C'est totalement en accord avec des données plus anciennes qui montraient qu'eIF4A s'associait et se dissociait du complexe au cours du balayage en consommant de l'ATP (Pause et al. 1994b). Cependant, comme nous l'avons déjà évoqué, l'activité processive d'eIF4A est modeste et il reste à déterminer si le facteur eIF4B pourrait intervenir dans ce mécanisme (Brito Querido et al. 2024b).

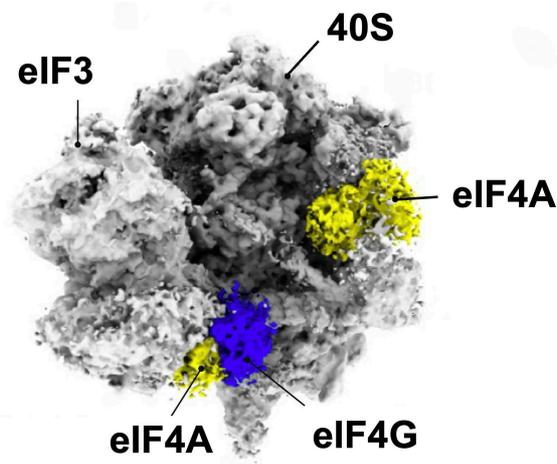


Figure 14. – Structure d'eIF4A en complexe avec la sous unité 40S

Structure reconstruite par Cryo-EM du complexe 48S montrant simultanément eIF4F et une seconde molécule de eIF4A au niveau de la face opposée de la sous unité. Adapté de (Brito Querido et al. 2024b)

4.5 Reconnaissance du codon AUG et assemblage du ribosome 80S

L'étape de balayage prend fin lorsque le complexe 48S identifie le site d'initiation ce qui induit un changement de conformation du ribosome permettant l'insertion complète de l'ARNt initiateur dans le site P du ribosome, favorisant un appariement complet entre la boucle anticodon de l'ARNt avec le triplet AUG (Hinnebusch 2017). Ces réorganisations conformationnelles (structure + eIF) favorisent le recrutement de la grande sous unité ribosomique 60 S pour constituer le ribosome 80 S et le début de l'élongation.

La reconnaissance du codon d'initiation de la traduction est influencée par les séquences nucléotidiques qui l'entourent. L'importance de cette séquence a été démontrée par Marilyn Kozak il y a plus de 40 ans (Kozak 1981) grâce à une étude comparative sur les séquences de 699 5'UTR qui a révélé une séquence consensus, appelée séquence Kozak (Kozak 1987). Cette séquence est GCCRCCAUGG avec une adénosine ou guanosine en position -3 et une guanosine en position +4 qui sont déterminants (l'adénosine du triplet AUG marque le +1) (Kozak 1987). Toutes variations par rapport à cette séquence ont souvent des conséquences sur l'efficacité traductionnelle et un mauvais contexte Kozak s'accompagne de *leaky scanning* dans lequel le ribosome balaye le codon sans le reconnaître et continue sa progression jusqu'au prochain codon AUG (Smith et al. 2005).

Différents eIF sont impliqués dans la reconnaissance du codon AUG et c'est majoritairement eIF1 en association avec eIF1A qui maintiennent le ribosome dans une conformation empêchant l'insertion totale de l'ARNt initiateur dans le site P du ribosome pendant le processus de balayage (Hinnebusch 2017). Lorsque le codon d'initiation est reconnu, l'ARNt se positionne alors favorablement dans le site P ce qui provoque un changement de conformation qui précipite le départ de eIF1 accompagné du phosphate inorganique issu de l'hydrolyse du GTP en GDP portée par eIF2. Le domaine N-terminal du facteur d'initiation eIF5B se positionne à la place laissée vacante par eIF1 ce qui change la configuration du 48S qui passe à une conformation fermée (Hussain et al. 2014; Hinnebusch 2017; Llácer et al. 2018; Yi et al. 2022; Brito Querido et al. 2024b). eIF5B interagit avec eIF1A pour changer l'orientation de l'ARNt initiateur (Lapointe et al. 2022) ce qui permet l'association de la sous unité 60S en éjectant eIF1A et eIF5B, avec l'hydrolyse de GTP portée par eIF5B, et la formation du ribosome 80S sur l'AUG.

5 Mécanismes alternatifs de traduction: quelques exemples

Il existe de très nombreux mécanismes qui divergent de l'initiation coiffe-dépendante décrite ci-dessus. Parmi ceux-ci, on peut citer la traduction des uORF, l'initiation sur des codons non-AUG, l'initiation à partir de bases modifiées (m6A) et l'initiation à partir

de séquences IRES (Internal ribosome entry segment) qui sera décrite dans le chapitre consacré à la traduction virale.

5.1 La traduction des uORF

La traduction des uORF implique la mise en place de plusieurs mécanismes tels que les processus de réinitiation et de leaky-scanning. La réinitiation de la traduction est un processus complexe au cours duquel une sous-unité ribosomique 40S, après avoir traduit une première courte phase de lecture, débute immédiatement la traduction d'une seconde ORF. Le mécanisme précis permettant la réinitiation contient encore de nombreuses zones d'ombre mais il nécessite le maintien d'eIF3 à proximité de la sous-unité 40S lors de la traduction de la première ORF (Pause et al. 1994b; Hronová et al. 2017). Dans le cas de la traduction des ARNm cellulaires qui sont monocistroniques (à l'opposé de certains ARN viraux), la réinitiation survient généralement après la traduction d'une uORF présente dans la région 5'UTR de l'ARNm considéré. Pour que la réinitiation soit efficace, l'uORF doit être courte et son positionnement par rapport à la phase de lecture principale est également déterminante (Russell et al. 2023).

Le mécanisme de traduction/réinitiation des uORF peut se résumer comme décrit sur la Figure 15 et selon trois principaux scénarios (Figure 15).

- Dans le premier cas, il y a traduction de l'uORF puis dissociation sur le codon stop des deux sous-unités du ribosome. De fait, le complexe ne pourra pas atteindre le codon AUG de la phase de lecture principale et on observe une baisse de taux de synthèse protéique.
- Dans le deuxième cas, notamment lorsque les uORF sont très courtes, le complexe 43S se dissocie, probablement de manière partielle, et démarre un nouveau cycle d'initiation en recrutant les facteurs d'initiation qui restent à proximité. C'est un mécanisme de réinitiation qui permet au ribosome d'atteindre l'AUG de la phase de lecture principale (Luukkonen et al. 1995; Kozak 2001). Dans ce cas, l'uORF n'a généralement pas, ou peu, d'effet inhibiteur sur la production de la protéine complète. Ce phénomène de réinitiation est d'autant plus efficace que l'uORF est courte et qu'elle se situe à une distance relativement importante de la phase de lecture principale (au

moins 80 nucléotides) afin que le ribosome puisse avoir le temps de recharger les facteurs d'initiation (Luukkonen et al. 1995; Kozak 2001).

- Il existe un troisième cas dans lequel le peptide néosynthétisé à partir de l'uORF interagit avec le ribosome et le bloque (ribosome stalling) sur l'uORF provoquant l'arrêt de la progression du ribosome. Ce blocage, en plus d'inhiber la traduction de l'ORF principale, génère fréquemment des collisions de ribosomes qui diminuent encore les chances de réinitiation (Hemmings-Mieszczak et al. 2000; Lin et al. 2019; Dever et al. 2020).

- Enfin, dans d'autres cas le codon initiateur de l'uORF n'est pas reconnu par la sous-unité ribosomique 40S qui poursuit le balayage jusqu'à ce qu'elle rencontre un autre codon AUG localisé en aval; on parle alors de *leaky scanning*.

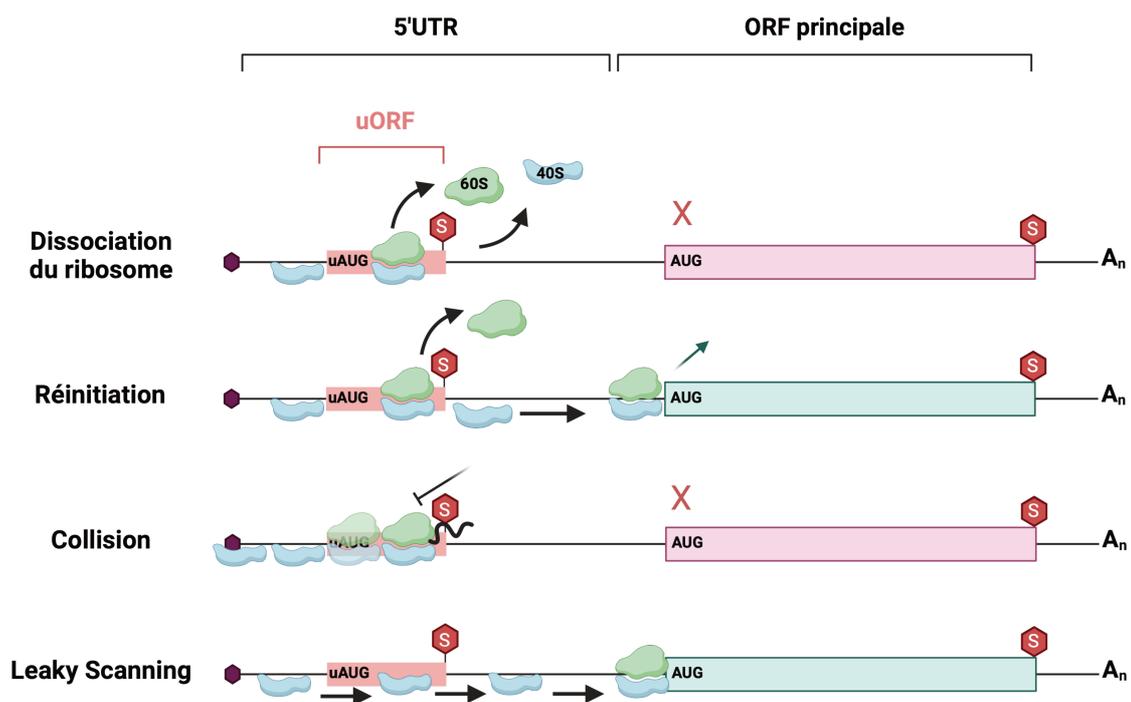


Figure 15. – Mécanisme régulateur des uORF

Les uORF influencent différemment l'initiation de la traduction au niveau de l'ORF principal. Dans certains cas, l'uORF n'est pas reconnue par la sous-unité ribosomique durant le balayage et, par conséquent, n'exerce aucun effet modulateur (*Leaky scanning*). Lorsqu'une uORF est reconnue par le ribosome, elle peut affecter l'ORF principale de différentes manières. Si le ribosome se dissocie entièrement après avoir traduit l'uORF, l'initiation sur l'ORF principale est diminuée, donnant à l'uORF un rôle inhibiteur. Parfois, le peptide nouvellement synthétisé peut également bloquer la progression du ribosome, entraînant son arrêt et des collisions ribosomiques, ce qui a également un effet inhibiteur. Enfin, il arrive que la réinitiation se produise après la traduction de l'uORF; dans ce cas, les uORF ont généralement un effet stimulateur sur la traduction de l'ORF principal. Dessinée sur BioRender.

Enfin, les uORF peuvent aussi avoir des répercussions importantes sur la stabilité de l'ARNm en induisant la voie du NMD (Non Sense Mediated Decay) qui est un contrôle qualité des ARNm (Rehwinkel et al. 2005; Russell et al. 2023). Dans ce cas, le codon stop de l'uORF est reconnu par la machinerie du NMD comme un codon stop prématuré ce qui déclenche la dégradation de l'ARNm et l'arrêt de la production protéique (Uchiyama-Kadokura et al. 2014).

L'impact des uORF sur la traduction est modulé en fonction du contexte physiologique de la cellule. Par exemple, lors d'un stress cellulaire (Young and Wek 2016) qui conduit à la phosphorylation du facteur eIF2 α , la formation du complexe ternaire va être affectée et cela a des conséquences directes sur les ARNm contenant des uORF. C'est par exemple le cas des gènes codant pour ATF4 (Young and Wek 2016), CHOP (Palam et al. 2011) ou encore GADD34 (Lee et al. 2009; Young et al. 2015). En conditions normales, les uORF présentes dans leurs 5'UTR sont traduites ce qui inhibe la production protéique à partir de l'ORF principale; en conditions de stress, la phosphorylation de eIF2 α réduit la formation des complexes ternaires actifs qui mettent plus de temps pour être chargés sur le ribosome qui va donc ignorer les premiers codons AUG rencontrés pour initier sur le codon AUG de la phase de lecture principale.

Dans d'autres cas, l'efficacité d'initiation au niveau des uORF peut être modulée en réponse à la disponibilité de certaines protéines en établissant un rétrocontrôle. Par exemple, l'ARNm codant pour la polyamine contient une uORF qui est régulée par la liaison de la polyamine sur l'ARNm. L'attachement de la protéine stimule le recrutement du ribosome ce qui augmente l'initiation à partir de l'uORF. À faible concentration de polyamine, l'uORF est moins reconnue par le ribosome, ce qui lève l'effet inhibiteur sur l'ORF principale et permet de nouveau la synthèse active de la polyamine; le processus s'inverse à nouveau lorsque la concentration en polyamine devient importante (Law et al. 2001) (Ivanov et al. 2008).

5.2 Initiation dépendante des codons non AUG

a) Présentation générale

La traduction peut débuter sur des codons non-AUG aussi bien chez la levure (Zitomer et al. 1984; Clements et al. 1988) que chez les mammifères (Peabody 1987; Peabody 1989). Dans la majorité des cas, ces codons diffèrent du codon AUG par un nucléotide et le mécanisme moléculaire reste le même avec la liaison de l'ARNt méthionine initiateur sur sa cible malgré une complémentarité imparfaite (Kearse and Wilusz 2017). Des études à large échelle ont mis en évidence que l'utilisation de codons non-AUG était un phénomène courant (Ingolia et al. 2009; Ingolia et al. 2011; Kearse et al. 2019; Chen et al. 2020) ce qui a été confirmé par des analyses protéomiques et peptidomiques par spectrométrie de masse (Slavoff et al. 2013; Vanderperre et al. 2013). Le nombre de protéines initiées à partir de codons non canoniques est toutefois sujet à débat en raison de l'absence fréquente de caractérisation moléculaire associée, à certains biais technologiques ou encore au manque d'évidences de sélection évolutive (Guttman et al. 2013; Gerashchenko and Gladyshev 2014; Kearse et al. 2019; Fedorova et al. 2022). Il n'en demeure pas moins que différentes protéines bien caractérisées sont exprimées à partir de codons non canoniques. À titre d'exemple, nous pouvons citer la protéine eIF4G2 qui est synthétisée exclusivement à partir d'un codon GUG (Takahashi et al. 2005; Liberman et al. 2015).

Un codon d'initiation non canonique peut aussi se trouver en amont du site d'initiation principal ce qui mène à la synthèse d'isoformes comportant une extension N-terminale. C'est par exemple le cas de l'oncogène Myc ou du suppresseur de tumeur PTEN qui possèdent, tous deux, un codon CUG dans leur région 5'UTR en phase avec le codon AUG de départ (Hann et al. 1988; Hopkins et al. 2013). Les isoformes synthétisées à partir de ces triplets CUG ont des fonctions et/ou des localisations cellulaires différentes (Hann et al. 1994; Hopkins et al. 2013). Notons que ces extensions N-terminales peuvent être relativement longues (171 acides aminés dans le cas de pTEN) (Hopkins et al. 2013) et conservées ce qui suggère qu'elles confèrent une fonction particulière aux protéines produites (Ivanov et al. 2011).

b) Mécanisme d'initiation de la traduction à partir des codons non canoniques

L'initiation à partir de codons non canoniques est, en règle générale, moins efficace qu'à partir d'un codon AUG et il a été observé que les CUG et GUG étaient les codons non canoniques les plus utilisés ce qui suggère que les positions U et G sont très importantes (Kozak 1989a; Kearsse and Wilusz 2017). Le contexte Kozak est tout aussi important pour leur reconnaissance par le complexe 43S (Diaz de Arce et al. 2018). Il a été montré qu'une hausse du niveau d'expression d'eIF5, qui est impliqué dans l'hydrolyse du GTP associée à eIF2, favorisait l'initiation sur des codons non canoniques (Loughran et al. 2012). A l'inverse, la présence d'eIF1 réprime l'initiation sur les codons non-AUG (Nanda et al. 2009; Ivanov et al. 2010) ce qui est cohérent avec son rôle lors du balayage du complexe 43S (Pestova et al. 1998a; Valásek et al. 2004). Ainsi l'équilibre entre eIF1 et eIF5 contribue probablement à la régulation de la sélection des codons non canoniques.

Bien qu'en règle générale le complexe ternaire soit impliqué à la fois dans la reconnaissance des codons d'initiation canoniques et non canoniques (Peabody 1989), les facteurs alternatifs eIF2A et eIF2D peuvent se substituer à eIF2 ce qui favorise l'initiation à partir des codons non canoniques (Dmitriev et al. 2010; Kim et al. 2011; Starck et al. 2012) et ils peuvent utiliser un ARNt leucine et de façon indépendante du GTP (Starck et al. 2012). Cela crée un avantage sélectif en conditions de stress dans lesquelles le facteur eIF2 est fréquemment inactif à cause de l'absence du recyclage du GDP en GTP. De façon intéressante, le niveau d'expression d'eIF2A augmente pendant le stress cellulaire (Starck et al. 2016) ce qui renforce l'importance du mécanisme de régulation des uORF dans cet état physiologique. D'ailleurs, la protéine BIP, effecteur central de la voie de réponse au stress du réticulum (Preissler and Ron 2019), contient 2 uORF qui utilisent des codons non AUG.

5.3 Initiation à partir de bases modifiées

Enfin nous terminerons ce chapitre en évoquant brièvement un dernier grand mécanisme de régulation de la traduction qui met en jeu la modification chimique de certaines bases. Elles sont retrouvées dans toutes les régions des ARNm, y compris

dans la 5'UTR (Perry et al. 1975). Les méthylations des adénines et cytosines font partie des modifications les plus courantes et sont assurées par une batterie d'enzymes et selon un processus complexe et réversible qui est bien décrit dans les revues suivantes (Trixl and Lusser 2019; Jiang et al. 2021). Ces modifications chimiques post-traductionnelles résultent de l'ajout d'un groupement méthyl par des enzymes cellulaires de type méthyltransférases, leurs actions sont réversibles et peuvent avoir des conséquences positives ou négatives sur l'efficacité traductionnelle (Nachtergaele and He 2018; Shi et al. 2019). Ainsi, la N6-méthyladénosine (m6A), qui est une des modifications les plus courantes chez les eucaryotes, joue un rôle important dans la régulation du métabolisme de plusieurs ARNm viraux et cellulaires (Rubio et al. 2018; McFadden et al. 2021).

Lorsque cette modification survient au sein de la région 5'UTR, le complexe 43S peut directement reconnaître cette m6A et s'y attacher comme il le ferait sur une séquence IRES (Meyer et al. 2015) (Figure 16). Dans ce cas de figure, eIF3 interagit à la fois avec la sous-unité ribosomique 40S et avec la méthyltransférase METTL3 qui est, elle-même, liée au m6A. Cela permet de recruter et échafauder le complexe de préinitiation 43S sur l'adénine méthylée qui se substitue fonctionnellement à la coiffe. (Choe et al. 2018). La protéine YTHDF3 qui se lie également aux adénosines méthylées a également été impliquée dans l'initiation de certains ARNm spécifiques (Li et al. 2017).

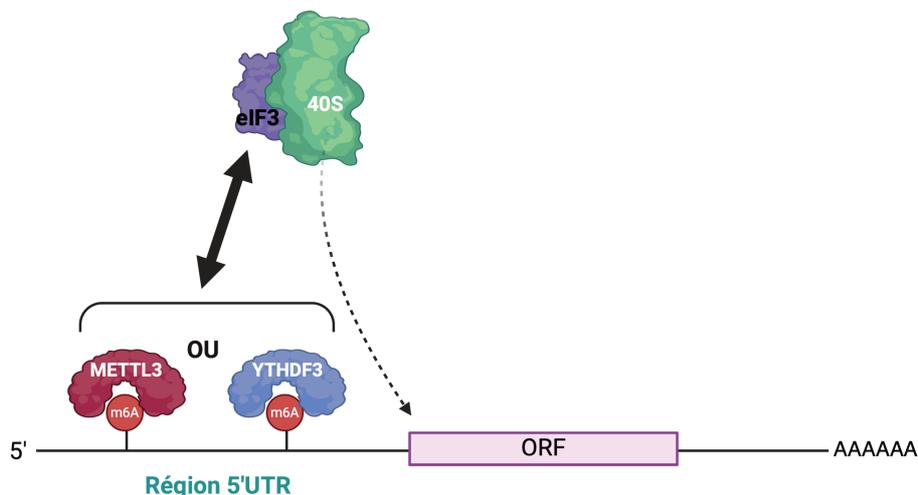


Figure 16. – Initiation de la traduction par l'intermédiaire des bases modifiées m6A

Le complexe 43S peut être recruté au niveau de modification N6-méthyladénosine dans les régions 5'UTR des ARNm par l'intermédiaire des protéines METTL3 ou YTHDF3. Ce recrutement permet une initiation de la traduction indépendante de la coiffe. Dessinée sur BioRender.

De façon parallèle, d'autres travaux ont montré que la présence de modifications m6A dans la partie codante des ARNm favorisait aussi la traduction en réduisant la formation de structures secondaires (Mao et al. 2019). Des expériences similaires chez les bactéries ont montré que la présence de modification m6A dans la partie codante perturbait l'élongation et changeait la dynamique de traduction (Choi et al. 2016; leong et al. 2021; Jain et al. 2023).

Ces modifications chimiques de l'ARNm peuvent aussi avoir des conséquences négatives comme la modification 5-Méthylcytosine (m5C) qui réduit considérablement l'efficacité traductionnelle (Schumann et al. 2020).

L'ensemble de ces données suggère que des mécanismes alternatifs d'initiation pourraient être mis en place pour contourner le recrutement sur la coiffe en cas de conditions physiologiques défavorables. Cela est en adéquation avec l'observation que le niveau de méthylation augmente en condition de stress cellulaire (Zhou et al. 2015). Enfin, il a récemment été suggéré que les ARN circulaires pouvaient être traduits (Legnini et al. 2017; Pamudurti et al. 2017). Cela pose donc la question du mécanisme de recrutement de la machinerie traditionnelle, puisqu'ils ne possèdent pas d'extrémité 5' libre, une des hypothèses émises postule le recours à un mécanisme dépendant des bases modifiées (Yang et al. 2017; Di Timoteo et al. 2020).

Chapitre 2 – La traduction virale

La diversité des virus est remarquable et reflète la multitude et la complexité de leurs capacités d'adaptation à l'hôte. Certains virus possèdent des génomes à ARN, tandis que d'autres sont dotés de génomes à ADN, ils peuvent être enveloppés ou non, cibler différentes voies d'entrée et de sortie, induire des infections rapides ou latentes et provoquer des pathologies diverses ou rester asymptomatiques. (Sharp 2002; Jones et al. 2021). Cependant, la caractéristique commune à tous les virus est leur relation de dépendance vis-à-vis de l'hôte. C'est particulièrement vrai si on considère cela par rapport à la machinerie traductionnelle de leur hôte. Lors d'une infection virale, les ARN viraux rivalisent avec les ARNm cellulaires pour cette machinerie traductionnelle, et au fil de l'évolution, chaque type de virus a développé des mécanismes distincts pour moduler cette compétition à son avantage. De plus, les génomes viraux, sont contraints en termes de taille de génome et ils contiennent souvent des caractéristiques structurales qui leur permettent d'optimiser l'intégration de multiples fonctions (Ho et al. 2021).

L'infection virale s'accompagne souvent d'une baisse du taux de synthèse des protéines cellulaires (appelé 'shutoff cellulaire') au profit des protéines virales (Stern-Ginossar et al. 2019). Cela suggère un mécanisme adaptif mis en place par les virus pour traduire dans des conditions défavorables aux ARNm cellulaires et cela peut être dû à la structure même de l'ARN viral et/ou à la manipulation des facteurs d'initiation. Dans ce chapitre, nous décrivons les caractéristiques structurales de certains virus qui leur permettent de traduire efficacement leurs protéines. Nous aborderons notamment les séquences IRES (Internal Ribosomal Entry Site) puis nous citerons des exemples de stratégies virales utilisées pour manipuler la machinerie de traduction de l'hôte et contourner les premières lignes de défense immunitaire.

1 CARACTERISTIQUE STRUCTURALES

La compaction des génomes viraux a pour conséquence une utilisation optimale de l'information génétique dans un minimum d'espace. Ainsi, tous les virus doivent au cours de leur cycle de réplication concilier la production d'un stock de protéines virales, l'encapsidation et la propagation de leur matériel génétique. Dans le cas des virus à ARN positif, c'est souvent la même molécule qui est utilisée pour ces différentes fonctions. Dans ce cas, les structures/motifs indispensables pour la réplication peuvent s'avérer délétères pour l'attachement et le balayage du ribosome.

Contrairement aux ARNm cellulaires qui sont monocistroniques, les ARNm viraux codent souvent pour plusieurs protéines virales à partir du même ARNm ce qui est incompatible avec le mécanisme de balayage canonique du ribosome.

De fait, les virus ont adapté des mécanismes traductionnels qui leur permettent d'accommoder ces contraintes structurales sans que cela soit au détriment de leur production protéique. Nous allons brièvement citer plusieurs exemples qui permettront de mieux comprendre les caractéristiques adaptatives des virus et nous décrirons en détail le mécanisme d'entrée interne du ribosome qui a été un véritable paradigme au cours des 40 dernières années pour mieux comprendre les mécanismes moléculaires du contrôle traductionnel.

1.1 L'expression de polyprotéines

Les ARNm cellulaires eucaryotes sont presque exclusivement monocistroniques c'est-à-dire qu'un ARNm code pour une protéine. En revanche ; il est possible que de nombreux ARNm viraux codent pour des polyprotéines à partir d'un seul ARNm, qui sont ensuite clivées à des sites spécifiques par des protéases, généralement virales, afin de libérer les différentes protéines matures avec des fonctions propres. Ce clivage peut être co-traductionnel ce qui signifie que le clivage des polyprotéines se produit simultanément à la traduction de la chaîne peptidique. Un tel processus est retrouvé chez le poliovirus, le virus de l'hépatite C ou encore le SARS-CoV-2 (Yao et al. 1999; Lee et al. 2022). En revanche, le précurseur Gag-Pol du virus de l'immunodéficience humaine est clivé de façon post-traductionnelle par sa protéase virale (Pettit et al.

2004). L'expression d'une polyprotéine permet ainsi aux virus de minimiser et d'optimiser l'espace occupé par les séquences régulatrices de l'expression des gènes.

1.2 Le ribosomal *Frameshifting*

Le "*frameshift*" est un mécanisme décrit comme un décalage de cadre de lecture du ribosome en cours d'élongation (Plant and Dinman 2006). Ce processus s'effectue au niveau d'une séquence appelée séquence glissante (« *slippery sequence* ») généralement composée de nucléotides répétés. Ce décalage du cadre de lecture peut être d'un ou de deux nucléotides et se produire dans les deux sens, bien que les décalages -1 soient plus souvent rencontrés (Caliskan et al. 2015). Il est souvent régulé par la présence de structures secondaires situées en aval de la séquence glissante qui permet de ralentir le ribosome et le faire changer de cadre. Les ARN génomiques du SARS-CoV-2 et du VIH-1 présentent chacun un mécanisme de *frameshift* qui a été bien étudié. Dans le cas du SARS-CoV-2, une structure de type pseudonœud est située à environ 8 nt en aval de la séquence de décalage et provoque un décalage (-1nt) du cadre de lecture du ribosome et permet la production d'autres protéines Nsp (Nsp12, Nsp13, Nsp14, Nsp15 et Nsp16) qui font partie de pp1ab (Ishimaru et al. 2013; Bhatt et al. 2021). Une structure de type tige boucle stable a été mise en évidence dans le cas de l'ARN génomique du VIH-1, où elle régule également un décalage de -1 du cadre de lecture ce qui permet l'expression de GAG-POL (Hung et al. 1998; Gaudin et al. 2005).

En plus de ces deux exemples, d'autres virus utilisent le *frameshifting*, tels que le virus du Nil occidental (WNV) (Melian et al. 2010), le virus de l'influenza (IAV) (Jagger et al. 2012), les picornavirus (Loughran et al. 2011), ou encore les gamma herpes virus (GHV) (Kwun et al. 2014). La fréquence du décalage de cadre peut être modulée par divers facteurs agissant comme transactivateurs. Ces transactivateurs peuvent être des protéines virales; on peut citer le *frameshifting* -1 du virus de l'encéphalomyocardite (EMCV) qui est catalysé par la protéine virale 2A (Loughran et al. 2011; Naphthine et al. 2017). Dans d'autres cas, ces transactivateurs sont des protéines cellulaires comme pour le *frameshifting* du virus du syndrome dysgénésique et respiratoire du porc (PRRSV) qui est modulé par un complexe formé par la protéine

cellulaire PCBP et la protéine virale Nsp1 β qui se lie au niveau de la séquence de glissement (Li et al. 2014; Naphtine et al. 2016).

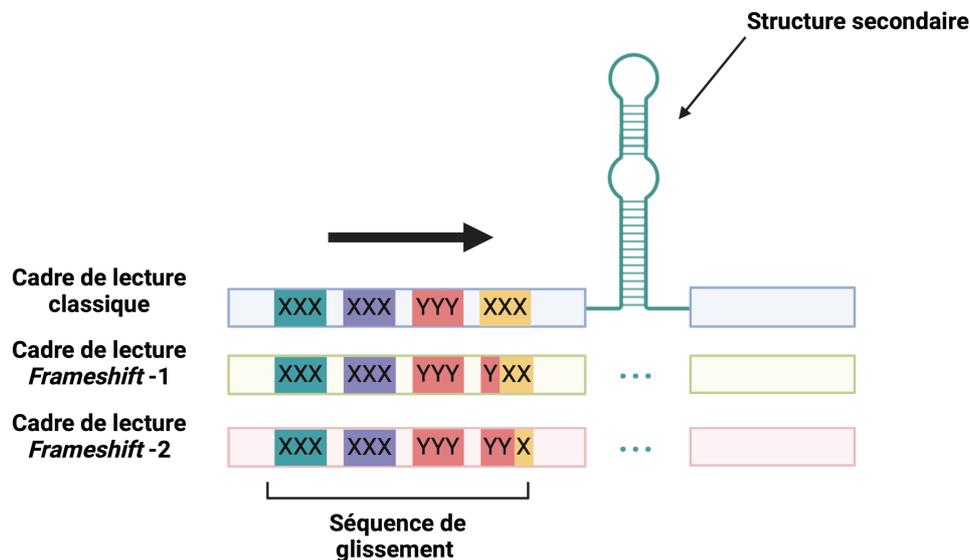


Figure 17. – Ribosomal *frameshifting*

Lors du mécanisme de *frameshifting*, le ribosome glisse sur une séquence spécifique se trouvant généralement en amont d'une structure secondaire stable. Cela entraîne un changement du cadre de lecture. Dessinée sur BioRender.

1.3 Initiation alternative par *leaky scanning*

Les virus utilisent le *leaky scanning* pour démarrer la traduction sur des sites d'initiation alternatifs ce qui permet de traduire plusieurs ORFs à partir d'un même ARN ou de produire différentes isoformes d'une protéine. Ainsi, le rétrovirus de la leucémie murine synthétise une forme allongée de GAG à partir d'un codon CUG (Prats et al. 1989) et le torovirus équin, virus proche des coronavirus, synthétise deux protéines U1 et U2 à partir de codons CUG (Stewart et al. 2018). Il y a aussi le virus mosaïque panicum, qui combine le *leaky scanning* et l'utilisation de codons d'initiation non canoniques pour produire quatre protéines virales à partir d'un seul et unique ARN (Turina et al. 2000; Pyle et al. 2019). Cela souligne une nouvelle fois l'habileté des virus à utiliser au maximum la capacité de codage de leur génome.

1.4 Les IRES

Les sites d'entrée interne du ribosome (IRES pour *Internal Ribosome Entry Site*) sont des structures secondaires et tertiaires formées par l'ARN qui permettent de recruter directement, ou indirectement, le complexe 43S de manière indépendante de la coiffe du messager (Jackson et al. 1990; Chen and Sarnow 1995; Lee et al. 2017). Les IRES ont dans un premier temps été identifiées au niveau des ARNm viraux (Jang et al. 1988; Pelletier and Sonenberg 1988; Jackson et al. 1990; Kühn et al. 1990; Bandyopadhyay et al. 1992), mais ils ont ensuite été découverts au sein de plusieurs ARNm cellulaires (Macejak and Sarnow 1991; Yang and Sarnow 1997). De récentes études à large échelle suggèrent qu'ils seraient même présents dans 10% des ARNm cellulaires chez les mammifères (Weingarten-Gabbay et al. 2016). Cependant, nous ne décrivons ici que les caractéristiques des IRES viraux.

La taille des IRES viraux est souvent assez importante, de l'ordre de plusieurs centaines de nucléotides et il n'existe pas de motif consensus qui permet de les identifier par simple lecture de la séquence (Chappell et al. 2000; Bonnal et al. 2003; Zhao et al. 2020). Les IRES sont très variables en termes de structure, mais on peut retrouver des parties communes notamment au sein des IRES appartenant à des familles virales proches. Lorsque c'est le cas, ces séquences ont des caractéristiques fonctionnelles conservées notamment au niveau de l'utilisation, ou non, des facteurs d'initiation. Alors que certains IRES n'ont besoin d'aucun facteur d'initiation pour recruter et localiser le ribosome sur le codon d'initiation, d'autres sont dépendants de plusieurs d'entre eux (Ali et al. 2001; Hellen and Sarnow 2001; Wek 2018). Par ailleurs, les IRES sont souvent accompagnés dans leurs fonctions par des protéines de support couramment appelées ITAF (pour *IRES Trans-Acting Factors*).

a) Classification des IRES

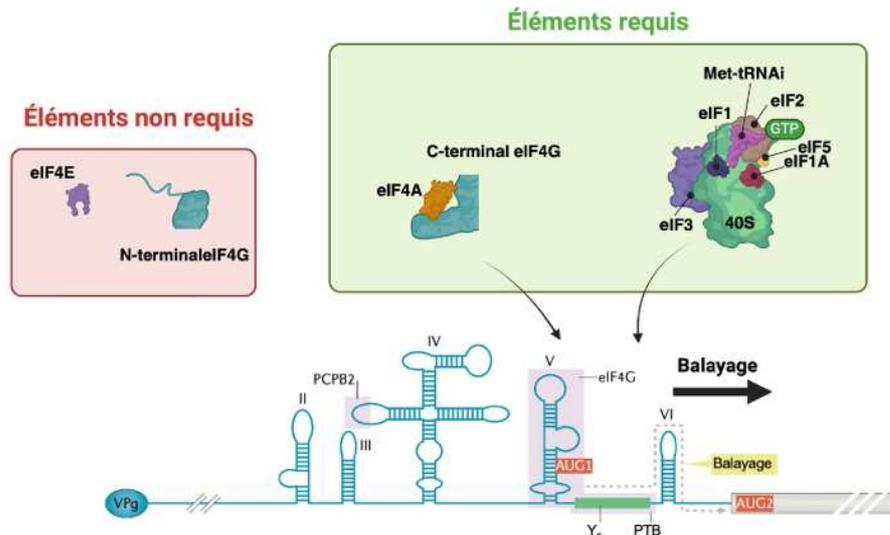
Malgré une composition très variable en termes de taille et de structure, il a été possible de classer de nombreux IRES en 4 grandes familles en se basant principalement sur leur dépendance dans les facteurs d'initiation (Lee et al. 2017). Notons que certains IRES viraux ne sont pas classifiés dans ces catégories, c'est par exemple le cas de l'IRES du génome viral du VIH-1 (Brasey et al. 2003; Gendron et

al. 2011; Deforges et al. 2017) ainsi que l'IRES de l'ARN codant pour la protéine vFLIP de l'herpès de type 8 (Othman et al. 2014).

b) Les IRES de classe I et II

Les séquences IRES ont été originellement découvertes chez les ARNm des picornavirus (Jang et al. 1988; Pelletier and Sonenberg 1988) parce que les 5'UTR de ces ARNm présentaient des caractéristiques incompatibles avec l'initiation canonique qui sont : (i) l'absence de coiffe remplacée par une structure spécifique reconnue par la protéine VPg (*viral protein genome-linked*) (Lee et al. 1977), (ii) la présence de nombreuses structures ARN complexes et énergétiquement stables, (iii) la présence de nombreux codons AUG silencieux. En revanche ces régions ont la capacité d'interagir directement avec le complexe d'initiation. Les IRES des picornavirus ont été essentiellement classés en deux catégories. Les IRES de classe I, tels que ceux du poliovirus (PV) (Pelletier and Sonenberg 1988; Haller et al. 1993), et du coxsackievirus (Yang et al. 1997) se caractérisent par le balayage des séquences en aval du site d'attachement du ribosome. Les IRES de type II, tels que ceux du virus de la fièvre aphteuse (*Foot and mouth disease virus*, FMDV) (Belsham and Brangwyn 1990) et du virus de l'encéphalomyocardite (EMCV) (Jang et al. 1988), initient la traduction sur un triplet AUG localisé à la bordure 3' de l'IRES sans balayage.

IRES de type I



IRES de type II

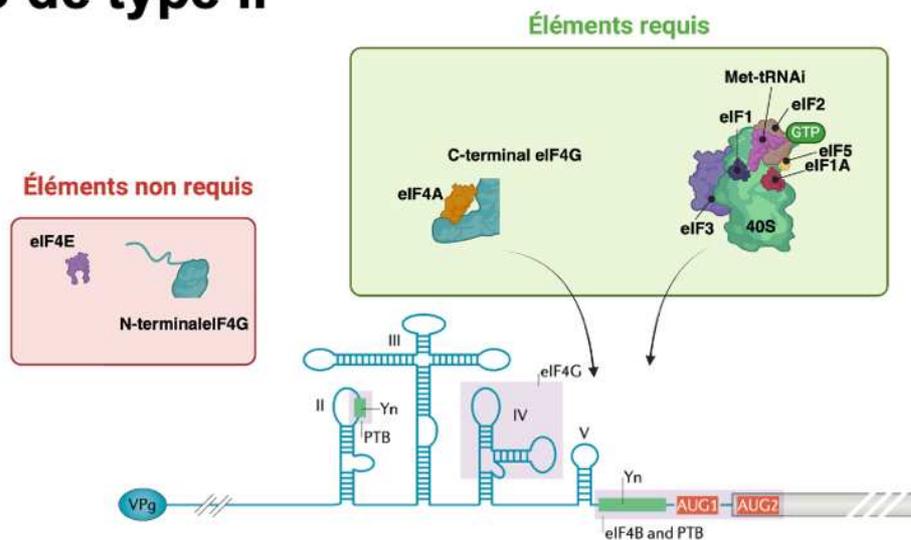


Figure 18. – Les IRES de type I et de type II

Les IRES de type I (haut) recrutent le ribosome 40S de 100 à 150 nucléotides en amont du codon d'initiation AUG. Un processus de balayage permet ensuite au ribosome de rejoindre le codon d'initiation. Dans le cas des IRES de type II (bas), le recrutement du ribosome 40S est directement effectué au niveau du codon AUG. Adaptée à partir de (Jaafar and Kieft 2019).

Ces IRES, localisés dans la 5'UTR, ont une taille d'environ 450 nt de long, (Jang et al. 1990; Lee et al. 2017; Zhao et al. 2020), et présentent des structures secondaires complexes de type tige-boucle qui constituent des domaines fonctionnels à la fois pour la réplication virale et la traduction des ARNm (Jang et al. 1990; Gao et al. 2016; Lee et al. 2017). Les IRES de type I et II sont caractérisés par un motif conservé (Yn-Xm-

AUG) riche en pyrimidine et localisé à leur extrémité 3'. Cette séquence, en complément des structures secondaires, contribue au recrutement de la machinerie traductionnelle.

Le recrutement du ribosome sur les IRES de type I et II requière généralement l'ensemble des eIF à l'exception d'eIF4E. Le couple formé par eIF4G et eIF4A est capable d'interagir avec la structure ARN située juste en amont du motif Yn-Xm-AUG (Kolupaeva et al. 1998; de Breyne et al. 2009) ce qui induit des changements conformationnels des structures de l'ARN et facilite l'attachement de la sous-unité 40S. La sélection du codon initiateur de la traduction diffère selon la catégorie d'IRES. En effet, chez les IRES de type II (EMCV), l'initiation s'effectue directement sur le codon AUG situé à proximité immédiate du site de recrutement du ribosome (Kaminski et al. 1994; Pestova et al. 1996b) et ne requiert pas de balayage de l'ARN. En revanche, l'initiation à partir d'un IRES de type I (PV) demande un processus de balayage additionnel afin de trouver le bon codon AUG situé à environ 150 nt du site d'entrée du ribosome. L'initiation à partir des IRES de classe II est facilitée par la présence de eIF1 ce qui est cohérent avec le rôle de ce dernier (Andreev et al. 2007; Sweeney et al. 2014).

c) Les IRES de type III

Les IRES de type III sont retrouvés chez les hepacivirus (ex. : virus de l'hépatite C (HCV) – (Otto et al. 2002)), les pestivirus (ex. : le virus de la peste porcine classique (CSFV) – (Fletcher and Jackson 2002)) et chez certains picornavirus (Pisarev et al. 2004; Bakhshesh et al. 2008; de Breyne et al. 2008b). La présence des IRES de type III dans des familles virales distinctes suggère qu'ils ont été transférés par des événements de recombinaison au cours de l'évolution (Hellen and de Breyne 2007).

IRES de type III

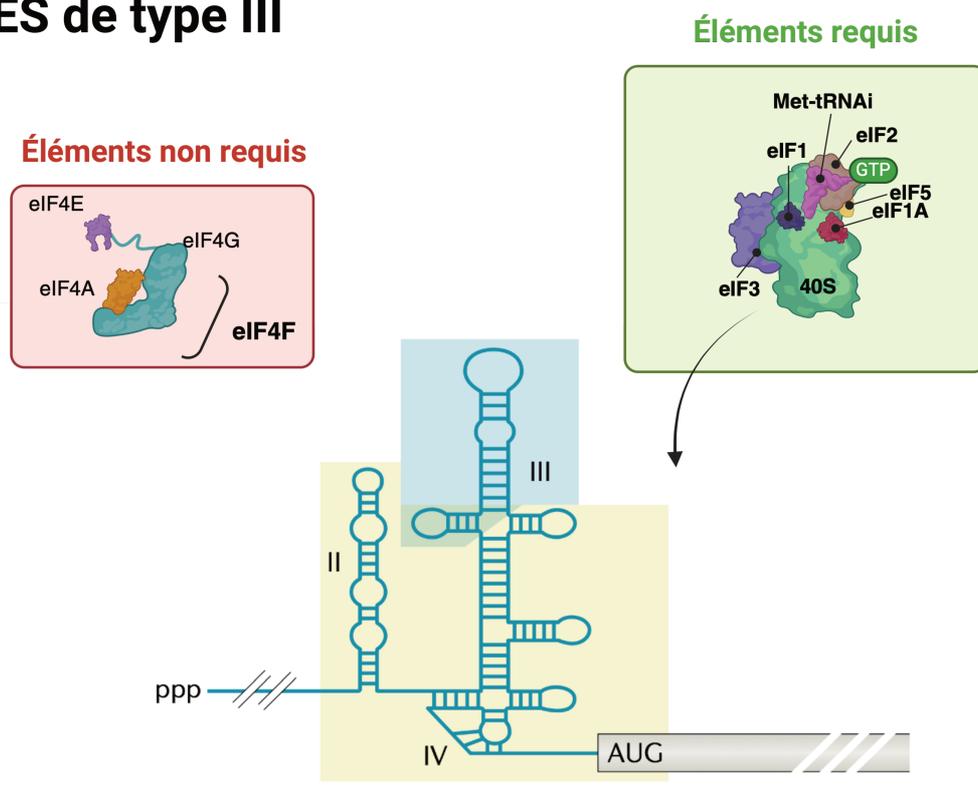


Figure 19. – IRES de type III

Les IRES de type III (HP et HP-like) permettent d'initier la traduction de façon indépendante du complexe eIF4F. Le recrutement du ribosome se fait directement au niveau du codon initiateur AUG sans processus de balayage. Adaptée à partir de (Jaafar and Kieft 2019).

Bien que des différences spécifiques puissent être observées, les IRES de type III possèdent des structures secondaires généralement conservées au niveau de leur domaine central (Hellen and de Breyne 2007; Filbin and Kieft 2009). Celui-ci est constituée d'un assemblage de plusieurs tiges-boucle organisées autour d'un pseudonœud (Yamamoto et al. 2014; Quade et al. 2015). *In vitro*, les IRES de type III sont en mesure d'interagir directement avec la sous-unité 40S du ribosome sans l'aide d'eIF (Pestova et al. 1998b); de manière intéressante, l'IRES interagit avec la sous-unité 40 S au niveau de son site de liaison à eIF3 ce qui pose la question d'une éventuelle compétition entre eIF3 et l'IRES (Kieft et al. 2001). Les IRES de type III sont également capables de se lier directement avec eIF3 par la partie apicale du domaine III (Sizova et al. 1998; de Breyne et al. 2008b). Cette interaction semble être

nécessaire pour induire une réorganisation et déplacer eIF3 du ribosome 40S afin de libérer le site de liaison de l'IRES (Hashem et al. 2013b; Sun et al. 2013).

Lors de son recrutement, la petite sous-unité ribosomique est directement positionnée quelques nucléotides en amont du codon AUG, la reconnaissance du codon AUG est assurée par la présence additionnelle du complexe ternaire eIF2.GTP.Met-tRNAⁱ et eIF1A (Khawaja et al. 2015; Quade et al. 2015; Jaafar et al. 2016). Contrairement aux IRES de type I et II, l'initiation de la traduction conduite par les IRES de type III est indépendante des facteurs du complexe eIF4F et ne nécessite pas un mécanisme de balayage (Otto and Puglisi 2004).

d) Les IRES de type IV

Les IRES de classe IV ont été caractérisés au sein de la famille des *Dicistroviridae* au niveau de la séquence intergénique de l'ARN du virus de la paralysie du cricket (CrPV) qui constitue l'exemple le mieux documenté (Pestova et al. 2004; Spahn et al. 2004; Deniz et al. 2009; Bonning and Miller 2010). Ces IRES sont relativement courts (150-200 nt) et adoptent une structure extrêmement compacte formée autour de trois pseudonœuds (PKI, PKII, PKIII) (Schüler et al. 2006; Murray et al. 2016). Certaines structures tige-boucle participent à l'interaction avec la sous-unité 40S (Pestova et al. 2004) ou à l'assemblage avec la sous-unité 60S (Fernández et al. 2014; Murray et al. 2016).

IRES de type IV

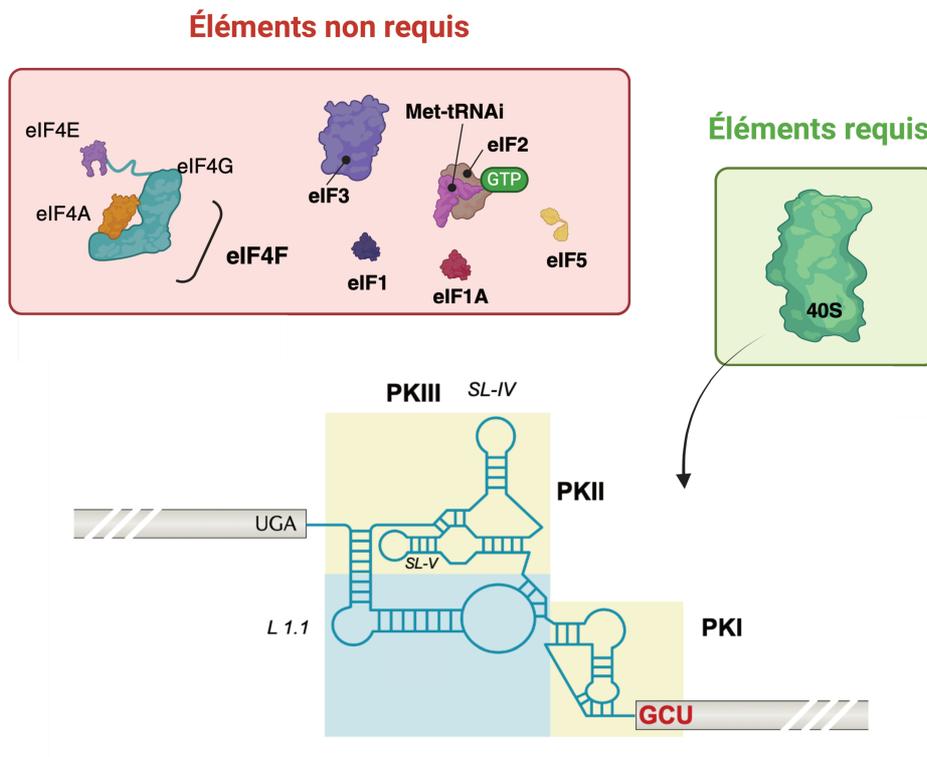


Figure 20. – IRES de type IV

Les IRES de type IV permettent d'initier la traduction de façon totalement indépendante des facteurs d'initiation de la traduction y compris l'ARNt-initiateur. Les IRES de type IV interagissent directement avec la sous unité ribosomique 40S. Adaptée à partir de (Jaafar and Kieft 2019).

Les IRES de type IV utilisent un mécanisme d'initiation particulier qui s'effectue en l'absence de tout eIF et d'ARNt-initiateur, et ne requiert pas de codon AUG (Wilson et al. 2000). Ils sont capables d'interagir directement avec le ribosome 40S et en particulier avec la protéine ribosomique RPS25 (Landry et al. 2009; Muhs et al. 2011). De plus, toujours en l'absence de tout facteur, un ribosome 80S compétent pour l'élongation peut se former sur l'IRES. Le domaine PKI de l'IRES occupe le site A du ribosome en mimant la structure du tRNAi (Costantino et al. 2008; Muhs et al. 2015; Butcher and Jan 2016; Pisareva et al. 2018). Le mécanisme spécifique des IRES de type IV leur permet d'initier la traduction dans des conditions de stress, leur conférant un avantage sélectif par rapport aux ARNm cellulaires (Fernandez et al. 2002).

1.5 Le shunt ribosomal

Le shunt ribosomal est un mécanisme particulier de l'initiation de la traduction utilisé par certains ARNm viraux et cellulaires (Zhang et al. 1989; Latorre et al. 1998; Ryabova and Hohn 2000; Yueh and Schneider 2000; Xi et al. 2005). Ce processus est généralement dépendant de la coiffe de l'ARNm. Cependant, le balayage de la 5'UTR n'est pas linéaire et le complexe de préinitiation, à un site donneur, contourne et ignore une partie de la 5'UTR ou de l'ARNm jusqu'à un site accepteur proche du codon d'initiation. Ce mécanisme permet (i) d'éviter (shunt) les structures secondaires/tertiaires stables qui pourraient stopper la progression du ribosome et/ou (ii) permettre l'expression de protéines situées plus proche de l'extrémité 3' de l'ARNm.

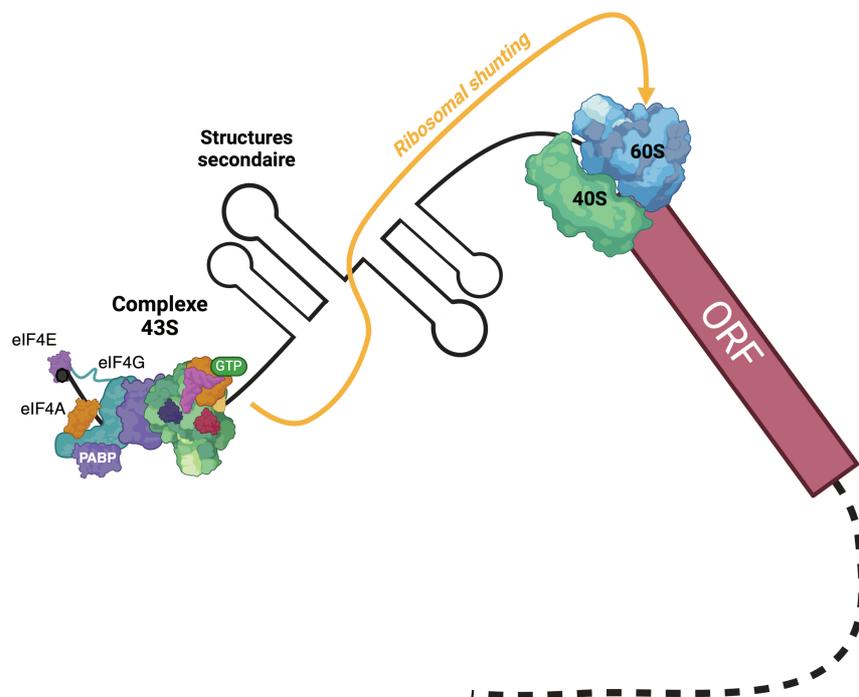


Figure 21. – Shunt ribosomal

Le complexe de préinitiation 43S se fixe à l'extrémité 5' de l'ARNm de façon coiffe-dépendante. Cependant, le ribosome ne rejoint pas le codon initiateur via un processus de balayage linéaire. À la place, la sous-unité ribosomique saute une région de la 5'UTR. Cette section est généralement riche en structures secondaires. Dessinée sur BioRender.

Le mécanisme moléculaire proprement dit du shunt ribosomal reste très largement méconnu. De plus, il est difficile de comparer les différents shunts ribosomaux entre eux parce qu'ils présentent chacun leurs spécificités. Le mieux décrit reste le shunt de l'ARNm 35S de CaMV qui est la résultante d'une réinitiation stimulée par le facteur

viral TAV et par la présence d'une tige-boucle rapprochant deux cadres de lecture (Park et al. 2001). Le shunt des ARNm tardifs d'adénovirus est actif dans des conditions défavorables pour l'initiation classique. Le shunt s'opère via une région structurée qui serait contournée par le complexe 43S grâce à des interactions entre la région structurée, l'ARNr 18S et p100k un facteur viral liant eIF4G (Yueh and Schneider 2000; Xi et al. 2004). Enfin, un mécanisme de shunt s'effectue pour exprimer les protéines Y du virus Sendai au cours duquel les ribosomes seraient directement transloqués à partir de la coiffe (Latorre et al. 1998) vers le site accepteur localisé à proximité des codons d'initiation. La séquence nucléotidique primaire ne semble pas être importante dans ce mécanisme (de Breyne et al. 2003) suggérant que c'est la structure de l'ARN qui favoriserait le shunt ribosomal.

2 Le contrôle de la traduction par les virus

2.1 Aperçu général

Durant l'infection, les virus déploient une variété très importante de tactiques sophistiquées pour manipuler les mécanismes de régulation de l'expression des gènes et induire une inhibition globale de la synthèse protéique. Deux approches principales se distinguent : la première s'attaque directement aux ARNm de la cellule hôte, cherchant à altérer leur structure, leur localisation ou leur synthèse, afin d'entraver leur reconnaissance et leur traduction par la machinerie cellulaire. La seconde interfère directement avec la machinerie de traduction, la rendant inopérante. Cette seconde stratégie sera explorée en détail dans la suite de ce chapitre, à travers l'analyse de trois exemples particuliers qui retiennent notre attention : (i) le contrôle des facteurs d'initiation de la traduction, (ii) le contrôle de la voie de régulation de eIF4E/4E-BP et (iii) le contrôle de la voie eIF2/PKR-PERK. Avant de plonger dans l'examen approfondi de ces régulations spécifiques, nous aborderons brièvement dans le paragraphe suivant les mécanismes de régulation pré-traductionnel afin de mieux saisir leur diversité.

Le contrôle de l'expression des gènes cellulaires par les virus commence dans le cas de certaines infections virales dès leur synthèse. En effet, plusieurs protéines virales entravent l'étape de transcription des ARNm en perturbant l'activité de l'ARN

polymérase II (Akhrymuk et al. 2012; Rutkowski et al. 2015; Wang et al. 2020) ou en ciblant spécifiquement les facteurs de transcription (Clark et al. 1993; Le May et al. 2004). D'autres virus synthétisent des endoribonucléases qui fragmentent les ARN cellulaires. C'est le cas, entre autres, de la protéine VHS de l'herpès simplex (Fenwick and McMenamin 1984; Elgadi et al. 1999), de la protéine SOX des gamma-herpèsvirus (Covarrubias et al. 2009) et de la protéine PA-X du virus de l'influenza (Hayashi et al. 2016). Les poxvirus, expriment quant à eux deux protéines capables de cliver la coiffe des ARN, accélérant leur dégradation (Parrish and Moss 2007; Peters et al. 2022). De façon unique, les coronavirus, via la protéine Nsp1, provoquent un clivage des ARNm sans posséder d'activité ribonucléase propre (Shehata and Parker 2023; Tardivat et al. 2023) ; ce point sera discuté en détail dans le chapitre 3 section 3.1 de cette thèse. Ces différents exemples sont très bien caractérisés et sont détaillés dans la revue suivante (Abernathy and Glaunsinger 2015).

De façon intéressante, diverses stratégies de contournement sont mises en place par les virus pour maintenir l'expression de leurs protéines dans ces conditions dégradées. À titre d'exemple nous pouvons citer la protéine PA-X du virus de l'influenza qui ne clive que les ARNm produits par l'ARN polymérase II, ce qui n'est pas le cas des ARN viraux, les protégeant ainsi de la dégradation (Khapersky et al. 2016; Gaucherand et al. 2019; Krischuns et al. 2021).

En outre plusieurs virus perturbent les voies d'export des ARNm vers le cytoplasme entraînant une accumulation nucléaire à distance de la machinerie de traduction. Cela implique généralement le ciblage des protéines du pore nucléaire ou des facteurs d'export des ARNm. (Qiu and Krug 1994; Castelló et al. 2009a; Zhang et al. 2019; Addetia et al. 2021).

2.2 Le contrôle des facteurs d'initiation

Plusieurs familles de virus ont développé des mécanismes qui leur permettent de cibler certains facteurs d'initiation en les clivant grâce à des protéases, ou en interférant avec leur compartimentalisation subcellulaire ou encore en modifiant leur expression.

a) Clivage protéolytique des facteurs d'initiation

Deux facteurs d'initiation sont généralement ciblés par les virus : eIF4G et PABP.

La protéase 2A du poliovirus est l'une des premières protéines produites lors de l'infection et induit le clivage d'eIF4G en deux fragments distincts (Etchison et al. 1982; Gradi et al. 1998b). Le fragment N-terminal contient le site de liaison à eIF4E et PABP alors que le fragment C-terminal (également appelé P100) arbore le domaine de liaison à eIF3 et à eIF4A. Une fois clivé, eIF4G n'est plus en mesure de jouer sa fonction d'intermédiaire entre la sous-unité 40S et l'ARNm via les interactions respectives à eIF3 et à eIF4E. Cela se traduit par une rapide et forte inhibition de la traduction cellulaire dès le début de l'infection (Gradi et al. 1998b). Les ARN viraux ne sont pas affectés par cette répression parce que l'initiation de la traduction s'effectue par l'intermédiaire d'un IRES. La partie C-terminale d'eIF4G est suffisante, en synergie avec eIF4A, pour interagir avec l'IRES (Kolupaeva et al. 1998; de Breyne et al. 2009) et assurer le recrutement du ribosome pour la synthèse des protéines virales (Ohlmann et al. 1996). Ce mode d'action est conservé chez différents picornavirus tels que EMCV, le rhinovirus ou encore le virus de la fièvre aphteuse (FMDV) (Medina et al. 1993; Haghighat et al. 1996). Le clivage de eIF4G a également été observé chez d'autres familles virales comme les calicivirus félins (Avanzino et al. 2017) et les rétrovirus, notamment les virus de l'immunodéficience humaine, (Ventoso et al. 2001; Alvarez et al. 2003).

La protéine PABP, qui lie la queue poly(A), est également ciblée par les protéases 2A et 3C des picornavirus (PV, EMCV, CBV3) (Bonderoff et al. 2008; Smith and Gray 2010). Il a été montré que le clivage de PABP par la protéase 3C du poliovirus provoque la diminution de la traduction virale ce qui permet de stopper la production protéique pour mobiliser le transcrit et former de nouvelles particules virales (Ichihashi et al. 2008; Boersma et al. 2020). Ce clivage de PABP a également été observé lors d'infections par les lentivirus et les calicivirus ou encore par le virus de l'hépatite A (Kuyumcu-Martinez et al. 2004; Zhang et al. 2007).

Quelques études ont mis en évidence un clivage d'eIF5B par la protéase 3C de certains entérovirus (de Breyne et al. 2008a). Ce clivage pourrait participer également dans le « shutoff » cellulaire des ARNm. La sous-unité du facteur eIF3 est ciblée par

la protéase du VIH-1 (Jäger et al. 2011). Le facteur eIF3 jouerait un rôle dans les premières étapes de l'infection potentiellement en inhibant la transcription inverse (Valente et al. 2009), un effet qui pourrait être contrecarré par le clivage d'eIF3d par la protéase virale.

b) Relocalisation des facteurs d'initiation

L'infection virale peut s'accompagner d'une perturbation de la localisation cellulaire des eIF avec des conséquences sur la traduction des ARN cellulaires et viraux. Voici quelques exemples employés par les virus. Il a été montré qu'eIF4E pouvait être relocalisé au noyau suite au clivage d'eIF4G par la protéase virale 2A du poliovirus ce qui inhibe encore davantage la traduction des ARNm cellulaires (Sukarieh et al. 2010). Lors de l'infection par le virus de Bunyamwera, la protéine PABP est adressée au noyau lors des phases tardives par un mécanisme impliquant la nucléocapside virale ce qui provoque la baisse de l'expression des ARNm cellulaires. De manière intéressante, la traduction des ARNm viraux, qui sont coiffés mais n'ont pas de queue poly(A), n'est pas affectée par la relocalisation de PABP (Blakqori et al. 2009). La localisation de PABP est également modulée pendant la phase lytique de l'infection par le virus Epstein-Barr. L'expression des protéines virales BGLF5 et ZEBRA, provoque sa relocalisation au noyau ce qui entraîne l'inhibition globale de la traduction (Park et al. 2014).

Certaines infections, telles que celles induites par les poxvirus ou bien le virus de la peste porcine africaine, présentent une relocalisation des facteurs eIF4E et eIF4G à proximité de compartiments cytoplasmiques qui concentrent à la fois les ARNm viraux et une population dense en ribosomes pour traduire activement et sélectivement les protéines virales (Katsafanas and Moss 2007; Walsh et al. 2008; Castelló et al. 2009b; Zaborowska et al. 2012).

c) Contrôle de l'expression des facteurs d'initiation

L'infection virale peut également agir directement sur le transcriptome de la cellule en régulant, entre autres, le niveau d'expression de certains facteurs d'initiation. Ce type de régulation est fréquent chez les virus à ADN qui transitent par le noyau de la cellule

hôte. Ainsi, dès 24 heures après infection par le cytomégalo virus, le niveau d'expression des composants du complexe eIF4F (eIF4E, eIF4A et eIF4G) et PABP est augmenté ce qui favorise la disponibilité du complexe eIF4F pour la traduction des ARN viraux (Walsh et al. 2005; McKinney et al. 2014). La protéine LMP1 du virus Epstein-Barr est un facteur de transcription qui stimule l'expression d'eIF4E et pourrait participer à l'activité oncogène induite par le virus Epstein-Barr (Zhao et al. 2014). La régulation de l'expression des facteurs d'initiation peut également s'exercer au niveau post-transcriptionnel par l'action de miRNAs. Par exemple, l'entérovirus 71 induit une hausse de l'expression du miR-141 qui est dirigé contre l'ARNm codant pour eIF4E. Il en résulte une diminution de eIF4E au cours de l'infection qui participe à l'inhibition globale de la traduction des ARNm sans affecter la production virale qui est sous le contrôle d'une séquence IRES (Ho et al. 2011).

2.3 Le contrôle de l'interaction 4E-BP/4E

4E-binding protein (4E-BP) est un acteur important de la régulation traductionnelle car elle séquestre eIF4E en interagissant avec ce dernier. Ainsi, eIF4E ne fait plus partie du complexe eIF4F ce qui induit une diminution de la traduction coiffe-dépendante (Pause et al. 1994a) (Figure 22). Il existe trois isoformes de 4E-BP chez les mammifères qui partagent 40 à 50% d'homologie de séquences (Poulin et al. 1998). 4E-BP1 est l'isoforme la mieux caractérisée et possède deux domaines de liaison à eIF4E. Le premier est similaire au site d'interaction de eIF4G sur eIF4E, permettant ainsi aux deux protéines (4E-BP et eIF4G) d'entrer en compétition pour la liaison à eIF4E (Mader et al. 1995). Le second domaine renforce l'affinité de 4E-BP pour eIF4E (Peter et al. 2015).

L'interaction de 4E-BP avec eIF4E est dépendante de l'état de phosphorylation de 4E-BP, qui possède 7 sites de phosphorylation identifiés. La forme hypo-phosphorylée a une affinité importante pour eIF4E et séquestre ce dernier (Khaleghpour et al. 1999; Peter et al. 2015) alors que la forme hyper-phosphorylée n'est plus en mesure de lier eIF4E (Pause et al. 1994a; Gingras et al. 2001). La liaison de 4E-BP n'induit pas de changement de conformation de eIF4E et n'entrave pas sa capacité à fixer la coiffe des ARNm mais, en revanche, cela l'empêche d'interagir avec eIF4G (Haghighat et

al. 1995). L'état de phosphorylation de 4E-BP régule donc la disponibilité de eIF4E pour initier la traduction coiffe dépendante.

Ce jeu subtil de phosphorylation est régulé par la voie AKT/PI3K/mTORC1 (Gingras et al. 1998). La molécule rapamycine agit en inhibant la kinase mTOR impliquée dans cette voie de signalisation cellulaire. Lorsque la voie mTOR est inhibée par la rapamycine, cela entraîne une diminution de la phosphorylation de 4E-BP et une inhibition de la traduction coiffe dépendante en présence de rapamycine (Beretta et al. 1996; Fadden et al. 1997; Gingras et al. 1998). Lorsqu'il est activé, le complexe mTORC1 recrute 4EBP via la sous unité raptor (Schalm et al. 2003), et la phosphoryle de façon hiérarchique (Gingras et al. 2001).

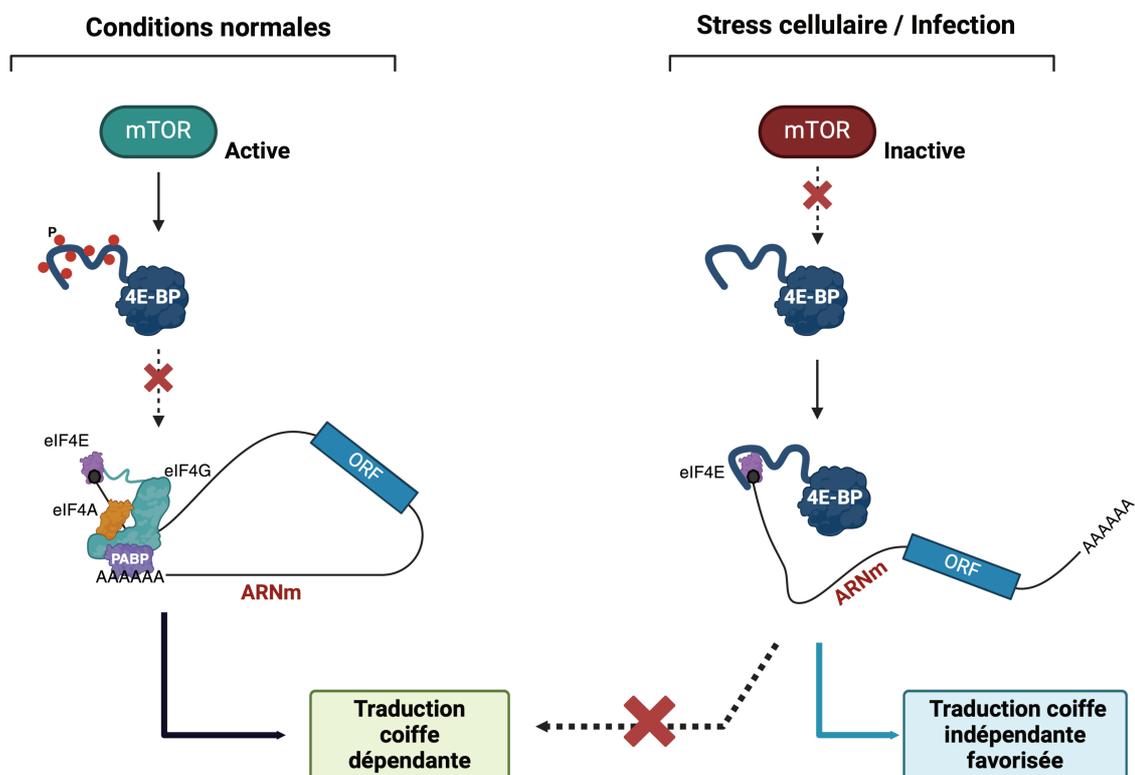


Figure 22. – Contrôle traductionnel médié par 4EBP

L'interaction de eIF4E avec eIF4G peut être interrompue par les protéines 4E-BP. Ces protéines entrent en compétition avec eIF4G pour la liaison à eIF4E, ce qui inhibe la formation du complexe eIF4F et, par conséquent, la traduction dépendante de la coiffe. En conditions normales, mTOR hyperphosphoryle 4E-BP, ce qui la maintient dans une forme inactive. À l'inverse, en conditions de stress, l'activité de mTOR est réprimée, et les protéines 4E-BP se retrouvent dans une forme hypophosphorylée, favorable à leur liaison inhibitrice à eIF4E. Dans ce cas, la traduction indépendante de la coiffe de certains ARN spécifiques peut être favorisée. Dessinée sur BioRender.

Ce n'est pas surprenant que toutes les infections virales affectent cette voie de signalisation afin de moduler l'expression des ARNm (Le Sage et al. 2016). Par exemple, les infections par EMCV et PV (picornavirus) sont associées à une hypophosphorylation de 4E-BP conduisant à une baisse de la traduction cellulaire. (Gingras et al. 1996). (Dunn and Connor 2011); le mécanisme précis reste à caractériser. D'autres infections, telles que celles du virus du papillome humain ou du virus de l'hépatite C, ciblent la voie mTOR afin d'induire une hyperphosphorylation de 4E-BP favorisant l'initiation de la traduction coiffe dépendante (Spangle and Mûnger 2010).

Enfin, l'activité d'eIF4E peut être directement modulée par la phosphorylation au niveau de la Sérine 209 par la kinase Mnk1 (Waskiewicz et al. 1999). Mnk1, de par son interaction avec eIF4G (Figure 4), a une localisation proche d'eIF4E et cette interaction est requise pour sa fonction. Cette kinase est stimulée de deux manières : par la voie des MAPK, stimulée par des stress cellulaires, et par la voie Ras/Raf répondant aux hormones de croissance, aux cytokines et aux mitogènes (Waskiewicz et al. 1997). La phosphorylation de eIF4E a des conséquences significatives sur la régulation de la traduction des ARNm. Cette modification post-traductionnelle peut favoriser la traduction de certains ARN, souvent associés à des gènes régulant la croissance cellulaire, la survie et le stress (Bramham et al. 2016). Les conséquences exactes de cette phosphorylation ne sont pas totalement élucidées.

Les adénovirus interfèrent dans cette régulation par l'intermédiaire de la protéine virale p100K qui interagit avec eIF4G et empêche son interaction avec Mnk1 (Cuesta et al. 2004). Cela induit une diminution de la phosphorylation d'eIF4E et la baisse de la traduction pendant la phase tardive de l'infection. À l'inverse, d'autres virus activent la phosphorylation d'eIF4E afin de réguler positivement la traduction de certains groupes d'ARN. C'est notamment le cas lors de l'infection par le norovirus murin (Royall et al. 2015), l'herpès simplex (Walsh and Mohr 2004), ou encore de l'herpès humain de type 8 (Arias et al. 2009).

2.4 Le contrôle de eIF2 α et PKR

a) Présentation générale

eIF2, composé de 3 sous unités α , β et γ , est impliqué dans la reconnaissance du codon AUG. Lors d'un stress cellulaire, la sous unité α est la cible de différentes kinases qui phosphorylent la sérine située en position 51 (Price et al., 1991; Ernst et al., 1980) ce qui empêche le recyclage du complexe eIF2-GDP en eIF2-GTP par eIF5B. Cela se traduit par une baisse globale de la traduction cellulaire et l'accumulation des ARNm dans des granules de stress (Ivanov et al. 2019).

Chez les humains, quatre kinases sont en mesure de phosphoryler eIF2 α : la kinase GCN2 (*General Control Nonderepressible 2*) activée par la privation en acides aminés (Harding et al. 2000; Sood et al. 2000); la kinase HRI (*Heme Regulated inhibitor*) induite par la baisse de la quantité d'hème disponible (Grayzel et al. 1966); la kinase PKR (*Proteine Kinase R*) en réponse à la détection d'ARN double brin, souvent signe d'une infection virale (Harding et al. 2000); et la kinase PERK (*Protein Kinase RNA-Like Kinase*) qui est activée suite à l'accumulation de protéines mal repliées au niveau du réticulum endoplasmique (Shi et al. 1998).

L'infection virale induit généralement l'expression des kinases PKR et PERK par la présence en grande quantité d'ARN double brin lors de la réplication virale et/ou de protéines incorrectement repliées. En conséquence, eIF2 α est phosphorylé et la traduction s'arrête (García et al. 2006); c'est pour contourner cette première ligne de défense que de nombreux virus ont mis en place des stratégies de contournement ou d'interférence.

b) Résistance à l'inactivation d'eIF2

Certains ARNm viraux ont la capacité d'initier la traduction en absence d'eIF2 ou bien dans des conditions où la sous-unité alpha d'eIF2 se retrouve phosphorylée. Nous pouvons citer l'IRES du CrPv dont l'initiation de la traduction ne requiert aucun eIF (Chapitre 2 section 1.4) et elle est donc peu sensible à la variation de l'activité d'eIF2 (Welnowska et al. 2011). Il a été montré que les IRES de HCV et de CSFV (IRES de type III) ont la capacité d'initier la traduction quand le facteur eIF2 est inactif (Pestova et al. 2008; Terenin et al. 2008). Des facteurs d'initiation alternatifs, tel que eIF2A ou

eIF2D, pourraient alors se substituer à eIF2 pour charger le tRNAi sur le ribosome (Robert et al. 2006; Kim et al. 2011). Les ARNm du virus Sindbis peuvent également être traduits efficacement lorsqu'eIF2 est inactif grâce à la présence d'une structure ARN de type tige-boucle localisée 24nt en aval du codon d'initiation. Cette structure ralentirait le ribosome 43S au niveau du codon AUG pour permettre l'appariement codon/anticodon (Ventoso et al. 2006; Toribio et al. 2016; Sanz et al. 2017).

Enfin, certains virus (virus de la peste porcine africaine et virus de l'herpès simplex) ont la capacité de contrebalancer la phosphorylation d'eIF2 α en recrutant spécifiquement la phosphatase cellulaire PP1 qui peut déphosphoryler eIF2 α et lui rendre son activité (Mulvey et al. 2003; Rojas et al. 2015).

c) Stratégie d'échappement à PKR

La kinase PKR est souvent ciblée par l'expression de protéines virales comme la protéine E3 du virus de la vaccine (White and Jacobs 2012) ou bien de la protéine ORF57 du virus de l'herpès qui l'empêche de se lier à sa protéine activatrice PACT (Sharma et al. 2017). Certains virus neutralisent PKR en exprimant des pseudosubstrats qui interfèrent avec son site de liaison à eIF2, comme la protéine PK2 des baculovirus (Li et al. 2015) ou la protéine K3L des poxvirus (Seo et al. 2008). Dans le cas du cytomégalo virus humain, l'expression de la protéine virale TRS1 induit la relocalisation de PKR au niveau du noyau (Hakki et al. 2006) tandis que la protéine NSs du virus de la fièvre de la vallée du Rift l'adresse au protéasome pour qu'elle soit dégradée (Mudhasani et al. 2016).

Une autre stratégie virale consiste à produire des ARN non codants qui interfèrent avec PKR. À titre d'exemple, on peut citer les ARN EBER synthétisés par le virus d'Epstein-Barr (Iwakiri et al. 2009), ou encore les ARN VA, exprimés par les adénovirus (Launer-Felty et al. 2015; Vachon and Conn 2016). Ces ARN produits par l'ARN polymérase III interagissent fortement avec PKR sans l'activer mais ils saturent le site de reconnaissance de la kinase qui ne reconnaît plus les ARN viraux (Howe and Shu 1989; Vuyisich et al. 2002; Iwakiri et al. 2009; Launer-Felty et al. 2015; Zerbe and Cole 2020).

Enfin, le virus hépatite murin (appartenant à la famille des coronavirus) adopte une stratégie alternative dans laquelle l'activité endoribonucléase, portée par Nsp15,

dégrade les ARN doubles brin produits avant qu'ils n'activent PKR (Kindler et al. 2017; Deng and Baker 2018). Un mécanisme similaire a été décrit pour l'endonucléase VHS du virus de l'herpès (Sciortino et al. 2013; Dauber et al. 2019).

d) Formation des granules de stress et stratégies virales d'échappement

Les granules de stress, sont des agrégats cytoplasmiques qui se forment en réponse au stress cellulaire et qui contiennent divers composants, dont des ARNm, des protéines de liaison à l'ARN et des composants de la machinerie traductionnelle (Ivanov et al. 2019) (Protter and Parker 2016). Cela s'accompagne souvent de la phosphorylation d'eIF2 alpha.

Les infections virales induisent très souvent la formation de granules de stress et beaucoup de virus ont mis en place des mécanismes de contournement en agissant sur l'activation de la voie PKR/eIF2 α (Finnen et al. 2014; Dauber et al. 2016; Rabouw et al. 2016; Amorim et al. 2017; Sharma et al. 2017) ou en ciblant directement des constituants clés des granules de stress. C'est le cas du virus Ébola et le SARS-CoV-2 dont les protéines respectives VP35 et Nsp5 interfèrent avec les protéines structurantes G3BP1 et G3BP2 des granules de stress (Le Sage et al. 2017) (Zheng et al. 2022). Enfin, certaines protéases picornavirales, telles que la protéase 3C du EMCV (Ng et al. 2013) et du PV (White et al. 2007) dégradent directement des composants des granules de stress incluant G3BP1

Chapitre 3 – Le contrôle traductionnel du SARS-CoV-2

Au début de ma thèse, j'ai activement contribué, avec le reste de notre équipe, à la rédaction d'une revue abordant le contrôle traductionnel exercé par les différents coronavirus (de Breyne et al. 2020). Celle-ci est présentée en annexe. Dans ce chapitre, je vais donc me concentrer plus spécifiquement sur les mécanismes liés au SARS-CoV-2.

1 Le virus du SARS-CoV-2

1.1 Présentation générale du SARS-CoV-2

Le SARS-CoV-2, qui appartient à la famille des *Coronaviridae* et a été identifié pour la première fois en décembre 2019. Il est rapidement devenu une préoccupation mondiale en raison de la pandémie dévastatrice qu'il a provoqué. À ce jour, presque 800 millions de personnes ont été infectées et 7 millions de personnes en sont décédées d'après les données de l'Organisation Mondiale de Santé (avril 2024).

Il est l'agent causal de la maladie COVID-19, caractérisée par une gamme étendue de symptômes allant de légers à graves. La maladie peut notamment se manifester par des symptômes respiratoires aigus mettant en péril la vie des personnes présentant ces tableaux cliniques sérieux. Il a été montré que le risque de développer une forme sévère de la maladie augmente avec l'âge et la présence de comorbidités.

Les coronavirus sont classés en quatre genres : alpha, bêta, gamma, et delta (Figure 23). Parmi eux, on retrouve sept coronavirus infectant les humains : deux alphacoronavirus (HCoV-229E et HCoV-NL63) et cinq bêta-coronavirus (HCoV-OC43, HCoV-HKU1, MERS-CoV, SARS-CoV-1, et SARS-CoV-2), responsables de maladies respiratoires à gravité variable. Le SARS-CoV-2 partage une importante homologie de séquence avec le SARS-CoV-1 (80%) et le MERS-CoV (50%) (Kim et al. 2020a; Kim et al. 2020b).

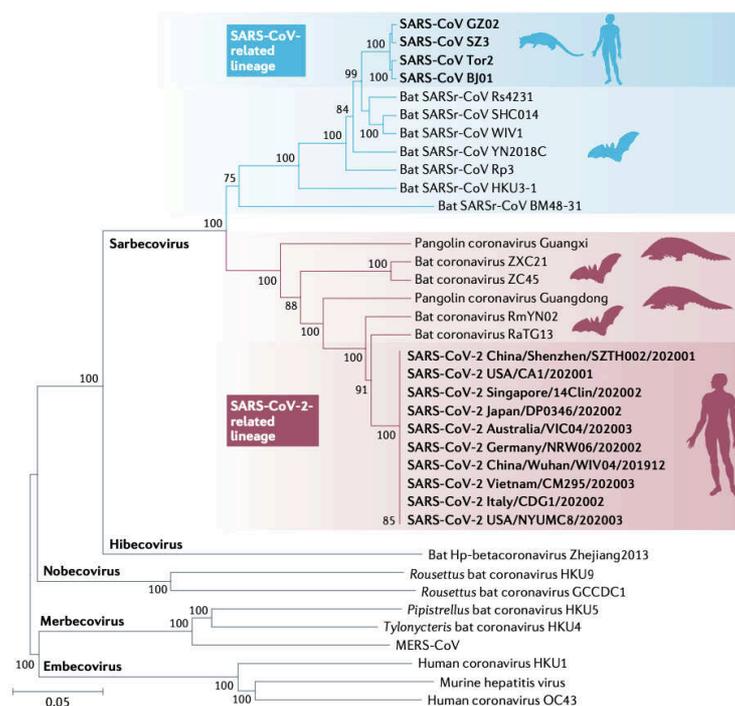


Figure 23. – Arbre phylogénétique

Arbre phylogénétique des séquences génomiques complètes du SARS-CoV-2, des SARS-CoV et d'autres β -coronavirus Adapté à partir de (Hu et al. 2021). La construction a été réalisée par la méthode de jonction de voisinage à l'aide du programme MEGA6, avec des valeurs de bootstrap calculées à partir de 1 000 arbres.

Le SARS-CoV-2 est un virus enveloppé possédant un génome à ARN simple brin positif d'environ 30 kilobases. Les deux tiers de la partie 5' de l'ARNg codent pour les polyprotéines ORF1a et ORF1ab qui sont ensuite clivées en protéines non structurales (Nsp1 à Nsp16) (Figure 24). L'ORF1b est produite par un événement de *frameshifting* à la fin de la synthèse de l'ORF1a. L'ARN génomique, qui est coiffé et polyadénylé, sert directement de support à la traduction des ORF1a et 1ab dès le début de l'infection. Les protéines structurales, Spike (S), Enveloppe (E), Membrane (M), et Nucléocapside (N) et les protéines accessoires, sont codées dans la partie 3' du génome viral et leur expression nécessite le passage par l'intermédiaire d'ARN sous-génomique. Au fil du temps, le SARS-CoV-2 a révélé un fort potentiel de mutation, aboutissant à l'apparition de multiples variants. Cette situation a complexifié l'élaboration de vaccins et de traitements antiviraux. Malgré cela, des vaccins innovants basés sur l'ARN messenger ont été développés, marquant une avancée significative dans la lutte contre la pandémie en adaptant rapidement les réponses immunitaires aux nouvelles souches virales.

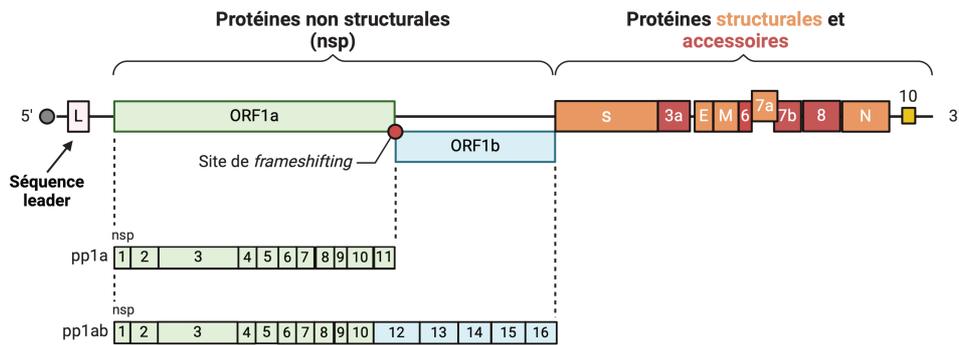


Figure 24. – Organisation du génome viral du SARS-CoV-2

Les deux tiers 5' du génome viral codent pour les protéines non structurales, qui sont traduites directement à partir de l'ARN génomique. Le tiers 3' du génome viral, encode quant à lui les protéines structurales et accessoires. Ces dernières sont synthétisées grâce à un mécanisme de transcription complexe impliquant des ARN sous-génomiques. Dessinée et adaptée sur BioRender.

1.2 Résumé du cycle de réplication virale

Comme pour tout virus, le cycle de réplication virale du SARS-CoV-2 est très complexe et il a été détaillé dans les revues suivantes (Hu et al. 2021; V'Kovski et al. 2021; V'kovski et al. 2021; Steiner et al. 2024). Un bref résumé du cycle est décrit ci-dessous (Figure 25).

La première étape du cycle viral implique l'attachement du virus à la cellule hôte ; ce processus est médié par une interaction entre la glycoprotéine virale Spike et son récepteur : l'enzyme de conversion de l'angiotensine 2 (ACE2). La protéine ACE2 est exprimée dans divers types de cellules humaines, mais elle est particulièrement abondante dans les cellules épithéliales des voies respiratoires (Hikmet et al. 2020; Beyerstedt et al. 2021).

L'entrée de la particule dans les cellules hôtes peut se dérouler selon deux mécanismes distincts, la fusion à la membrane ou la voie endosomale. Bien que moins efficace, l'entrée médiée par les endosomes peut jouer un rôle important dans l'infection, en particulier dans les conditions cellulaires où les protéases de surface nécessaires à la fusion membranaire ne sont pas abondantes. (Zhao et al. 2021b; Fraser et al. 2022a; Jackson et al. 2022; Lavie et al. 2022).

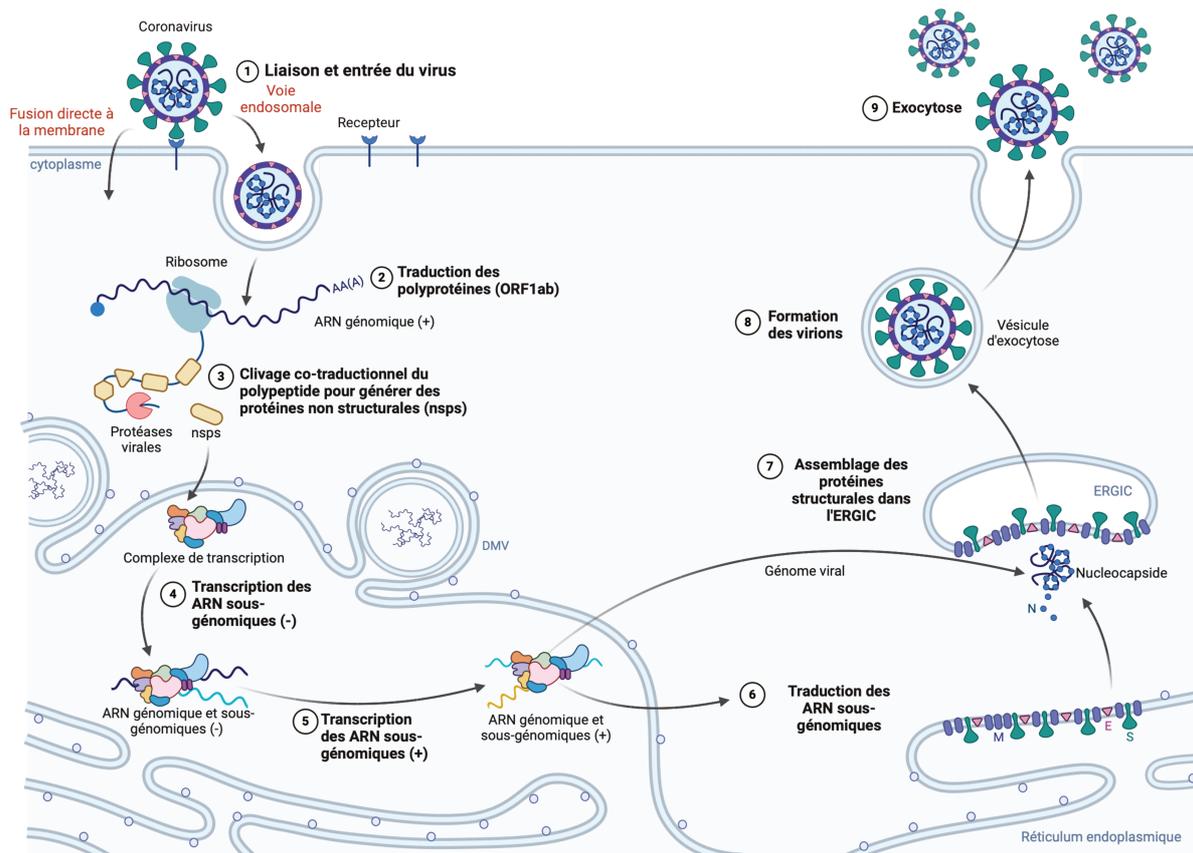


Figure 25. – Cycle de réplication virale du SARS-CoV-2

Le cycle de réplication du SARS-CoV-2 débute par l'entrée du virus dans la cellule hôte, suivie de la libération de son ARN génomique. Celui-ci est traduit pour produire les polyprotéines. Ces dernières sont clivées en protéines non structurales, qui jouent un rôle dans la réplication de l'ARN viral et la formation d'ARN sous-génomiques. Ces ARN servent de matrices pour la synthèse des protéines virales structurales et accessoires. Enfin, les virions assemblés sont relâchés hors de la cellule par exocytose, prêts à infecter d'autres cellules. Dessinée et adaptée sur BioRender.

Quel que soit la voie d'entrée utilisée, deux clivages successifs de la protéine Spike par des enzymes cellulaires sont nécessaires (Figure 26). Le premier est médié par la furine et se déroule dans les cellules productrices (Lavie et al. 2022). Le second dépend de la voie d'entrée. Il est médié soit par la protéase transmembranaire à sérine 2 (TMPRSS2) pour la voie de fusion à la membrane (Jackson et al. 2022) soit par la cathepsine L pour la voie endosomale (Zhao et al. 2021b; Fraser et al. 2022b). Ces clivages permettent de dévoiler le domaine « peptide fusion » de spike pour qu'il s'insère dans la membrane cellulaire (ou celle de l'endosome). Cette étape facilite ainsi la fusion des membranes virales et cellulaires pour permettre, l'entrée dans la cellule hôte.

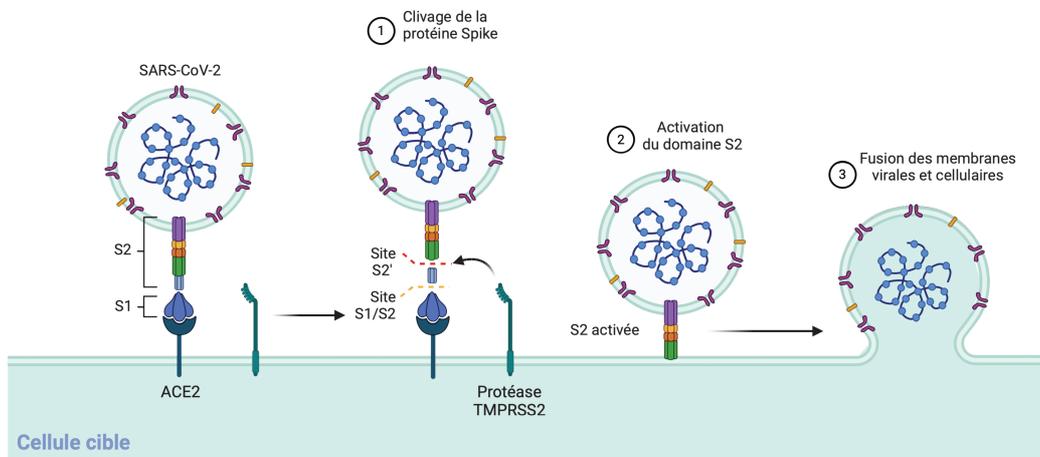


Figure 26. – Mécanisme de fusion du SARS-CoV-2

Pour permettre l'entrée du SARS-CoV-2, deux étapes de clivage de la protéine Spike sont requises. Initialement, un premier clivage survient à la jonction S1/S2, effectué au sein même des cellules productrices. Ensuite, un second clivage se produit au niveau du site S2', orchestré par la protéase TMPRSS2, suite à la liaison du virion avec ACE2 dans le cadre de la voie d'entrée directe. Dessinée et adaptée sur BioRender.

Le génome à ARN positif est directement reconnu dans le cytoplasme par la machinerie traductionnelle pour la synthèse des deux polyprotéines virales pp1a et pp1ab. Grâce à un mécanisme de clivage protéolytique orchestré par la protéase Mpro (également appelée 3CLpro) et la protéase de papain-like (PLpro), 16 protéines non structurales (nsp 1 à 16) sont générées (de Breyne et al. 2020) (Figure 27).

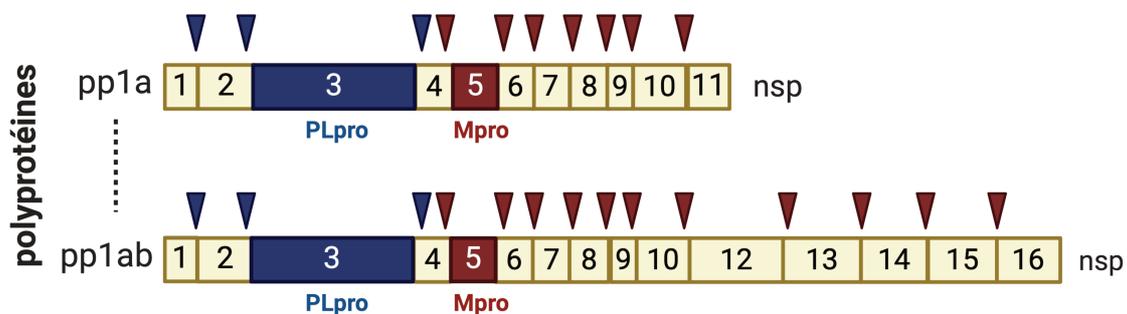


Figure 27. – Représentation des polyprotéines du SARS-CoV-2

Lors de l'infection les deux polyprotéines encodées par les ORF1a et ORF1ab sont autoclivées par les protéines virales PLpro et Mpro. Dessinée sur BioRender.

La réplication du génome viral et la transcription des ARN sous-génomiques sont orchestrées par un complexe de réplication-transcription (RTC) formé par les protéines Nsp néosynthétisées et ont lieu dans des compartiments spécifiques dérivés de l'appareil de Golgi et du réticulum endoplasmique. Ces structures membranaires créent un environnement propice à la réplication virale, tout en protégeant le génome des défenses immunitaires de l'hôte (Klein et al. 2020). Au sein du complexe RTC, il va y avoir la synthèse d'anti-génomomes ou ARN complémentaires de brin négatif à partir de l'ARN génomique qui seront utilisés comme matrice pour la synthèse de nouveaux brins d'ARN génomique (ARNg) (réplication) et sous-génomiques (ARNsg) (transcription). Les différents ARNsg codent pour les protéines de structure S, E, M et N ainsi que pour les protéines accessoires ou régulatrices 3a, 6, 7a, 7b, 8, 9b et 10. Nous n'aborderons pas le rôle individuel de chaque protéine virale mais uniquement celles qui influencent la traduction cellulaire et virale; les autres sont présentées à titre d'information dans les tableaux 2, 3 et 4 (voir plus bas).

Pour finir, l'assemblage des nouveaux virions se produit à l'intérieur d'un compartiment dérivé du réticulum endoplasmique (appelé ERGIC) où la protéine de nucléocapside N s'associe à l'ARNg pour le diriger au sein des particules virales en formation grâce aux autres protéines de structure. Les particules virales formées sont ensuite transportées à la surface cellulaire et libérées par exocytose pour entamer un nouveau cycle d'infection (Scherer et al. 2022).

Tableau 2. – Tableau récapitulatif des protéines structurales du SARS-CoV-2

	Protéines	Rôle	Longueur en acides aminés
Protéines structurales	Protéine S	Impliquée dans l'entrée du virus en se liant au récepteur ACE2	1273 aa
	Protéine E	Petite protéine impliquée dans l'assemblage et la libération des virions	75 aa
	Protéine M	Facilite l'assemblage des virions	222 aa
	Protéine N	Se lie à l'ARN viral et forme un complexe ribonucléoprotéique. Essentiel à la réplication et transcription virale	419 aa

Tableau 3. – Tableau récapitulatif des protéines Nsp du SARS-CoV-2

Protéines	Rôle	Fonction moléculaire	Poids moléculaire (en kDa)	
Protéines non structurales (Nsp)	Nsp1	Inhibition de la traduction et dégradation des ARNm de l'hôte. Inhibition de la réponse immunitaire.	Liaison à la sous-unité 40S du ribosome	19,8
	Nsp2	Rôle incertain. Potentiellement impliqué dans la réplication et la transcription virale.	//	70,5
	Nsp3	Clivage des polyprotéines virales. Essentielle pour la réplication et la transcription du virus. Inhibition de la réponse immunitaire de l'hôte.	Activité protéasique de type papaine (PLpro)	217,3
	Nsp4	Formation et maintien des organelles de réplication-transcription du virus en modifiant les membranes cellulaires.	Formation de vésicules à double membrane	56,2
	Nsp5	Clivage des polyprotéines pp1a et pp1ab en Nsp matures. Inhibe la signalisation de la voie de l'interféron.	Activité protéasique de type 3C (Mpro)	33,8
	Nsp6	Formation des complexes de réplication-transcription en modifiant les membranes cellulaires. Inhibe la formation d'autolysosomes.	Formation de vésicules à double membrane	33
	Nsp7	Co-facteur de l'ARN polymérase ARN-dépendante (RdRp, Nsp12).	//	9,2
	Nsp8	Co-facteur de l'ARN polymérase ARN-dépendante (RdRp, Nsp12).	//	21,9
	Nsp9	Implication potentielle dans la réplication de l'ARN viral.	Protéine de liaison à l'ARN	12,4
	Nsp10	Association et activation des protéines nsp14 et nsp16.	//	14,8
	Nsp11	Non connu, Expression non avérée	//	1,3
	Nsp12	ARN polymérase ARN-dépendante essentielle pour la réplication du génome viral. Implication dans la formation de la coiffe virale	ARN polymérase ARN-dépendante Guanylyltransférase	106,7
	Nsp13	Nécessaire pour la réplication et la transcription de l'ARN viral en déroulant les structures d'ARN double brin. Implication dans l'étape initiale de l'ajout de la coiffe par l'hydrolyse de l'extrémité 5' de l'ARN.	Hélicase 5' triphosphatase	66,9
	Nsp14	Exonucléase corrigeant les erreurs dans l'ARN viral lors de la réplication, assure la fidélité de la réplication virale. Implication dans le processus de la formation de la coiffe virale	Activité de relecture 3'-5' exonucléase N7-méthyltransférase	59,8
	Nsp15	Impliquée dans la dégradation des ARN viraux et l'échappement à la détection par le système immunitaire	Endoribonucléase virale spécifique des U	38,8
	Nsp16	Méthyle l'extrémité 5' des ARNs viraux pour générer la coiffe de type 1	Ribose 2'-O-méthyltransférase (2'-O-Mtase)	33,3

Tableau 4. – Tableau récapitulatif des protéines accessoires du SARS-CoV-2

Protéines	Rôle	Longueur en acides aminés	
Protéines accessoires	ORF 3a	Inhibition de la formation d'autophagosomes, induction de la mort cellulaire et effet antiprolifératif	275 aa
	ORF 6	Interférence avec l'import nucléaire de facteurs de transcription Implication dans la réponse antivirale de l'hôte	61 aa
	ORF 7a	Implication dans la modulation de la réponse immunitaire	121 aa
	ORF 7b	Inhibition la croissance cellulaire et induction la l'apoptose Inhibition de la signalisation de l'interféron	43aa
	ORF8	Inhibition de l'expression des molécules du complexe majeur d'histocompatibilité type I Contribution à l'évasion du système immunitaire	121 aa
	ORF 10	Fonction incertaine	38 aa

2 Organisation du génome viral

2.1 Les caractéristiques des transcrits du SARS-CoV-2

L'ensemble des transcrits viraux produits lors de l'infection sont coiffés et polyadénylés grâce à l'action enzymatique de plusieurs protéines Nsp (Park et al. 2022). Ils ont tous en commun une région appelée leader longue de 75 nt à l'extrémité 5' du messager formant le début de la 5'UTR et une séquence à l'extrémité 3' longue de 229 nt qui constitue la région 3'UTR. Cette caractéristique résulte de l'étape de

réplication/transcription assurée entre autre par la protéine virale Nsp12 qui est une ARN polymérase ARN-dépendante.

En effet, durant cette étape, l'ARNg sert de matrice pour sa réplication et pour la transcription des ARNsg. Chaque processus nécessite des modèles dédiés de brin négatif : l'anti-génome (ou ARN complémentaire) et un ensemble d'ARNsg (ou ARN complémentaires sous génomiques), respectivement. Cette première étape copie l'ARNg à partir de son extrémité 3' vers l'extrémité 5'. Ainsi, la synthèse de l'ARNg implique un mécanisme de transcription continue de l'anti-génome. En revanche, la production des ARNsg requiert un mécanisme complexe de transcription discontinue (figure 28) dans laquelle le RTC met en pause la synthèse d'ARN après avoir copié l'une des séquences régulatrices de transcription appelées TRS-B, et se détache du modèle. Le RTC se déplace ensuite vers une position proche de l'extrémité 5' du modèle de l'ARNg, où le complément du TRS-B (anti-TRS-B) de l'ARNsg de brin négatif naissant s'apparie avec les bases du motif TRS-L (L pour leader) et implique des interactions intra génomiques longue distance. Ce saut du RTC est très important parce qu'il conduit à l'ajout du complément de la séquence de leader de l'ARNg (anti-leader) à l'extrémité 3' de chaque ARNsg de brin négatif, qui seront utilisés comme modèles pour la production des ARNsg. Cela explique pourquoi tous les ARNsg du SARS-CoV-2 partagent une séquence leader 5'-terminale d'environ 75 nucléotides qui est commune aussi à la séquence 5'-terminale de l'ARNg (Malone et al. 2022).

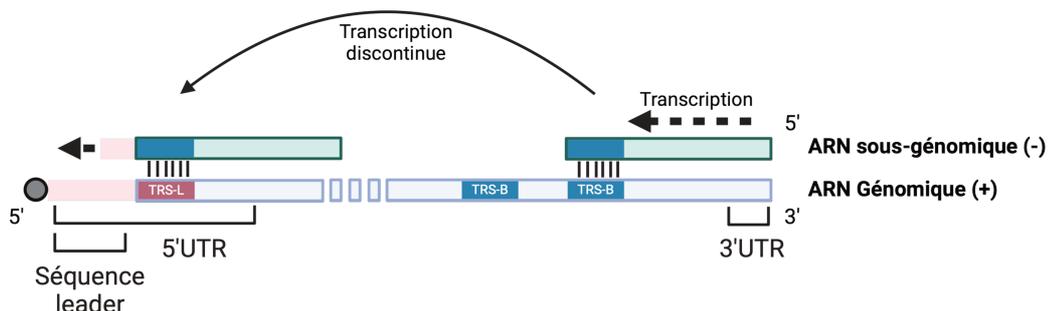


Figure 28. – Modèle de la transcription discontinue

Les ARN sous-génomiques négatifs sont transcrits par un mécanisme de synthèse discontinue. Le complexe de transcription initie la synthèse à l'extrémité 3' de l'ARN génomique et se dirige vers l'extrémité 5'. Rencontrant une séquence régulatrice TRS-B, le complexe peut s'arrêter ou continuer. S'il s'arrête, les interactions ARN/ARN longue distance entre TRS-B et le TRS-L près de la séquence leader en 5' permettent à la polymérase de changer de modèle et d'ajouter la séquence leader aux ARN sous-génomiques. Dessinée sur BioRender.

Ce processus permet d'exprimer les protéines de structure et accessoires, qui sont situées dans la partie 3' du génome viral (à plus de 20kb de l'extrémité 5'), en aval de la région leader (Figure 29). Ce mécanisme conditionne également la quantité de chaque ARNsg produit ; en effet l'abondance de chacun d'entre eux est proportionnelle à la proximité du motif TRS-B propre à chaque région codante avec l'extrémité 3' de l'ARNg (Wang et al. 2021). En effet, plus le motif TRS-B permettant d'initier la transcription discontinue est loin de l'extrémité 3' de l'ARNg, plus la probabilité que la polymérase s'y rende sans identifier d'autres motifs TRS-B est faible. C'est également ce principe qui explique qu'à 24h post-infection, l'ARNg dont la transcription nécessite une transcription continue de l'anti-génome, ne représente que 5% des transcrits totaux du SARS-CoV-2 (Wang et al. 2021).

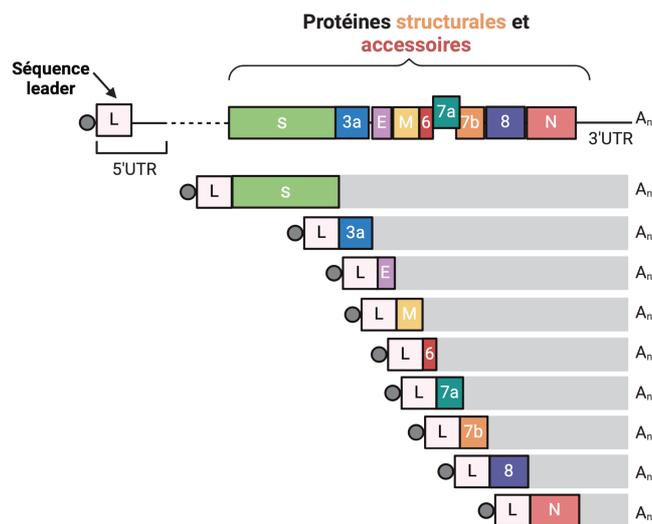


Figure 29. – Représentation des ARN sous-génomiques du SARS-CoV-2

Schéma des ARN sous-génomiques et de l'ARN génomique du SARS-CoV-2. La coiffe et la queue poly(A) sont représentées. La région 5'UTR de chaque ARNm viral est composée d'une séquence leader commune à tous les ARN viraux. Dessinée et adaptée sur BioRender.

La conservation de la région leader sur l'ensemble des transcrits du SARS-CoV-2 suggère que cette séquence a un rôle majeur pour les étapes post-transcriptionnelles (Sola et al. 2015). Il est à noter que les différents motifs TRS-B ne sont pas forcément localisés immédiatement avant le codon d'initiation de chaque ORF. Cette distance varie de 0 à 150 nucléotides et elle est propre à chaque ARNm sous génomique (Figure 30). Ainsi, la 5'UTR de chaque ARNsg est formée par l'association de la

séquence leader de 75nt et d'une région variable propre à chaque ARNsg. La taille de ces 5'UTR est résumée dans le tableau 1 de l'introduction de l'article Conde et al 2021 (section résultat 1). En ce qui concerne la longueur de la 5'UTR de l'ARNg, elle est de 265 nt en incluant les 75 nt de la séquence leader.

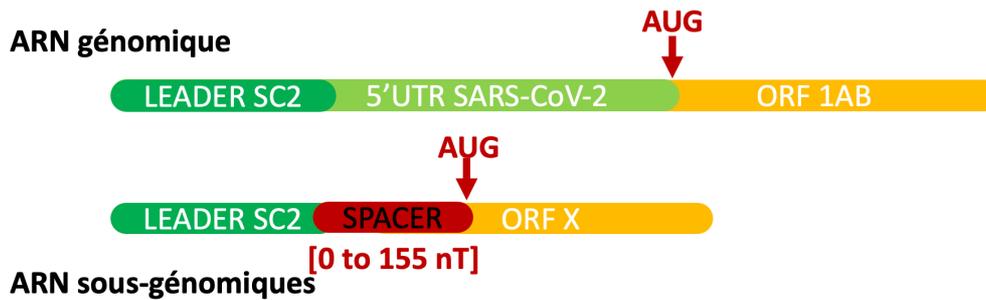


Figure 30. – Organisation des ARN génomiques et sous génomiques du SARS-CoV-2

L'ARN génomique et les ARN sous-génomiques du SARS-CoV-2 présentent une organisation spécifique de leur région 5'UTR. Cette dernière comprend une séquence leader commune de 75 nucléotides. Cette séquence leader est couplée à une séquence variable, qui diffère entre les différents ARN, tant en longueur qu'en composition. (Cette séquence est appelée « spacer » sur ce schéma).

La traduction des ARNm peut être fortement influencée par la région 5'UTR dont les caractéristiques varient entre chaque ARN viral ce qui peut conduire à des mécanismes de contrôle traductionnel spécifiques. Au cours de ma thèse, nous avons porté une attention particulière à l'étude de ces différences et à leur impact sur la traduction des ARNm viraux. En outre, la région 3'UTR des ARN viraux du SARS-CoV-2 est conservée et elle peut également moduler leur traduction ainsi que leur stabilité. Cet aspect n'a toutefois pas été exploré en détail au cours de ma thèse.

2.2 Les éléments fonctionnels des régions 5 'UTR

Comme nous l'avons décrit dans le chapitre 1, la région 5'UTR d'un ARNm peut renfermer des structures de type tige-boucle et/ou pseudonœud, des uORF, ou encore des motifs ARN reconnus par des protéines spécifiques qui régulent la traduction. La structure de la 5'UTR du SARS-CoV-2 a été déterminée initialement par des études *in*

silico (Rangan et al. 2020) puis confirmée par des techniques classiques d’empreinte chimique et enzymatique (Miao et al. 2021) et également à partir d’ARN isolé de virions (Cao et al. 2021) et de cellules infectées (Lan et al. 2022b). La 5’UTR de l’ARNg (figure 31) est constituée de 5 structures en forme de tige-boucle et nommées SL1, SL2, SL3, SL4 et SL5. La séquence leader, commune à tous les ARN viraux, comporte uniquement les structures SL1, SL2 et SL3. SL1 est positionnée à 6 nucléotides de l’extrémité 5’ du messager et a une taille de 27nt incluant une boucle de 4nt et un bulge au milieu de la tige. Cette structure est nécessaire pour que les ARN viraux soit exprimés (Bujanic et al. 2022) dans les conditions d’inhibition globale de la traduction imposées par l’expression de la protéine virale Nsp1 (ceci est discuté dans la suite de ce chapitre).

Les structures SL2 et SL3 sont plus petites et participent à la réplication/transcription du génome (Chen and Olsthoorn 2010; Wacker et al. 2020a). SL3 renferme le motif TRS-L qui assure la présence de la séquence leader sur les transcrits viraux.

SL4 est formée par une longue tige qui est relativement stable d’un point de vue thermodynamique avec un sous-structure (une petite tige boucle SL4.5) qui reste controversée car elle n’a pas été observée dans toutes les publications (Wacker et al. 2020b; Miao et al. 2021; Lan et al. 2022a).

Enfin, SL5 est la plus complexe et appelée “four-way junction” (jonction à quatre voies) formée par SL5a, SL5b et SL5c. Il a été suggéré que les boucles de SL5 pourraient être impliquées dans l’empaquetage du génome dans le virion (Bosch et al. 2003; Masters 2019). Le codon d’initiation des deux polyprotéines est localisé en position 266 et il est imbriqué dans la longue tige formant l’élément SL5 ce qui ne semble pas affecter son utilisation. Le fait que l’ARNg soit coiffé et polyadénylé, suggère l’utilisation d’un mécanisme d’initiation de la traduction dépendant de la coiffe qui permet de dérouler l’ensemble des SL avec une bonne efficacité. Cependant la présence d’une structure ARN de type “jonction à quatre voies” à proximité du codon d’initiation ressemble aux structures d’ARN décrites au niveau de certains IRES dont HCV, EMCV et PV (voir chapitre 2 section 1.4)

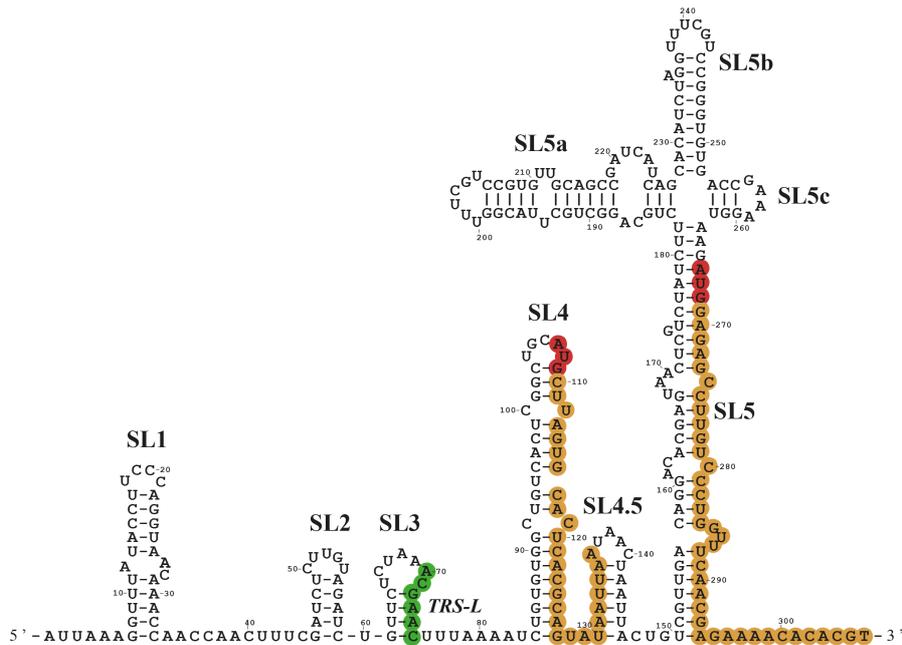


Figure 31. – Structure de la région 5'UTR du SARS-CoV-2

Modélisation de la structure de la région 5'UTR de l'ARNg du SARS-CoV-2. Les différentes structures en tige-boucle (SL1 à SL5) de la région 5'UTR sont illustrées. La position des codons AUG de l'uORF et de l'ORF principal est identifiée en rouge, la région codante en orange et le TRS-L en vert. Cette région 5'UTR contient un codon d'initiation non canonique (CUG-59) entre les structures en tige-boucle SL1 et SL2. Adapté de (Miao et al. 2021)

De manière intéressante, la boucle de SL4 renferme un triplet AUG en position 107 ce qui crée une uORF longue de 27nt. Des expériences de ribosome profiling sur des cellules infectées indiquent que ce codon AUG¹⁰⁷ est reconnu par la machinerie traductionnelle (Finkel et al. 2021a; Kim et al. 2022) et il est conservé chez plusieurs coronavirus incluant le SARS-CoV, le MERS-CoV et le MHV (Wu et al. 2014). Ces mêmes études de ribosome profiling ont montré que des complexes de préinitiation marquent un arrêt sur le triplet CUG localisé à la position 59 et qu'il pourrait donc être utilisé comme codon d'initiation non canonique (Finkel et al. 2021b; Andreev et al. 2022). Ce dernier est en phase avec le codon AUG¹⁰⁷ et ces deux sites potentiels d'initiation pourraient réguler la traduction de pp1 et pp1a ajoutant un degré de complexité dans les mécanismes traductionnels s'opérant sur la 5'UTR de l'ARNg. De par sa localisation dans la région leader, ce codon d'initiation non canonique est présent sur l'ensemble des transcrits viraux. Étant donné que chaque ARNsg possède une séquence variable qui lui est propre, ce codon CUG ne se trouve pas toujours dans le même cadre de lecture par rapport à l'ORF principale (Figure 32). Une étude

récente suggère que ce codon CUG est fréquemment reconnu par les ribosomes avec une fréquence d'utilisation de 20% à 48 heures post infection pour certains ARN sous-génomiques (Finkel et al. 2021b; Kim et al. 2021). Les ORF potentiels, initiés à partir du CUG, présentent des conformations variées. Dans le cas de plusieurs ARNsg (ORF3a, E, M, ORF6, ORF7a, ORF8, et N), ces ORF sont chevauchants. Concernant l'ORF7b, l'uORF créée par le codon CUG est entièrement intégré dans la région 5'UTR. En revanche, pour l'ARNsg qui code pour la protéine Spike, l'uORF est en phase avec l'ORF principale (Bujanic et al. 2022).

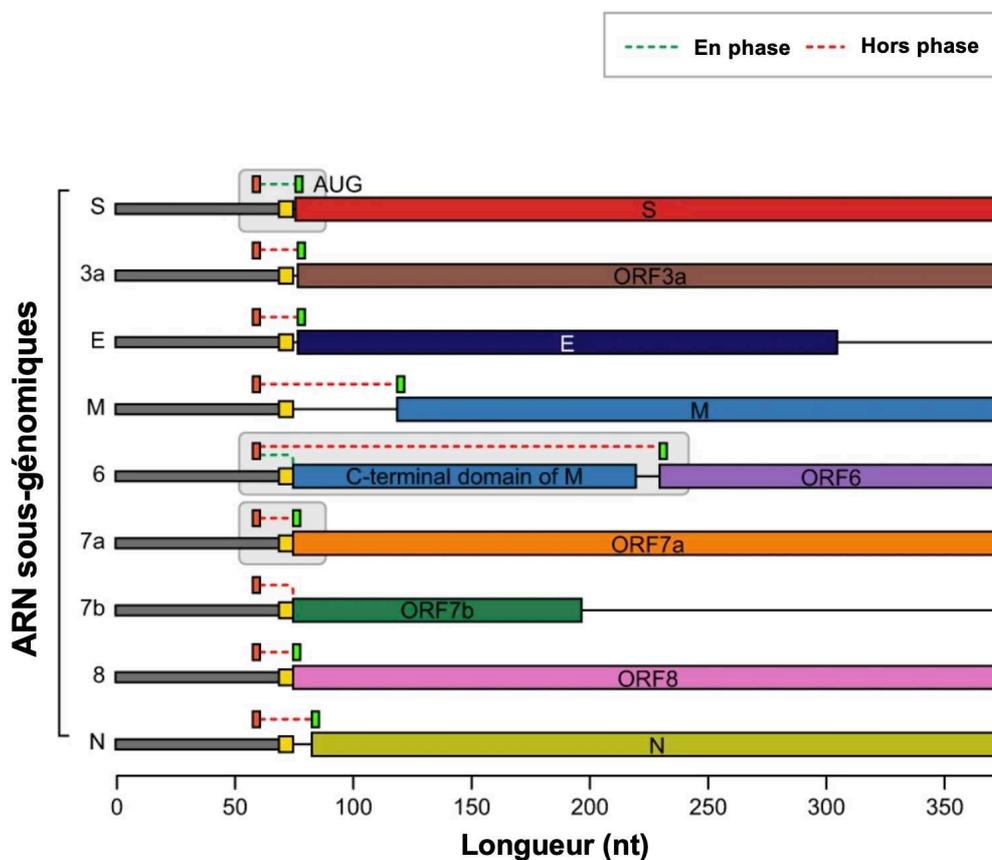


Figure 32. – Position du codon non canonique CUG-59 dans les différentes 5'UTR du SARS-CoV-2

Vue d'ensemble de la position du codon non canonique CUG-59 inclus dans la séquence leader commune par rapport au codon AUG des différents ARN sous-génomiques du SARS-CoV-2. Pour chacun des ARN, le cadre de lecture hypothétique résultant d'une initiation à partir du codon CUG-59 est représenté par une ligne pointillée. Les lignes vertes correspondent à un codon CUG en phase, et celles en rouge, à un codon hors phase, en comparaison avec le cadre de lecture de l'ORF annoté. Adaptée de (Kim et al. 2021)

3 Contrôle de la traduction pendant l'infection par le SARS-CoV-2

Avant de discuter plus en détail du mécanisme de traduction des ARN du SARS-CoV-2 pendant l'infection je vais dans un premier temps aborder l'effet global de l'infection sur le processus de traduction cellulaire.

3.1 Nsp1 : protéine centrale de l'inhibition traductionnelle

Le SARS-CoV-2 a développé des stratégies pour inhiber la traduction des ARNm cellulaires (Schubert et al. 2023) qui permettent à la fois de détourner la machinerie traductionnelle pour la production des protéines virales et de contrer la synthèse des protéines antivirales cellulaires telles que les interférons et celles induites par les interférons (Thoms et al. 2020; Fisher et al. 2022).

Nsp1 est la première protéine virale exprimée lors de l'infection par le SARS-CoV-2, elle est impliquée non seulement dans l'inhibition globale de la traduction des ARNm cellulaires (Schubert et al. 2020b; Thoms et al. 2020) mais également dans leur dégradation (Finkel et al. 2021a) et le blocage de l'export nucléaire des ARNm (Banerjee et al. 2020). Il n'est guère surprenant que sa surexpression en cellules humaines soit toxique (Yuan et al. 2020).

Nsp1 est une protéine de 180 acides aminés (20 kDa) relativement conservée chez les β -coronavirus et partage 84% d'identité avec son homologue du SARS-CoV (Gordon et al. 2020). Nsp1 est constituée de trois domaines distincts : sa partie N-terminale (aa 1-128) présentant une structure globulaire (Clark et al. 2021), suivie d'un linker de 20 aa et son domaine C-terminal plutôt désordonné en solution (Agback et al. 2021; Zhao et al. 2021a; Wang et al. 2023) (Figure 33). Le fragment C-terminal (aa 153 à 179) adopte une conformation structurale sous la forme de deux hélices α qui interagissent avec la sous unité 40S au niveau du site d'entrée des ARNm ce qui bloque son accessibilité pour les ARNm cellulaires et provoque l'inhibition de la traduction (Schubert et al. 2020b; Thoms et al. 2020; Lapointe et al. 2021). Des études de cryo-EM du complexe 40S-Nsp1 ont montré que les deux hélices étaient connectées par une petite boucle avec un motif KH (Lysine/Histidine) qui interagit spécifiquement avec l'hélice h18 de l'ARNr 18S (Schubert et al. 2020b; Thoms et al.

2020). Les résidus des deux hélices interagissent également avec ceux des protéines ribosomiques uS5 et uS3 (Schubert et al. 2020a; Thoms et al. 2020). La mutation du motif KH en alanine (K164A et H164A) empêche l'association de Nsp1 avec la sous unité ribosomique 40S et abroge l'effet inhibiteur sur la traduction (Schubert et al. 2020b; Thoms et al. 2020; Shen et al. 2021). De plus, les données de la structure du complexe 40S/Nsp1 indiquent que le ribosome 40S adopte une conformation fermée qui rend impossible le recrutement d'un ARNm (Yuan et al. 2020) (Figure 34).

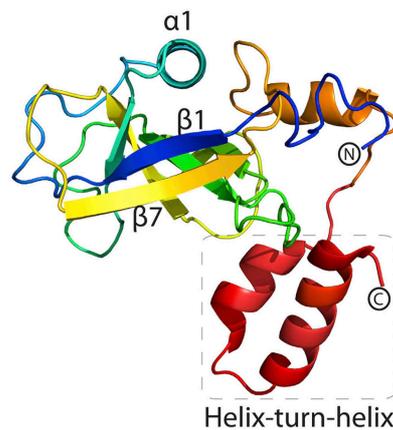


Figure 33. – Structure de la protéine virale Nsp1

Représentation schématique de la structure de la protéine Nsp1 du SARS-CoV-2. La région C-terminale impliquée dans le blocage du canal d'entrée du ribosome est colorée en rouge. Adaptée de (Semper et al. 2021). Référence PDB: 7K3N

De manière intéressante, un variant du SARS-CoV-2 contient une délétion dans la région C-terminale de Nsp1 qui est associée à une faible activation de la voie de l'interféron γ et des formes moins graves de la maladie (Benedetti et al. 2020). Par ailleurs d'autres études suggèrent que des mutations spécifiques du domaine N-terminal de Nsp1 réduiraient l'activité répressive de la traduction de Nsp1 (Mendez 2021)(Schubert et al. 2023; Shehata and Parker 2023). Bien que ce domaine ne soit pas inséré dans le canal d'entrée de l'ARNm du ribosome, il pourrait stabiliser l'interaction de Nsp1 avec la sous-unité 40S (Shehata and Parker 2023). De manière intéressante, une étude cryo-EM a montré que la partie N-terminale de la protéine Nsp1 du coronavirus Bat-Hp-CoV (chauve-souris) est localisée au niveau du site A du centre de décodage du ribosome 40S et perturberait la liaison d'eIF1A (Schubert 2023) ce qui suggère une capacité d'évolution et d'adaptation de Nsp1 entre les différentes espèces (Schubert et al. 2023).

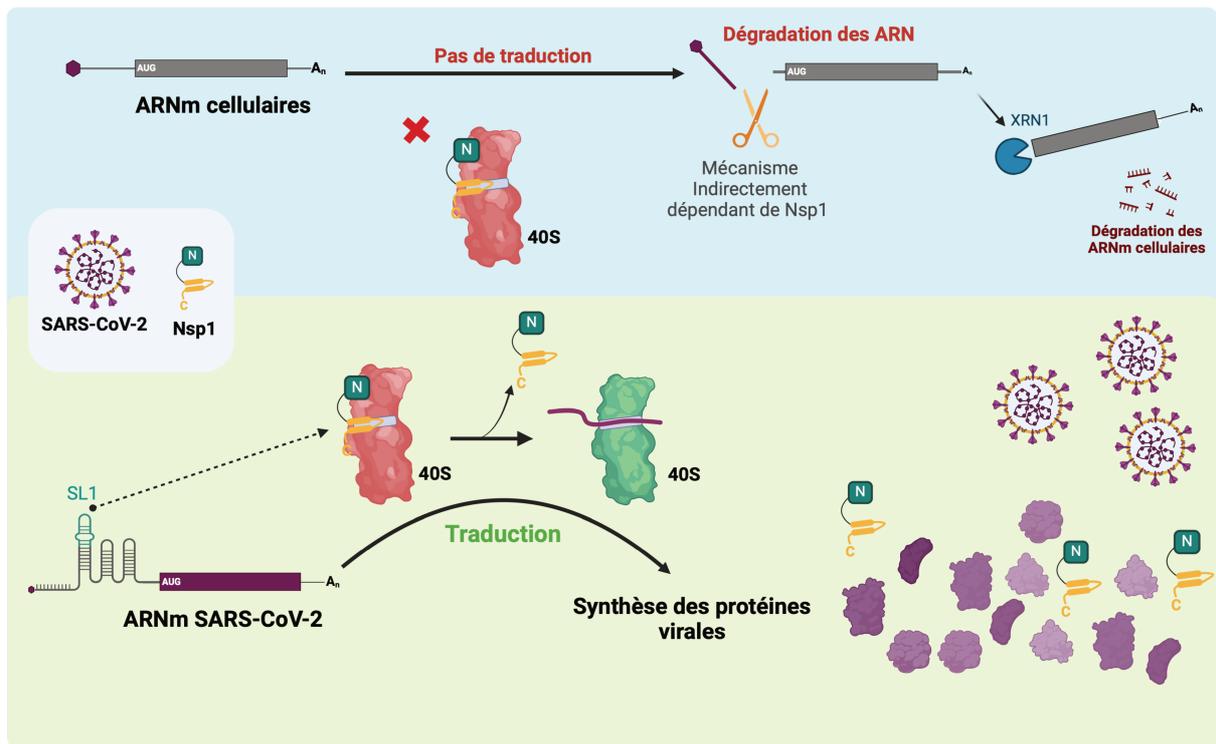


Figure 34. – Mécanisme d'action de la protéine Nsp1

Lors de l'infection, la protéine Nsp1 bloque le canal d'entrée du ribosome via son domaine C-terminal, entravant ainsi la traduction des ARNm cellulaires. En outre, ces derniers sont soumis à un clivage dans un processus indirectement médié par Nsp1, nécessitant leur association préalable avec le ribosome. En parallèle, la structure en tige-boucle SL1, présente dans la séquence leader commune des ARN viraux, interagit avec Nsp1 et induit son déplacement. Cette interaction permet la traduction des ARNm viraux, favorisant de ce fait l'expression des protéines virales et la réplication du virus. Dessinée sur BioRender.

En plus de blocage traductionnel, l'expression de Nsp1 est associée à une dégradation des ARNm de la cellule hôte au cours de l'infection et c'est aussi observé en cellules transfectées et en traduction *in vitro* (Abaeva et al. 2023; Shehata and Parker 2023). Cette dégradation est dépendante de l'association de Nsp1 avec le ribosome 40S (Shehata and Parker 2023; Tardivat et al. 2023) et c'est encore une fois le motif KH qui est impliqué (Shehata and Parker 2023). Cependant, les ARN non codants ne sont pas ciblés par Nsp1 (Shehata and Parker 2023) indiquant que l'ARNm doit être, à un moment donné, associé ou engagé au ribosome. Par des stratégies de mutagenèse dirigée de Nsp1, il a été possible de dissocier la fonction 'blocage du ribosome' de la fonction 'dégradation des ARNm'. En effet, c'est la région N-terminale de Nsp1 qui est responsable du clivage des ARNm (Mendez et al. 2021; Shehata and Parker 2023). Nsp1 ne possède pas d'activité nucléase et est incapable, seule, d'interagir avec les

ARNm cellulaires (Nakagawa and Makino 2021) suggérant la participation additionnelle d'au moins une protéine cellulaire avec une activité nucléase. Il est possible d'induire la dégradation des ARNm par Nsp1 dans des systèmes d'étude *in vitro* de la traduction tels que le lysat de réticulocytes de lapin (Tidu et al. 2020) et surtout dans le système reconstitué de la traduction à partir de facteurs purifiés (Abaeva et al. 2023) suggérant qu'un facteur d'initiation est impliqué dans ces clivages. L'équipe des Dr Pestova et Hellen a ainsi exploité les caractéristiques de l'IRES du CrPV afin d'identifier le facteur incriminé. L'addition de Nsp1 au complexe CrPV IRES/40S provoque des clivages de l'ARN CrPV uniquement en présence d'eIF3 et en particulier de la sous-unité eIF3g (Abaeva et al. 2023). Des expériences de mutagénèse dirigées ont confirmé l'importance du domaine N-terminal de Nsp1 et du domaine RRM (RNA recognition motif) de eIF3g dans le clivage des ARNm (Abaeva et al. 2023).

Enfin, bien que Nsp1 ait une localisation cytoplasmique, on le retrouve aussi au niveau du complexe du pore nucléaire où elle interagit avec le facteur NXF1 et inhibe l'export des ARNm cellulaires (Zhang et al. 2021).

3.2 Comment les ARN viraux échappent à l'inhibition induite par Nsp1

Les ARN viraux sont très bien exprimés pendant l'infection malgré l'expression de Nsp1 (Banerjee et al. 2020; Tidu et al. 2020; Kim et al. 2021; Bujanic et al. 2022; Vora et al. 2022; Schubert et al. 2023), ce qui indique qu'ils échappent à la restriction imposée par cette dernière. Cet échappement repose sur l'interaction de la tige-boucle SL1 avec des résidus spécifiques de la protéine Nsp1, en particulier dans sa région N-terminale aux positions G98, R99, R124 et K125 (Bujanic et al. 2022; Schubert et al. 2023). La tige boucle SL1 est nécessaire pour contourner le blocage du ribosome et les cytosines situées aux positions 15, 19 et 20 sont requis dans ce mécanisme d'échappement (Bujanic et al. 2022). De plus, la structure de la SL1 doit également être à minima conservée pour que le mécanisme d'échappement soit observé (Tidu et al. 2020; Bujanic et al. 2022).

Les ARN viraux échappent également au clivage des ARN induit par Nsp1 même si ce point est sujet à controverse. Il a été montré *in vitro* qu'en présence d'une faible concentration de Nsp1 seuls les ARNm cellulaires sont clivés alors qu'en présence d'une forte concentration de Nsp1 les ARNm viraux le sont également. Cependant, le patron de clivage n'est pas le même avec une coupure qui intervient dans les 18 premiers nucléotides pour les ARN cellulaires (Abaeva et al. 2023; Tardivat et al. 2023) alors que les ARN viraux sont clivés à des positions plus lointaines et bien précises (45, 46, 49) (Tardivat et al. 2023).

3.3 Implication des autres protéines virales

L'activité méthyltransférase de Nsp14 joue un rôle dans le blocage des ARNm cellulaires au niveau des étapes d'épissage et d'export (Zaffagni et al. 2022; Katahira et al. 2023). Cela serait dû à une production de m^7GTP (un analogue de la coiffe) au cours de l'infection ce qui détournerait le complexe de liaison de la coiffe (CBP20/CBP80) de sa fonction initiale (Katahira et al. 2023). Enfin, la protéase virale Nsp5 du SARS-CoV-2 cible eIF4G, un facteur clé dans la traduction canonique, afin d'inhiber la réponse immunitaire (Koudelka et al. 2021; Liang et al. 2023) sans que l'impact sur la traduction cellulaire et virale ne fût réellement étudié.

Certaines protéines virales favorisent l'expression des ARN viraux comme c'est le cas du domaine SUD de la protéine Nsp3 qui stimule leur traduction. Bien que le mécanisme précis n'ait pas été caractérisé, il a été montré que ce domaine interagissait par l'intermédiaire de sa région N-terminale avec la région centrale de la protéine stimulatrice de la traduction Paip1. Cette interaction entraîne une augmentation de l'association des protéines Paip1 et PABPC et la stimulation de la traduction des ARN viraux (Lavigne et al. 2021).

Enfin, la protéine Nsp4 active la voie de stress du réticulum pendant l'infection par l'intermédiaire des voies ATF6 et PERK ce qui a pour effet collatéral de favoriser le bon repliement des protéines virales. Ce mécanisme est finement contrebalancé par une action antagoniste de la protéine Nsp3 afin d'éviter les effets délétères de l'activation de ces voies cellulaires sur le long terme (Davies et al. 2024).

Projet de thèse

Pour tous les virus, la synthèse des protéines virales dépend exclusivement de l'appareil traductionnel de la cellule hôte. Cependant, lors de l'infection par le SARS-CoV-2, la protéine virale Nsp1 induit une inhibition globale de la traduction des ARNm cellulaires. Malgré ces conditions défavorables, l'ensemble des ARNm du SARS-CoV-2 reste efficacement exprimé par un mécanisme qui reste encore non compris.

Tous les transcrits de SARS-CoV-2 commencent avec les mêmes 75 premiers nucléotides. Cette séquence, appelée leader, est imposée par la réplication/transcription de l'ARN génomique et elle forme le début de la 5'UTR. Cependant, cette séquence leader associée à une séquence variable est propre à chaque ARN viral. Les différences au sein de ces 5'UTR ouvrent la possibilité à des contrôles traductionnels spécifiques à chacun des ARN viraux. Nous proposons que ces 5'UTR permettent une expression régulée des différentes protéines virales.

Mon projet de thèse vise à caractériser les éléments importants présents dans la 5'UTR de l'ARN génomique et dans les 5'UTR des ARN sous-génomiques qui régulent leur traduction, tout particulièrement au niveau de l'étape d'initiation. Dans un premier temps, j'ai déterminé la capacité, de ces différentes 5'UTR présentes sur les ARNm du SARS-CoV-2, à initier la traduction sur un ARNm rapporteur. Puis, je me suis intéressé aux mécanismes par lesquels le ribosome était recruté sur l'ARNm en déterminant quels sont les facteurs de l'initiation requis. Je me suis également focalisé sur les différents éléments *cis*-régulateurs qui sont communs mais aussi uniques à chaque 5'UTR afin de déterminer leur importance dans le contrôle traductionnel. Enfin, pour mieux comprendre comment les ARN viraux restent fortement exprimés pendant l'infection, nous avons étudié la traduction initiée à partir des différentes 5'UTR du SARS-CoV-2 en présence de la protéine Nsp1.

Résultats : partie 1

Dans la première partie de mon travail de doctorat, je me suis intéressé à l'efficacité de traduction des différents ARNm du SARS-CoV-2 ainsi qu'à leur régulation. Cette étude a permis une première publication dans le journal *Viruses* en juillet 2022 et le dépôt d'un brevet sur l'utilisation de la 5'UTR du SARS-CoV-2 pour exprimer des protéines d'intérêt.

Les régions 5'UTR des différents ARNm du SARS-CoV-2 possèdent toutes une région commune (la séquence leader) et une région variable propre à chaque ARN (**Figure 1**). En construisant 9 ARNm rapporteurs correspondant à la séquence de chacune des 5'UTR des ARNm du SARS-CoV-2 suivie d'un gène codant pour la luciférase, nous avons mesuré leur efficacité de traduction en lysat de réticulocytes de lapin et en cellules HEK293T. Nos résultats ont non seulement montré que tous les ARNm du SARS-CoV-2 étaient efficacement traduits, mais aussi qu'ils l'étaient à des niveaux variables suggérant l'existence d'un contrôle traductionnel spécifique pour chaque ARNm viral. De façon surprenante, la 5'UTR de l'ARN génomique du SARS-CoV-2, s'est révélée être la plus efficace pour initier la traduction (**Figure 3**) malgré une séquence ARN longue, complexe et riche en structures secondaires (Figure 2).

Cette propriété est conservée chez le SARS-CoV-1 mais pas chez le MERS-CoV (**Figure 4**) ce qui s'explique par la présence d'une uORF fortement inhibitrice présente dans la 5'UTR du MERS-CoV (**Figure 7**).

Nous avons aussi montré que la 5'UTR de l'ARNg SARS-CoV-2 possédait également une uORF mais qu'elle n'inhibait que légèrement l'efficacité translationnelle. (**Figure 7**).

Enfin, nous avons mis en évidence que l'initiation de la traduction des ARNm du SARS-CoV-2 se déroulait selon un mécanisme qui était dépendant de la coiffe (**Figure 5**) et du complexe d'initiation de la traduction eIF4F (**Figure 6**).

Article

Translation of SARS-CoV-2 gRNA Is Extremely Efficient and Competitive despite a High Degree of Secondary Structures and the Presence of an uORF

Lionel Condé, Omran Allatif , Théophile Ohlmann *  and Sylvain de Breyne * 

CIRI, Centre International de Recherche en Infectiologie, (Team Ohlmann), Univ Lyon, Inserm U1111, Université Claude Bernard Lyon 1, CNRS UMR5308, ENS de Lyon, F-69007 Lyon, France; lionel.conde@ens-lyon.fr (L.C.); omran.allatif@ens-lyon.fr (O.A.)

* Correspondence: theophile.ohlmann@ens-lyon.fr (T.O.); sylvain.de.breyne@ens-lyon.fr (S.d.B.)

Abstract: The SARS-CoV-2 infection generates up to nine different sub-genomic mRNAs (sgRNAs), in addition to the genomic RNA (gRNA). The 5'UTR of each viral mRNA shares the first 75 nucleotides (nt.) at their 5' end, called the leader, but differentiates by a variable sequence (0 to 190 nt. long) that follows the leader. As a result, each viral mRNA has its own specific 5'UTR in term of length, RNA structure, uORF and Kozak context; each one of these characteristics could affect mRNA expression. In this study, we have measured and compared translational efficiency of each of the ten viral transcripts. Our data show that most of them are very efficiently translated in all translational systems tested. Surprisingly, the gRNA 5'UTR, which is the longest and the most structured, was also the most efficient to initiate translation. This property is conserved in the 5'UTR of SARS-CoV-1 but not in MERS-CoV strain, mainly due to the regulation imposed by the uORF. Interestingly, the translation initiation mechanism on the SARS-CoV-2 gRNA 5'UTR requires the cap structure and the components of the eIF4F complex but showed no dependence in the presence of the poly(A) tail in vitro. Our data strongly suggest that translation initiation on SARS-CoV-2 mRNAs occurs via an unusual cap-dependent mechanism.

Keywords: translation initiation; SARS-CoV-2; eIF4F; mRNA; uORF



Citation: Condé, L.; Allatif, O.;

Ohlmann, T.; de Breyne, S.

Translation of SARS-CoV-2 gRNA Is Extremely Efficient and Competitive despite a High Degree of Secondary Structures and the Presence of an uORF. *Viruses* **2022**, *14*, 1505.

<https://doi.org/10.3390/v14071505>

Academic Editor: Oliver Schildgen

Received: 3 June 2022

Accepted: 7 July 2022

Published: 8 July 2022

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



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In eukaryotes, translation is a highly regulated step of gene expression in which mRNAs are decoded by ribosomes into proteins. This mechanism can be divided into four steps: initiation, elongation, termination and ribosome recycling. The initiation process is the rate-limiting step that promotes the formation of ribosomal complexes at the initiation codon [1]. The majority of the cellular and viral mRNAs use the scanning model to engage the ribosome on the mRNAs. This starts with the formation of a 43S complex, which is composed of the 40S small ribosomal subunit associated with several eukaryotic initiation factors (eIFs), notably eIF1, eIF1A, eIF3 and the ternary complex (TC) [1,2]. This TC results in the association of eIF2 with a GTP molecule and the Met-tRNA_i^{Met} [3]. Most eukaryotic mRNAs are capped at their 5' extremity and polyadenylated at their 3' end in order to ensure integrity, stability and efficient translation of the mRNA. The cap structure consists of a 7-methylguanylate residue connected to the mRNA via an unusual 5' to 5' triphosphate linkage [4]. The cap interacts with the eIF4F complex, which is composed of the cap-binding protein eIF4E, the scaffold protein eIF4G and the RNA helicase eIF4A. The eIF4F complex plays a crucial role during initiation because it establishes a physical and functional link between the mRNA (through a cap–eIF4E interaction) and the 43S complex (through an eIF4G–eIF3 interaction) [5–7]. Once bound to the 5' end of the mRNA, the 43S complex scans the 5' untranslated region (UTR) in a 5' to 3' direction to reach the cognate AUG start codon. This is catalyzed by the RNA and ATP-dependent helicase activity of eIF4A, which

unfolds the RNA structures encountered within the 5'UTR. Recognition of the translation initiation site (TIS) is ensured by a perfect base pairing between the AUG codon and the anticodon of the tRNA_i^{Met}, which triggers GTP hydrolysis and eIFs release, followed by association of the 40S ribosomal subunit with the 60S ribosomal subunit to form the 80S ribosome competent for elongation [1,2,8].

The efficiency of translation initiation is largely regulated by both eIF2 and eIF4F [9–11], but the intrinsic characteristics of the mRNA, such as length, RNA folding and upstream AUG codons, can also modulate the flow and quantity of ribosomes on the transcript [12]. In general, initiation on most eukaryotic mRNAs adheres to the first AUG rule, in which the AUG triplet nearest the 5' end is the beginning of the reading frame [13–16]. It is noteworthy that the nucleotides that surround the initiation codon are critical for efficient recognition with the following sequence—GCCA/GCCAUGG—being optimal and defined as the Kozak context [13,14]. It has been reported that the nucleotides upstream of the AUG codon, especially the purine at position -3 (according to the position of the AUG codon), are important for efficient initiation [13,17–19]. UV cross-linking experiments in the 48S pre-initiation complex (PIC) strongly suggested an interaction between eIF2 α and the -3 purine [20]. This was confirmed later by cryo-EM analyses of PIC indicating that the arginine residues (Arg55 and Arg57) of eIF2 α are located within close proximity of the mRNA nucleotides upstream to the AUG codon, including the -3 position [21]. These data are consistent with the role of eIF2 in AUG codon recognition [22]. However, the influence of the nucleotide in the $+4$ position is rather controversial. Some experiments showed that mutation of the guanosine reduces translation initiation efficiency [13], but it was shown by analysis of all possible combinations using a FACS-seq approach that the $+4$ G was not always the optimum [23]. In addition, an alternative explanation was suggested and it was based on the amino acid constraint on the second codon that participates in post-translational modification, such as the N-terminus methionine excision [24,25]. Thus, according to the robustness of this Kozak context, the AUG codon can be more or less recognized as a translation initiation site or even completely ignored. In addition to these features, the presence of short upstream open reading frames (uORFs) within the 5'UTR can also have an impact on translational efficiency of the main gene due to the fact that some ribosomes will translate these uORFs and then dissociate from the mRNA before reaching the main reading frame [26–28]. This is why the presence of these uORFs is often associated with a reduction of production of the main protein product. Interestingly, in some particular physiological conditions, the level of active TC is reduced and this favors the bypass of these uORFs, resulting in efficient translation of the main ORF [29–31].

Another key feature of the mRNA that can regulate ribosome recruitment is the RNA structure adopted by the 5'UTR. It is now well known that a thermodynamically stable hairpin structure at the 5' end can completely block 43S complex attachment to the cap structure [32–36] and it usually requires additional RNA helicases to overcome the energy barrier imposed by these structures [37–39]. In some cases, these RNA motifs can serve as anchors for the binding of specific proteins, creating an RNA-protein complex that can arrest the scanning of the 43S ribosome. The best-known example remains the mRNA coding for ferritin, in which the 5'UTR contains an iron-responsive element, which is a short and conserved stem-loop that is recognized when the endogenous iron concentration becomes low in order to down-regulate translation of ferritin [40–43]. In RNA positive-strand viruses, there are often many stable RNA structures that are involved in different steps of viral replication (dimerization, packaging, export, etc.) [44–46] and can interfere with viral translation. As such, viruses often develop strategies to council replication and translation. The best example comes from picornaviruses, which contain uncapped mRNAs harboring a very long and highly structured 5'UTR with multiple upstream AUG codons that are not used for translation [47,48]. Indeed, the RNA folding adopted by the 5'UTR has the capacity to create a functional structure that directly promotes ribosome recruitment and was defined as internal ribosome entry segment (IRES) [49]. IRESes are not restricted to picornaviruses and are present in other viral families [50,51] and cellular mRNAs [51–53].

Viral IRESes remain the most studied and despite differences in length, sequence and structures, they usually have the characteristics to directly interact with some components of the 43S initiation complex to bring the ribosome close to the initiation codon [54–63]. On the other hand, eukaryotic IRESes concern ~10% of the mammalian mRNAs and they are generally less structured than viral IRESes [51]. Overall, this mode of initiation is cap- and eIF4E-independent and can become the predominant mode of translation initiation mechanism during cellular stresses and viral translational escape [51,64–67].

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [68] contains a long and positive RNA genome (gRNA), which is capped and polyadenylated. Once in the cytoplasm, the gRNA is immediately translated into two large polyproteins, ORF1a and ORF1ab, which generate 16 non-structural proteins (nsp) after proteolytic cleavage. Expression of these non-structural proteins is essential for replication and this starts with the transcription of a full-length negative-sense genomic copy that serves as template for the production of positive-sense genomic RNAs. One of the peculiarities of the coronaviruses is that they harbor a conserved leader sequence located at the 5' end of the transcript that serves as a *cis*-acting element for the transcription of all sub-genomic mRNAs (sgRNA). This leader sequence is added to all sgRNAs through a polymerase template switching that occurs when the viral polymerase encounters the transcriptional regulatory sequence (TRS). The TRS precedes every single viral gene and is called the body TRS (TRS-B), with a conserved core sequence that is thought to hybridize with the leader TRS (TRS-L) [69–71]. As a result, all 5'UTRs of SARS-CoV-2 mRNAs (sgRNAs and gRNAs) share an absolutely conserved 75-nucleotide (nt.) stretch, which is called the leader sequence. This leader sequence is then followed by a variable RNA length (from 0 to 190 nt.) that terminates onto an AUG start codon (Figure 1). Viral transcription generates up to nine sgRNAs coding for the structural (S, E, M and N) and accessory (ORF-3a, ORF-6, ORF-7a, ORF-7b and ORF-8) proteins (Figure 1) [71,72]. The 5'UTRs of ORF-7a and ORF-8 gRNA are identical and harbor only the leader sequence (Figure 1 and Table 1). The 5'UTRs of ORF-S, ORF-3a, ORF-E and ORF-N are composed of the leader sequence followed by only a few nucleotides (from 1 to 8 nt.) (Figure 1 and Table 1). In contrast, the 5'UTR of ORF-M, ORF-6 and ORF-7b are relatively long, with respective lengths of 119, 230 and 151 nt. including the leader sequence (Figure 1 and Table 1). Thus, every single sgRNA harbors its own 5'UTR and the nucleotide surrounding context of each AUG start codon is also unique (Table 1); these structural features could potentially modulate translational efficiency.

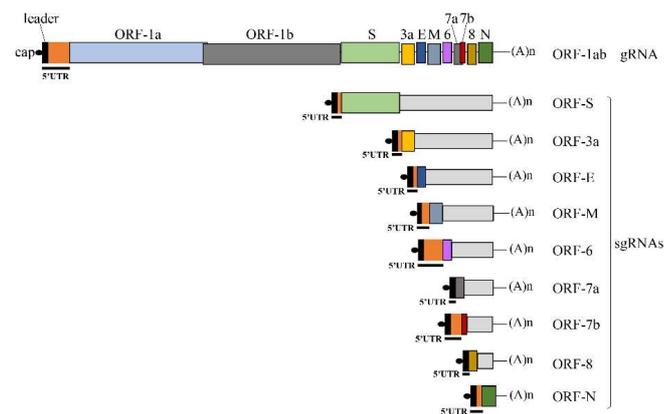


Figure 1. Scheme of the SARS-CoV-2 gRNA and sgRNAs. The cap and poly(A) tail structures are indicated. The 5'UTR of each viral mRNA is underlined. The 5'UTR of each viral mRNA is underlined and is composed of the common leader sequence (black square) and the variable region (orange square). The other different colors correspond to the different ORFs as indicated on the top.

Table 1. Length, Kozak context and %GC of the SARS-CoV-2 gRNA and sgRNA 5'UTRs.

5'UTR	Length (nt.)	Kozak Context ¹	%GC
ORF-1ab	265	GGU <u>A</u> AG AUG <u>G</u> AG	45
ORF-S	76	CGA <u>A</u> CA AUG <u>U</u> UU	37
ORF-3a	77	GAA <u>C</u> UU AUG <u>G</u> AU	36
ORF-E	77	GAA <u>C</u> UU AUG <u>U</u> AC	36
ORF-M	119	UUA <u>G</u> CC AUG <u>G</u> CA	31
ORF-6	230	CAA <u>C</u> AG AUG <u>U</u> UU	41
ORF-7a	75	ACG <u>A</u> AC AUG <u>A</u> AA	37
ORF-7b	151	GAC <u>A</u> GA AUG <u>A</u> UU	34
ORF-8	75	ACG <u>A</u> AC AUG <u>A</u> AA	37
ORF-N	83	ACU <u>A</u> AA AUG <u>U</u> CU	35

¹ The AUG codons are represented in bold and the nucleotides at positions -3 and +4 are underlined.

The 5'UTR of the gRNA is the longest (265 nt.) of all the viral mRNAs and contains five stem-loop structures (SL1 to SL5) that have recently been modeled by chemical probing analyses (Figure 2) [73,74]. These SL are critical for replication, RNA synthesis and escape from nsp1-mediated translation inhibition [75–81]; however, they could potentially represent hurdles for translation, notably at the level of ribosome recruitment and scanning. Although the viral genome of SARS-CoV-2 is capped and polyadenylated, the translation initiation mechanism used to locate the AUG start codon inside the SL5 remains undetermined. In particular, one can predict that the high level of RNA secondary structures within the 5'UTR could be a serious hurdle for ribosome scanning and this could suggest another mechanism for translation initiation. Such a hypothesis is also strengthened by the fact that the AUG start codon is located at the vicinity and downstream of a four-way junction structure (Figure 2), a situation that is reminiscent of similar motifs found in some viral IRESes [49,82–85]. Last but not least, the SARS-CoV-2 5'UTR also contains an uORF starting at AUG¹⁰⁷, which is located within the SL4 (Figure 2) and seems to be used as an initiation site, as suggested by ribosome profiling analyses in infected cells [72,86]. Interestingly and curiously, ribosome profiling analyses have also highlighted other uORFs initiating at near-cognate initiation codons [72,86]. One is of particular interest because it is located on a CUG codon at position 59 just 10 nt. upstream of the TRS leader, which means that it will be present on all the sgRNAs. However, one should bear in mind that ribosome occupancy observed in ribosome profiling analysis remains only indicative, as it is often difficult to distinguish between a real TIS and the pausing of the 48S preinitiation complex [72].

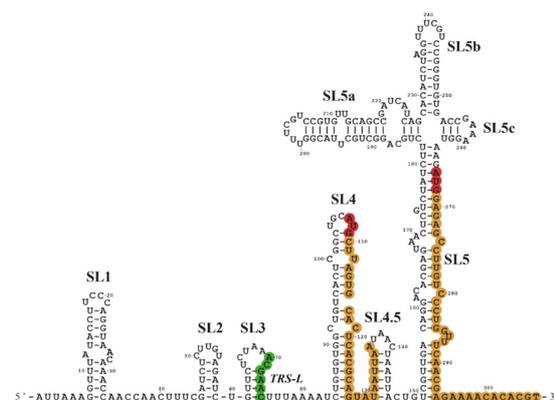


Figure 2. Modelization of the structure of the SARS-CoV-2 gRNA 5'UTR. This model, adapted from Miao et al. (2021) [73], shows the different stem-loop (SL) RNA structures. Position of the AUG codons of uORF and main ORF are highlighted in red, the coding region in orange and the TRS-L in green.

All these unusual structural features can have a great impact on the translational efficiency of the genomic and sub-genomic SARS-CoV-2 RNAs and this set the rationale for this study. By using a series of *in vitro* and *ex vivo* functional translational assays coupled with specific inhibitors of the cap-dependent mechanism, we have investigated and compared translation of all SARS-CoV-2 transcripts. Our data show some unexpected characteristics of the genomic RNA, which is extremely efficient to recruit ribosomes for protein synthesis in all systems assayed despite its length, complex RNA structure and the presence of a short upstream reading frame; these translational features are conserved with the SARS-CoV-1 but not with MERS-CoV.

2. Materials and Methods

2.1. DNA Constructs

The pGlo-Renilla, pEMCV-Renilla, pHCV-Renilla and pMM1-L-protease constructs have been previously described [87–89]. The leader sequence and the 5'UTR (5'UTR^{ORF1ab}) of SARS-CoV-1 (NC_004718.3), MERS-CoV (NC_038294) and SARS-CoV-2 (NC_045512.2) gRNA have been generated by consecutive PCRs with specific primers (Table S1). The resulting PCR products were digested with the BamHI restriction enzyme and inserted into the pRenilla vector between the BglII and BamHI restriction sites. The 5'UTRs of each SARS-CoV-2 sgRNA associated to the coding region of the Renilla luciferase were produced by consecutive PCRs with specific primers (Table S2), digested with BamHI and EcoRV restriction enzymes and inserted into the pRenilla vector between the BglII and EcoRV restriction sites. Mutation of the AUG¹⁰⁷ and AUG¹⁵⁸ in the SARS-CoV-2 and MERS-CoV gRNA, respectively, and the mutation of the UAA triplet at position 133 in SARS-CoV-2 have been performed by consecutive PCRs with specific primers (Table S3) to generate the mut-AUG¹⁰⁷, mut-AUG¹⁵⁸ and uORF-phase constructs. The resulting PCR products were digested with the BamHI restriction enzyme and inserted into the pRenilla vector between the BglII and BamHI restriction sites. All DNA constructs were verified by DNA sequencing.

2.2. In Vitro Transcription

The four combinations—capped/polyadenylated (+/+), capped/non-polyadenylated (+/–), uncapped/polyadenylated (–/+) and uncapped/non-polyadenylated (–/–)—were obtained by *in vitro* transcription using the linear template of pRenilla vector by digestion either at the XbaI (non-polyadenylated RNAs) or at the EcoRI restriction site (polyadenylated RNAs). RNA transcription was performed from 1 µg of linear DNA template in transcription buffer (40 mM Tris-HCl (pH 7.5), 6 mM MgCl₂, 2 mM spermidine and 10 mM NaCl) in the presence of 40U of T7 RNA polymerase (Promega, Madison, WI, USA), 40 U of RNAsin (Promega), 1.6 mM of each ribonucleotide triphosphate and 10 mM DTT. For capped mRNAs, the rGTP concentration was reduced to 0.32 mM, and m⁷-GpppG cap analog (New England Biolabs) was added at a concentration of 1.28 mM. The transcription reaction was carried out at 37 °C for 1 h 30 min, then treated for 30 min with the RQ1 DNase (Promega). mRNAs were isolated using NucleoSpin RNA XS purification columns (Macherey Nagel, Düren, Germany). The integrity of RNA was checked by electrophoresis on non-denaturing 1% agarose gel and the concentration was determined by reading the absorbance using Nanodrop technology.

2.3. In-Vitro Translation Assays

Nuclease-treated Flexi Rabbit Reticulocytes lysate (RRL) or untreated RRL (Promega), supplemented as previously described [90,91], were programmed with 5 or 10 fmol of *in vitro* synthesized mRNAs in the presence of 75 mM KCl, 0.5 mM MgOAc, 2 mM DDT and 20 µM of each amino acid for 30 min at 30 °C. For radioactive labeling, 0.25 mCi/mL of [³⁵S] methionine was added to the reaction mix. Translation products were then stopped in either Renilla lysis-juice buffer (PJK Biotech, Kleinblittersdorf, Germany) for luciferase assay or in Laemmli sample buffer (4% SDS, 10% β-mercaptoethanol, 20% glycerol, 0.004%

bromophenol Blue, 0.125 M Tris-HCl pH 6.8) and resolved by 12% SDS-PAGE. The gel was dried and quantified with a Molecular Dynamics PhosphorImager.

The L-protease from Foot-and-Mouth Disease Virus (FMDV) was prepared in RRL as previously described [87]. L-protease (0.2 μ L and 0.4 μ L), hippuristanol (0.15, 0.3 and 0.6 μ M) and the cap analog (1 μ M) were incubated 10 min in the RRL before RNA addition.

2.4. Renilla Luciferase Activity

Renilla luciferase activity was measured using the Renilla-Juice luciferase assay (PJK GmbH, Kleinblittersdorf, Germany) following the manufacturer instructions on a LUMIstar apparatus (BMG LABTECH, Champigny-sur-Marne, France) or on a Mithras LB940 (Berthold, Thoiry, France).

2.5. Western-Blot against eIF4G

Proteins from mock or L-protease-treated RRL extracts were separated by a 7.5% SDS-PAGE. The proteins were transferred onto polyvinylidene difluoride membranes (Boehringer Mannheim, Ingelheim am Rhein, Germany). Blots were incubated first with antibody against eIF4G (#2617 Cell Signaling Technology, Danvers, MA, USA) for 16 h at 4 °C, then with an anti-Rabbit IgG–Peroxidase antibody (Sigma Aldrich, Saint-Louis, MO, USA) for 2 h. The chemiluminescence signal was detected using an Pierce™ ECL reagent (ThermoFisher, Waltham, MA, USA) in the Chemidoc Imager (Biorad, Hercules, CA, USA).

2.6. Cell Culture and RNA Transfection

Jurkat T-cell line, obtained from ATCC, was cultured in RPMI-1640 medium (Life Technologies), supplemented with 10% fetal calf serum, penicillin, streptomycin, Hepes pH 7.5, sodium pyruvate and L-glutamine. Cells were incubated at 37 °C in a 5% CO₂ atmosphere.

For RNA transfection, 5.10(5) cells were resuspended in 10 μ L of buffer R (Life technology) in the presence of 500 fmol of capped and polyadenylated mRNAs, and immediately transfected using the Neon transfection kit (Life technology, Carlsbad, CA, USA) according to the manufacturer instructions. Cells were incubated in RPMI medium for 90 min at 37 °C, washed twice in PBS and split in two pools to measure luciferase activities and to extract cytoplasmic RNAs.

2.7. RNA Extraction and RT-qPCR

RNAs were isolated using the NucleoSpin RNA extraction kit according to the manufacturer instructions (Macherey-Nagel, Düren, Germany) and quantified using a nanodrop 2000 spectrophotometer (ThermoFisher, Waltham, MA, USA). Reverse transcription of 250 ng of cytoplasmic RNAs was performed using qScript kit (Quanta Bio, Beverly, MA, USA). mRNA quantification was performed by quantitative PCR with the ONEGreen® FAST QPCR PREMIX (Ozyme, Saint-Cyr-l'École, France) according to the manufacturer instructions. Renilla luciferase (Renilla forward TGGACAATAACTTCTTCGTGAAAAC/Renilla reverse GCTGCAAATCTTCTGGTTCTAA) was amplified in parallel with the endogenous housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (GAPDH forward (Human) CGACAGTCAGCCGCATCTT/ GAPDH reverse (Human) CC-CATGGTGTCTGAGCG). The relative copy numbers of Renilla cDNAs were compared to GAPDH using $x-\Delta^{Ct}$ (where x corresponds to the experimentally calculated amplification efficiency of each primer couple).

2.8. Sequence Alignment

The 5'UTRs of the SARS-CoV-1 (NC_004718.3), MERS-CoV (NC_038294) and SARS-CoV-2 (NC_045512.2) gRNAs were analyzed with a Multiple Sequence Comparison by Log-Expectation (MUSCLE) using the default parameters [92].

3. Results

3.1. Translation Initiation Efficiency Mediated by the SARS-CoV-2 5'UTRs

During infection, SARS-CoV-2 generates at least 10 different mRNAs that harbor unique 5'UTRs in terms of length, structure and context of their AUG initiation site. In order to investigate further the role of these 5'UTRs in translational control, we have cloned each of the viral 5'UTR upstream of a reporter gene coding for the Renilla luciferase into the pRenilla vector [88,89,93] (Figure 3A). The nucleotide triplet downstream to the TIS has been conserved to maintain the SARS-CoV-2 original Kozak context. As a control, we have used the 5'UTR of the globin mRNA, which is known to promote translation initiation in a very efficient manner. Capped and polyadenylated RNAs have been produced *in vitro* to program the rabbit reticulocytes lysate (RRL) in the presence of S³⁵-methionine. Protein production was quantified on a phosphoImager after an SDS-PAGE resolution (Figure 3B). For most of the transcripts, expression was as efficient as that generated by the 50 nt. of the β -globin 5'UTR (NC_000011.10; Figure 3B compare lanes 4–7, 9 and 11 with lane 2). Interestingly, the 5'UTRs derived from ORF1-ab, ORF-6 and ORF-7b promoted a higher level of expression than from the globin mRNA reporter gene (Figure 3B lanes 3, 8 and 10); this was not expected considering that these viral 5'UTRs are the longest and the most structured of the series (Figure 2 and Table 1). We were also surprised to find that translation of ORF-3a, ORF-S, ORF-E and ORF-6 was still very efficient despite the fact that their AUG start codons are in a poor Kozak context (with a cytosine in -3 and/or a uracil in +4; Table 1).

Quantification of protein production was also performed by measurement of the luciferase activity (Figure 3C). It should be remembered that the first amino acid downstream to the methionine codon was variable for each construct as we conserved the Kozak context of the original TIS (Figure S1). Thus, this N-terminal modification could have an impact on the activity of the luciferase; the same samples have also been quantified by luciferase assay. As the results obtained by reading of the luciferase assays were consistent with S³⁵-methionine quantification (compare Figures 3B and 3C), we assumed that this N-terminal modification did not impact significantly the enzymatic activity of luciferase and the latter has been used as a readout for protein production for the remainder of this study.

Next, we translated the different mRNAs into the untreated RRL, which represents an *in vitro* competitive translation system due to the presence of endogenous mRNAs that compete for the translational apparatus [89]. In such a system, translation initiation mediated by the majority of the SARS-CoV-2 5'UTRs was further enhanced in comparison to the expression mediated by the globin 5'UTR (Figure 3D). Once again, and consistent with the RRL results, translation from the ORF-1ab 5'UTR was the most efficient. This very high level of translation of viral RNAs in the untreated RRL suggests that they have the intrinsic ability to recruit ribosomes in a competitive environment. To go further, the mRNAs have been electroporated into cells and the protein expression level has been measured by luciferase assay and normalized to the quantity of RNA by RT-qPCR. The results showed in Figure 3E are consistent with the data obtained in the RRL systems. Taken together, our data show that the 5'UTRs of the SARS-CoV-2 have the ability to promote a high level of translation both *in vitro* and *ex-vivo*, with the 5'UTR^{ORF1ab} (gRNA) sequence being the most efficient of all.

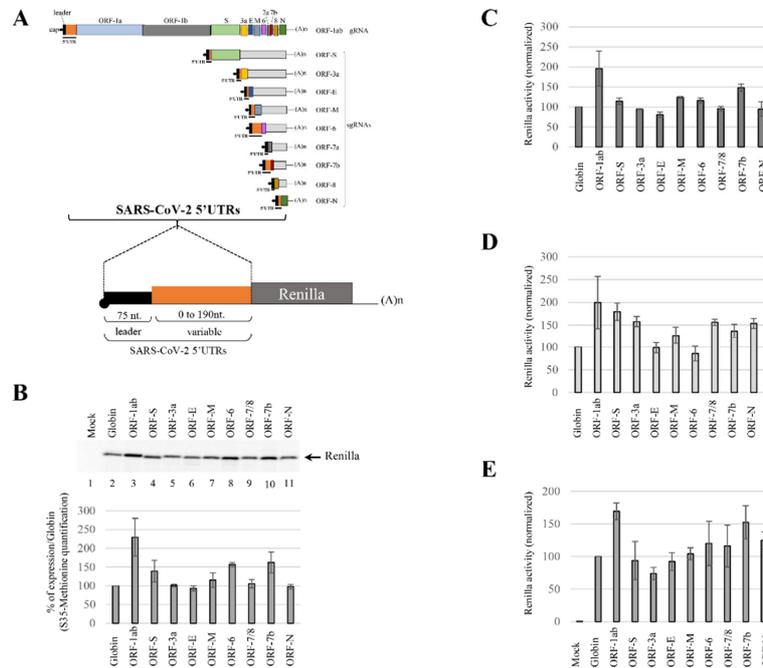


Figure 3. Translation efficiency of the SARS-CoV-2 5'UTRs. (A) Scheme of the gRNA and the different sgRNAs, with the leader sequence in black and the variable region in orange, which represent each 5'UTR (this panel was obtained from Figure 1). The sequences corresponding to each 5'UTR have been cloned upstream of the Renilla ORF. (B,C) RRL and (D) untreated RRL have been programmed with capped and polyadenylated mRNAs harboring specifically the 5'UTR of the ORF1ab, ORF-S, ORF-3a, ORF-E, ORF-M, ORF-6, ORF-7a, ORF-7b, ORF-8 or ORF-N. The resulting protein products have been resolved on (B) an SDS-PAGE and quantified with a phosphoImager, or (C,D) analyzed by luciferase assay. (E) Jurkat cells were electroporated with capped and polyadenylated mRNAs as indicated. After 90 min of incubation, the cellular extracts were used for RT-qPCR and luciferase assay. All the results are normalized to the globin 5'UTR expression (set at 100%). Values are the mean (+/− S.D.) for three independent experiments (panels (B,C)) or (+/− S.E.M.) for four independent experiments (panels (D,E)).

3.2. Translation Efficiency from MERS-CoV gRNA 5'UTR Differs from SARS-CoV gRNA 5'UTRs

We next wondered whether a high level of translation could also be observed with RNAs derived from closed relative strains, i.e., the SARS-CoV-1 [94–96] and the MERS-CoV [97–99]. The 5'UTR of the SARS-CoV-1 gRNA is 264 nt. long and shares 90% of its identity with the sequence of the SARS-CoV-2 (Figure 4A and Figure S2). Once more, the global RNA structure of the 5'UTR is also conserved [73,74,77], which could indicate that expression initiated from the 5'UTR of the gRNA from both SARS-CoV strains could be similar. In contrast, the MERS-CoV gRNA's 5'UTR is slightly longer (277 nt.) and more distant from SARS-CoV sequences as it shares only ~60% of its identity with type 1 and type 2 (Figure 4A and Figure S2), which could also suggest a divergence in translational efficiency. To directly test this hypothesis, we have used the pRenilla reporter gene harboring either the leader sequence, which represents the shorter 5'UTR (named leader), or the 5'UTR of the SARS- and the MERS-CoV gRNA (named 5'UTR^{ORF1ab} in the rest of the manuscript) (Figure 4B). Capped and polyadenylated mRNAs were generated by in vitro transcription and translated in the RRL (Figure 4C) or electroporated into Jurkat cells (Figure 4D). The products of the expression were quantified by luciferase assays and normalized to the

expression mediated by the globin 5'UTR, which served as a positive control. Consistent with our previous results, the 5'UTR^{ORF1ab} of SARS-CoV-2 gRNA promoted strong protein production (Figure 4C,D). As expected, a similar expression profile was observed when translation was driven by the 5'UTR^{ORF1ab} of the SARS-CoV-1, both in the RRL and in cells (Figure 4C,D), confirming that the translational characteristics are well conserved between the two strains. In contrast, although the leader sequence of MERS-CoV was also efficient to drive translation in the RRL, it still remained lower than that of the SARS-CoV transcripts (Figure 4C). This difference was even sharper with the 5'UTR^{ORF1ab} of MERS-CoV, which showed a decrease of up to 30% when compared to globin expression (Figure 4C). These differences were also observed when the reporter mRNAs were transfected in cell lines (Figure 4D). Our data show that the translation characteristics of the 5'UTR^{ORF1ab} are conserved between the two SARS-CoV strains but differ with MERS-CoV. This indicates that divergences in sequence conservation are also observed in terms of translational efficiency.

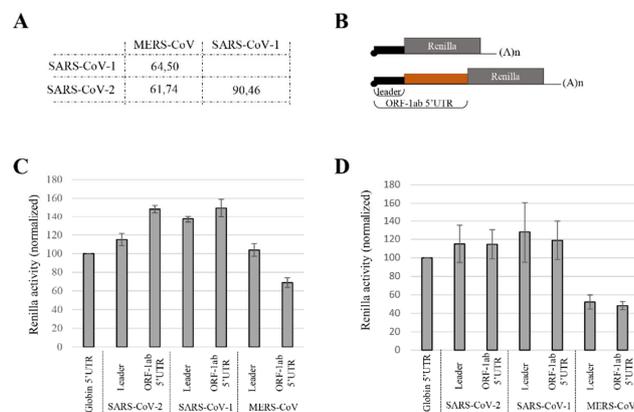


Figure 4. Divergence in translation efficiency between the 5'UTRs of MERS-CoV and SARS-CoV gRNA. **(A)** Percentage of identity between the 5'UTRs of MERS-CoV, SARS-CoV-1 and SARS-CoV-2 gRNAs resulting from a MUSCLE alignment (Figure S2). **(B)** Scheme of capped and polyadenylated mRNAs coding for the Renilla reporter gene with the leader or the 5'UTR^{ORF1ab} from SARS-CoV-1/2 and MERS-CoV. **(C)** RRL was programmed with the corresponding in vitro transcribed reporter mRNAs. Expression products were quantified by luciferase assays and normalized to globin 5'UTR expression (set at 100%). **(D)** Jurkat cells were electroporated with capped and polyadenylated mRNAs as indicated. After 90 min of incubation, the cellular extracts were used for luciferase assay and RT-qPCR. All the results are normalized to the globin 5'UTR expression (set at 100%). Values shown are the mean (+/− S.E.M.) for 6 **(C)** and 5 **(D)** independent experiments.

3.3. SARS-CoV-2 Translation Initiation Requires the Cap Structure and the Components of eIF4F Complex

Data obtained so far have pointed out that translation of the gRNA of the SARS-CoV-2 is very efficient, both in vitro and ex vivo, but the molecular mechanism by which this occurs remains to be determined. Surprisingly, the gRNA is the longest and most structured transcript, which should, a priori, represent a physical barrier for linear ribosomal scanning. On the other hand, the viral gRNA is capped and polyadenylated, which is a strong argument for a cap-dependent mechanism of translation initiation. This is also in agreement with previous studies showing that addition of cap analog to the RRL resulted in a strong reduction of translation driven by the 5'UTR of the bovine CoV strain [100,101]. Furthermore, the destabilization of the eIF4F complex also results in the inhibition of viral replication of MERS-CoV and HCoV-229E [102,103]. Although these studies suggest a cap-dependent mechanism of translation for coronavirus mRNAs, this has not been experimentally established. On the other hand, the structural features of the 5'UTR^{ORF1ab} (gRNA), as well as the presence of a uORF and an AUG start codon located in a four-way

junction RNA structure, could suggest the presence of an active IRES motif within the 5'UTR [49,82–85]. Therefore, in order to distinguish between these two possibilities, we decided to further investigate the molecular mechanism by which translation initiation occurs on the SARS-CoV-2 gRNA.

We first looked at the effects of both the cap and the poly(A) tail structures in the translation of SARS-CoV-2 mRNAs. In order to avoid the bias of RNA stability caused by the absence of both structures, we performed our study in *in vitro* translation systems. For this, we produced different combinations of capped, uncapped, polyadenylated and non-polyadenylated transcripts harboring the 5'UTR of the globin, the 5'UTR^{ORF1ab} or the leader sequence of SARS-CoV-2. Both RRL (Figure 5A) and untreated RRL (Figure 5B) were programmed for 30 min and expression of the Renilla was quantified by luciferase assays. As expected, the synergic effect of the combination of the cap structure and the poly(A) tail was observed with the RNA control (globin 5'UTR). Enhancement of translation provided by these two elements was even amplified in the competitive untreated RRL, which is an experimental system known to recapitulate faithfully cap/poly(A) synergy [89]. Expression initiated by the leader sequence or the 5'UTR^{ORF1ab} was highly dependent on the presence of the cap structure with a 10-fold stimulation over the uncapped transcripts (Figure 5B). In contrast, we were surprised to find that the addition of the poly(A) tail on the SARS-CoV-2 mRNAs did not add any supplementary translational advantage over the non-polyadenylated mRNAs. This shows an unambiguous requirement for the m⁷GTP cap structure but no significant effect of the poly(A) tail. In addition, this denotes a major difference with canonical globin translation regarding the requirement for the 3' poly(A) tail.

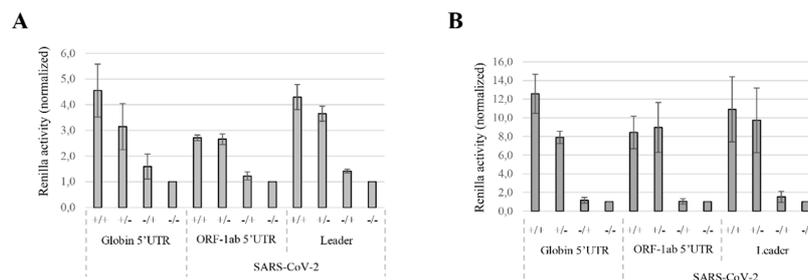


Figure 5. The effects of the cap and poly(A) tail on viral translation. Capped/polyadenylated (+/+), capped/non-polyadenylated (+/-), uncapped/polyadenylated (-/+) or uncapped/non-polyadenylated (-/-) Globin Renilla, 5'UTR^{ORF1ab} Renilla and leader Renilla RNAs were translated in (A) the RRL or in (B) the untreated RRL. Translational products were analyzed by luciferase assays and normalized to the (-/-) condition (arbitrary set at 1). Values shown are the mean +/- S.D. for three independent experiments.

We decided to continue our investigation to look at the role of the initiation factors that are required for SARS-CoV-2 translation with a particular focus on the cap-binding complex eIF4F (Figure 6A). This complex is formed by the association of three initiation factors: eIF4E, eIF4A and eIF4G and its role is to bind both to the m⁷GTP cap structure and to the ribosome [2]. Addition of cap analog to the RRL chelates eIF4F by competitive binding and sequesters it from binding to the 5' end of the mRNAs (Figure 6A) [104]. As expected, the addition of the cap analog in our translation assay resulted in a strong repression of globin translation (Figure 6B) and a similar effect was observed on both the 5'UTR^{ORF1ab} and leader of the SARS-CoV-2 transcripts, confirming that the eIF4F complex is needed for their expression.

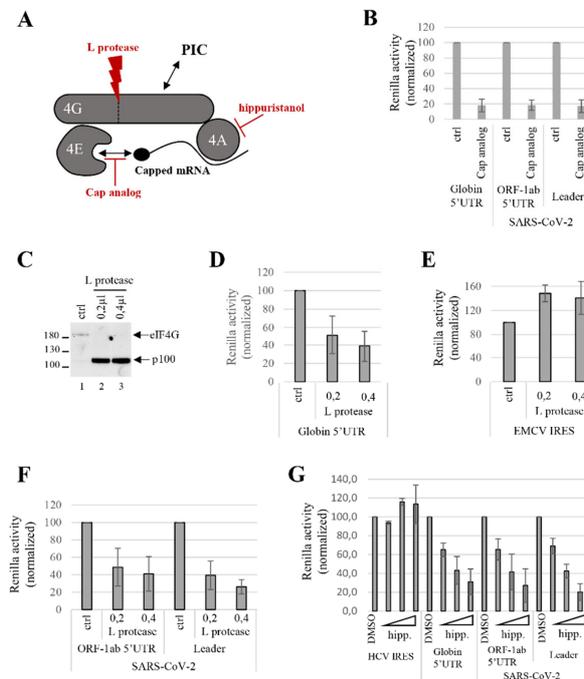


Figure 6. Requirement for the eIF4F complex. (A) Schematic cartoon of the eIF4F complex, which is composed of eIF4E, eIF4G and eIF4A. The interactions with the PIC, cap analog, L protease and hippuristanol are depicted on the figure. (B) RRL was treated with cap analog (1 μ M) or with water (ctrl) and then programmed with capped and polyadenylated globin-Renilla and SARS-CoV-2 leader or 5'UTR^{ORF1ab}-Renilla mRNAs. Products of expression were quantified by luciferase assays and normalized to the control condition (set at 100%). (C–F) Untreated RRL was pre-incubated alone (ctrl) or in presence of different amounts of L protease (μ L) as indicated, and then programmed with (D) globin-Renilla, (E) EMCV IRES-Renilla, (F) SARS-CoV-2 leader or 5'UTR^{ORF1ab}-Renilla mRNAs. Translation products were (C) resolved by a 7.5% SDS-PAGE and then transferred to a PVDF membrane for Western blotting analysis against eIF4G, or (D–F) analyzed by luciferase assay and normalized to the control condition (set at 100%). (G) RRL pre-treated with DMSO or increasing amounts (μ M) of hippuristanol (hipp.) was programmed for 30 min with globin-Renilla, HCV-IRES Renilla, SARS-CoV-2 leader or ORF-1ab 5'UTR-Renilla mRNAs. Expression products were analyzed by luciferase assay and normalized to the control condition (set at 100%). Values shown are the mean (+/– S.D.) for three independent experiments.

Next, in order to distinguish whether the whole eIF4F, or parts of it, are needed, we decided to specifically focus on the independent role of both eIF4A and eIF4G, which are core subunits of eIF4F. eIF4G can be targeted by viral proteases and eIF4A can be specifically inhibited by some chemicals (Figure 6A). We first studied the dependence on eIF4G by using the L protease from the foot-and-mouth-disease virus [105,106]. This enzyme cleaves eIF4G to yield a N-terminal domain able to bind to eIF4E and a C-terminal domain (called p100), which harbors the interacting domains with eIF4A and eIF3 [107]. Functionally, the cleavage of eIF4G inhibits cap-dependent transcripts but does not affect, and even stimulates, translation of mRNAs bearing an IRES [108]. We have added in-vitro-translated L protease to the untreated RRL that resulted in the cleavage of eIF4G (Figure 6C) and we observed different effects on the reporter constructs assayed. As expected, cap-dependent translation driven by globin 5'UTR was inhibited (Figure 6D), whereas the encephalomyocarditis virus (EMCV) IRES-mediated protein synthesis was slightly stimulated (Figure 6E) as previously described [87]. Cleavage of eIF4G resulted

in the inhibition of expression from both SARS-CoV-2 5'UTR^{ORF1ab} and leader sequence mRNAs (Figure 6F) and to a similar magnitude as that for globin, suggesting that translation initiation from SARS-CoV-2 5'UTRs depends on a mechanism that requires the integrity of eIF4G.

Finally, we evaluated the requirement for the DEAD box RNA helicase eIF4A, which is the third partner of the eIF4F complex and is involved in RNA unwinding and scanning [109]. Indeed, chemical and enzymatic probing analysis of the SARS-CoV-2 5'UTR^{ORF1ab} [73,74] has revealed a high level of RNA secondary structures (Figure 2) that may interfere with ribosome scanning. To test this hypothesis, we used a natural chemical inhibitor of the eIF4A RNA-binding activity called hippuristanol (Figure 6A) [110]. Addition of increasing amounts of hippuristanol in the RRL inhibited globin cap-dependent translation in a dose-dependent manner, whereas it did not interfere with expression of a mRNA driven by the Hepatitis C Virus (HCV) IRES (Figure 6G), which does not require eIF4A [111]. The RRL was then programmed with reporter genes containing SARS-CoV-2 5'UTR^{ORF1ab} or leader sequence RNA constructs and this showed that initiation on both of these RNAs was repressed upon addition of hippuristanol (Figure 6G) and the extent of translation repression was comparable to that obtained for globin mRNAs.

Taken together, our data indicate that translation from the SARS-CoV-2 5'UTR^{ORF1ab} requires the activity of each component of the eIF4F complex, namely eIF4E, eIF4G and eIF4A, which is consistent with a cap-dependent initiation mechanism.

3.4. uORF Downregulates the Translation Efficiency of gRNA

The 5'UTR^{ORF1ab} of the SARS-CoV-2 (gRNA) harbors a potential uORF of nine amino acids that starts at the AUG¹⁰⁷ triplet and finishes at the UAA triplet located at position 134 (Figures 2 and 7A). uORFs are *cis*-acting elements located in the 5'UTR that impact on the translation efficiency of the mRNA. When these uORFs are translated by incoming ribosomes from the 5' end, this often results in the decrease of translation initiation from the main downstream ORF [27,112–115]. However, in some particular cases, the presence of uORFs can also act as enhancers of protein production, as it was described with some stress-related mRNAs [29,114,116,117].

In the case of the SARS-CoV-2 gRNA, ribosomal profiling data have suggested that some ribosomes recognize and pause at the AUG¹⁰⁷ codon [86]. However, formal proof that this uORF is translated and the level at which it is expressed remain to be determined. To test this hypothesis, we introduced two mutations in the pSARS-CoV-2 5'UTR^{ORF1ab}-Renilla to change the AUG¹⁰⁷ triplet into an UAA triplet (Mut-AUG¹⁰⁷) and to delete the stop codon (UAA¹³⁴) in order to align AUG¹⁰⁷ in the same reading frame as the Renilla ORF (uORF phase). The resulting RNAs have been synthesized by *in vitro* transcription and translated into the RRL in the presence of S³⁵-methionine. Data are presented in Figure 7 and show that mutation of the AUG¹⁰⁷ triplet resulted in a slight stimulation of expression of the main ORF (Renilla) (Figure 7B lane 3), confirming that this AUG triplet was recognized as an initiation site by the scanning ribosome. This was confirmed by the second mutation (deletion of the UAA), which resulted in an increased expression of Renilla (Figure 7B lane 4, denoted by the star). Altogether, our data unambiguously show that the uORF located in the 5'UTR^{ORF1ab} was translated in the RRL, but the effect on protein expression, i.e., both production of the short peptide and increased production of Renilla, remained relatively modest. In order to better appreciate the magnitude of this effect, we have performed similar experiments on the 5'UTR^{ORF1ab} of the MERS-CoV gRNA as it also contains an uORF starting at the AUG¹⁵⁸ triplet, which is located closer to the AUG start codon of ORF-1ab (Figure 7A). To evaluate its role on mRNA translation, we have also mutated the AUG¹⁵⁸ triplet. The resulting mRNA was translated in the RRL and protein products were analyzed on SDS-PAGE and quantified with a phosphoImager. We observed that mutation of the AUG¹⁵⁸ triplet induced a strong expression of the main ORF, which was stimulated by up to three times when compared to the parental 5'UTR (Figure 7C). This stimulation was also observed when the reporter mRNAs were transfected in cell

lines (Figure 7D). These results confirmed important differences in translational control exerted by the uORFs in the coronavirus family with a moderate attenuation of translation in SARS-CoV-2 compared to a more pronounced impact on the MERS-CoV 5'UTR.

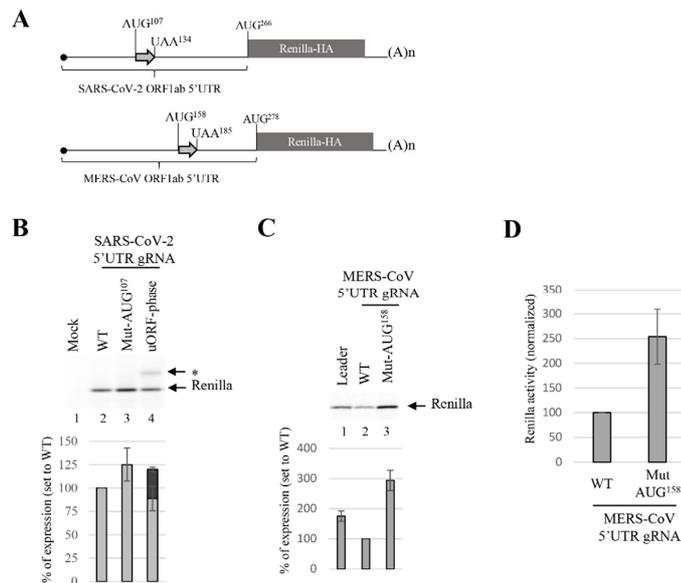


Figure 7. The 5'UTR uORF differentially regulates translation from SARS-CoV-2 and MERS-CoV. (A) Scheme of the reporter mRNA with the 5'UTR^{ORF1ab} of SARS-CoV-2 or MERS-CoV gRNA with their respective uORF and positions. (B) Capped and polyadenylated mRNA harboring the SARS-CoV-2 5'UTR gRNA WT or the Mut-AUG¹⁰⁷ or uORF-phase mutations were translated in the RRL. Translation products were resolved by 12% SDS-PAGE and quantified with a phosphoImager. (C) Capped and polyadenylated mRNA harboring the MERS-CoV 5'UTR gRNA WT or the Mut-AUG¹⁵⁸ mutation, were translated in the RRL. Translation products were resolved by 12% SDS-PAGE and quantified with a phosphoImager. (D) Jurkat cells were electroporated with capped and polyadenylated mRNAs as indicated. After 90 min of incubation, the cellular extracts were used for luciferase assay and RT-qPCR. Values shown are the mean (+/− S.D.) for three independent experiments.

4. Discussion

Infection by SARS-CoV-2 generates a gRNA and at least nine sgRNAs, which all start with a 75 nt. common leader sequence [71]. Each of these transcripts harbor a unique 5'UTR, which is composed of the shared leader sequence followed by an RNA region that varies in length (between 0 to 190 nt. (Table 1)). Thus, the resulting viral 5'UTRs have different degrees of RNA structures and some of them contain *cis*-acting elements, such as an uORF, which is known to have an impact on translation initiation efficiency. This set the rationale for this study aimed at investigating the translational properties of all virally encoded mRNAs. To answer to this question, we showed that viral 5'UTRs show significant variations of expression although they all exhibit a high level of activity compared to globin 5'UTR control. Importantly, the high and relative levels of expression between viral mRNAs were maintained regardless of the experimental systems used (RRL, competitive RRL and cells) (Figure 3), with the 5'UTR^{ORF1ab} (gRNA) being the most efficient of all. This is rather surprising considering that the 5'UTR^{ORF1ab} is the longest of the series (265 nt.), with multiple stable hairpin RNA structures and, above all, the presence of the uORF (Figure 2). A priori, these structural features should represent hurdles for ribosomal binding and scanning and one would predict a poor level of protein

expression. However, this is clearly not the case as expression was revealed to be very efficient and conserved between the two SARS-CoV strains (Figure 4). This suggested to us that the SARS-CoV-2 5'UTR^{ORF1ab} could use an alternative cap-independent mechanism; such an assumption was also strengthened by the presence of the AUG start codon in a four-way junction, which is a structural conformation that was previously described for HCV IRES [118]. This prompted us to look for the requirement of the viral gRNA in the initiation factors involved in cap-driven translation initiation. We showed that both the cap structure and the eIF4F complex are necessary to initiate translation, which strongly suggest a cap-dependent mechanism (Figures 5 and 6). We can assume that, after ribosomal attachment to the cap structure, the PIC scans the viral 5'UTR to reach the initiation codon as eIF4A is required (Figure 6G) and, above all, the uORF is translated (Figure 7). However, we cannot rule out that ribosomal shunting could occur through the 5'UTR to bypass some hairpin RNA structures [119–121] and contribute to this high level of translation. This mechanism usually needs additional proteins [121] and this will require further studies to determine its implication in SARS-CoV-2 translation initiation.

In the canonical translation initiation mechanism, the interaction between the poly(A) binding protein (PABP) and eIF4G [122,123] brings the 5'- and 3'-ends of the mRNA together and forms a mRNA's circularization called the closed-loop model [124,125], which promotes translation [89,126]. Interestingly, in our system, which is able to recapitulate the cap/poly(A) synergy [89], the poly(A) tail did not provide any translational advantage to the mRNAs that harbor the SARS-CoV-2 5'UTR^{ORF1ab} when compared to globin 5'UTR, which uses a strict canonical cap-dependent translation initiation mechanism (Figure 5). This particularity is quite unusual in the eukaryotic kingdom and this raises the question of whether this closed-loop conformation takes place on mRNAs that harbor the SARS-CoV-2 5'UTR. The poly(A) tail requirement should be further investigated in the future to determine if the PABP is necessary to promote initiation and still interacts with eIF4G. To our knowledge, a neutral influence of the poly(A) tail, in translation, has never been reported before. Viral and cellular mRNAs without a poly(A) tail have been described but they use alternative and specific *cis*-acting elements to bring the 5'- and 3'-ends of the mRNA together [127–132]. Analyses of the formation of 48S PIC in RRL through a sucrose gradient have indicated that it is almost absent on mRNAs harboring the SARS-CoV-2 5'UTR but not with the EMCV IRES [80], suggesting that the initiation process on SARS-CoV-2 is not canonical. In SARS-CoV-2 infected cells, the size of the poly(A) tail is quite short (47 nt. for median length) and is longer on the gRNA than on the sgRNAs [71]. Interestingly, during bovine CoV infection, the length of the poly(A) tail changes from ~45 nt. immediately after virus entry to ~65 nt. at 6 to 9 h post-infection [133]. These quick modifications of the length are reminiscent of the translational control that takes place during early oocyte/embryo development [134]. Although we did not notice a requirement for the poly(A) tail in SARS-CoV-2 translation in the RRL, this does not exclude that the poly(A) tail of SARS-CoV-2 mRNAs may regulate translation efficiency, mRNA stability and/or mRNA localization in infected cells.

The exact mechanism by which the SARS-CoV-2 5'UTRs drive the high levels of expression is puzzling considering the structural features of the gRNA; nevertheless, this characteristic could turn out to be a real benefit for viral replication in infected cells. Indeed, in the early phases of infection, the viral nsp1 interacts with the 18S ribosomal RNA in the mRNA entry channel of the ribosome and imposes a blockage by steric hindrance, leading to a global inhibition of mRNA translation [135–137]. However, viral mRNAs can escape this nsp1-mediated repression via a complex interplay between the leader sequence, especially the SL1 RNA structure [78,80,81,137], which is able to interact with nsp1 [80,138]. It has been shown that, at low concentrations, nsp1 is able to stimulate the expression of mRNA reporters that harbor the SL1 sequence in cells [81] but not in the RRL [80], which is not a competitive system. This, along with our data, suggests that SARS viral RNAs have the ability to compete very efficiently with cellular transcripts for the translational

apparatus. Thus, it would be very interesting to assay the role of nsp1 in the untreated RRL as well as in infected cells.

Finally, the expression of the gRNA is also regulated by an uORF localized within its 5'UTR, some 120 nucleotides upstream from the authentic AUG initiation codon. We first showed that this uORF was recognized by scanning ribosomes and translated into a short peptide whose role remains to be determined. In addition, mutation of the start AUG codon of the uORF induced only a mild stimulation of the expression of the main reading frame. This contrasts with the situation obtained with the MERS-CoV, in which we showed that this uORF exerts a strong control on translation (Figure 7C,D). This is puzzling as the uORF is conserved, suggesting its importance in the control and regulation of protein synthesis despite having different effects on the flux of incoming ribosomes.

Genome-wide sequencing of 5'UTR has revealed that uORFs are much more pervasive in the genome than initially described, with an occurrence of about 50% in mammalian mRNAs [139,140]. These uORFs can be located within non-coding RNAs, within the 5' and 3' UTRs or they can also be overlapping with the main ORF, either in-frame or out-of-frame with the main protein product [141]. According to their localization, these uORFs regulate gene expression in at least three different ways: (i) by imposing a sharp translational control on the main ORF; (ii) by producing functional micro-peptides and (iii) generating truncated/extended isoforms of the main protein product.

In the case of the uORF present in the SARS/MERS coronavirus, it is clearly located within the 5'UTR^{ORF1ab} and encodes a putative short peptide of 9–11 amino acids, which does not appear conserved amongst the different members (Figure S3). Thus, it suggests that this uORF regulates expression of the main protein product by throttling down the flux of incoming ribosomes. This can happen by either stalling the 80S ribosomal complex that forms on the encoded short reading frame, thus creating a roadblock for scanning ribosomes or, alternatively, the short reading frame can also be a waylay for scanning pre-initiation complexes [22]. As the same conserved uORF has a different impact on expression from the SARS-CoV-1 and -2 compared to the MERS-CoV, this could reflect the fact that two distinct mechanisms are at play. This is an important point that will be addressed in the future. In any case, this uORF must be a key regulator of viral translation and replication as it is widely conserved. Indeed, this uORF is highly conserved from different coronaviruses [142] and recent ribosome profiling analyses have confirmed that its AUG initiation codon is recognized by ribosomal complexes during both Mouse Hepatitis Virus (MHV) and SARS-CoV-2 infection [86,143]. One study has focused on the physiological function of the peptide issued by this uORF in MHV infection. For that, the authors mutated the AUG-initiated uORF in the viral genome, showing that the virus was viable in cell culture. However, serial passage of this mutant variant resulted in a reversion to an intact uORF, indicating that this uORF must play a beneficial role in virus survival in cell culture [142].

To summarize, our data strongly suggest that translation of the SARS-CoV-2 mRNAs involves an unconventional cap-dependent mechanism that involves *cis*-acting RNA elements such as the structure of the leader, the length of the poly(A) tail and the role of the uORF. Taken together, these elements must probably functionally interact with each other in order to regulate the flow of ribosomes on the viral mRNAs.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/v14071505/s1>, Figure S1: N-terminal modification of the Renilla luciferase coding region; Figure S2: SARS CoV-1/2 and MERS CoV gRNA 5'UTRs conservation; Figure S3: 5'UTR^{ORF1ab} uORFs conservation Table S1: List of primers used for the cloning of the SARS-CoV 1/2 and MERS-CoV leader and 5'UTR^{ORF1ab}; Table S2: List of primers used for the cloning of the different 5'UTRs of the SARS-CoV-2 gRNA and sgRNAs; Table S3: List of primers used for the cloning of the uORFs mutations.

Author Contributions: Conceptualization, L.C., S.d.B. and T.O.; methodology, L.C. and S.d.B.; validation, L.C. and S.d.B.; formal analysis, L.C., O.A. and S.d.B.; investigation, L.C. and S.d.B.;

writing—original draft preparation, L.C. and S.d.B.; writing—review and editing, L.C., S.d.B. and T.O.; visualization, O.A. and S.d.B.; supervision, S.d.B. and T.O.; project administration, S.d.B. and T.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Requests for further information about resources, reagents, and data availability should be directed to the corresponding author.

Acknowledgments: We thank J. Pelletier (McGill University, Canada) and B. Sargueil (Paris Descartes University, France) for the generous gift of hippuristanol.

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

References

- Merrick, W.C.; Pavitt, G.D. Protein Synthesis Initiation in Eukaryotic Cells. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a033092. [[CrossRef](#)] [[PubMed](#)]
- Pelletier, J.; Sonenberg, N. The Organizing Principles of Eukaryotic Ribosome Recruitment. *Annu. Rev. Biochem.* **2019**, *88*, 307–335. [[CrossRef](#)] [[PubMed](#)]
- Kapp, L.D.; Lorsch, J.R. GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. *J. Mol. Biol.* **2004**, *335*, 923–936. [[CrossRef](#)] [[PubMed](#)]
- Pelletier, J.; Schmeing, T.M.; Sonenberg, N. The multifaceted eukaryotic cap structure. *Wiley Interdiscip. Rev. RNA* **2021**, *12*, e1636. [[CrossRef](#)]
- LeFebvre, A.K.; Korneeva, N.L.; Trutschl, M.; Cvek, U.; Duzan, R.D.; Bradley, C.A.; Hershey, J.W.; Rhoads, R.E. Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J. Biol. Chem.* **2006**, *281*, 22917–22932. [[CrossRef](#)]
- Villa, N.; Do, A.; Hershey, J.W.; Fraser, C.S. Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. *J. Biol. Chem.* **2013**, *288*, 32932–32940. [[CrossRef](#)]
- Kumar, P.; Hellen, C.U.; Pestova, T.V. Toward the mechanism of eIF4F-mediated ribosomal attachment to mammalian capped mRNAs. *Genes Dev.* **2016**, *30*, 1573–1588. [[CrossRef](#)]
- Hinnebusch, A.G. Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem. Sci.* **2017**, *42*, 589–611. [[CrossRef](#)]
- Advani, V.M.; Ivanov, P. Translational Control under Stress: Reshaping the Translatome. *Bioessays* **2019**, *41*, e1900009. [[CrossRef](#)]
- Proud, C.G. Phosphorylation and Signal Transduction Pathways in Translational Control. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a033050. [[CrossRef](#)]
- Wek, R.C.; Jiang, H.Y.; Anthony, T.G. Coping with stress: eIF2 kinases and translational control. *Biochem. Soc. Trans.* **2006**, *34 Pt 1*, 7–11. [[CrossRef](#)] [[PubMed](#)]
- Leppek, K.; Barna, M. An rRNA variant to deal with stress. *Nat. Microbiol.* **2019**, *4*, 382–383. [[CrossRef](#)] [[PubMed](#)]
- Kozak, M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **1986**, *44*, 283–292. [[CrossRef](#)]
- Kozak, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **1987**, *15*, 8125–8148. [[CrossRef](#)]
- Kozak, M. The scanning model for translation: An update. *J. Cell Biol.* **1989**, *108*, 229–241. [[CrossRef](#)]
- Kozak, M. Initiation of translation in prokaryotes and eukaryotes. *Gene* **1999**, *234*, 187–208. [[CrossRef](#)]
- Kozak, M. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* **1987**, *196*, 947–950. [[CrossRef](#)]
- Kozak, M. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res.* **1981**, *9*, 5233–5252. [[CrossRef](#)]
- Kozak, M. Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo. *Nature* **1984**, *308*, 241–246. [[CrossRef](#)]
- Pisarev, A.V.; Kolupaeva, V.G.; Pisareva, V.P.; Merrick, W.C.; Hellen, C.U.; Pestova, T.V. Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev.* **2006**, *20*, 624–636. [[CrossRef](#)]
- Hussain, T.; Llacer, J.L.; Fernandez, I.S.; Munoz, A.; Martin-Marcos, P.; Savva, C.G.; Lorsch, J.R.; Hinnebusch, A.G.; Ramakrishnan, V. Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **2014**, *159*, 597–607. [[CrossRef](#)] [[PubMed](#)]
- Hinnebusch, A.G.; Ivanov, I.P.; Sonenberg, N. Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science* **2016**, *352*, 1413–1416. [[CrossRef](#)] [[PubMed](#)]

23. Noderer, W.L.; Flockhart, R.J.; Bhaduri, A.; Diaz de Arce, A.J.; Zhang, J.; Khavari, P.A.; Wang, C.L. Quantitative analysis of mammalian translation initiation sites by FACS-seq. *Mol. Syst. Biol.* **2014**, *10*, 748. [[CrossRef](#)] [[PubMed](#)]
24. Xia, X. The +4G site in Kozak consensus is not related to the efficiency of translation initiation. *PLoS ONE* **2007**, *2*, e188. [[CrossRef](#)]
25. Volkova, O.A.; Kochetov, A.V. Interrelations between the nucleotide context of human start AUG codon, N-end amino acids of the encoded protein and initiation of translation. *J. Biomol. Struct. Dyn.* **2010**, *27*, 611–618. [[CrossRef](#)]
26. Chen, H.H.; Tarn, W.Y. uORF-mediated translational control: Recently elucidated mechanisms and implications in cancer. *RNA Biol.* **2019**, *16*, 1327–1338. [[CrossRef](#)] [[PubMed](#)]
27. Johnstone, T.G.; Bazzini, A.A.; Giraldez, A.J. Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J.* **2016**, *35*, 706–723. [[CrossRef](#)]
28. Wethmar, K. The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip. Rev. RNA* **2014**, *5*, 765–778. [[CrossRef](#)]
29. Palam, L.R.; Baird, T.D.; Wek, R.C. Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. *J. Biol. Chem.* **2011**, *286*, 10939–10949. [[CrossRef](#)]
30. Vattem, K.M.; Wek, R.C. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11269–11274. [[CrossRef](#)]
31. Zhou, D.; Palam, L.R.; Jiang, L.; Narasimhan, J.; Staschke, K.A.; Wek, R.C. Phosphorylation of eIF2 directs ATF5 translational control in response to diverse stress conditions. *J. Biol. Chem.* **2008**, *283*, 7064–7073. [[CrossRef](#)] [[PubMed](#)]
32. Kozak, M. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol. Cell. Biol.* **1989**, *9*, 5134–5142. [[PubMed](#)]
33. Babendure, J.R.; Babendure, J.L.; Ding, J.H.; Tsien, R.Y. Control of mammalian translation by mRNA structure near caps. *RNA* **2006**, *12*, 851–861. [[CrossRef](#)]
34. Pelletier, J.; Sonenberg, N. Photochemical cross-linking of cap binding proteins to eucaryotic mRNAs: Effect of mRNA 5' secondary structure. *Mol. Cell. Biol.* **1985**, *5*, 3222–3230.
35. Pelletier, J.; Sonenberg, N. Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell* **1985**, *40*, 515–526. [[CrossRef](#)]
36. Kozak, M. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 2850–2854. [[CrossRef](#)]
37. Parsyan, A.; Svitkin, Y.; Shahbazian, D.; Gkogkas, C.; Lasko, P.; Merrick, W.C.; Sonenberg, N. mRNA helicases: The tacticians of translational control. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 235–245. [[CrossRef](#)] [[PubMed](#)]
38. Lai, M.C.; Wang, S.W.; Cheng, L.; Tarn, W.Y.; Tsai, S.J.; Sun, H.S. Human DDX3 interacts with the HIV-1 Tat protein to facilitate viral mRNA translation. *PLoS ONE* **2013**, *8*, e68665.
39. Soto-Rifo, R.; Rubilar, P.S.; Limousin, T.; de Breyne, S.; Decimo, D.; Ohlmann, T. DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. *EMBO J.* **2012**, *31*, 3745–3756. [[CrossRef](#)]
40. Hentze, M.W.; Caughman, S.W.; Rouault, T.A.; Barriocanal, J.G.; Dancis, A.; Harford, J.B.; Klausner, R.D. Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. *Science* **1987**, *238*, 1570–1573. [[CrossRef](#)]
41. Hentze, M.W.; Rouault, T.A.; Caughman, S.W.; Dancis, A.; Harford, J.B.; Klausner, R.D. A cis-acting element is necessary and sufficient for translational regulation of human ferritin expression in response to iron. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 6730–6734. [[CrossRef](#)] [[PubMed](#)]
42. Gray, N.K.; Hentze, M.W. Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs. *EMBO J.* **1994**, *13*, 3882–3891. [[CrossRef](#)] [[PubMed](#)]
43. Muckenthaler, M.; Gray, N.K.; Hentze, M.W. IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. *Mol. Cell* **1998**, *2*, 383–388. [[CrossRef](#)]
44. Friebe, P.; Lohmann, V.; Krieger, N.; Bartenschlager, R. Sequences in the 5' nontranslated region of hepatitis C virus required for RNA replication. *J. Virol.* **2001**, *75*, 12047–12057. [[CrossRef](#)] [[PubMed](#)]
45. Bieniasz, P.; Telesnitsky, A. Multiple, Switchable Protein:RNA Interactions Regulate Human Immunodeficiency Virus Type 1 Assembly. *Annu. Rev. Virol.* **2018**, *5*, 165–183. [[CrossRef](#)]
46. Malone, B.; Urakova, N.; Snijder, E.J.; Campbell, E.A. Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat. Rev. Mol. Cell Biol.* **2022**, *23*, 21–39. [[CrossRef](#)]
47. Jang, S.K.; Krausslich, H.G.; Nicklin, M.J.; Duke, G.M.; Palmenberg, A.C.; Wimmer, E. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J. Virol.* **1988**, *62*, 2636–2643. [[CrossRef](#)]
48. Pelletier, J.; Sonenberg, N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **1988**, *334*, 320–325. [[CrossRef](#)]
49. Martinez-Salas, E.; Francisco-Velilla, R.; Fernandez-Chamorro, J.; Embarek, A.M. Insights into Structural and Mechanistic Features of Viral IRES Elements. *Front. Microbiol.* **2017**, *8*, 2629. [[CrossRef](#)]
50. Mailliot, J.; Martin, F. Viral internal ribosomal entry sites: Four classes for one goal. *Wiley Interdiscip. Rev. RNA* **2018**, *9*. [[CrossRef](#)]
51. Kwan, T.; Thompson, S.R. Noncanonical Translation Initiation in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032672. [[CrossRef](#)] [[PubMed](#)]

52. Komar, A.A.; Hatzoglou, M. Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states. *Cell Cycle* **2011**, *10*, 229–240. [[CrossRef](#)] [[PubMed](#)]
53. Yang, T.H.; Wang, C.Y.; Tsai, H.C.; Liu, C.T. Human IRES Atlas: An integrative platform for studying IRES-driven translational regulation in humans. *Database* **2021**, *2021*, baab025. [[CrossRef](#)] [[PubMed](#)]
54. De Breyne, S.; Yu, Y.; Unbehauen, A.; Pestova, T.V.; Hellen, C.U. Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9197–9202. [[CrossRef](#)]
55. Pestova, T.V.; Hellen, C.U.; Shatsky, I.N. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell. Biol.* **1996**, *16*, 6859–6869. [[CrossRef](#)]
56. Sizova, D.V.; Kolupaeva, V.G.; Pestova, T.V.; Shatsky, I.N.; Hellen, C.U. Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. *J. Virol.* **1998**, *72*, 4775–4782. [[CrossRef](#)]
57. Pestova, T.V.; Lomakin, I.B.; Hellen, C.U. Position of the CrPV IRES on the 40S subunit and factor dependence of IRES/80S ribosome assembly. *EMBO Rep.* **2004**, *5*, 906–913. [[CrossRef](#)]
58. Kolupaeva, V.G.; Pestova, T.V.; Hellen, C.U.; Shatsky, I.N. Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. *J. Biol. Chem.* **1998**, *273*, 18599–18604. [[CrossRef](#)]
59. Wilson, J.E.; Pestova, T.V.; Hellen, C.U.; Sarnow, P. Initiation of protein synthesis from the A site of the ribosome. *Cell* **2000**, *102*, 511–520. [[CrossRef](#)]
60. Spahn, C.M.; Jan, E.; Mulder, A.; Grassucci, R.A.; Sarnow, P.; Frank, J. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: The IRES functions as an RNA-based translation factor. *Cell* **2004**, *118*, 465–475. [[CrossRef](#)]
61. De Breyne, S.; Yu, Y.; Pestova, T.V.; Hellen, C.U. Factor requirements for translation initiation on the Simian picornavirus internal ribosomal entry site. *RNA* **2008**, *14*, 367–380. [[CrossRef](#)] [[PubMed](#)]
62. Hashem, Y.; des Georges, A.; Dhote, V.; Langlois, R.; Liao, H.Y.; Grassucci, R.A.; Pestova, T.V.; Hellen, C.U.; Frank, J. Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **2013**, *503*, 539–543. [[CrossRef](#)] [[PubMed](#)]
63. Chamond, N.; Deforges, J.; Ulryck, N.; Sargueil, B. 40S recruitment in the absence of eIF4G/4A by EMCV IRES refines the model for translation initiation on the archetype of Type II IRESs. *Nucleic Acids Res.* **2014**, *42*, 10373–10384. [[CrossRef](#)]
64. Stern-Ginossar, N.; Thompson, S.R.; Mathews, M.B.; Mohr, I. Translational Control in Virus-Infected Cells. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a033001. [[CrossRef](#)] [[PubMed](#)]
65. Sajjanar, B.; Deb, R.; Raina, S.K.; Pawar, S.; Brahmane, M.P.; Nirmale, A.V.; Kurade, N.P.; Manjunathareddy, G.B.; Bal, S.K.; Singh, N.P. Untranslated regions (UTRs) orchestrate translation reprogramming in cellular stress responses. *J. Biol.* **2017**, *65*, 69–75. [[CrossRef](#)]
66. Godet, A.C.; David, F.; Hantelys, F.; Tatin, F.; Lacazette, E.; Garmy-Susini, B.; Prats, A.C. IRES Trans-Acting Factors, Key Actors of the Stress Response. *Int. J. Mol. Sci.* **2019**, *20*, 924. [[CrossRef](#)]
67. Jaafar, Z.A.; Kieft, J.S. Viral RNA structure-based strategies to manipulate translation. *Nat. Rev. Microbiol.* **2019**, *17*, 110–123. [[CrossRef](#)]
68. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382*, 727–733. [[CrossRef](#)]
69. Sola, I.; Almazan, F.; Zuniga, S.; Enjuanes, L. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu. Rev. Virol.* **2015**, *2*, 265–288. [[CrossRef](#)]
70. Brant, A.C.; Tian, W.; Majerciak, V.; Yang, W.; Zheng, Z.M. SARS-CoV-2: From its discovery to genome structure, transcription, and replication. *Cell Biosci.* **2021**, *11*, 136. [[CrossRef](#)]
71. Kim, D.; Lee, J.Y.; Yang, J.S.; Kim, J.W.; Kim, V.N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020**, *181*, 914–921.e10. [[CrossRef](#)] [[PubMed](#)]
72. Kim, D.; Kim, S.; Park, J.; Chang, H.R.; Chang, J.; Ahn, J.; Park, H.; Park, J.; Son, N.; Kang, G.; et al. A high-resolution temporal atlas of the SARS-CoV-2 transcriptome and transcriptome. *Nat. Commun.* **2021**, *12*, 5120. [[CrossRef](#)] [[PubMed](#)]
73. Miao, Z.; Tidu, A.; Eriani, G.; Martin, F. Secondary structure of the SARS-CoV-2 5'-UTR. *RNA Biol.* **2021**, *18*, 447–456. [[CrossRef](#)] [[PubMed](#)]
74. Lan, T.C.T.; Allan, M.F.; Malsick, L.E.; Woo, J.Z.; Zhu, C.; Zhang, F.; Khandwala, S.; Nyeo, S.S.Y.; Sun, Y.; Guo, J.U.; et al. Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. *Nat. Commun.* **2022**, *13*, 1128. [[CrossRef](#)]
75. Madhugiri, R.; Fricke, M.; Marz, M.; Ziebuhr, J. Coronavirus cis-Acting RNA Elements. *Adv. Virus Res.* **2016**, *96*, 127–163.
76. Madhugiri, R.; Karl, N.; Petersen, D.; Lamkiewicz, K.; Fricke, M.; Wend, U.; Scheuer, R.; Marz, M.; Ziebuhr, J. Structural and functional conservation of cis-acting RNA elements in coronavirus 5'-terminal genome regions. *Virology* **2018**, *517*, 44–55. [[CrossRef](#)]
77. Yang, D.; Leibowitz, J.L. The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res.* **2015**, *206*, 120–133. [[CrossRef](#)]
78. Vora, S.M.; Fontana, P.; Mao, T.; Leger, V.; Zhang, Y.; Fu, T.M.; Lieberman, J.; Gehrke, L.; Shi, M.; Wang, L.; et al. Targeting stem-loop 1 of the SARS-CoV-2 5' UTR to suppress viral translation and Nsp1 evasion. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2117198119. [[CrossRef](#)]

79. Sosnowski, P.; Tidu, A.; Eriani, G.; Westhof, E.; Martin, F. Correlated sequence signatures are present within the genomic 5'UTR RNA and NSP1 protein in coronaviruses. *RNA* **2022**, *28*, 729–741. [[CrossRef](#)]
80. Tidu, A.; Janvier, A.; Schaeffer, L.; Sosnowski, P.; Kuhn, L.; Hammann, P.; Westhof, E.; Eriani, G.; Martin, F. The viral protein NSP1 acts as a ribosome gatekeeper for shutting down host translation and fostering SARS-CoV-2 translation. *RNA* **2020**, *27*. [[CrossRef](#)]
81. Bujanic, L.; Shevchuk, O.; von Kugelgen, N.; Kalinina, A.; Ludwik, K.; Koppstein, D.; Zerna, N.; Sickmann, A.; Chekulaeva, M. The key features of SARS-CoV-2 leader and NSP1 required for viral escape of NSP1-mediated repression. *RNA* **2022**, *28*, 766–779. [[CrossRef](#)] [[PubMed](#)]
82. Lukavsky, P.J. Structure and function of HCV IRES domains. *Virus Res.* **2009**, *139*, 166–171. [[CrossRef](#)] [[PubMed](#)]
83. Fletcher, S.P.; Jackson, R.J. Pestivirus internal ribosome entry site (IRES) structure and function: Elements in the 5' untranslated region important for IRES function. *J. Virol.* **2002**, *76*, 5024–5033. [[CrossRef](#)] [[PubMed](#)]
84. Gross, L.; Vicens, Q.; Einhorn, E.; Noireterre, A.; Schaeffer, L.; Kuhn, L.; Imler, J.L.; Eriani, G.; Meignin, C.; Martin, F. The IRES5'UTR of the dicistrovirus cricket paralysis virus is a type III IRES containing an essential pseudoknot structure. *Nucleic Acids Res.* **2017**, *45*, 8993–9004. [[CrossRef](#)] [[PubMed](#)]
85. Fernandez-Miragall, O.; Martinez-Salas, E. Structural organization of a viral IRES depends on the integrity of the GNRA motif. *RNA* **2003**, *9*, 1333–1344. [[CrossRef](#)] [[PubMed](#)]
86. Finkel, Y.; Mizrahi, O.; Nachshon, A.; Weingarten-Gabbay, S.; Morgenstern, D.; Yahalom-Ronen, Y.; Tamir, H.; Achdout, H.; Stein, D.; Israeli, O.; et al. The coding capacity of SARS-CoV-2. *Nature* **2021**, *589*, 125–130. [[CrossRef](#)]
87. De Breyne, S.; Chamond, N.; Decimo, D.; Traub, M.A.; Andre, P.; Sargueil, B.; Ohlmann, T. In vitro studies reveal that different modes of initiation on HIV-1 mRNA have different levels of requirement for eukaryotic initiation factor 4F. *FEBS J.* **2012**, *279*, 3098–3111. [[CrossRef](#)]
88. Dizin, E.; Gressier, C.; Magnard, C.; Ray, H.; Decimo, D.; Ohlmann, T.; Dalla Venezia, N. BRCA1 interacts with poly(A)-binding protein: Implication of BRCA1 in translation regulation. *J. Biol. Chem.* **2006**, *281*, 24236–24246. [[CrossRef](#)]
89. Soto Rifo, R.; Ricci, E.P.; Decimo, D.; Moncorge, O.; Ohlmann, T. Back to basics: The untreated rabbit reticulocyte lysate as a competitive system to recapitulate cap/poly(A) synergy and the selective advantage of IRES-driven translation. *Nucleic Acids Res.* **2007**, *35*, e121. [[CrossRef](#)]
90. Jackson, R.J.; Hunt, T. Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. *Methods Enzymol.* **1983**, *96*, 50–74.
91. Pelham, H.R.; Jackson, R.J. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur. J. Biochem.—FEBS* **1976**, *67*, 247–256. [[CrossRef](#)]
92. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.; Clamp, M.; Barton, G.J. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [[CrossRef](#)] [[PubMed](#)]
93. Deforges, J.; de Breyne, S.; Ameer, M.; Ulryck, N.; Chamond, N.; Saaidi, A.; Ponty, Y.; Ohlmann, T.; Sargueil, B. Two ribosome recruitment sites direct multiple translation events within HIV1 Gag open reading frame. *Nucleic Acids Res.* **2017**, *45*, 7382–7400. [[CrossRef](#)] [[PubMed](#)]
94. Rota, P.A.; Oberste, M.S.; Monroe, S.S.; Nix, W.A.; Campagnoli, R.; Icenogle, J.P.; Penaranda, S.; Bankamp, B.; Maher, K.; Chen, M.H.; et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **2003**, *300*, 1394–1399. [[CrossRef](#)] [[PubMed](#)]
95. Marra, M.A.; Jones, S.J.; Astell, C.R.; Holt, R.A.; Brooks-Wilson, A.; Butterfield, Y.S.; Khattra, J.; Asano, J.K.; Barber, S.A.; Chan, S.Y.; et al. The Genome sequence of the SARS-associated coronavirus. *Science* **2003**, *300*, 1399–1404. [[CrossRef](#)] [[PubMed](#)]
96. Ksiazek, T.G.; Erdman, D.; Goldsmith, C.S.; Zaki, S.R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J.A.; Lim, W.; et al. A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* **2003**, *348*, 1953–1966. [[CrossRef](#)]
97. Bermingham, A.; Chand, M.A.; Brown, C.S.; Aarons, E.; Tong, C.; Langrish, C.; Hoschler, K.; Brown, K.; Galiano, M.; Myers, R.; et al. Severe respiratory illness caused by a novel coronavirus, in a patient transferred to the United Kingdom from the Middle East, September 2012. *Eurosurveillance* **2012**, *17*, 20290. [[CrossRef](#)]
98. Lu, G.; Liu, D. SARS-like virus in the Middle East: A truly bat-related coronavirus causing human diseases. *Protein Cell* **2012**, *3*, 803–805. [[CrossRef](#)]
99. Zaki, A.M.; van Boheemen, S.; Bestebroer, T.M.; Osterhaus, A.D.; Fouchier, R.A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* **2012**, *367*, 1814–1820. [[CrossRef](#)]
100. Senanayake, S.D.; Brian, D.A. Bovine coronavirus I protein synthesis follows ribosomal scanning on the bicistronic N mRNA. *Virus Res.* **1997**, *48*, 101–105. [[CrossRef](#)]
101. Senanayake, S.D.; Brian, D.A. Translation from the 5' untranslated region (UTR) of mRNA 1 is repressed, but that from the 5' UTR of mRNA 7 is stimulated in coronavirus-infected cells. *J. Virol.* **1999**, *73*, 8003–8009. [[CrossRef](#)] [[PubMed](#)]
102. Muller, C.; Schulte, F.W.; Lange-Grunweller, K.; Obermann, W.; Madhugiri, R.; Pleschka, S.; Ziebuhr, J.; Hartmann, R.K.; Grunweller, A. Broad-spectrum antiviral activity of the eIF4A inhibitor silvestrol against corona- and picornaviruses. *Antivir. Res.* **2018**, *150*, 123–129. [[CrossRef](#)] [[PubMed](#)]
103. Cencic, R.; Desforges, M.; Hall, D.R.; Kozakov, D.; Du, Y.; Min, J.; Dingleline, R.; Fu, H.; Vajda, S.; Talbot, P.J.; et al. Blocking eIF4E-eIF4G interaction as a strategy to impair coronavirus replication. *J. Virol.* **2011**, *85*, 6381–6389. [[CrossRef](#)]

104. Jackson, R.J.; Hunt, S.L.; Reynolds, J.E.; Kaminski, A. Cap-dependent and cap-independent translation: Operational distinctions and mechanistic interpretations. In *Cap-Independent Translation; Current Topics in Microbiology and Immunology*; Springer: Berlin/Heidelberg, Germany, 1995; Volume 203, pp. 1–29.
105. Ohlmann, T.; Rau, M.; Morley, S.; Pain, V.M. Effect of cleavage of the p220 subunit of eukaryotic translation initiation factor eIF-4F on protein synthesis in vitro. *Biochem. Soc. Trans.* **1995**, *23*, 315S. [[CrossRef](#)]
106. Ohlmann, T.; Rau, M.; Morley, S.J.; Pain, V.M. Proteolytic cleavage of initiation factor eIF-4 gamma in the reticulocyte lysate inhibits translation of capped mRNAs but enhances that of uncapped mRNAs. *Nucleic Acids Res.* **1995**, *23*, 334–340. [[CrossRef](#)] [[PubMed](#)]
107. Prevot, D.; Darlix, J.L.; Ohlmann, T. Conducting the initiation of protein synthesis: The role of eIF4G. *Biol. Cell Under Auspices Eur. Cell Biol. Organ.* **2003**, *95*, 141–156.
108. Borman, A.M.; Kirchweber, R.; Ziegler, E.; Rhoads, R.E.; Skern, T.; Kean, K.M. eIF4G and its proteolytic cleavage products: Effect on initiation of protein synthesis from capped, uncapped, and IRES-containing mRNAs. *RNA—A Publ. RNA Soc.* **1997**, *3*, 186–196.
109. Linder, P.; Fuller-Pace, F.V. The biology of RNA helicases- modulation for life. Preface. *Biochim. Biophys. Acta* **2013**, *1829*, 749. [[CrossRef](#)]
110. Cencic, R.; Pelletier, J. Hippuristanol—A potent steroid inhibitor of eukaryotic initiation factor 4A. *Translation* **2016**, *4*, e1137381. [[CrossRef](#)]
111. Pestova, T.V.; Shatsky, I.N.; Fletcher, S.P.; Jackson, R.J.; Hellen, C.U. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev.* **1998**, *12*, 67–83. [[CrossRef](#)]
112. Morris, D.R.; Geballe, A.P. Upstream open reading frames as regulators of mRNA translation. *Mol. Cell. Biol.* **2000**, *20*, 8635–8642. [[CrossRef](#)] [[PubMed](#)]
113. Calvo, S.E.; Pagliarini, D.J.; Mootha, V.K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7507–7512. [[CrossRef](#)] [[PubMed](#)]
114. Silva, J.; Fernandes, R.; Romao, L. Translational Regulation by Upstream Open Reading Frames and Human Diseases. *Adv. Exp. Med. Biol.* **2019**, *1157*, 99–116. [[PubMed](#)]
115. Zhang, H.; Wang, Y.; Lu, J. Function and Evolution of Upstream ORFs in Eukaryotes. *Trends Biochem. Sci.* **2019**, *44*, 782–794. [[CrossRef](#)] [[PubMed](#)]
116. Spriggs, K.A.; Bushell, M.; Willis, A.E. Translational regulation of gene expression during conditions of cell stress. *Mol. Cell* **2010**, *40*, 228–237. [[CrossRef](#)] [[PubMed](#)]
117. Young, S.K.; Wek, R.C. Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. *J. Biol. Chem.* **2016**, *291*, 16927–16935. [[CrossRef](#)]
118. Berry, K.E.; Waghray, S.; Mortimer, S.A.; Bai, Y.; Doudna, J.A. Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning. *Structure* **2011**, *19*, 1456–1466. [[CrossRef](#)]
119. Haimov, O.; Sinvani, H.; Dikstein, R. Cap-dependent, scanning-free translation initiation mechanisms. *Biochim. Biophys. Acta* **2015**, *1849*, 1313–1318. [[CrossRef](#)]
120. Futterer, J.; Kiss-Laszlo, Z.; Hohn, T. Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* **1993**, *73*, 789–802. [[CrossRef](#)]
121. Sorokin, I.I.; Vassilenko, K.S.; Terenin, I.M.; Kalinina, N.O.; Agol, V.I.; Dmitriev, S.E. Non-Canonical Translation Initiation Mechanisms Employed by Eukaryotic Viral mRNAs. *Biochemistry* **2021**, *86*, 1060–1094. [[CrossRef](#)]
122. Tarun, S.Z., Jr.; Sachs, A.B. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J.* **1996**, *15*, 7168–7177. [[CrossRef](#)] [[PubMed](#)]
123. Imataka, H.; Gradi, A.; Sonenberg, N. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* **1998**, *17*, 7480–7489. [[CrossRef](#)] [[PubMed](#)]
124. Kahvejian, A.; Roy, G.; Sonenberg, N. The mRNA closed-loop model: The function of PABP and PABP-interacting proteins in mRNA translation. *Cold Spring Harb. Symp. Quant. Biol.* **2001**, *66*, 293–300. [[CrossRef](#)] [[PubMed](#)]
125. Jacobson, A. Poly(A) Metabolism and Translation: The Closed-loop Model. In *Translational Control*; Hershey, J.W.B.M., Sonenberg, N.M.B., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring, NY, USA, 1996; pp. 451–480.
126. Gallie, D.R. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev.* **1991**, *5*, 2108–2116. [[CrossRef](#)] [[PubMed](#)]
127. Poncet, D.; Laurent, S.; Cohen, J. Four nucleotides are the minimal requirement for RNA recognition by rotavirus non-structural protein NSP3. *EMBO J.* **1994**, *13*, 4165–4173. [[CrossRef](#)]
128. Vende, P.; Piron, M.; Castagne, N.; Poncet, D. Efficient translation of rotavirus mRNA requires simultaneous interaction of NSP3 with the eukaryotic translation initiation factor eIF4G and the mRNA 3' end. *J. Virol.* **2000**, *74*, 7064–7071. [[CrossRef](#)]
129. Grot, C.M.; Burley, S.K. Recognition of eIF4G by rotavirus NSP3 reveals a basis for mRNA circularization. *Mol Cell* **2002**, *9*, 1273–1283. [[CrossRef](#)]
130. Michel, Y.M.; Borman, A.M.; Paulous, S.; Kean, K.M. Eukaryotic initiation factor 4G-poly(A) binding protein interaction is required for poly(A) tail-mediated stimulation of picornavirus internal ribosome entry segment-driven translation but not for X-mediated stimulation of hepatitis C virus translation. *Mol. Cell. Biol.* **2001**, *21*, 4097–4109. [[CrossRef](#)]

131. Von Moeller, H.; Lerner, R.; Ricciardi, A.; Basquin, C.; Marzluff, W.F.; Conti, E. Structural and biochemical studies of SLIP1-SLBP identify DBP5 and eIF3g as SLIP1-binding proteins. *Nucleic Acids Res.* **2013**, *41*, 7960–7971. [[CrossRef](#)]
132. Neusiedler, J.; Mocquet, V.; Limousin, T.; Ohlmann, T.; Morris, C.; Jalinot, P. INT6 interacts with MIF4GD/SLIP1 and is necessary for efficient histone mRNA translation. *RNA* **2012**, *18*, 1163–1177. [[CrossRef](#)]
133. Wu, H.Y.; Ke, T.Y.; Liao, W.Y.; Chang, N.Y. Regulation of coronavirus poly(A) tail length during infection. *PLoS ONE* **2013**, *8*, e70548. [[CrossRef](#)] [[PubMed](#)]
134. Teixeira, F.K.; Lehmann, R. Translational Control during Developmental Transitions. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032987. [[CrossRef](#)] [[PubMed](#)]
135. Thoms, M.; Buschauer, R.; Ameisemeier, M.; Koepke, L.; Denk, T.; Hirschenberger, M.; Kratzat, H.; Hayn, M.; Mackens-Kiani, T.; Cheng, J.; et al. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* **2020**, *369*, 1249–1255. [[CrossRef](#)] [[PubMed](#)]
136. Schubert, K.; Karousis, E.D.; Jomaa, A.; Scaiola, A.; Echeverria, B.; Gurzeler, L.A.; Leibundgut, M.; Thiel, V.; Muhlemann, O.; Ban, N. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 956–966. [[CrossRef](#)]
137. Banerjee, A.K.; Blanco, M.R.; Bruce, E.A.; Honson, D.D.; Chen, L.M.; Chow, A.; Bhat, P.; Ollikainen, N.; Quinodoz, S.A.; Loney, C.; et al. SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. *Cell* **2020**, *183*, 1325–1339.e21. [[CrossRef](#)]
138. Vankadari, N.; Jeyasankar, N.N.; Lopes, W.J. Structure of the SARS-CoV-2 Nsp1/5'-Untranslated Region Complex and Implications for Potential Therapeutic Targets, a Vaccine, and Virulence. *J. Phys. Chem. Lett.* **2020**, *11*, 9659–9668. [[CrossRef](#)]
139. Ingolia, N.T.; Ghaemmaghami, S.; Newman, J.R.; Weissman, J.S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **2009**, *324*, 218–223. [[CrossRef](#)]
140. Lee, S.; Liu, B.; Lee, S.; Huang, S.X.; Shen, B.; Qian, S.B. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2424–E2432. [[CrossRef](#)]
141. Wright, B.W.; Yi, Z.; Weissman, J.S.; Chen, J. The dark proteome: Translation from noncanonical open reading frames. *Trends Cell Biol.* **2022**, *32*, 243–258. [[CrossRef](#)]
142. Wu, H.Y.; Guan, B.J.; Su, Y.P.; Fan, Y.H.; Brian, D.A. Reselection of a genomic upstream open reading frame in mouse hepatitis coronavirus 5'-untranslated-region mutants. *J. Virol.* **2014**, *88*, 846–858. [[CrossRef](#)]
143. Irigoyen, N.; Firth, A.E.; Jones, J.D.; Chung, B.Y.; Siddell, S.G.; Brierley, I. High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. *PLoS Pathog.* **2016**, *12*, e1005473. [[CrossRef](#)] [[PubMed](#)]

Résultats : partie 2

1 Caractérisation de la fonction de la tige boucle SL1

1.1 La tige boucle SL1 module l'efficacité de traduction

La région 5'UTR de l'ARN génomique du SARS-CoV-2 possède de nombreuses régions structurées qui forment 5 tiges-boucles (SL1 à SL5) (Figure 1A). Les trois premières sont communes à l'ensemble des 9 ARNm sous-génomiques produits pendant l'infection. Dans le but de déterminer l'impact de ces structures secondaires sur la traduction des ARN viraux du SARS-CoV-2, nous avons réalisé plusieurs mutants de délétion (Figure 1 A) que nous avons clonés en amont d'un gène rapporteur codant pour la Renilla (Figure 1B). Tous nos mutants possèdent un codon AUG dans le même contexte Kozak afin que ce paramètre ne rentre pas en compte dans nos observations. Nous avons généré 11 ARNm contenant des délétions dans chacune des 5 tiges-boucles qui ont été transcrits *in vitro* puis traduits dans le lysat de réticulocytes de lapin (RRL). L'efficacité de traduction est déterminée en mesurant l'activité enzymatique de la Renilla luciférase. Les résultats sont présentés dans la figure 1C et on s'aperçoit que la plupart des ARNm sont bien traduits si on compare avec l'ARNm contenant la 5'UTR de la β -globine qui est utilisée comme contrôle interne. De façon intéressante, la délétion de la séquence leader (Δ Lead) ou seulement de la tige-boucle SL1(Δ SL1) entraîne une baisse de l'efficacité de traduction par rapport à la séquence sauvage (5'UTR-SARS-CoV-2) de l'ordre de 50 % et 20 % respectivement. Les délétions individuelles des SL2 à SL5 de la région 5'UTR ne provoquent pas de diminution significative de l'expression dans notre système. Afin de s'assurer que l'effet observé ne résulte pas du changement de longueur de la 5'UTR due à la délétion de SL1, celle-ci a été remplacée par quatre différentes séquences aléatoires de longueur similaire à SL1 (25 nucléotides) dans une nouvelle série d'ARNm qui ont été traduits dans les mêmes conditions que ci-dessus (Figure 1D). On observe, par rapport au contrôle 5'UTR SARS-CoV-2, une diminution comparable à celle observée précédemment. Notons toutefois des différences entre les séquences aléatoires utilisées ce qui reflète probablement des

niveaux de structure différents. Nous avons ensuite électroporé des cellules HEK293T avec les ARNm contrôles et mutants, après 3 heures les cellules ont été récoltées, lysées et l'activité de la luciférase a été évaluée. De façon parallèle, afin de normaliser les résultats le niveau d'expression de chaque ARN rapporteur a été déterminé par RT-qPCR; les résultats sont présentés dans les Figure 1E et F et confirment les données obtenues en lysat de réticulocytes qui montrent que la délétion de la tige-boucle SL1 entraîne une baisse d'expression et que le remplacement de cette séquence par un segment de même longueur ne rétablit pas l'activité traductionnelle.

Nous avons ensuite regardé si des régions spécifiques de la tige-boucle SL1 étaient impliquées dans l'effet observé. Pour ce faire, deux mutants tronqués de SL1 ont été générés : le premier possède la partie basse de la tige-boucle (mutant shortening SL1 #1) et le second, sa partie haute (mutant shortening SL1 #2). Après transcription et purification, ces 2 ARNm ont été électroporés dans des cellules HEK293T. Les résultats (Figure 1G) montrent que la délétion partielle de la tige-boucle SL1, qu'elle concerne la partie haute ou la partie basse, impacte l'efficacité de traduction de ces ARNm ce qui confirme l'importance de celle-ci dans ce mécanisme.

1.2 Plusieurs éléments de la tige-boucle SL1 sont impliqués dans l'efficacité traductionnelle médiée par la 5'UTR du SARS-CoV-2

Nous avons ensuite voulu identifier les sous-domaines de cette structure qui étaient impliqués dans sa fonction. SL1 est composée de deux tiges de 5 paires de bases, d'un bulge formé par 3 nt et d'une boucle de 4 nt (Figure 2 A). Nous avons réalisé au total plus de 100 mutants différents, mais nous ne décrivons que les mutants qui ont montré des différences significatives. Nous avons remplacé la boucle par 4 nt complémentaires (mutant mLoop #1), muté la région des bulges, soit en les supprimant (mutant mBulge #1), soit en créant un appariement complémentaire des nucléotides impliqués (mutant mBulges #2) (Figure 2A). Toutes ces modifications ont été créées, dans le contexte de la 5'UTR de l'ARNg du SARS-CoV-2 qui dirige la synthèse de la protéine Renilla. Après transcription, ces 3 ARNm ont été traduits dans le RRL et comparés au contrôle 5'UTR sauvage de l'ARNg du SARS-CoV-2 (Figure

2A). Nous pouvons remarquer que les trois différentes mutations de la boucle et du bulge de la structure SL1 diminuent l'efficacité de traduction de la Renilla indiquant que ces deux éléments de la tige-boucle SL1 sont importants.

Nous avons poursuivi nos investigations en remplaçant la partie supérieure de la tige-boucle SL1 par sa séquence complémentaire, créant ainsi le mutant dénommé SL1-FlipA et remplacé la totalité de la tige-boucle SL1 par sa séquence complémentaire, pour générer le mutant SL1-FlipA-B (Figure 2B). Afin de mieux comprendre l'éventuelle action synergique des différentes parties de la tige-boucle SL1 et en cohérence avec nos observations précédentes, nous avons aussi construit un mutant combinant l'inversion de la tige et la modification de la boucle déjà caractérisée à la figure 2A (SL1_FlipA-B_mLoop). Ces trois ARNm ont été synthétisés *in vitro*, traduits en RRL et les résultats sont présentés dans la figure 2B. On observe que l'inversion de la partie haute de la tige-boucle (mutant SL1-FlipA) est suffisante pour affecter la traduction de ce transcrit. De manière intéressante, nous pouvons observer que les mutations simultanées de la tige et de la boucle se potentialisent, atteignant ainsi une diminution de l'efficacité de traduction de l'ordre de 40% par rapport à l'ARNm non muté. Cela confirme une nouvelle fois le rôle essentiel de la tige-boucle SL1 et en particulier sa partie apicale.

La position 15 de la SL1 a été précédemment identifiée comme une région importante dans l'échappement des ARN viraux à l'effet inhibiteur de la protéine Nsp1 (Bujanic et al. 2022). Nous avons introduit une mutation de sauvetage (Figure 2C) qui restaure le G15 en C15 sur le mutant SL1_FlipA-B_mLoop (SL1_FlipA-B_mLoop_C15). La traduction de ces constructions en RRL montre une augmentation de l'expression avec la mutation de sauvetage G15C à la fois en lysat de réticulocytes et en cellules HEK293T (Figure 2C).

Nous avons ensuite substitué la séquence de la boucle inversée AGGG par la boucle sauvage UCCC ce qui crée un nouveau mutant (SL1-FlipA-B_C15). La correction de la position 15 combinée au remplacement de la boucle AGGG par la boucle sauvage UCCC permet de restaurer intégralement le niveau de traduction au niveau du contrôle positif (Figure 2D).

Cela confirme que la position C15 dans la tige-boucle SL1, ainsi que la composition en nucléotides de la boucle, sont importantes pour maintenir l'expression observée avec la 5'UTR de l'ARN génomique du SARS-CoV-2.

Nous avons ensuite étudié le rôle des séquences ARN qui jouxtent la tige-boucle SL1 en mutant les 6 nucléotides situés immédiatement en aval (mutant mSeqDownstream_SL1) et en amont (mutant mSeqUpstream_SL1) et en veillant à ne pas introduire de codons d'initiation (Figure 3A). Ces ARN ont été transcrits *in vitro*, traduits en RRL (Figure 3B), et en cellules HEK293T (Figure 3C). Les résultats montrent que les mutations, tant en amont qu'en aval, entraînent une baisse de l'expression *in vitro* et *ex-vivo* avec un effet plus prononcé dans les cellules. Il est important de noter que pour écarter l'impact potentiel de la stabilité des ARNm dans les cellules électroporées, les transcrits ont été quantifiés par RT-qPCR et les résultats sont donc normalisés par rapport aux quantités d'ARNm mesurées.

Nous pouvons donc conclure de ces expériences que l'intégralité de la séquence de SL1, ainsi que le contexte qui jouxte la tige boucle en amont et en aval sont requis pour assurer une traduction efficace.

1.3 Effet de la protéine Nsp1 sur la traduction des ARNm du SARS-CoV-2: effet de la tige boucle SL1

Lors de l'infection, l'expression de la protéine Nsp1 réprime fortement la synthèse des protéines cellulaires par plusieurs mécanismes que nous avons décrits au chapitre III de l'introduction. L'un d'eux est la répression traductionnelle, répression qui ne semble pas, ou peu, affecter la synthèse des protéines virales. Cependant le mécanisme moléculaire qui permet l'échappement des ARNm viraux à l'inhibition n'est pas élucidé bien qu'il ait été montré que la structure et la composition nucléotidique de la 5'UTR du SARS-CoV-2 soit importante, notamment au niveau de la tige-boucle SL1 (Bujanic et al. 2022).

De fait, nous avons exploité le potentiel du système de traduction *in vitro* pour mesurer les effets de Nsp1 sur la traduction virale.

Pour ce faire, nous avons utilisé les constructions décrites dans le chapitre précédent (Résultats 1, Condé et Al) qui codent pour la luciférase sous le contrôle des séquences des 5'UTR des 8 ARN sous-génomiques du SARS-CoV-2 (N, ORF7a-8, ORF 7b, ORF6, M, E, ORF 3a, S) et de l'ARN génomique (5'UTR SARS-CoV-2). Le gène codant pour la protéine Nsp1 WT et le mutant KH ont été clonés dans un vecteur d'expression procaryote (pET28a) leur conférant une étiquette histidine. Le mutant KH décrit dans la littérature est fonctionnellement inactif (Bujanic et al. 2022). Des bactéries BL21 ont été transformées avec le plasmide pET28-Nsp1 WT et mutant puis induites avec l'IPTG pendant 3 heures avant lyse et purification sur colonne de cobalt. Les protéines sauvages et mutées ont été resuspendues dans un tampon (25 mM Tris-Cl pH 7.5, 50 mM NaCl, 0.1mM MgCl₂, 25% glycérol) et leur concentration déterminée par mesure de l'absorbance (λ 260) au Nanodrop ; leur intégrité et leur taille moléculaire a été vérifiée sur gel SDS-PAGE.

Une quantité équivalente (0,1 μ g) de la protéine sauvage ou mutée a été ajoutée au lysat de réticulocytes en présence de chacun des ARNm du SARS-CoV-2. Après 40 minutes d'incubation, l'expression des transcrits viraux a été déterminée en mesurant l'activité luciférase. Pour évaluer l'impact de la protéine Nsp1 sur l'expression des différents rapporteurs, nous avons calculé le rapport d'expression des ARNm en présence de la protéine Nsp1 sauvage comparé à leur expression en présence de la version mutée inactive de Nsp1 (mutant KH) en fixant arbitrairement à 1 le rapport obtenu avec la 5'UTR de l'ARNg.

Les résultats de la figure 4A montrent que la traduction de tous les ARNm subgénomiques, à l'exception de M et 7b, est inhibée par l'ajout de la protéine Nsp1 avec une intensité variable. Les ARN subgénomiques dont la traduction est dépendante des 5'UTR de l'ORF E et de l'ORF 3a sont les plus affectés avec une diminution significative de leur activité de l'ordre de 30 % à 40% par rapport au contrôle (5'UTR SARS-CoV-2). Néanmoins, ils restent significativement au-dessus de l'activité de l'ARNm de la globine qui est, comme prévu, très fortement impactée par l'ajout de Nsp1. On retrouve une situation intermédiaire avec les constructions contenant la 5'UTR des ORF S, ORF8 et N.

Nos résultats sont en accord avec l'hypothèse que la tige boucle SL1, qui est présente sur chacun des transcrits viraux, permet de résister à l'effet inhibiteur de Nsp1 qui est

beaucoup moins prononcé que sur l'ARNm globine. En revanche, il existe une grande variabilité d'effet parmi les transcrits viraux ce qui indique que la présence de SL1 n'est pas le seul élément déterminant.

Nous avons ensuite voulu savoir si les autres tiges-boucles, notamment SL2, SL3, SL4 et SL5 pouvaient, elles aussi, contrer l'effet inhibiteur de Nsp1. Pour ce faire, des délétions de chacune des 5 SL ont été effectuées au sein de la 5'UTR ainsi que les délétions des sous-domaines de SL4 (SL4A et SL4B) et SL5 (5A, 5B et 5C) qui sont les segments les plus longs et les plus structurés. Chacun de ces 10 ARNm mutants a été traduit dans le lysat de réticulocytes en présence de la protéine recombinante sauvage ou mutée et les résultats sont normalisés par le rapport de l'activité obtenue comme décrit ci-dessus. On observe sur la Figure 4B que seules les délétions de SL1 (Δ Leader et Δ SL1) ont un effet important (50 % d'inhibition) sur l'activité traductionnelle. On note également que cette diminution reste moins importante que celle observée avec l'ARNm globine comme c'était le cas précédemment. Les délétions de tout, ou partie, des SL 2, 3, 4 et 5 n'ont pas, ou peu, de conséquences significatives sur l'activité traductionnelle.

Nos résultats, présentés dans les figures 1 à 4, confirment que tige-boucle SL1 est essentielle pour que les ARN du SARS-CoV-2 contournent l'inhibition traductionnelle induite par la protéine Nsp1. Ils révèlent également la contribution de cette structure à l'efficacité de traduction de la région 5'UTR de l'ARN génomique du SARS-CoV-2, et ce, même en l'absence de Nsp1. De plus, nos observations suggèrent que la tige-boucle SL1 n'est pas le seul élément qui contribue à l'échappement des ARN viraux face à l'inhibition de la traduction induite par Nsp1.

2 Des uORF régulent l'expression des ARN viraux

Une des particularités du SARS-CoV-2 réside dans le fait que pour chacun de ses ARN, aussi bien génomiques que sous-génomiques, la séquence leader est ajoutée à une distance variable du codon AUG initiateur ce qui rend unique chacun des transcrits viraux. Le leader commun de 75 nucléotides contient un codon CUG à la position 59 et des études récentes de ribosome profiling *in vivo* ont montré que ce codon pourrait servir de site d'initiation pour les ribosomes (Kim et al. 2021). Selon l'ARN

subgénomique considéré, ce codon non canonique peut être en phase ou hors phase avec le codon d'initiation principal créant ainsi potentiellement une uORF différente dans chaque configuration. Celle-ci pourrait être totalement comprise dans la région 5'UTR (ORF1ab et ORF6), chevauchante (ORF7a/8) ou encore en phase (ORF S).

2.1 Impact sur la traduction de l'ARN génomique (ORF 1a/b)

Nous avons cloné la région 5'UTR de l'ORF 1a/b dans laquelle le codon CUG-59 est en phase avec l'uORF débutant à la position 107. Dans ce contexte, le codon CUG-59 peut démarrer une courte phase de lecture qui s'achève avec le codon (stop) à la position 134 (Figure 5A). Nous avons construit différents mutants avec une sélection de mutations qui changent le codon CUG-59 en codon AUG, d'autres qui modifient son contexte nucléotidique ou qui suppriment les codons stop en aval pour mettre en phase les codons CUG-59 et AUG-107 avec le cadre de lecture de la Renilla. Afin de visualiser tous les produits de traduction en fonction des sites d'initiation utilisés, nous avons réalisé ces expériences en présence de méthionine radiomarquée au S³⁵ puis séparé les produits de traduction sur gel de polyacrylamide révélé par autoradiographie.

L'expression de l'ARN contenant la 5'UTR WT du SARS-CoV-2 montre une seule bande correspondant à la synthèse de Renilla (Figure 5B piste 1). Lorsque les 3 codons stop aux positions 134, 137, 252 sont supprimés (Figure 5 B, piste 2), on voit apparaître une protéine supplémentaire qui résulte de l'initiation au codon AUG-107. Ceci est confirmé par la mutation de ce triplet en codon CCG (piste 8). L'initiation sur le codon CUG-59 n'est pas détectable bien qu'il soit aussi en phase avec la Renilla (pistes 2 et 5). En revanche, lorsque celui-ci est remplacé par un codon AUG, il est alors reconnu par le complexe ribosomal (pistes 7 et 8). Cependant, de manière surprenante, la mutation du codon CUG-59 en CCG ou la modification de son contexte nucléotidique produit un modeste, mais reproductible, impact sur la traduction de la Renilla (comparer les pistes 3 et 4 avec la piste 1). Ces résultats suggèrent que le codon CUG à la position 59, bien qu'il ne soit pas en mesure de servir de codon d'initiation, semble avoir un effet régulateur indirect sur l'efficacité de traduction de l'ORF principale (AUG-266). De manière intéressante, nous constatons que la mise en phase induit systématiquement une baisse d'expression de l'ORF principale (pistes

2, 5, 7 et 8) alors que la simple mutation du codon AUG-107 en triplet CCG stimulait l'expression de la Renilla (Figure 7B de Condé et al 2022). Ces résultats confirment le rôle régulateur de cet uORF dans l'expression de l'ORF principale.

Enfin, les évènements d'initiation sur les codons AUG-59 et AUG-107 (pistes 2, 5, 7, 8) confirment aussi que la 5'UTR de l'ARN génomique du SARS-CoV-2 est balayée de manière linéaire par les ribosomes qui peuvent reconnaître un site d'initiation fort en amont. C'est en accord avec nos résultats précédents montrant un mécanisme coiffe-dépendant et cela s'oppose aux résultats de nos collègues qui suggèrent la présence d'une séquence IRES au sein de l'ARNg du SARS-CoV-2 (Slobodin et al. 2022).

2.2 Impact sur l'ORF 7a et 8

Nous avons ensuite examiné le rôle du codon CUG-59 dans le contexte des régions 5'UTR des ARN sous-génomiques des ORF 7a et 8 qui sont similaires et se résument à la séquence leader. Dans cette configuration, le codon CUG à la position 59 ne se trouve pas en phase avec le codon AUG principal à la position 76. Dans un premier temps, nous avons muté le triplet CUG-59 en codon CCG. Cet ARNm exprime la Renilla à un niveau proche de l'ARN WT (Figure 6b piste 1 vs 5). Nous avons ensuite muté le triplet CUG-59 en codon AUG afin d'être sûr qu'il peut être reconnu par le complexe traductionnel. Comme nous pouvons le voir à la Figure 6B piste 2, cette mutation inhibe fortement la synthèse de la protéine principale initiée au codon AUG-76 ce qui suggère que le codon AUG-59 est pris comme site d'initiation. Pour confirmer cela, nous avons construit un nouvel ARN rapporteur dans lequel le codon AUG-59 est dans le même cadre de lecture que le codon AUG-76. Pour ce faire, nous avons introduit un décalage de phase en ajoutant 2 nt (+2), ou en retirant 1 nt (-1) en amont du codon AUG-76. Les pistes 3 et 4 de la figure 6B montrent que la synthèse de la protéine s'effectue majoritairement à partir du codon localisé à la position 59.

Nous avons ensuite inséré une séquence de 30 nt entre la position 59 et le codon d'initiation 76 pour mieux détecter les produits de synthèse. Cela permet de visualiser de manière très claire que la mutation CUG en AUG à la position 59 entraîne une très forte baisse de l'initiation à l'AUG-76 (Figure 6C, piste 2) et que la mise en phase de

ce codon permet de produire une protéine avec une extension N-terminale qui correspond à l'initiation des ribosomes au niveau du codon AUG-59 (Figure 6C, piste 3 et 4).

Enfin, nous avons changé le contexte du codon CUG-59 pour un bon contexte nucléotidique (acc-CUG-guc) (figure 6D, piste 2) ou pour un mauvais contexte nucléotidique (uuu-CUG-cuc) (figure 6D, piste 3). En comparant la traduction de ces deux rapporteurs au contrôle (figure 6D, piste 1), nous constatons qu'un bon contexte Kozak entraîne une légère baisse de l'initiation à l'AUG-76.

En conclusion, l'ensemble des expériences de la figure 6 nous permet de conclure que le codon CUG-59 semble être peu utilisé pour initier la traduction dans son contexte normal.

2.3 Impact sur l'ORF 6

Nous avons également étudié le rôle du codon CUG-59 dans le contexte de la région 5'UTR de l'ORF 6. Cette région qui est relativement complexe, représente, après celle de l'ARN génomique, la plus longue des 5'UTR du SARS-CoV-2. Elle contient la fin de la région codante de l'ORF M qui la précède dans le génome viral (Figure 7A) et le codon CUG-59 crée potentiellement une uORF longue de 162 nt.

Lorsque nous avons mis en phase le codon CUG-59 avec la Renilla, nous n'avons pas observé d'extension N-terminale de la protéine ce qui indique que le codon CUG n'est pas reconnu par le ribosome (Figure 7B, ligne 2). De même, la mutation du triplet CUG en CCG (lignes 3 et 4) ou les mutations qui changent son contexte nucléotidique en un bon, ou mauvais, contexte Kozak (lignes 7 et 8) ne changent pas vraiment le profil de traduction ce qui confirme que le codon CUG n'est pas reconnu par la machinerie traductionnelle dans le contexte de l'ORF 6.

En revanche, la substitution du triplet CUG-59 en un codon AUG à la même position inhibe de manière drastique la traduction de la Renilla sur l'AUG-231 (Renilla) (ligne 5 et 6) confirmant que le ribosome balaye la 5'UTR.

2.4 Impact sur l'ORF S

Enfin, nous avons évalué l'impact du codon CUG-59 sur l'initiation de la traduction dans le contexte de l'ARN sous-génomique codant pour la protéine Spike. Dans la

5'UTR de cet ARN, le codon CUG-59 est dans le même cadre de lecture que l'AUG principal (AUG-77) (figure 8A).

La traduction dans le lysat de réticulocytes d'un ARNm où le gène rapporteur est sous le contrôle de la 5'UTR de Spike semble indiquer la présence d'une bande de faible intensité juste au-dessus du produit principal (Figure 8B, piste 1) qui semble disparaître lorsque le contexte du CUG-59 est muté (piste 3) ou remplacé par un codon CAG (piste 2). Conformément avec les 5'UTR des autres ORFs (ORF 1a/b, ORF 6, ORF7a-8, FIG 5, 6 et 7), la substitution du codon CUG-59 par un codon AUG (piste 5) induit un léger décalage vers le haut de la Renilla confirmant un changement du site d'initiation au profit de ce dernier. De manière intéressante, la délétion de l'AUG principal (piste 4) révèle clairement la présence d'une bande qui résulte d'une initiation au triplet CUG-59 et qui s'accompagne de produits de traduction plus courts correspondant à des événements d'initiation en aval de la position normale du codon AUG-77.

Afin de mieux visualiser les protéines produites à partir des codons CUG-59 et AUG-77, nous avons inséré une séquence aléatoire de 30 nucléotides entre ces deux positions et nous avons introduit différentes mutations qui améliorent (note Good sur la Figure 8C) le contexte nucléotidique des codons CUG-59 ou AUG-77 ou qui dégradent ce contexte (note Bad sur la Figure 8C). On note tout d'abord (Figure 8C piste 2) que l'ajout de l'extension de 30 nucléotides permet de visualiser un produit de traduction de taille plus importante ce qui confirme que le codon CUG-59 est reconnu par le ribosome. De plus, les mutations qui améliorent le contexte du CUG-59 s'accompagnent d'une meilleure traduction sur ce dernier (pistes 3 à 5); au contraire, lorsque le codon CUG-59 se trouve dans un contexte Kozak défavorable (piste 6, 7 et 8), on perd complètement l'initiation sur ce site.

Ces résultats indiquent que le codon CUG-59 de la 5'UTR de l'ARNsg codant pour la protéine spike est reconnu par la machinerie traductionnelle comme un site d'initiation ce qui implique que, dans le contexte de l'infection virale, la protéine Spike possède très probablement une isoforme qui comprend cette extension N-terminale et qui pourrait avoir un rôle dans le cycle viral.

L'ensemble des résultats présentés aux figures 5, 6, 7 et 8 montre que le codon CUG-59 peut être utilisé comme codon initiateur lorsqu'il se trouve dans le contexte de l'ORF

codant pour Spike. Pour tous les ARNm testés, la position 59 est bien reconnue comme un site initiateur fort si le codon CUG est muté en codon AUG ce qui indique clairement que les ribosomes balayent cette région et peuvent reconnaître un site initiateur potentiel.

FIGURE 1

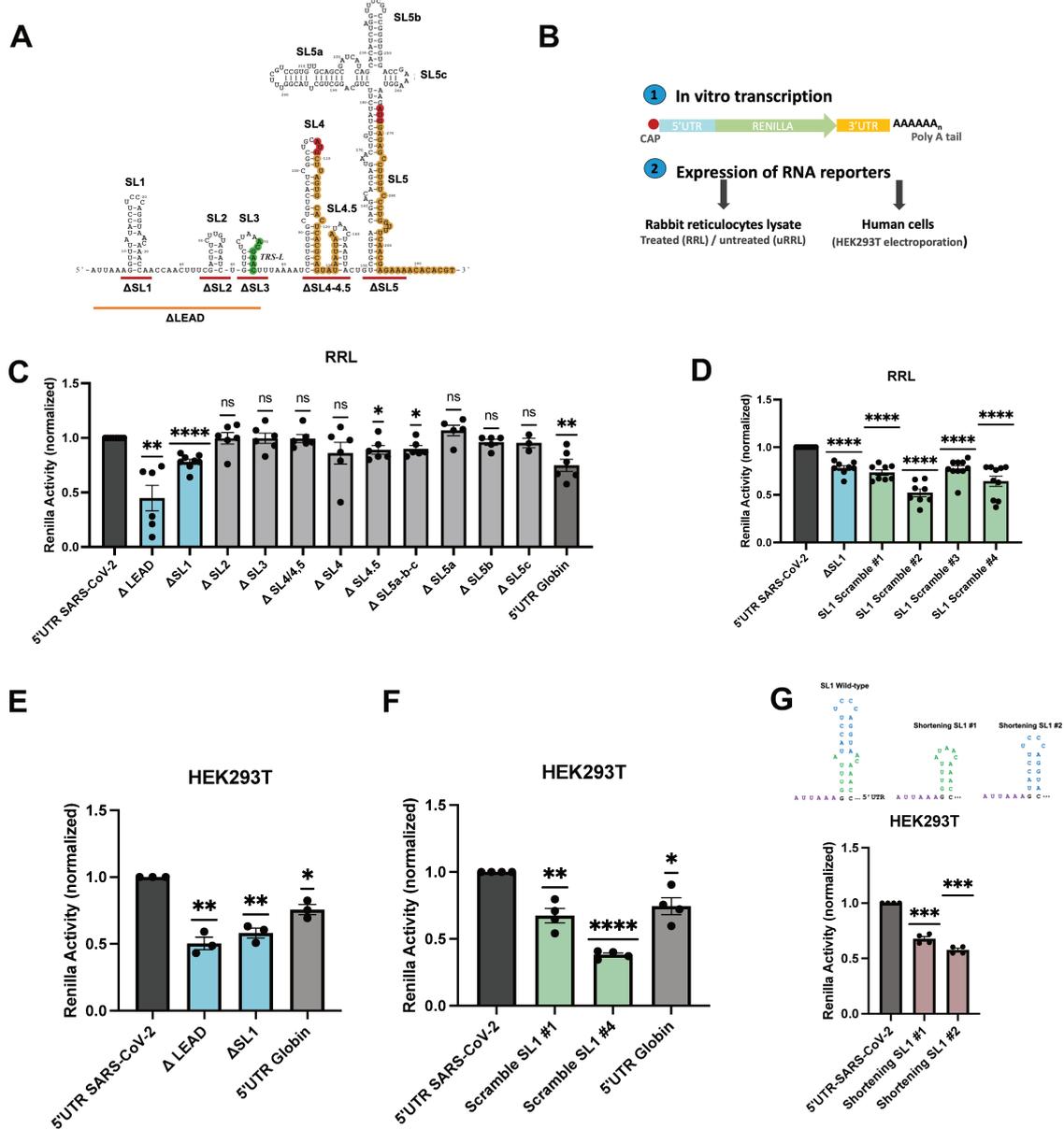


Figure 1: La tige-boucle SL1 est importante pour l'efficacité de traduction

(A) Schéma du protocole expérimental d'expression des ARN rapporteurs codant pour la Renilla luciférase en lysat de réticulocytes (RRL) et en cellules HEK293T. (B) Modélisation de la région 5'UTR du SARS-CoV-2 adaptée de Miao et al. 2021. Le schéma représente les différentes structures tige-boucle (SL1 à SL5) de la région 5'UTR du SARS-CoV-2. Les codons AUG de l'uORF et de l'ORF principale sont indiqués en rouge. Le TRS-L est indiqué en vert. Les mutations de délétion utilisées dans cette étude sont indiquées sur le schéma. (C) Les ARN mutants schématisés en (B) ont été traduits en lysat de réticulocytes pendant 30 minutes. Le niveau d'expression de chacun d'entre eux a ensuite été mesuré par un essai luciférase. (D) Des ARN rapporteurs, pour lesquels la tige-boucle SL1 a été remplacée par des séquences aléatoires, ont été traduits en lysat de réticulocytes, et leur niveau d'expression a été mesuré en suivant l'activité de la luciférase. (E, F) Les ARN mutants utilisés respectivement en (B et D), ont été électroporés dans des cellules HEK293T. Après 3 heures d'incubation, les cellules ont été lysées et l'efficacité de traduction des différents rapporteurs a été quantifiée par mesure de l'activité de la luciférase. (G) Des ARN rapporteurs pour lesquels la tige-boucle SL1 a été partiellement délétée ont été électroporés dans des cellules HEK293T. Leur efficacité de traduction a été mesurée après 3h d'incubation. L'ensemble des ARN utilisés possèdent une coiffe en 5' et une queue poly(A) à leur extrémité 3'. Tous les résultats sont normalisés par rapport à l'efficacité de traduction médiée par la région 5'UTR sauvage du SARS-CoV-2. Les valeurs représentent la moyenne (+/-SD) d'au minimum trois expériences indépendantes. Test statistique : t-test bilatéral.

FIGURE 2

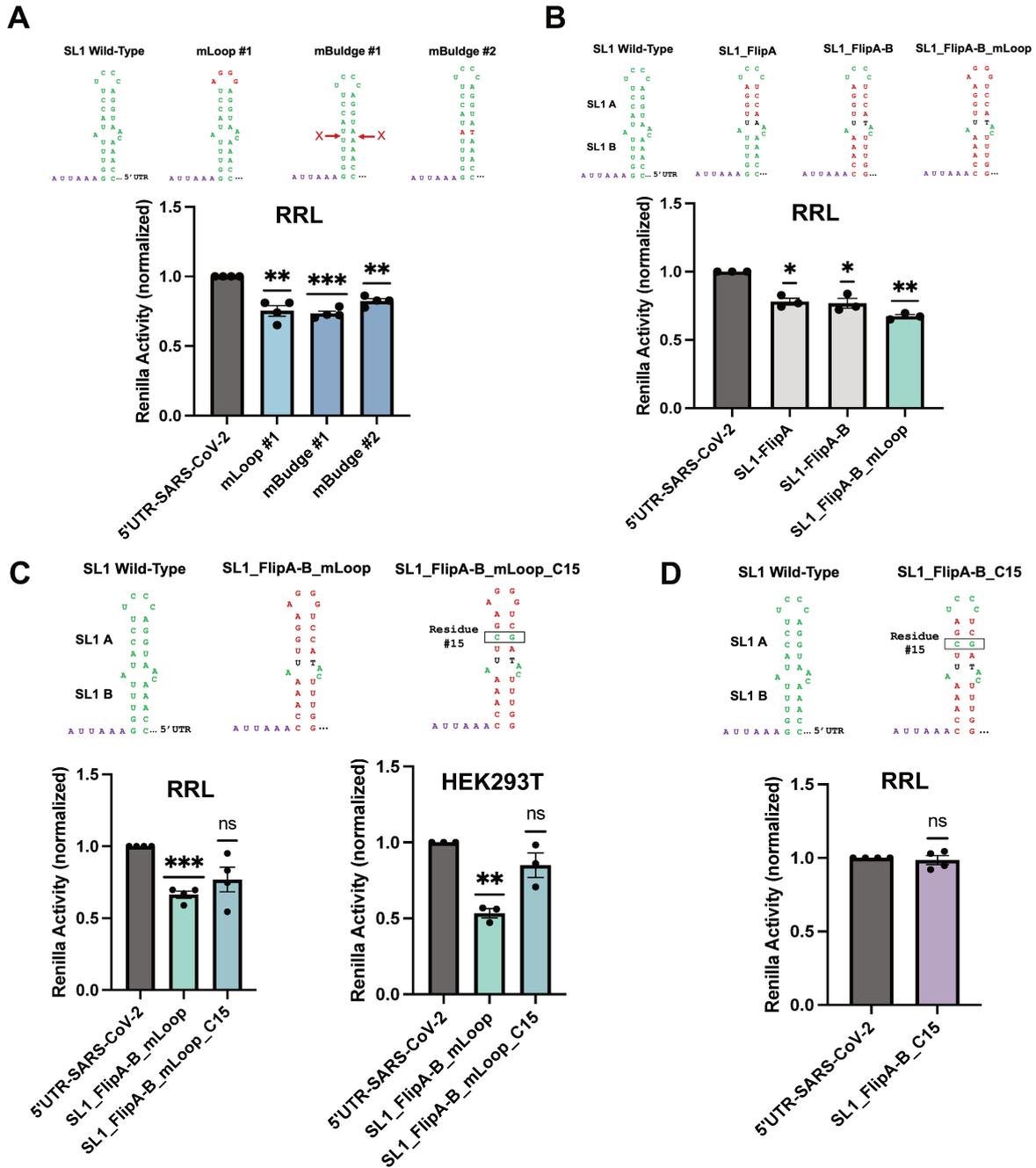


Figure 2: La partie apicale de la tige boucle SL1 est nécessaire pour maintenir une haute efficacité de traduction

(A, B) Les ARN rapporteurs représentés avec les mutations spécifiées au sein de la tige-boucle SL1 ont été traduits en lysat de réticulocytes (RRL). L'activité luciférase a été mesurée après 30 minutes d'incubation. (C) Les ARN rapporteurs représentés avec les mutations indiquées de la tige-boucle SL1 ont été traduits en lysat de réticulocytes ou électroporés dans des cellules HEK293T. L'efficacité de la traduction a été évaluée par un test de luciférase après, respectivement, 30 minutes et 90 minutes d'incubation. (D) Les ARN rapporteurs représentés avec les mutations dans la tige-boucle SL1 ont été traduits en lysat de réticulocytes, et l'activité luciférase a été mesurée après 30 minutes de traduction. Tous les ARN rapporteurs codent pour la Renilla luciférase et possèdent une coiffe en 5' et d'une queue poly(A) en 3'. Les résultats ont été normalisés par rapport à l'efficacité de traduction obtenue avec un ARN rapporteur sous le contrôle de la région 5'UTR sauvage du SARS-CoV-2. Lorsqu'une mutation compensatrice a été introduite, celle-ci est indiquée en noir. Les valeurs représentent la moyenne (+/- SD) d'au minimum trois expériences indépendantes. Test statistique : t-test bilatéral.

FIGURE 3

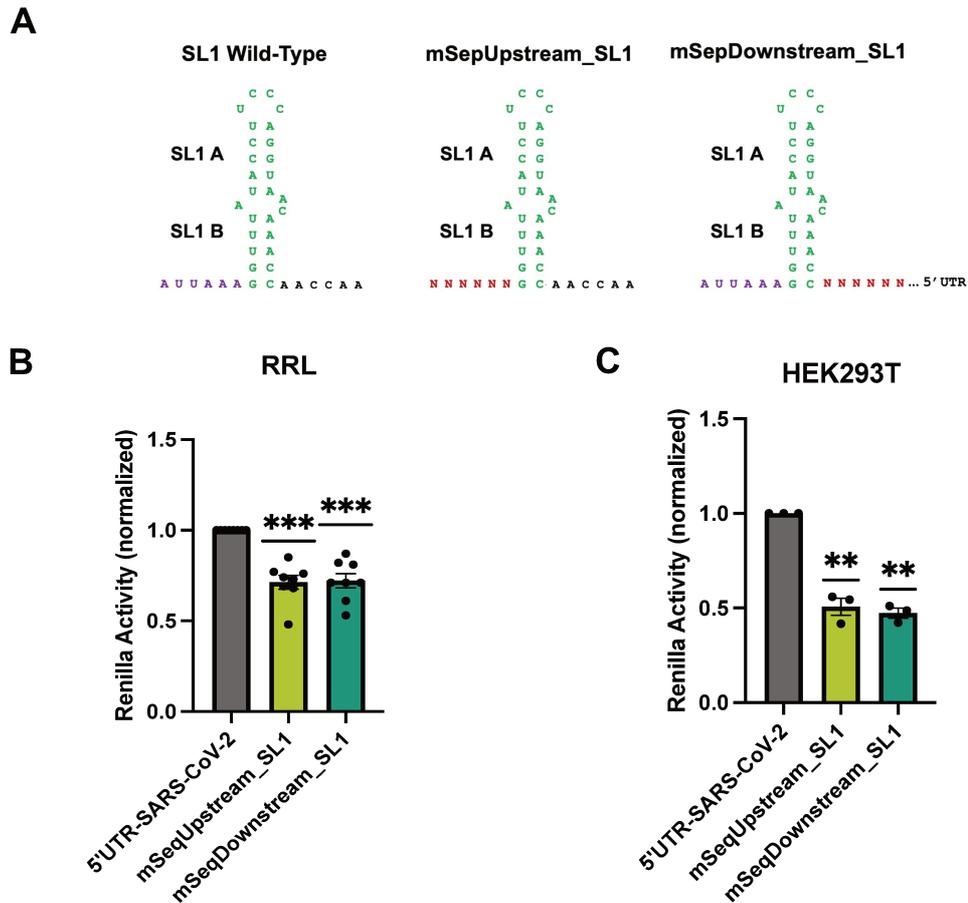


Figure 3: Les régions en amont et en aval de la tige boucle SL1 contribuent à l'efficacité de traduction

(A) Schéma des modifications apportées aux séquences adjacentes de la tige-boucle SL1 du SARS-CoV-2. Les ARN rapporteurs mutants, illustrés en (A), ont été traduits soit en lysat de réticulocytes (B), soit dans des cellules HEK293T (C). L'activité de la luciférase a été évaluée après des périodes d'incubation de 30 minutes pour le lysat de réticulocytes et de 90 minutes pour les cellules HEK293T. Chaque résultat a été normalisé par rapport au niveau d'expression de l'ARN rapporteur exprimé par la 5'UTR sauvage du SARS-CoV-2. Dans le cas des cellules HEK293T les résultats ont été normalisés sur le niveau d'expression de chaque ARN électroporés évalué par RT-qPCR. Les valeurs représentent la moyenne (+/-SD) d'au minimum trois expériences indépendantes. Test statistique : t-test bilatéral.

FIGURE 4

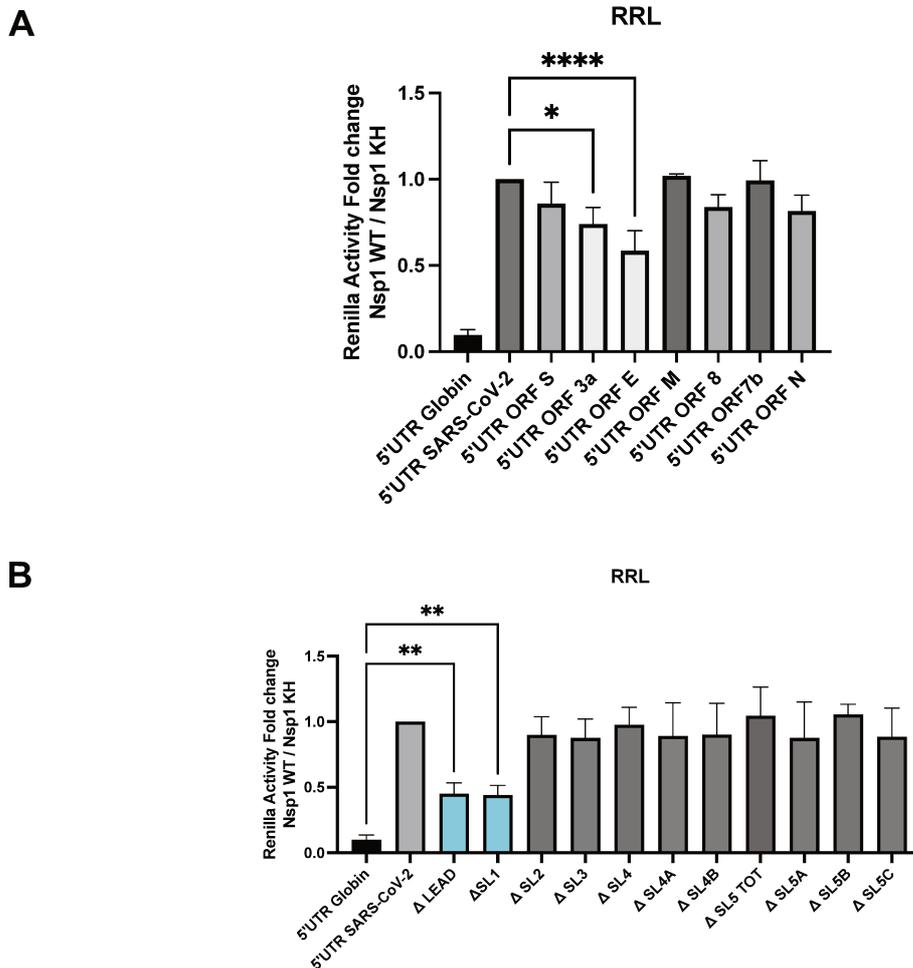


Figure 4: La tige-boucle SL1 permet de contrer partiellement l'inhibition médiée par Nsp1

(A) Des ARN rapporteurs contenant les régions 5'UTR des différents ARN du SARS-CoV-2 ont été traduits en lysat de réticulocyte en présence de la protéine Nsp1 recombinante sauvage ou mutante (KH). L'efficacité de traduction de chaque ARN rapporteur a été évaluée après 30 minutes d'incubation par la mesure de l'activité luciférase. Le graphique représente le rapport de l'expression des différents rapporteurs en présence du Nsp1 sauvage par rapport à l'efficacité de traduction en présence de la protéine Nsp1 Mutante (KH) (B) Les ARN rapporteurs contenant des délétions des différentes tige-boucle de la région 5'UTR du SARS-CoV-2 ont été traduits en lysat de réticulocyte en présence de la protéine Nsp1 sauvage ou mutante (KH). L'efficacité de traduction de chacun des ARN rapporteur a été évaluée après 30 minutes d'incubation par mesure de l'activité luciférase. Le graphique représente le rapport d'expression des différents rapporteurs en présence du Nsp1 sauvage par rapport à l'efficacité de traduction en présence de Nsp1 Mutante (KH). Tous les ARN utilisés sont coiffés et polyadénylés. Les valeurs représentent la moyenne (+/-SD) d'au minimum trois expériences indépendantes. Test statistique : t-test bilatéral

FIGURE 5

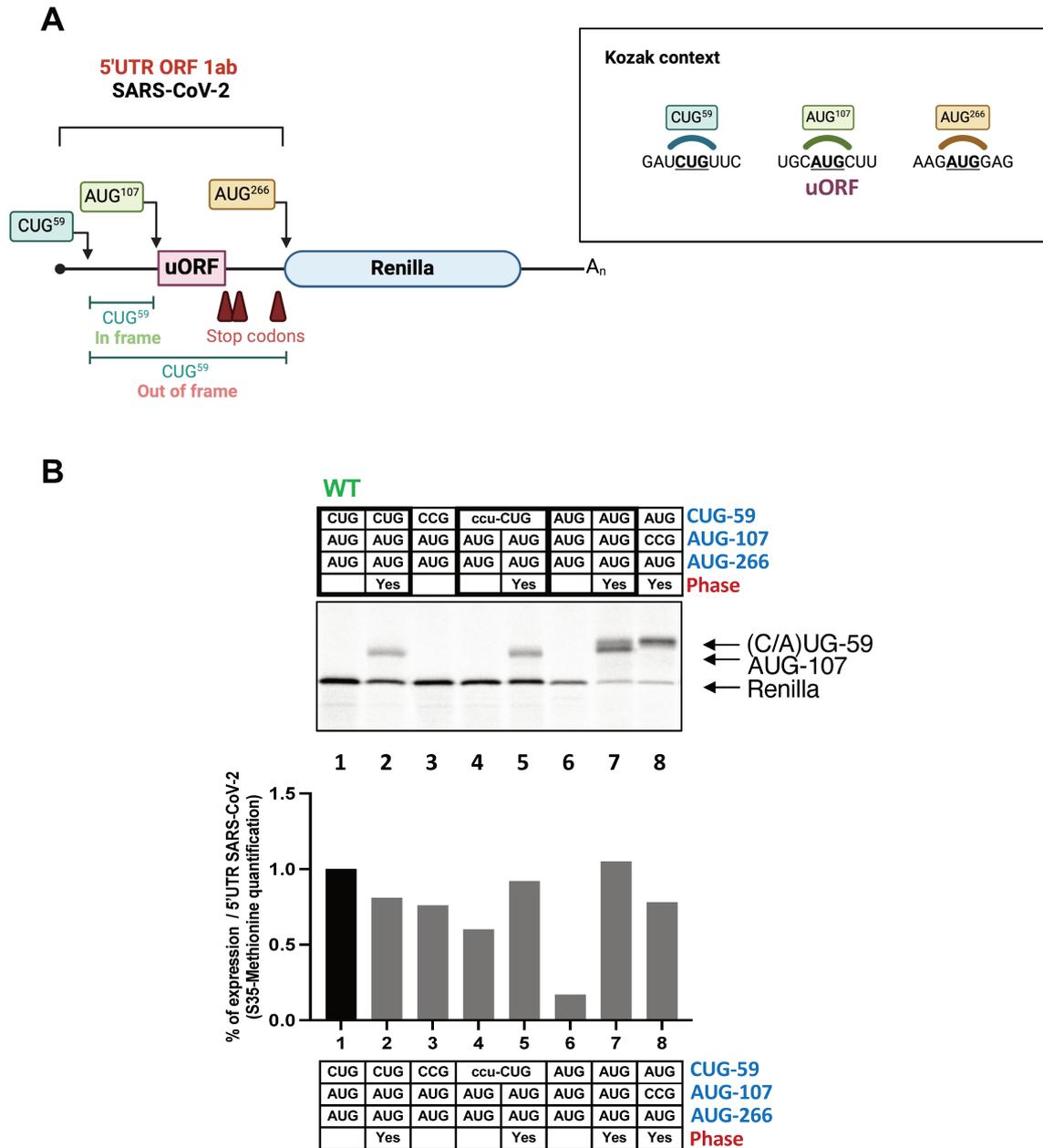


Figure 5: CUG 59 dans le contexte de l'ORF 1ab

(A) Représentation schématique de la région 5'UTR de l'ARN génomique du SARS-CoV-2. La position des différents codons d'initiation AUG et du codon CUG-59 sont indiqués. L'uORF contenue dans la région 5'UTR du SARS-CoV-2 est représentée. (B) Des ARN rapporteurs coiffés et polyadénylés contenant les mutations indiquées au niveau des positions CUG-59, AUG-107 et AUG-265 de la région 5'UTR de l'ARN génomique (ORF1ab) du SARS-CoV-2 ont été traduits en lysat de réticulocytes (RRL). La traduction a été effectuée en présence de méthionine S35 radiomarquée. Les protéines produites ont été résolues sur un gel de polyacrylamide à 12% et analysées à l'aide d'un *phosphoimager*. L'évaluation du niveau d'expression de chaque construit a été réalisée à l'aide du logiciel ImageQuant, en prenant en compte toutes les formes protéiques détectées. Les résultats sont représentatifs de trois expériences indépendantes.

FIGURE 6

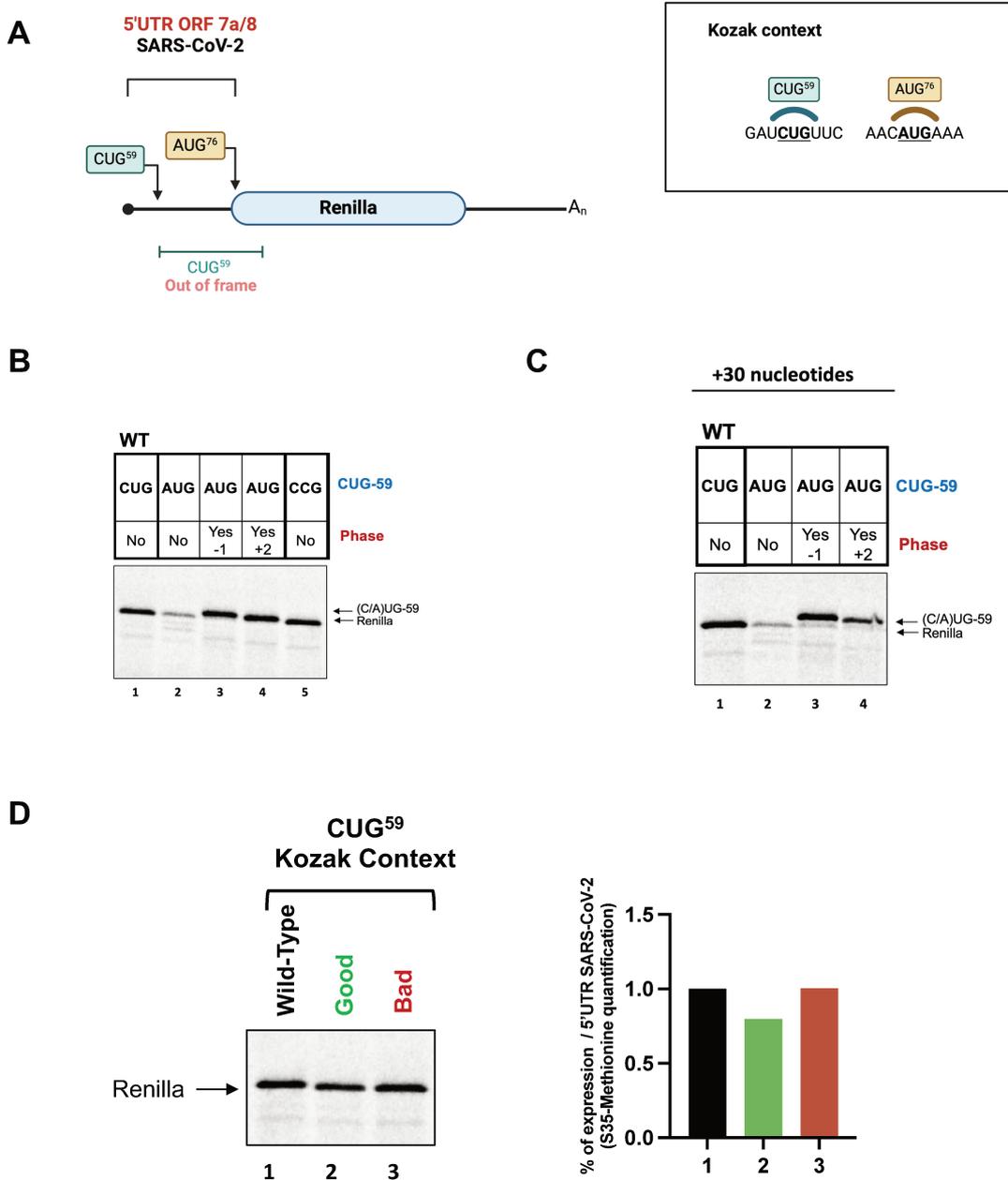


Figure 6: CUG 59 dans le contexte de ORF 7a/8

(A) Représentation schématique de la région 5'UTR de l'ARN sous-génomique codant pour les ORF 7a/8 du SARS-CoV-2. La position du codon d'initiation AUG-76 et du codon CUG-59 ainsi que la phase de ce dernier sont indiqués. (B) Des ARN rapporteurs contenant les mutations indiquées au niveau des positions CUG-59, AUG-76 de l'ARN sous-génomique codant pour les ORF7a/8 du SARS-CoV-2 ont été traduits en lysat de réticulocytes (RRL). (C) Une séquence aléatoire de 30 nucléotides a été ajoutée entre les positions 59 et 77 des constructions utilisées en (A). (D) Des ARN rapporteurs contenant un bon (xxx) ou mauvais (xxx) contexte Kozak autour du codon CUG-59 ont été traduits en lysat de réticulocytes. Tous les ARN rapporteurs sont coiffés et polyadénylés et codent pour la Renilla luciférase. La traduction a été effectuée en présence de méthionine S35 radiomarquée. Les protéines produites ont été résolues sur un gel de polyacrylamide à 12% et analysées à l'aide d'un *phosphoimager*. Les résultats sont représentatifs de trois expériences indépendantes. L'évaluation du niveau d'expression des construits présentés à la figure D a été réalisée à l'aide du logiciel ImageQuant, en prenant en compte toutes les formes protéiques détectées. Les résultats sont représentatifs de trois expériences indépendantes.

FIGURE 8

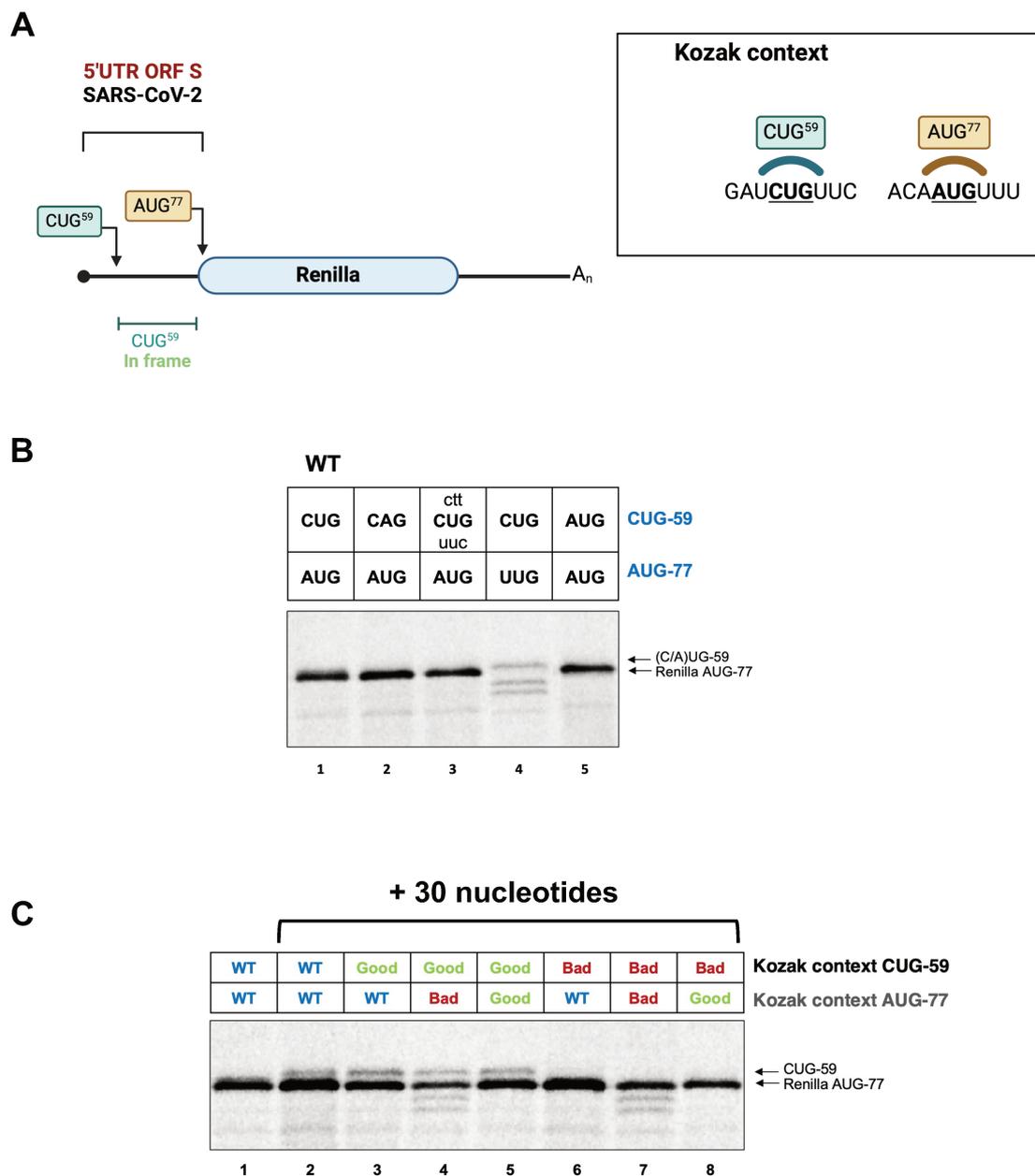


Figure 8: CUG 59 dans le contexte de ORF Spike

(A) Représentation schématique de la région 5'UTR de l'ARN sous-génomique codant pour la protéine Spike du SARS-CoV-2. La position du codon d'initiation AUG-77 et du codon CUG-59 ainsi que la phase de ce dernier sont indiqués. (B) Des ARN rapporteurs coiffés et polyadénylés contenant les mutations indiquées au niveau des positions CUG-59, AUG-77 de l'ARN sous-génomique codant pour la protéine Spike du SARS-CoV-2 ont été traduits en lysat de réticulocytes (RRL). (C) Une séquence aléatoire de 30 nucléotides a été ajoutée entre les positions 59 et 77 des constructions utilisées en (B). La traduction a été réalisée en présence de méthionine S35 radiomarquée. Les protéines produites ont été résolues sur un gel de polyacrylamide à 12% et analysées à l'aide d'un *phosphorimager*. Les résultats sont représentatifs de trois expériences indépendantes.

Matériel et Méthodes

Cette section "Matériel et Méthodes" se veut un complément à celle exposée dans l'article de Condé et al., 2021. Compte tenu des similitudes importantes avec les expériences décrites précédemment, l'objectif ici est de fournir des détails sur les nouvelles expériences introduites et les éventuelles modifications.

Purification des protéines recombinantes

Des bactéries E.Coli BL21 exprimant le vecteur pET28a-Nsp1 sauvage et pET28a-Nsp1 KH ont été cultivées jusqu'à ce que l'absorbance à 600 nm atteigne une valeur comprise entre 0,6–0,8, puis elles ont été induites avec 0,5 mM d'IPTG pendant 5 h à 30°C. Les culots bactériens ont été resuspendus dans un tampon de lyse B-PER (Thermo Scientific, 78243) complété avec 1 mg/ml de Lysozyme (Sigma), 30 U/ml de DNaseI (Promega) et un cocktail d'inhibiteurs de protéases (Roche). Après incubation le lysat a été clarifié à 6000g pendant 20 min. Le surnageant a été récupéré et incubé avec de la résine Cobalt HisPur (préalablement équilibrée dans le tampon de lyse natif ; Thermo Scientific, 89965), à 4°C pendant au moins 6 h sous agitation rotative douce. Le mélange a ensuite été appliqué sur une colonne de 5 ml (Bio-Rad) puis s'est écoulé par gravité. La résine a ensuite été lavée quatre fois avec cinq volumes de tampon de lavage natif (50 mM de NaH₂PO₄ pH 8.0; 300 mM de NaCl et 20 mM d'imidazole), et la protéine a été éluée avec le tampon d'éluion (50 mM de NaH₂PO₄ pH 8.0; 300 mM de NaCl; 500 mM d'imidazole et un comprimé de cocktail d'inhibiteurs de protéases, Roche). Le tampon d'éluion a par la suite été remplacé par dialyse dans 2 litres de tampon de stockage (25 mM de Tris-Cl pH 7.5, 50 mM de NaCl, 0.1 mM de MgCl₂, 25% de glycérol et un cocktail d'inhibiteurs de protéases). Des fractions ont été prélevées à chaque étape intermédiaire et analysées par électrophorèse sur gel SDS-PAGE à 10% suivi d'une coloration au bleu de Coomassie.

Traduction *in vitro*

Du lysat de réticulocytes de lapin Flexi traité à la nucléase micrococcale (RRL, Promega) a été programmé avec 5 à 15 fmol d'ARNm synthétisés *in vitro*, en présence de 75 mM de KCl, 0,5 mM de MgOAc, 2 mM de DTT, et 20 µM de chaque acide aminé,

pendant 30 minutes à 30 °C. Pour le marquage radioactif, 0,30 mCi/mL de méthionine-^[35S] a été incorporé dans le mélange réactionnel. Les réactions de traduction ont été arrêtées soit dans du tampon de lyse Renilla (PJK Biotech) pour une analyse par essai luciférase, soit dans du tampon de chargement Laemmli (4% SDS, 10% β-mercaptoéthanol, 20% glycérol, 0.004% bleu de bromophénol, 0.125 M Tris-HCl pH 6.8) pour une résolution sur gel SDS-PAGE à 12%. Le gel a ensuite été séché et analysé quantitativement avec un PhosphorImager de Molecular Dynamics. Des traductions ont également été réalisées en présence de diverses concentrations de Nsp1, sauvage ou mutante. Dans ces expériences, le lysat a été préincubé avec les protéines dans les quantités spécifiées pour chaque expérience pendant 10 minutes à 30°C, avant l'ajout des ARN synthétisés *in vitro* au mélange réactionnel.

Culture cellulaire et électroporation d'ARN

Des cellules HEK293T ont été cultivées dans du milieu DMEM (Dulbecco's Modified Eagle Medium, Invitrogen) complété avec 10 % de sérum de veau fœtal, 100 µg/ml de streptomycine et 100 unités/ml de pénicilline (Invitrogen). Les cellules ont été maintenues à 37 °C dans une atmosphère contenant 5 % de CO₂ et l'absence de mycoplasmes a été régulièrement testées.

Pour la transfection d'ARN, 3 à 5 x 10⁵ cellules ont été resuspendues dans 10 µL de tampon R (Life Technologies) en présence de 100 à 500 fmol d'ARNm coiffés et polyadénylés, puis immédiatement transfectées en utilisant l'appareil d'électroporation Neon (Life Technologies) selon les instructions du fabricant. Après l'électroporation, les cellules ont été incubées dans le milieu DMEM pendant 90 minutes à 37 °C, puis lavées deux fois avec du PBS et réparties en deux lots. Le premier lot a été utilisé pour mesurer les activités de luciférase, comme décrit dans Conde et al. 2021, tandis que le second a servi à l'extraction des ARN cytoplasmiques.

Extraction des ARN totaux et RT-qPCR

L'ARN total a été isolé à partir d'extraits cellulaires en utilisant le kit d'extraction d'ARN total de Macherey-Nagel. Ce kit inclut une étape d'élimination des ADN contaminants. La concentration d'ARN a été mesurée par spectrophotométrie sur un NanoDrop 2000 (Thermo Fisher). Les réactions de transcription inverse ont été réalisées avec 250 ng

d'ARN en utilisant le kit qScript (QuantaBio), conformément à la procédure du fabricant. La réaction de qPCR a été effectuée avec le réactif ONEGreen® FAST QPCR PREMIX (Ozyme) sur un appareil StepOnePlus (Thermo Fisher Scientific). Les échantillons ont été analysés en triplicats et normalisés par rapport à l'expression de l'ARN ribosomal 18S. Les amorces utilisées sont listées ci-dessous : Renilla luciférase (Amorce Renilla sens : TGGACAATAACTTCTTCGTGAAAAC ; Amorce Renilla antisens : GCTGCAAATTCTTCTGGTTCTAA). ARN ribosomal 18S (Amorce 18S sens : TCGGTCCTATTTTGCTGGTTT ; Amorce 18S antisens : ATCGCTAGTTGGCATCGTTT).

Génération des constructions plasmidiques

Les constructions pGlo-Renilla et pSARS-CoV-2-Renilla ont été décrites antérieurement (Conde et al. 2021). Les construits codant pour les mutants de la tige-boucle SL1 et de ses régions adjacentes ont été produits en utilisant un plasmide intermédiaire conçu pour le clonage Golden Gate. Pour ce faire, une réaction PCR a été effectuée sur le plasmide pSARS-CoV-2-Renilla, introduisant des extensions pour l'ajout de deux sites BsmBI, encadrées par les sites BstBI et AgeI. Ce produit de PCR a par la suite été inséré dans le plasmide pSF2-Renilla-Poly(A)100, préalablement digéré avec les enzymes AgeI et BstBI. Les sites BsmBI sont localisés de part et d'autre de la tige-boucle SL1, permettant son remplacement en une unique réaction Golden Gate. Pour chaque mutant de la tige-boucle SL1, deux oligonucléotides portant les mutations souhaitées et encadrés par des extrémités compatibles avec les sites BsmBI de notre plasmide receveur ont été hybridés, puis intégrés dans le plasmide via une réaction de digestion-ligation.

Les construits codant pour les mutants des tiges-boucles du SARS-CoV-2 ont été obtenus en clonant directement les différentes 5'UTR tronquées commandées auprès de Twist Bioscience, entre les sites BglII et BstBI du plasmide pGlo-Renilla.

Afin de générer les construits mutants des régions 5'UTR de l'ORF1AB, de l'ORF S, de l'ORF6 et des ORF7a/8, les inserts ont été amplifiés par des PCR consécutives avec des amorces spécifiques (Annexe 3). Les produits de PCR obtenus ont été digérés par l'enzyme de restriction BamHI et insérés dans le vecteur pRenilla entre les

sites de restriction BglII et BamHI. L'ensemble des construits ont été vérifiés par séquençage Sanger de toutes les régions d'intérêt.

Transcription *in vitro*

Les réactions de transcription *in vitro* ont été réalisées conformément à la méthode décrite dans Conde et al. 2021, avec cependant une modification. En plus des constructions dans le plasmide pRenilla précédemment décrit, des constructions dans le plasmide pSF2 ont été utilisées comme matrices pour la synthèse des ARN *in vitro*. Tout comme pour le plasmide pRenilla, ces constructions ont été digérées avec l'enzyme de restriction EcoRI afin de linéariser la matrice dans une configuration permettant ainsi l'ajout d'une queue poly(A) lors de l'étape de transcription subséquente.

Discussion et Perspectives

Le contrôle de l'expression des gènes est un élément central de l'interaction hôte-pathogène lors de l'infection virale. D'un côté, les virus dépendent de la machinerie de traduction cellulaire afin de synthétiser leurs protéines, une étape indispensable à la réplication virale. D'un autre côté, les cellules doivent se défendre et exprimer les acteurs de la réponse antivirale. Prendre le contrôle des voies de la traduction offre une possibilité intéressante pour chacun des protagonistes. Ainsi, les virus doivent limiter la traduction des ARNm cellulaires pour contrer la réponse antivirale cellulaire et détourner la machinerie traductionnelle à leur profit. À l'opposé, les cellules doivent inhiber la synthèse des protéines virales pour empêcher la réplication. Lors de l'infection par le SARS-CoV-2, 10 ARN viraux sont exprimés possédant tous une séquence leader commune de 75 nucléotides associée à une séquence variable propre à chacun d'entre eux.

J'ai étudié l'expression des ARNm du SARS-CoV-2 pour mieux comprendre comment et par quels mécanismes, ils étaient capables d'optimiser leur traduction.

Cela a permis de démontrer que les différents ARNm du SARS-CoV-2 n'étaient pas tous traduits avec la même efficacité. C'est la 5'UTR de l'ARN génomique qui, malgré sa taille et sa structure, s'exprime le mieux à la fois *in vitro* et *ex-vivo*; on retrouve cette caractéristique chez le SARS-CoV-1. Nous avons montré que la structure en tige-boucle SL1 ainsi que les séquences adjacentes sont nécessaires pour une expression optimale. L'initiation de la traduction à partir de cette 5'UTR emprunte un mécanisme dépendant de la coiffe et nécessite le complexe eIF4F, mais elle n'est pas stimulée par la présence de la queue poly(A).

Mon travail a aussi permis de mettre en évidence le rôle régulateur de courtes phases de lecture (uORF) au sein des régions 5'UTR virales qui exercent des effets variables selon leur position et les transcrits qui les contiennent.

Enfin nous avons montré que les différents ARN du SARS-CoV-2 n'étaient pas tous régulés de la même façon par la protéine Nsp1.

Les ARN du SARS-CoV-2 possèdent des efficacités de traduction variables

En comparant leur expression en lysat de réticulocytes et en cellules HEK 293T, nous avons mis en évidence que les différents ARNm du SARS-CoV2 avaient un niveau d'expression variable. C'est la 5'UTR de l'ARN génomique du SARS-CoV-2 qui est la plus efficace bien qu'elle soit pourtant longue et contienne 5 structures secondaires de type tige-boucle (SL1 à SL5) relativement stables (Miao et al. 2021; Richter et al. 2021) ce qui devrait, *a priori*, inhiber le processus de balayage du ribosome (Pelletier and Sonenberg 1985a). Ce n'est pas le cas et on retrouve des caractéristiques similaires chez le SARS-CoV-1 ce qui suggère que ces structures d'ARN représentent un avantage pour le virus en contenant probablement des éléments régulateurs. Des recherches ont montré le rôle important de la région 3'UTR qui est commune aux ARN du SARS-CoV-2 et qui possède des motifs qui stimulent la traduction des ARN viraux de façon spécifique (Khan et al. 2023). Il sera donc pertinent d'explorer si les régions 5'UTR et 3'UTR coopèrent pour réguler la production des protéines virales. En outre, il est important de rappeler que le contrôle de la traduction peut être influencé par le stress cellulaire (Liu and Qian 2014; Jia et al. 2024) ou l'expression de protéines virales (Stern-Ginossar et al. 2019) et que ces paramètres devront aussi être pris en considération.

La tige boucle SL1 est un élément central du contrôle traductionnel des ARN du SARS-CoV-2

La tige-boucle SL1 est un élément conservé formées de 25 nucléotides qui se trouve à l'extrémité 5' de l'ensemble des ARN viraux et qui a un rôle central dans leur traduction des ARN du SARS-CoV-2. (Bujanic et al. 2022) Dans le contexte de l'infection par le SARS-CoV-2, SL1 permet l'échappement des ARN viraux à l'inhibition traductionnelle produite par l'expression de la protéine Nsp1 (Kim et al. 2021). Plusieurs nucléotides de la tige-boucle sont requis, en particulier les cytosines aux positions 15, 19 et 20 car ils permettent à SL1 d'interagir avec le ribosome ce qui est primordial pour échapper à l'inhibition induite par Nsp1 (Bujanic et al. 2022).

Nos résultats suggèrent que SL1 contribue à l'expression de l'ARN génomique (en l'absence de Nsp1) et que toute délétion/mutation/substitution a un impact négatif sur la traduction. Nous avons montré que la cytosine en position 15, ainsi que les résidus

apicaux (incluant les positions C19 et C20), étaient indispensables pour maintenir l'efficacité de traduction et ce, en absence de la protéine virale Nsp1. Nous pouvons émettre l'hypothèse que SL1 pourrait recruter une protéine de liaison à l'ARN (RBP) ou un facteur d'initiation qui faciliterait le recrutement du complexe ribosomal. À notre connaissance aucun exemple de ce type n'a été décrit dans le cas de l'ARN génomique du SARS-CoV-2. Cependant un mécanisme similaire existe dans le cas de l'ARN génomique du VIH-1, qui recrute la protéine DDX3 sur la tige boucle TAR ce qui facilite l'initiation de la traduction en interagissant avec le complexe eIF4F (Soto-Rifo et al. 2012; Soto-Rifo et al. 2013). L'implication potentielle d'une RBP dans le contrôle de l'expression de l'ARNg du SARS-CoV-2 est séduisante car elle permettrait de moduler la traduction en fonction du cycle viral et pouvoir orienter l'ARNg au niveau de la traduction ou de l'empaquetage. Il sera important de déterminer l'importance de la tige boucle SL1 dans le contexte de l'infection mais cela reste techniquement complexe car les résidus impliqués dans le maintien de l'efficacité de la traduction sont les mêmes que ceux qui permettent l'échappement à l'inhibition globale de la traduction.

L'expression des ARN du SARS-CoV-2 implique-t-elle un mécanisme d'initiation de la traduction non canonique?

La région 5'UTR du SARS-CoV-2 possède une structure relativement complexe et le codon AUG principal (AUG 266) se trouve en aval d'une jonction à quatre voies. Ce type de configuration est retrouvé dans l'IRES de HCV (Melcher et al. 2003) et nous avons pensé à la possibilité d'un mécanisme d'initiation par entrée interne du ribosome. Cependant nos expériences ont déterminé que la présence de la coiffe était nécessaire à la traduction efficace des ARN du SARS-CoV-2 et que l'intégrité du complexe eIF4F (eIF4E, eIF4A et eIF4G) était requise. De plus, l'introduction d'un codon AUG à la position 59 permet d'initier la traduction de manière très efficace ce qui renforce la notion d'un mécanisme de traduction avec un balayage linéaire des ribosomes.

Nos résultats contrastent fortement avec les observations de Boris Slobodin et collègues qui décrivent la présence d'une séquence IRES au sein de la région 5'UTR du SARS-CoV-2 (Slobodin et al. 2022). Leurs travaux montrent que la 5'UTR de l'ARN

génomique peut initier la traduction dans un système bicistronique mais le niveau d'expression demeure extrêmement faible et non compatible avec la production virale observée en conditions d'infection. Une explication qui pourrait concilier les deux hypothèses serait l'utilisation en alternance des deux mécanismes d'initiation, en fonction des conditions cellulaires. Cela restera à déterminer dans de futures expériences.

Tout comme les ARNm cellulaires, les ARN du SARS-CoV-2 sont polyadénylés à leur extrémité 3'. Dans le modèle de traduction canonique de l'initiation dépendante de la coiffe, la protéine eIF4G établit une liaison entre la coiffe par le biais de eIF4E et la queue poly(A) via la PABP ce qui permet la pseudocircularisation des ARNm (Imataka et al. 1998) et active l'expression des transcrits (Munroe and Jacobson 1990; Gallie 1991). Afin de tester l'influence de la queue poly(A), nous avons utilisé un système *in vitro* qui est en mesure de récapituler la synergie entre coiffe et la queue poly(A) (lysat de réticulocytes non traité) (Soto Rifo et al. 2007). De manière surprenante, nous n'avons pas observé de dépendance à la queue poly(A) pour les ARN dont la traduction est dirigée par la 5'UTR de l'ARN génomique du SARS-CoV-2. Cela contraste avec le modèle classique de traduction qui requiert l'interaction entre la queue poly(A) et la coiffe. Il existe toutefois des exemples d'ARN viraux ou cellulaires dépourvus de queue poly(A) mais dans ces cas, des mécanismes alternatifs de circularisation sont mis en place et ils impliquent d'autres interactions 5'-3' qui permettent la pseudo circularisation. De façon intéressante des études structurales indiquent que les ARN des coronavirus (Ziv et al. 2020), y compris du SARS-CoV-2, sont capables de se circulariser, notamment par une interaction entre la région 3'UTR et les tiges boucles SL2, SL3 et SL4 de la région 5'UTR (Huston et al. 2021). Dans le cadre du coronavirus murin (MHV) il a même été montré que la tige boucle SL1 était indispensable à ce processus (Li et al. 2008). Bien que ces événements de circularisation aient été principalement décrits dans le cadre de la transcription des ARN viraux, il est envisageable qu'ils se produisent également pendant l'étape de traduction ce qui expliquerait l'absence d'effet de la queue poly(A).

Les régions 5'UTR du SARS-CoV-2 possèdent des uORF qui régulent leur expression.

Nous avons identifié et caractérisé deux uORF, une localisée dans l'ARN génomique (ORF-1) et une autre qui commence à partir d'un codon CUG (ORF-2) et qui est présente dans la séquence leader commune à tous les ARN du SARS-CoV-2. Des études récentes de ribosome profiling ont confirmé que ces uORF pouvaient être reconnues par les ribosomes (Lee et al. 2021; Liu et al. 2023)

Les uORF sont des éléments présents dans les 5'UTR des ARN et les techniques de ribosome profiling suggèrent qu'elles pourraient être retrouvées dans plus de 50% des ARN du génome (Ingolia et al. 2011; Lee et al. 2012). Elles peuvent exercer différents niveaux de contrôle sur la traduction de l'ORF principale mais, le plus généralement, elles ont une action inhibitrice en empêchant le ribosome d'atteindre le codon AUG principal (Young and Wek 2016). Dans le cas de la première uORF (uORF-1) nous avons montré que sa délétion augmentait la traduction à partir de l'AUG principal mais que cet effet restait modeste. La position de l'uORF1 est très conservée chez les différents coronavirus bien que la séquence nucléotidique puisse varier. Ainsi, la délétion de cette même uORF chez le MERS a des conséquences beaucoup plus prononcées avec une forte stimulation de la synthèse protéique à partir de l'AUG authentique. Nous n'avons pas mis en évidence d'éléments régulateurs permettant d'expliquer cette différence entre les deux homologues. Il est à noter que l'uORF de la région 5'UTR de l'hépatite murine, qui est également un coronavirus, est la mieux caractérisée. Sa délétion n'entraîne pas d'effets visibles immédiats et il y a production de particules virales infectieuses; cependant, après plusieurs cycles de réplication, le séquençage des virus montre une réversion vers l'état sauvage et la réapparition de cette uORF suggérant qu'elle est nécessaire pour une réplication optimale.

La deuxième uORF (uORF-2) débute par un codon CUG à la position 59 de la séquence leader commune et son utilisation a aussi été observée par les expériences de ribosome profiling (Finkel et al. 2021b; Kim et al. 2021). Les codons CUG peuvent dans certains contextes nucléotidiques favorables être utilisés pour initier la traduction avec des efficacités variables (Dasso and Jackson 1989; Lee et al. 2012). Dans le cas des ARN du SARS-CoV-2 ce codon CUG est tout particulièrement intéressant car il

se retrouve dans des cadres de lecture différents pour chacun des ARN sous génomiques. Nous avons étudié l'impact de ce CUG dans le contexte de l'ARN génomique et des ARN sous-génomiques codant pour les protéines S, ORF6 et ORF7a-8. Cela a permis de mettre en évidence que, seulement dans le contexte de la 5'UTR de l'ARN codant pour la protéine spike, ce CUG-59 était utilisé et ce avec une fréquence modeste (environ 10%). En conditions d'infection, ce codon CUG est mieux reconnu par les ribosomes avec une fréquence de l'ordre de 50 % mesurée par ribosome profiling (Finkel et al. 2021b; Kim et al. 2021). Ces variations pourraient s'expliquer par la différence de systèmes expérimentaux (*in vitro* vs *in vivo*) et pourrait refléter un besoin en protéines cellulaires ou virales qui ne se trouvent pas dans le modèle du lysat de réticulocytes. En effet, l'initiation à des positions non canoniques (codon CUG / uORF) est souvent modulée par différents facteurs cellulaires tels que le stress (Vattem and Wek 2004; Andreev et al. 2015; Zhang et al. 2015; Starck et al. 2016). C'est le cas par exemple de l'ARN codant pour la protéine L du virus Ébola qui contient une uORF inhibitrice en conditions normales mais qui devient inactive lorsqu'eIF2 α est phosphorylé, ce qui permet l'expression de la polymérase virale L (Shabman et al. 2013).

De manière intéressante, la traduction de l'uORF-2 de l'ARN codant pour spike rajoute une extension N-terminale de six acides aminés dont le rôle biologique n'a pas été caractérisé. Il est très probable que cette extension N-terminale soit clivée au cours de sa translocation au réticulum endoplasmique (Kim et al. 2021; Xia 2021) mais il n'en demeure pas moins qu'elle pourrait jouer un rôle au niveau de la réponse immunitaire en étant reconnue comme antigène.

Nous avons été surpris de constater que le codon CUG-59 n'était pas décodé par les ribosomes parcourant l'ARN génomique ou l'ORF 6 ou l'ORF7a/8 et cela est d'autant plus étonnant que le contexte nucléotidique de ce codon est identique pour tous les ARN sous génomiques. En revanche, son remplacement par un codon AUG permet sa reconnaissance chez tous les transcrits étudiés (ARNg, Spike, ORF6, ORF7a/8) et change même radicalement le patron d'initiation en le désignant comme site principal de départ de la synthèse protéique. Cela confirme que les ribosomes balayent la position 59 et que le contexte nucléotidique est favorable à l'initiation à ce site. Cela contraste de nouveau avec les résultats d'un autre groupe avançant la présence d'une

séquence IRES (Slobodin et al. 2022) et nous conforte dans notre hypothèse d'un mécanisme cap-dépendant à partir de l'extrémité 5' des ARN viraux.

Impact de la protéine virale Nsp1 sur les ARN viraux

Les ARN du SARS-CoV-2 sont capables de contourner l'inhibition globale de la traduction médiée par la protéine Nsp1 et ceci repose sur la présence de la tige boucle SL1 (Bujanic et al. 2022). La protéine Nsp1 exerce sa répression traductionnelle de deux manières distinctes : (i) son domaine C-terminal obstrue le canal d'entrée du ribosome et bloque l'accès pour le décodage des transcrits; (2) Nsp1 induit le clivage des ARN cellulaires par un mécanisme dépendant de son association à la sous unité ribosomique 40S.

Notre travail a montré des différences notables de traduction *in vitro* entre les différents ARN sous génomiques du SARS-CoV-2 en présence de la protéine Nsp1 recombinante. Ainsi, les ARN les plus longs échappent le plus efficacement à l'inhibition médiée par Nsp1. À l'inverse, les deux ARN viraux les plus sensibles à l'effet de Nsp1 possèdent une région 5' UTR de 77 nucléotides dans un cadre de lecture identique (ORF E et ORF 3a). De fait, le codon CUG-59 est dans le même contexte nucléotidique au sein de ces deux 5'UTR. Nous ne disposons pas d'autres évidences pour affirmer que le codon CUG-59 puisse être impliqué dans ce mécanisme mais ce serait en accord avec des résultats récents qui suggèrent un rôle pour Nsp1 dans la sélection du site d'initiation en coopérant avec les facteurs d'initiation eIF1 et eIF1a (Aviner et al. 2024).

Le contrôle traductionnel des ARN du SARS-CoV-2 est-il modulé par des protéines de liaison à l'ARN?

Pendant mon travail de doctorat, j'ai généré un très grand nombre de mutants des régions 5'UTR ce qui permet de déterminer le rôle des éléments fonctionnels de ces ARN sur la traduction. Cependant, ces méthodes ont toutefois leurs limites ; elles nécessitent beaucoup de travail de clonage et ne permettent pas, par exemple, d'identifier facilement si des facteurs cellulaires ou viraux s'associent spécifiquement au niveau de ces positions pour moduler leur fonction. Elles ne permettent pas, non

plus, de déterminer les dynamiques d'association/dissociation de ces protéines avec les complexes d'initiation.

C'est pour répondre à ces questions que j'ai décidé de développer un outil biotechnologique innovant dans les derniers mois de mon travail de thèse et que j'aimerais présenter très brièvement dans le chapitre suivant 'nouvelles perspectives' sous forme de résultats préliminaires.

Nouvelles perspectives et Projets futurs

Projet de biotinylation de proximité des protéines de liaison à l'ARN

J'ai choisi de développer et adapter la technique de biotinylation de proximité pour déterminer précisément la composition et la dynamique d'assemblage des complexes d'initiation de la traduction sur un ARNm donné, dans notre cas les différents ARNm du SARS-CoV-2. Nous souhaitons également caractériser leur évolution au fil de l'infection. Jusqu'à présent, il était techniquement difficile, voire impossible, de piéger et de marquer ces complexes sur une région précise de l'ARN. Afin de surmonter ces limitations, nous avons couplé la protéine de biotinylation de proximité Apex (Hung et al. 2016; Nguyen et al. 2020) à un système d'édition génomique de type CRISPR/dCas13, permettant de cibler des régions précises de l'ARN avec une résolution de l'ordre du nucléotide (Yan et al. 2018). En tirant parti des propriétés des deux enzymes impliquées, notre outil peut se localiser à des positions précises de l'ARN ciblé et ajouter des groupements biotine sur les protéines se trouvant à proximité immédiate de ces ARN. Celles-ci peuvent ensuite être identifiées par une analyse de spectrométrie de masse (Figure 1).

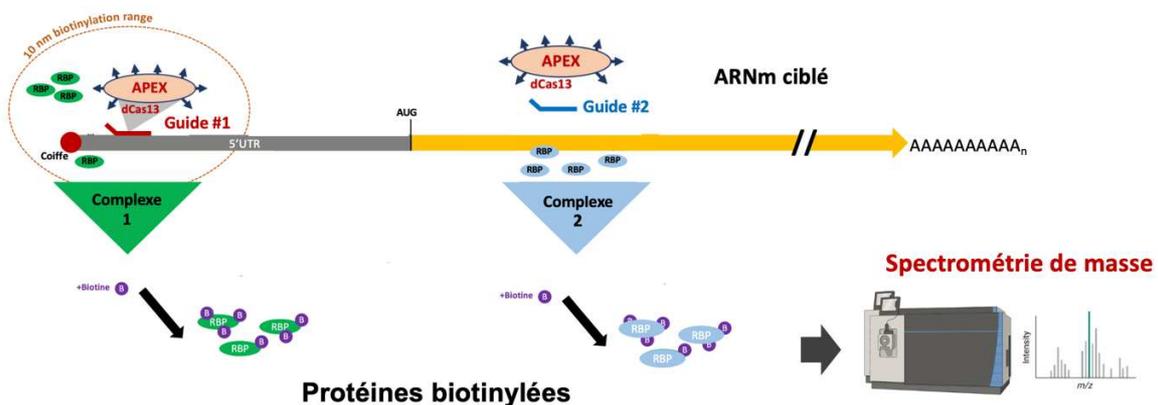


Figure 1 : Dispositif expérimental

En couplant l'enzyme Apex avec un système d'édition du génome contenant l'enzyme Cas13, il est possible de cibler des régions précises de l'ARN pour déterminer les complexes protéiques associés qui sont ensuite identifiés par spectrométrie de masse.

L'enzyme Apex est une biotine peroxydase qui est activée par l'ajout de peroxyde d'hydrogène sur des temps très brefs de l'ordre de la minute (Pfeiffer et al. 2022). Cela permet de pouvoir concilier un système inductible avec une période d'activité faible pour minimiser les effets non désirables tout en conservant une haute spécificité d'action.

Parmi les différentes Cas13 disponibles, nous avons choisi d'utiliser Cas13d car elle présente moins de cibles non spécifiques que ses homologues (Abudayyeh et al. 2017; Cox et al. 2017). Aussi, des versions catalytiquement inactives ont été développées (mutants dCas13) qui reconnaissent et se fixent sur leurs ARN cibles sans induire leur clivage (Abudayyeh et al. 2019). De plus, la dCas13 peut être fusionnée à différentes protéines d'intérêt et être facilement livrée dans les cellules y compris en utilisant des VLP (Mangeot et al. 2019).

Nous avons d'ores et déjà testé la capacité du complexe Apex-dCas13 à biotinyler le partenaire connu d'un ARN (Figure 2). Pour cela nous avons construit un rapporteur contenant la 5'UTR de la bêta-globine suivie de 8 tiges-boucles MS2. Cet outil nous permet de fixer Apex-dCas13 sur un ARN cible via l'utilisation d'un guide spécifique. Des cellules HEK293T ont été co-transfectées avec cet ARN rapporteur et un guide dirigé contre la région 5'UTR de la globine. La construction Apex-dCas13 ainsi qu'un plasmide codant pour la protéine de liaison à MS2 fusionnée avec une protéine fluorescente (GFP ou mCh) sont co-transfectés. Les extraits cellulaires sont analysés par Western-Blot à l'aide d'un marquage HRP-Streptavidine, permettant de révéler les protéines biotinylées.

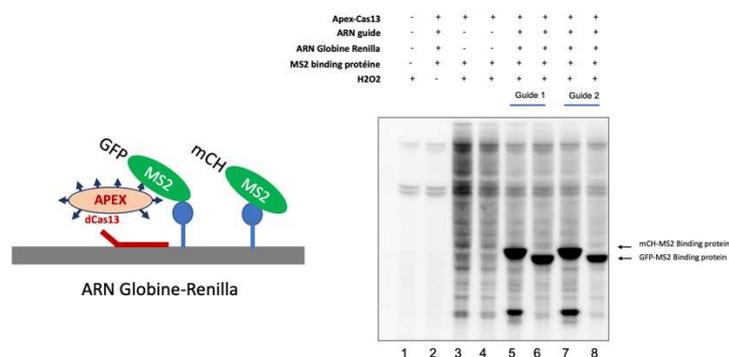


Figure 2 : Biotinylation de proximité.

L'ARN Globine Rénilla contenant 8 tiges-boucles MS2 a été transfecté en présence de l'enzyme Apex-dCas13, les protéines MS2 fusionnées à GFP ou mCh et 2 guides ARN complémentaires comme indiqué sur la figure. La biotinylation des protéines de fusion MS2 est indiquée sur la figure.

Nos résultats indiquent que l'addition de l'ARN guide change le patron de biotinylation observé, ce qui confirme la relocalisation de Apex-dCas13 sur l'ARN cible en présence du guide. De plus, nos résultats montrent que lorsque nous contraignons la MS2 binding protéine fusionnée à la GFP ou à mCH à se localiser sur l'ARN rapporteur sur lequel la dCas13 est amenée, ces deux protéines se retrouvent biotinylées ce qui démontre à la fois, la fonctionnalité et la spécificité de notre système. L'expérience de la figure 2 (ci-dessus) amène la preuve de concept que nous pouvions guider l'enzyme biotinylante sur une position donnée de l'ARN avec une résolution de l'ordre du nucléotide.

Cependant, il persiste un important bruit de fond (Figure 2 pistes 3 et 4) lorsque la protéine Apex est exprimée seule, c'est à dire en absence de guide et en absence de cible. Nous avons réalisé que ce bruit de fond était trop important pour être en mesure d'obtenir des données de protéomique exploitables.

Afin de réduire ce bruit de fond, nous avons adapté notre système. Nous avons notamment remplacé Apex2 par une version « split TurboID ». Ce système est constitué de deux domaines (N-terminal et C-terminal), qui peuvent être exprimés à partir de deux plasmides distincts. Avec ce nouvel outil, c'est seulement lorsque les deux parties de l'enzyme s'associent au niveau de l'ARN cible grâce à la dCas13 et que le TurboID actif est reconstitué (Figure 3).

Nous avons évalué l'efficacité du split TurboID sur un ARN, introduit au préalable dans des cellules HEK293T, qui porte la 5'UTR de la globine et médie l'expression de la GFP (5'UTR globine-GFP). Cette approche réduit considérablement le niveau du bruit de fond. Cela est montré sur la Figure 3 (panel de gauche) sur laquelle on peut voir un spectre de biotinylation spécifique au niveau de la piste 1 alors qu'il n'y a que très peu de protéines biotinylées lorsque l'une des extrémités de TurboID est manquante (piste 2) ou que les ARN guides ne sont pas ajoutés (piste 3 et piste 4). Il est aussi à noter que l'ARN cible est non modifié, c'est à dire qu'il ne contient pas les tiges-boucles MS2 comme c'était le cas dans l'expérience préliminaire présentée à la figure 2.

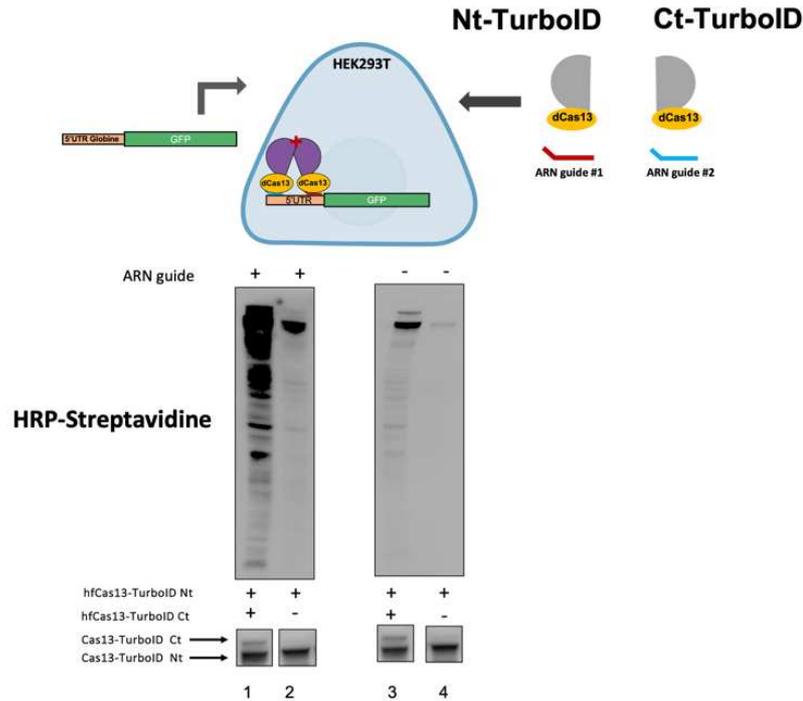


Figure 3: Biotinylation d'un ARN exogène et endogène dans des cellules HEK293T

Un plasmide codant pour un ARN contenant la 5'UTR de la globine et qui dirige l'expression de la GFP a été transfecté dans des cellules HEK293T. Le domaine Nt du système split TurboID a ensuite été exprimé avec, ou sans, le domaine Ct comme indiqué sur la figure. Les guides ARN spécifiques de la région 5'UTR de globine ont été ajoutés, ou non, aux cellules. Après addition de biotine dans le milieu, les protéines biotinylées ont été récupérées, éluées et résolues sur gel. La forme active du Split TurboID est schématisée par une croix rouge.

Pour finir, le dernier obstacle technique était d'évaluer si notre dispositif de biotinylation était à la fois efficace et capable de différencier les complexes protéiques en fonction de la région spécifique de l'ARN visé. Pour ce faire, nous avons réalisé notre première expérience de protéomique en ciblant un 5'UTR globine-GFP, transfecté dans des cellules HEK293T. Nous avons utilisé des guides spécifiques de la 5'UTR (condition 1) et d'autres qui ciblent la région codante GFP (condition 2). Après addition de biotine dans le milieu suivie d'une période d'incubation de 30 minutes, nous avons récupéré les complexes protéiques obtenus pour chaque condition avec des billes couplées à la streptavidine. Ces échantillons ont ensuite été analysés par spectrométrie de masse (Figure 4).

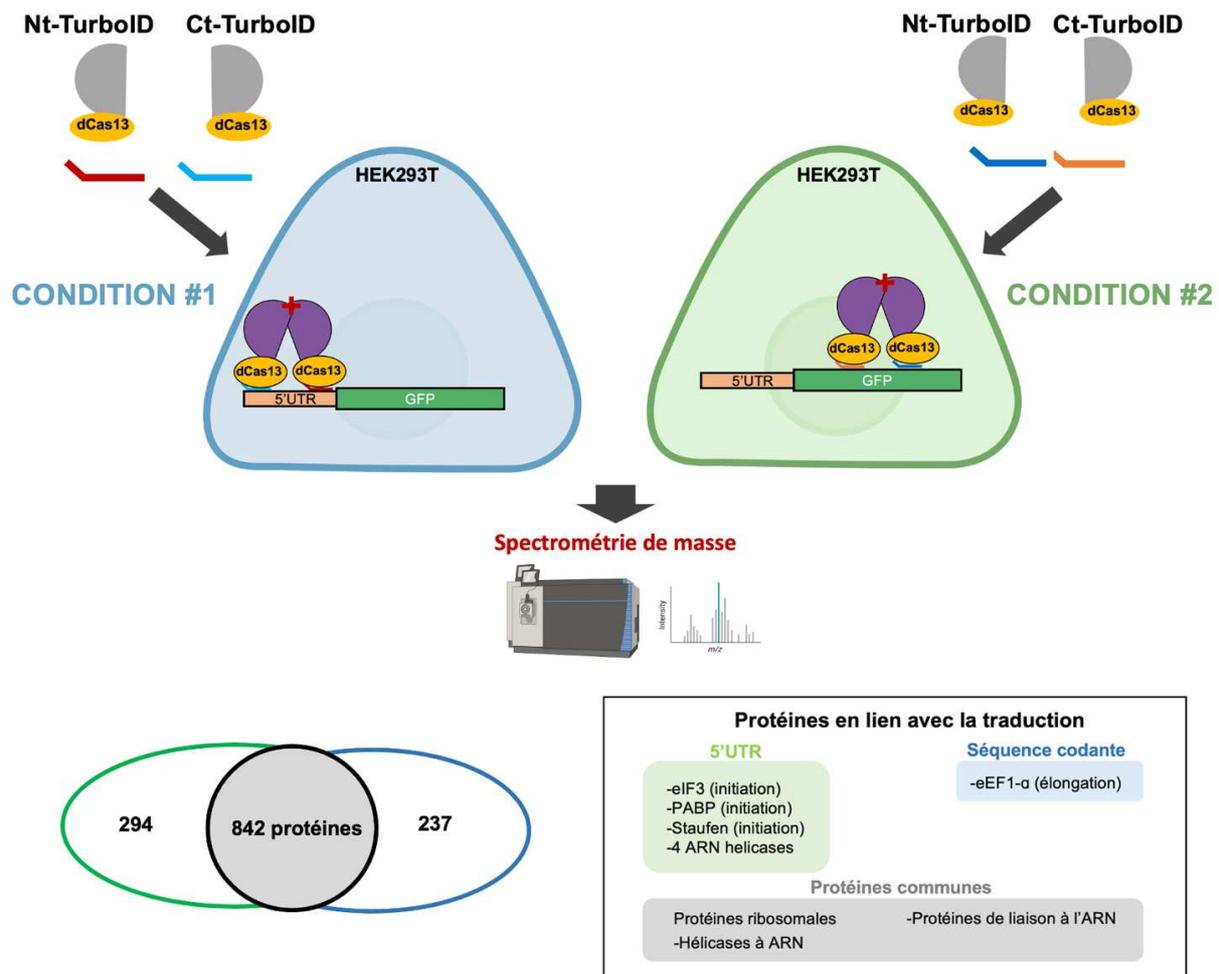


Figure 4: Validation de l'approche expérimentale

Un plasmide codant pour un ARN contenant la 5'UTR de la globine et qui dirige l'expression de la GFP a été transfecté dans des cellules HEK293T. Le système split TurboID a ensuite été exprimé dans ces mêmes cellules en présence d'ARN guides spécifiques de la région 5'UTR (condition 1) ou de la GFP (condition 2). Après activation de l'enzyme par ajout de biotine libre, les protéines biotinylées dans chacune des deux conditions ont été isolées et analysées par spectrométrie de masse. La forme active du Split TurboID est schématisée par une croix rouge.

Les premiers résultats montrent qu'il y a 1136 et 1079 protéines biotinylées dans les conditions 1 et 2 respectivement. Parmi celles-ci, 842 sont communes aux deux régions ciblées, 294 sont uniquement retrouvées dans la région 5'UTR et 237 sont spécifiquement retrouvées dans la région codante. Cela montre d'ores et déjà une composition différente des complexes protéiques en fonction de leur position sur l'ARNm (5'UTR vs région codante) et que notre approche expérimentale est assez résolutive pour les mettre en évidence. Un travail plus approfondi d'analyse des données est en cours mais on peut déjà affirmer que nous retrouvons un

enrichissement de protéines impliquées dans l'initiation de la traduction au niveau de la région 5'UTR tel que le facteur d'initiation eIF3, la poly(A)-binding protein (PABP) qui lie la coiffe et assure la jonction 5'-3' de l'ARN au cours de l'initiation ou encore la protéine Staufen 1. Il est à noter que nous avons réalisé cette expérience avec une série de contrôles adéquats (absence/présence de guides ARN, guides ARN aléatoires, absence/présence de l'ARN cible et protéine TurboID couplée à la GFP à la place de dCas13). Par souci de clarté et de concision, nous n'avons présenté que les résultats concernant les complexes protéiques piégés sur la 5'UTR et sur la région codante.

Conclusion:

Ces résultats bien que préliminaires sont encourageants et seront affinés et les expériences répétées dans les prochains mois. Nous allons utiliser cet outil dans le cadre de l'étude de la traduction des ARN de trois virus différents. Premièrement, le VSV, pour optimiser l'outil de façon simple avec un système répliatif. Deuxièmement, le SARS-CoV-2, pour répondre à plusieurs questions abordées largement au cours de cette thèse. Et enfin, le VIH-1, très étudié au laboratoire, pour identifier les partenaires protéiques présents au niveau de trois sites d'initiation de la traduction distincts sur l'ARN génomique. Ce dernier projet, auquel je participe également activement, a été récemment financé par deux grants successifs de l'ANRS, soutien important de ce développement technologique.

Références bibliographiques

- Abaeva IS, Arhab Y, Miścicka A, Hellen CUT, Pestova TV. 2023. In vitro reconstitution of SARS-CoV-2 Nsp1-induced mRNA cleavage reveals the key roles of the N-terminal domain of Nsp1 and the RRM domain of eIF3g. *Genes Dev* **37**: 844-860.
- Abernathy E, Glaunsinger B. 2015. Emerging roles for RNA degradation in viral replication and antiviral defense. *Virology* **479-480**: 600-608.
- Abudayyeh OO, Gootenberg JS, Essletzbichler P, Han S, Joung J, Belanto JJ, Verdine V, Cox DBT, Kellner MJ, Regev A et al. 2017. RNA targeting with CRISPR-Cas13. *Nature* **550**: 280-284.
- Abudayyeh OO, Gootenberg JS, Franklin B, Koob J, Kellner MJ, Ladha A, Joung J, Kirchgatterer P, Cox DBT, Zhang F. 2019. A cytosine deaminase for programmable single-base RNA editing. *Science* **365**: 382-386.
- Addetia A, Lieberman NAP, Phung Q, Hsiang TY, Xie H, Roychoudhury P, Shrestha L, Loprieno MA, Huang ML, Gale M, Jr. et al. 2021. SARS-CoV-2 ORF6 Disrupts Bidirectional Nucleocytoplasmic Transport through Interactions with Rae1 and Nup98. *mBio* **12**.
- Adivarahan S, Livingston N, Nicholson B, Rahman S, Wu B, Rissland OS, Zenklusen D. 2018. Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway. *Mol Cell* **72**: 727-738.e725.
- Agback T, Dominguez F, Frolov I, Frolova EI, Agback P. 2021. 1H, 13C and 15N resonance assignment of the SARS-CoV-2 full-length nsp1 protein and its mutants reveals its unique secondary structure features in solution. *PLoS One* **16**: e0251834.
- Akhrymuk I, Kulemzin SV, Frolova EI. 2012. Evasion of the innate immune response: the Old World alphavirus nsP2 protein induces rapid degradation of Rpb1, a catalytic subunit of RNA polymerase II. *J Virol* **86**: 7180-7191.
- Algire MA, Maag D, Lorsch JR. 2005. Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol Cell* **20**: 251-262.
- Ali IK, McKendrick L, Morley SJ, Jackson RJ. 2001. Activity of the hepatitis A virus IRES requires association between the cap-binding translation initiation factor (eIF4E) and eIF4G. *J Virol* **75**: 7854-7863.
- Altmann M, Handschin C, Trachsel H. 1987. mRNA cap-binding protein: cloning of the gene encoding protein synthesis initiation factor eIF-4E from *Saccharomyces cerevisiae*. *Mol Cell Biol* **7**: 998-1003.
- Altmann M, Müller PP, Wittmer B, Ruchti F, Lanker S, Trachsel H. 1993. A *Saccharomyces cerevisiae* homologue of mammalian translation initiation factor 4B contributes to RNA helicase activity. *Embo j* **12**: 3997-4003.
- Alvarez E, Menéndez-Arias L, Carrasco L. 2003. The eukaryotic translation initiation factor 4G1 is cleaved by different retroviral proteases. *J Virol* **77**: 12392-12400.
- Amorim R, Temzi A, Griffin BD, Mouland AJ. 2017. Zika virus inhibits eIF2 α -dependent stress granule assembly. *PLoS Negl Trop Dis* **11**: e0005775.

- Amrani N, Ghosh S, Mangus DA, Jacobson A. 2008. Translation factors promote the formation of two states of the closed-loop mRNP. *Nature* **453**: 1276-1280.
- Andreev DE, Fernandez-Miragall O, Ramajo J, Dmitriev SE, Terenin IM, Martinez-Salas E, Shatsky IN. 2007. Differential factor requirement to assemble translation initiation complexes at the alternative start codons of foot-and-mouth disease virus RNA. *Rna* **13**: 1366-1374.
- Andreev DE, Loughran G, Fedorova AD, Mikhaylova MS, Shatsky IN, Baranov PV. 2022. Non-AUG translation initiation in mammals. *Genome Biology* **23**: 111.
- Andreev DE, O'Connor PB, Fahey C, Kenny EM, Terenin IM, Dmitriev SE, Cormican P, Morris DW, Shatsky IN, Baranov PV. 2015. Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *Elife* **4**: e03971.
- Archer SK, Shirokikh NE, Beilharz TH, Preiss T. 2016. Dynamics of ribosome scanning and recycling revealed by translation complex profiling. *Nature* **535**: 570-574.
- Arias C, Walsh D, Harbell J, Wilson AC, Mohr I. 2009. Activation of host translational control pathways by a viral developmental switch. *PLoS Pathog* **5**: e1000334.
- Asano K, Krishnamoorthy T, Phan L, Pavitt GD, Hinnebusch AG. 1999. Conserved bipartite motifs in yeast eIF5 and eIF2Bepsilon, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. *Embo j* **18**: 1673-1688.
- Asano K, Phan L, Anderson J, Hinnebusch AG. 1998. Complex formation by all five homologues of mammalian translation initiation factor 3 subunits from yeast *Saccharomyces cerevisiae*. *J Biol Chem* **273**: 18573-18585.
- Avanzino BC, Fuchs G, Fraser CS. 2017. Cellular cap-binding protein, eIF4E, promotes picornavirus genome restructuring and translation. *Proc Natl Acad Sci U S A* **114**: 9611-9616.
- Aviner R, Lidsky PV, Xiao Y, Tassetto M, Kim D, Zhang L, McAlpine PL, Elias J, Frydman J, Andino R. 2024. SARS-CoV-2 Nsp1 cooperates with initiation factors EIF1 and 1A to selectively enhance translation of viral RNA. *PLoS Pathog* **20**: e1011535.
- Baer BW, Kornberg RD. 1983. The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *J Cell Biol* **96**: 717-721.
- Bakhshesh M, Gropelli E, Willcocks MM, Royall E, Belsham GJ, Roberts LO. 2008. The picornavirus avian encephalomyelitis virus possesses a hepatitis C virus-like internal ribosome entry site element. *J Virol* **82**: 1993-2003.
- Bandyopadhyay PK, Wang C, Lipton HL. 1992. Cap-independent translation by the 5' untranslated region of Theiler's murine encephalomyelitis virus. *J Virol* **66**: 6249-6256.
- Banerjee AK, Blanco MR, Bruce EA, Honson DD, Chen LM, Chow A, Bhat P, Ollikainen N, Quinodoz SA, Loney C et al. 2020. SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. *Cell* **183**: 1325-1339.e1321.
- Basavappa R, Sigler PB. 1991. The 3 A crystal structure of yeast initiator tRNA: functional implications in initiator/elongator discrimination. *Embo j* **10**: 3105-3111.
- Beaudoin JD, Perreault JP. 2010. 5'-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res* **38**: 7022-7036.
- Belsham GJ, Brangwyn JK. 1990. A region of the 5' noncoding region of foot-and-mouth disease virus RNA directs efficient internal initiation of protein synthesis

- within cells: involvement with the role of L protease in translational control. *J Virol* **64**: 5389-5395.
- Ben-Asouli Y, Banai Y, Pel-Or Y, Shir A, Kaempfer R. 2002. Human interferon-gamma mRNA autoregulates its translation through a pseudoknot that activates the interferon-inducible protein kinase PKR. *Cell* **108**: 221-232.
- Benedetti F, Snyder GA, Giovanetti M, Angeletti S, Gallo RC, Ciccozzi M, Zella D. 2020. Emerging of a SARS-CoV-2 viral strain with a deletion in nsp1. *J Transl Med* **18**: 329.
- Beretta L, Gingras AC, Svitkin YV, Hall MN, Sonenberg N. 1996. Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. *Embo j* **15**: 658-664.
- Bernstein P, Peltz SW, Ross J. 1989. The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. *Mol Cell Biol* **9**: 659-670.
- Beyerstedt S, Casaro EB, Rangel É B. 2021. COVID-19: angiotensin-converting enzyme 2 (ACE2) expression and tissue susceptibility to SARS-CoV-2 infection. *Eur J Clin Microbiol Infect Dis* **40**: 905-919.
- Bhatt PR, Scaiola A, Loughran G, Leibundgut M, Kratzel A, Meurs R, Dreos R, O'Connor KM, McMillan A, Bode JW et al. 2021. Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. *Science* **372**: 1306-1313.
- Blakqori G, van Knippenberg I, Elliott RM. 2009. Bunyamwera orthobunyavirus S-segment untranslated regions mediate poly(A) tail-independent translation. *J Virol* **83**: 3637-3646.
- Boersma S, Rabouw HH, Bruurs LJM, Pavlovič T, van Vliet ALW, Beumer J, Clevers H, van Kuppeveld FJM, Tanenbaum ME. 2020. Translation and Replication Dynamics of Single RNA Viruses. *Cell* **183**: 1930-1945.e1923.
- Bonderoff JM, Larey JL, Lloyd RE. 2008. Cleavage of poly(A)-binding protein by poliovirus 3C proteinase inhibits viral internal ribosome entry site-mediated translation. *J Virol* **82**: 9389-9399.
- Bonnal S, Boutonnet C, Prado-Lourenço L, Vagner S. 2003. IRESdb: the Internal Ribosome Entry Site database. *Nucleic Acids Res* **31**: 427-428.
- Bonning BC, Miller WA. 2010. Dicistroviruses. *Annu Rev Entomol* **55**: 129-150.
- Borman AM, Michel YM, Kean KM. 2000. Biochemical characterisation of cap-poly(A) synergy in rabbit reticulocyte lysates: the eIF4G-PABP interaction increases the functional affinity of eIF4E for the capped mRNA 5'-end. *Nucleic Acids Res* **28**: 4068-4075.
- Borman AM, Michel YM, Malnou CE, Kean KM. 2002. Free poly(A) stimulates capped mRNA translation in vitro through the eIF4G-poly(A)-binding protein interaction. *J Biol Chem* **277**: 36818-36824.
- Bosch BJ, van der Zee R, de Haan CA, Rottier PJ. 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J Virol* **77**: 8801-8811.
- Bramham CR, Jensen KB, Proud CG. 2016. Tuning Specific Translation in Cancer Metastasis and Synaptic Memory: Control at the MNK-eIF4E Axis. *Trends Biochem Sci* **41**: 847-858.
- Brasey A, Lopez-Lastra M, Ohlmann T, Beerens N, Berkhout B, Darlix JL, Sonenberg N. 2003. The leader of human immunodeficiency virus type 1 genomic RNA

- harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. *J Virol* **77**: 3939-3949.
- Brito Querido J, Sokabe M, Díaz-López I, Gordiyenko Y, Fraser CS, Ramakrishnan V. 2024a. The structure of a human translation initiation complex reveals two independent roles for the helicase eIF4A. *Nat Struct Mol Biol*.
- . 2024b. The structure of a human translation initiation complex reveals two independent roles for the helicase eIF4A. *Nat Struct Mol Biol* **31**: 455-464.
- Brito Querido J, Sokabe M, Kraatz S, Gordiyenko Y, Skehel JM, Fraser CS, Ramakrishnan V. 2020. Structure of a human 48S translational initiation complex. *Science* **369**: 1220-1227.
- Bujanic L, Shevchuk O, von Kügelgen N, Kalinina A, Ludwik K, Koppstein D, Zerna N, Sickmann A, Chekulaeva M. 2022. The key features of SARS-CoV-2 leader and NSP1 required for viral escape of NSP1-mediated repression. *Rna* **28**: 766-779.
- Burd CG, Matunis EL, Dreyfuss G. 1991. The multiple RNA-binding domains of the mRNA poly(A)-binding protein have different RNA-binding activities. *Mol Cell Biol* **11**: 3419-3424.
- Butcher SE, Jan E. 2016. tRNA-mimicry in IRES-mediated translation and recoding. *RNA Biol* **13**: 1068-1074.
- Caliskan N, Peske F, Rodnina MV. 2015. Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. *Trends Biochem Sci* **40**: 265-274.
- Calviello L, Venkataramanan S, Rogowski KJ, Wyler E, Wilkins K, Tejura M, Thai B, Krol J, Filipowicz W, Landthaler M et al. 2021. DDX3 depletion represses translation of mRNAs with complex 5' UTRs. *Nucleic Acids Res* **49**: 5336-5350.
- Calvo SE, Pagliarini DJ, Mootha VK. 2009. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci U S A* **106**: 7507-7512.
- Cao C, Cai Z, Xiao X, Rao J, Chen J, Hu N, Yang M, Xing X, Wang Y, Li M et al. 2021. The architecture of the SARS-CoV-2 RNA genome inside virion. *Nat Commun* **12**: 3917.
- Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, Curk T, Krijgsveld J, Hentze MW. 2016. Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol Cell* **63**: 696-710.
- Castelló A, Izquierdo JM, Welnowska E, Carrasco L. 2009a. RNA nuclear export is blocked by poliovirus 2A protease and is concomitant with nucleoporin cleavage. *J Cell Sci* **122**: 3799-3809.
- Castelló A, Quintas A, Sánchez EG, Sabina P, Nogal M, Carrasco L, Revilla Y. 2009b. Regulation of host translational machinery by African swine fever virus. *PLoS Pathog* **5**: e1000562.
- Chappell SA, Edelman GM, Mauro VP. 2000. A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. *Proc Natl Acad Sci U S A* **97**: 1536-1541.
- Chaudhuri J, Chowdhury D, Maitra U. 1999. Distinct functions of eukaryotic translation initiation factors eIF1A and eIF3 in the formation of the 40 S ribosomal preinitiation complex. *J Biol Chem* **274**: 17975-17980.
- Chen CY, Chen ST, Juan HF, Huang HC. 2012. Lengthening of 3'UTR increases with morphological complexity in animal evolution. *Bioinformatics* **28**: 3178-3181.

- Chen CY, Sarnow P. 1995. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* **268**: 415-417.
- Chen J, Brunner AD, Cogan JZ, Nuñez JK, Fields AP, Adamson B, Itzhak DN, Li JY, Mann M, Leonetti MD et al. 2020. Pervasive functional translation of noncanonical human open reading frames. *Science* **367**: 1140-1146.
- Chen J, Tresenrider A, Chia M, McSwiggen DT, Spedale G, Jorgensen V, Liao H, van Werven FJ, Ünal E. 2017. Kinetochores inactivated by expression of a repressive mRNA. *Elife* **6**.
- Chen SC, Olsthoorn RC. 2010. Group-specific structural features of the 5'-proximal sequences of coronavirus genomic RNAs. *Virology* **401**: 29-41.
- Cheng Z, Otto GM, Powers EN, Keskin A, Mertins P, Carr SA, Jovanovic M, Brar GA. 2018. Pervasive, Coordinated Protein-Level Changes Driven by Transcript Isoform Switching during Meiosis. *Cell* **172**: 910-923.e916.
- Choe J, Lin S, Zhang W, Liu Q, Wang L, Ramirez-Moya J, Du P, Kim W, Tang S, Sliz P et al. 2018. mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* **561**: 556-560.
- Choi J, Jeong KW, Demirci H, Chen J, Petrov A, Prabhakar A, O'Leary SE, Dominissini D, Rechavi G, Soltis SM et al. 2016. N(6)-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat Struct Mol Biol* **23**: 110-115.
- Clark LK, Green TJ, Petit CM. 2021. Structure of Nonstructural Protein 1 from SARS-CoV-2. *J Virol* **95**.
- Clark ME, Lieberman PM, Berk AJ, Dasgupta A. 1993. Direct cleavage of human TATA-binding protein by poliovirus protease 3C in vivo and in vitro. *Mol Cell Biol* **13**: 1232-1237.
- Clements JM, Laz TM, Sherman F. 1988. Efficiency of translation initiation by non-AUG codons in *Saccharomyces cerevisiae*. *Mol Cell Biol* **8**: 4533-4536.
- Cohen-Chalamish S, Hasson A, Weinberg D, Namer LS, Banai Y, Osman F, Kaempfer R. 2009. Dynamic refolding of IFN-gamma mRNA enables it to function as PKR activator and translation template. *Nat Chem Biol* **5**: 896-903.
- Costantino DA, Pflingsten JS, Rambo RP, Kieft JS. 2008. tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat Struct Mol Biol* **15**: 57-64.
- Covarrubias S, Richner JM, Clyde K, Lee YJ, Glaunsinger BA. 2009. Host shutoff is a conserved phenotype of gammaherpesvirus infection and is orchestrated exclusively from the cytoplasm. *J Virol* **83**: 9554-9566.
- Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, Joung J, Zhang F. 2017. RNA editing with CRISPR-Cas13. *Science* **358**: 1019-1027.
- Cramer P. 2019. Organization and regulation of gene transcription. *Nature* **573**: 45-54.
- Cuesta R, Xi Q, Schneider RJ. 2004. Structural basis for competitive inhibition of eIF4G-Mnk1 interaction by the adenovirus 100-kilodalton protein. *J Virol* **78**: 7707-7716.
- Damoc E, Fraser CS, Zhou M, Videler H, Mayeur GL, Hershey JW, Doudna JA, Robinson CV, Leary JA. 2007. Structural characterization of the human eukaryotic initiation factor 3 protein complex by mass spectrometry. *Mol Cell Proteomics* **6**: 1135-1146.
- Dasso MC, Jackson RJ. 1989. Efficient initiation of mammalian mRNA translation at a CUG codon. *Nucleic Acids Res* **17**: 6485-6497.

- Dauber B, Poon D, Dos Santos T, Duguay BA, Mehta N, Saffran HA, Smiley JR. 2016. The Herpes Simplex Virus Virion Host Shutoff Protein Enhances Translation of Viral True Late mRNAs Independently of Suppressing Protein Kinase R and Stress Granule Formation. *J Virol* **90**: 6049-6057.
- Dauber B, Saffran HA, Smiley JR. 2019. The herpes simplex virus host shutoff (vhs) RNase limits accumulation of double stranded RNA in infected cells: Evidence for accelerated decay of duplex RNA. *PLoS Pathog* **15**: e1008111.
- Davies JP, Sivadas A, Keller KR, Roman BK, Wojcikiewicz RJH, Plate L. 2024. Expression of SARS-CoV-2 Nonstructural Proteins 3 and 4 Can Tune the Unfolded Protein Response in Cell Culture. *Journal of Proteome Research* **23**: 356-367.
- de Breyne S, Bonderoff JM, Chumakov KM, Lloyd RE, Hellen CU. 2008a. Cleavage of eukaryotic initiation factor eIF5B by enterovirus 3C proteases. *Virology* **378**: 118-122.
- de Breyne S, Simonet V, Pelet T, Curran J. 2003. Identification of a cis-acting element required for shunt-mediated translational initiation of the Sendai virus Y proteins. *Nucleic Acids Res* **31**: 608-618.
- de Breyne S, Vindry C, Guillin O, Condé L, Mure F, Gruffat H, Chavatte L, Ohlmann T. 2020. Translational control of coronaviruses. *Nucleic Acids Res* **48**: 12502-12522.
- de Breyne S, Yu Y, Pestova TV, Hellen CU. 2008b. Factor requirements for translation initiation on the Simian picornavirus internal ribosomal entry site. *Rna* **14**: 367-380.
- de Breyne S, Yu Y, Unbehaun A, Pestova TV, Hellen CU. 2009. Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proc Natl Acad Sci U S A* **106**: 9197-9202.
- Decroly E, Ferron F, Lescar J, Canard B. 2012. Conventional and unconventional mechanisms for capping viral mRNA. *Nature Reviews Microbiology* **10**: 51-65.
- Deforges J, de Breyne S, Ameer M, Ulryck N, Chamond N, Saaidi A, Ponty Y, Ohlmann T, Sargueil B. 2017. Two ribosome recruitment sites direct multiple translation events within HIV1 Gag open reading frame. *Nucleic Acids Res* **45**: 7382-7400.
- Deng X, Baker SC. 2018. An "Old" protein with a new story: Coronavirus endoribonuclease is important for evading host antiviral defenses. *Virology* **517**: 157-163.
- Deniz N, Lenarcic EM, Landry DM, Thompson SR. 2009. Translation initiation factors are not required for Dicistroviridae IRES function in vivo. *Rna* **15**: 932-946.
- des Georges A, Dhote V, Kuhn L, Hellen CU, Pestova TV, Frank J, Hashem Y. 2015. Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* **525**: 491-495.
- Dever TE, Ivanov IP, Sachs MS. 2020. Conserved Upstream Open Reading Frame Nascent Peptides That Control Translation. *Annu Rev Genet* **54**: 237-264.
- Di Timoteo G, Dattilo D, Centrón-Broco A, Colantoni A, Guarnacci M, Rossi F, Incarnato D, Oliviero S, Fatica A, Morlando M et al. 2020. Modulation of circRNA Metabolism by m(6)A Modification. *Cell Rep* **31**: 107641.
- Diaz de Arce AJ, Noderer WL, Wang CL. 2018. Complete motif analysis of sequence requirements for translation initiation at non-AUG start codons. *Nucleic Acids Res* **46**: 985-994.

- Dmitriev SE, Andreev DE, Terenin IM, Olovnikov IA, Prassolov VS, Merrick WC, Shatsky IN. 2007. Efficient translation initiation directed by the 900-nucleotide-long and GC-rich 5' untranslated region of the human retrotransposon LINE-1 mRNA is strictly cap dependent rather than internal ribosome entry site mediated. *Mol Cell Biol* **27**: 4685-4697.
- Dmitriev SE, Terenin IM, Andreev DE, Ivanov PA, Dunaevsky JE, Merrick WC, Shatsky IN. 2010. GTP-independent tRNA delivery to the ribosomal P-site by a novel eukaryotic translation factor. *J Biol Chem* **285**: 26779-26787.
- Drabkin HJ, Helk B, RajBhandary UL. 1993. The role of nucleotides conserved in eukaryotic initiator methionine tRNAs in initiation of protein synthesis. *J Biol Chem* **268**: 25221-25228.
- Duncan R, Hershey JW. 1983. Identification and quantitation of levels of protein synthesis initiation factors in crude HeLa cell lysates by two-dimensional polyacrylamide gel electrophoresis. *J Biol Chem* **258**: 7228-7235.
- Dunn EF, Connor JH. 2011. Dominant inhibition of Akt/protein kinase B signaling by the matrix protein of a negative-strand RNA virus. *J Virol* **85**: 422-431.
- Elfakess R, Dikstein R. 2008. A translation initiation element specific to mRNAs with very short 5'UTR that also regulates transcription. *PLoS One* **3**: e3094.
- Elfakess R, Sinvani H, Haimov O, Svitkin Y, Sonenberg N, Dikstein R. 2011. Unique translation initiation of mRNAs-containing TISU element. *Nucleic Acids Res* **39**: 7598-7609.
- Elgadi MM, Hayes CE, Smiley JR. 1999. The herpes simplex virus vhs protein induces endoribonucleolytic cleavage of target RNAs in cell extracts. *J Virol* **73**: 7153-7164.
- Erickson FL, Hannig EM. 1996. Ligand interactions with eukaryotic translation initiation factor 2: role of the gamma-subunit. *Embo j* **15**: 6311-6320.
- Erzberger JP, Stengel F, Pellarin R, Zhang S, Schaefer T, Aylett CHS, Cimermančič P, Boehringer D, Sali A, Aebersold R et al. 2014. Molecular architecture of the 40S-eIF1-eIF3 translation initiation complex. *Cell* **158**: 1123-1135.
- Etchison D, Milburn SC, Edery I, Sonenberg N, Hershey JW. 1982. Inhibition of HeLa cell protein synthesis following poliovirus infection correlates with the proteolysis of a 220,000-dalton polypeptide associated with eucaryotic initiation factor 3 and a cap binding protein complex. *J Biol Chem* **257**: 14806-14810.
- Fadden P, Haystead TA, Lawrence JC, Jr. 1997. Identification of phosphorylation sites in the translational regulator, PHAS-I, that are controlled by insulin and rapamycin in rat adipocytes. *J Biol Chem* **272**: 10240-10247.
- Farruggio D, Chaudhuri J, Maitra U, RajBhandary UL. 1996. The A1 x U72 base pair conserved in eukaryotic initiator tRNAs is important specifically for binding to the eukaryotic translation initiation factor eIF2. *Mol Cell Biol* **16**: 4248-4256.
- Fedorova AD, Kiniry SJ, Andreev DE, Mudge JM, Baranov PV. 2022. Thousands of human non-AUG extended proteoforms lack evidence of evolutionary selection among mammals. *Nature Communications* **13**: 7910.
- Fekete CA, Applefield DJ, Blakely SA, Shirokikh N, Pestova T, Lorsch JR, Hinnebusch AG. 2005. The eIF1A C-terminal domain promotes initiation complex assembly, scanning and AUG selection in vivo. *Embo j* **24**: 3588-3601.

- Fenwick ML, McMenamin MM. 1984. Early virion-associated suppression of cellular protein synthesis by herpes simplex virus is accompanied by inactivation of mRNA. *J Gen Virol* **65 (Pt 7)**: 1225-1228.
- Fernández IS, Bai XC, Murshudov G, Scheres SH, Ramakrishnan V. 2014. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* **157**: 823-831.
- Fernandez J, Yaman I, Sarnow P, Snider MD, Hatzoglou M. 2002. Regulation of internal ribosomal entry site-mediated translation by phosphorylation of the translation initiation factor eIF2alpha. *J Biol Chem* **277**: 19198-19205.
- Filbin ME, Kieft JS. 2009. Toward a structural understanding of IRES RNA function. *Curr Opin Struct Biol* **19**: 267-276.
- Finkel Y, Gluck A, Nachshon A, Winkler R, Fisher T, Rozman B, Mizrahi O, Lubelsky Y, Zuckerman B, Slobodin B et al. 2021a. SARS-CoV-2 uses a multipronged strategy to impede host protein synthesis. *Nature* **594**: 240-245.
- Finkel Y, Mizrahi O, Nachshon A, Weingarten-Gabbay S, Morgenstern D, Yahalom-Ronen Y, Tamir H, Achdout H, Stein D, Israeli O et al. 2021b. The coding capacity of SARS-CoV-2. *Nature* **589**: 125-130.
- Finnen RL, Hay TJ, Dauber B, Smiley JR, Banfield BW. 2014. The herpes simplex virus 2 virion-associated ribonuclease vhs interferes with stress granule formation. *J Virol* **88**: 12727-12739.
- Fisher T, Gluck A, Narayanan K, Kuroda M, Nachshon A, Hsu JC, Halfmann PJ, Yahalom-Ronen Y, Tamir H, Finkel Y et al. 2022. Parsing the role of NSP1 in SARS-CoV-2 infection. *Cell Rep* **39**: 110954.
- Fletcher SP, Jackson RJ. 2002. Pestivirus internal ribosome entry site (IRES) structure and function: elements in the 5' untranslated region important for IRES function. *J Virol* **76**: 5024-5033.
- Fraser BJ, Beldar S, Seitova A, Hutchinson A, Mannar D, Li Y, Kwon D, Tan R, Wilson RP, Leopold K et al. 2022a. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nat Chem Biol* **18**: 963-971.
- Fraser BJ, Beldar S, Seitova A, Hutchinson A, Mannar D, Li Y, Kwon D, Tan R, Wilson RP, Leopold K et al. 2022b. Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nature Chemical Biology* **18**: 963-971.
- Galicia-Vázquez G, Cencic R, Robert F, Agenor AQ, Pelletier J. 2012. A cellular response linking eIF4A1 activity to eIF4A1 transcription. *Rna* **18**: 1373-1384.
- Gallie DR. 1991. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev* **5**: 2108-2116.
- Gao Y, Sun SQ, Guo HC. 2016. Biological function of Foot-and-mouth disease virus non-structural proteins and non-coding elements. *Viol J* **13**: 107.
- García MA, Gil J, Ventoso I, Guerra S, Domingo E, Rivas C, Esteban M. 2006. Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action. *Microbiol Mol Biol Rev* **70**: 1032-1060.
- García-García C, Frieda KL, Feoktistova K, Fraser CS, Block SM. 2015. RNA BIOCHEMISTRY. Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science* **348**: 1486-1488.
- Gaucherand L, Porter BK, Levene RE, Price EL, Schmaling SK, Rycroft CH, Kevorkian Y, McCormick C, Khapersky DA, Gaglia MM. 2019. The Influenza A Virus

- Endoribonuclease PA-X Usurps Host mRNA Processing Machinery to Limit Host Gene Expression. *Cell Rep* **27**: 776-792.e777.
- Gaudin C, Mazauric MH, Traïkia M, Guittet E, Yoshizawa S, Fourmy D. 2005. Structure of the RNA signal essential for translational frameshifting in HIV-1. *J Mol Biol* **349**: 1024-1035.
- Gendron K, Ferbeyre G, Heveker N, Brakier-Gingras L. 2011. The activity of the HIV-1 IRES is stimulated by oxidative stress and controlled by a negative regulatory element. *Nucleic Acids Res* **39**: 902-912.
- Gerashchenko MV, Gladyshev VN. 2014. Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res* **42**: e134.
- Gingras AC, Kennedy SG, O'Leary MA, Sonenberg N, Hay N. 1998. 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. *Genes Dev* **12**: 502-513.
- Gingras AC, Raught B, Gygi SP, Niedzwiecka A, Miron M, Burley SK, Polakiewicz RD, Wyslouch-Cieszynska A, Aebersold R, Sonenberg N. 2001. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. *Genes Dev* **15**: 2852-2864.
- Gingras AC, Raught B, Sonenberg N. 1999. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem* **68**: 913-963.
- Gingras AC, Svitkin Y, Belsham GJ, Pause A, Sonenberg N. 1996. Activation of the translational suppressor 4E-BP1 following infection with encephalomyocarditis virus and poliovirus. *Proc Natl Acad Sci U S A* **93**: 5578-5583.
- Gordon DE Jang GM Bouhaddou M Xu J Obernier K White KM O'Meara MJ Rezelj VV Guo JZ Swaney DL et al. 2020. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**: 459-468.
- Gorgoni B, Gray NK. 2004. The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: a developmental perspective. *Brief Funct Genomic Proteomic* **3**: 125-141.
- Görlach M, Burd CG, Dreyfuss G. 1994. The mRNA poly(A)-binding protein: localization, abundance, and RNA-binding specificity. *Exp Cell Res* **211**: 400-407.
- Gradi A, Imataka H, Svitkin YV, Rom E, Raught B, Morino S, Sonenberg N. 1998a. A novel functional human eukaryotic translation initiation factor 4G. *Mol Cell Biol* **18**: 334-342.
- Gradi A, Svitkin YV, Imataka H, Sonenberg N. 1998b. Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. *Proc Natl Acad Sci U S A* **95**: 11089-11094.
- Gray NK, Hentze MW. 1994. Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs. *Embo j* **13**: 3882-3891.
- Grayzel AI, Hörchner P, London IM. 1966. The stimulation of globin synthesis by heme. *Proc Natl Acad Sci U S A* **55**: 650-655.
- Grifo JA, Abramson RD, Satler CA, Merrick WC. 1984. RNA-stimulated ATPase activity of eukaryotic initiation factors. *J Biol Chem* **259**: 8648-8654.
- Grüner S, Peter D, Weber R, Wohlbold L, Chung MY, Weichenrieder O, Valkov E, Igreja C, Izaurralde E. 2016. The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation. *Mol Cell* **64**: 467-479.

- Gu Y, Mao Y, Jia L, Dong L, Qian SB. 2021. Bi-directional ribosome scanning controls the stringency of start codon selection. *Nat Commun* **12**: 6604.
- Guttman M, Russell P, Ingolia NT, Weissman JS, Lander ES. 2013. Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell* **154**: 240-251.
- Haghighat A, Mader S, Pause A, Sonenberg N. 1995. Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. *Embo j* **14**: 5701-5709.
- Haghighat A, Sonenberg N. 1997. eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. *J Biol Chem* **272**: 21677-21680.
- Haghighat A, Svitkin Y, Novoa I, Kuechler E, Skern T, Sonenberg N. 1996. The eIF4G-eIF4E complex is the target for direct cleavage by the rhinovirus 2A proteinase. *J Virol* **70**: 8444-8450.
- Hakki M, Marshall EE, De Niro KL, Geballe AP. 2006. Binding and nuclear relocalization of protein kinase R by human cytomegalovirus TRS1. *J Virol* **80**: 11817-11826.
- Halder K, Wieland M, Hartig JS. 2009. Predictable suppression of gene expression by 5'-UTR-based RNA quadruplexes. *Nucleic Acids Res* **37**: 6811-6817.
- Haller AA, Nguyen JH, Semler BL. 1993. Minimum internal ribosome entry site required for poliovirus infectivity. *J Virol* **67**: 7461-7471.
- Hann SR, Dixit M, Sears RC, Sealy L. 1994. The alternatively initiated c-Myc proteins differentially regulate transcription through a noncanonical DNA-binding site. *Genes Dev* **8**: 2441-2452.
- Hann SR, King MW, Bentley DL, Anderson CW, Eisenman RN. 1988. A non-AUG translational initiation in c-myc exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. *Cell* **52**: 185-195.
- Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M, Ron D. 2000. Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol Cell* **6**: 1099-1108.
- Hashem Y, des Georges A, Dhote V, Langlois R, Liao HY, Grassucci RA, Hellen CU, Pestova TV, Frank J. 2013a. Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153**: 1108-1119.
- Hashem Y, des Georges A, Dhote V, Langlois R, Liao HY, Grassucci RA, Pestova TV, Hellen CU, Frank J. 2013b. Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **503**: 539-543.
- Hashem Y, Frank J. 2018. The Jigsaw Puzzle of mRNA Translation Initiation in Eukaryotes: A Decade of Structures Unraveling the Mechanics of the Process. *Annual Review of Biophysics* **47**: 125-151.
- Hayashi T, Chaimayo C, McGuinness J, Takimoto T. 2016. Critical Role of the PA-X C-Terminal Domain of Influenza A Virus in Its Subcellular Localization and Shutoff Activity. *J Virol* **90**: 7131-7141.
- Hellen CU, de Breyne S. 2007. A distinct group of hepacivirus/pestivirus-like internal ribosomal entry sites in members of diverse picornavirus genera: evidence for modular exchange of functional noncoding RNA elements by recombination. *J Virol* **81**: 5850-5863.
- Hellen CU, Sarnow P. 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev* **15**: 1593-1612.

- Hemmings-Mieszczak M, Hohn T, Preiss T. 2000. Termination and peptide release at the upstream open reading frame are required for downstream translation on synthetic shunt-competent mRNA leaders. *Mol Cell Biol* **20**: 6212-6223.
- Hentze MW, Caughman SW, Rouault TA, Barriocanal JG, Dancis A, Harford JB, Klausner RD. 1987. Identification of the iron-responsive element for the translational regulation of human ferritin mRNA. *Science* **238**: 1570-1573.
- Hernández G, Altmann M, Lasko P. 2010. Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. *Trends Biochem Sci* **35**: 63-73.
- Hershey JW, Sonenberg N, Mathews MB. 2012. Principles of translational control: an overview. *Cold Spring Harb Perspect Biol* **4**.
- Hikmet F, Méar L, Edvinsson Å, Micke P, Uhlén M, Lindskog C. 2020. The protein expression profile of ACE2 in human tissues. *Mol Syst Biol* **16**: e9610.
- Hinnebusch AG. 2005. Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* **59**: 407-450.
- . 2006. eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem Sci* **31**: 553-562.
- . 2017. Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem Sci* **42**: 589-611.
- Hinnebusch AG, Ivanov IP, Sonenberg N. 2016. Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science* **352**: 1413-1416.
- Ho BC, Yu SL, Chen JJ, Chang SY, Yan BS, Hong QS, Singh S, Kao CL, Chen HY, Su KY et al. 2011. Enterovirus-induced miR-141 contributes to shutoff of host protein translation by targeting the translation initiation factor eIF4E. *Cell Host Microbe* **9**: 58-69.
- Ho JSY, Zhu Z, Marazzi I. 2021. Unconventional viral gene expression mechanisms as therapeutic targets. *Nature* **593**: 362-371.
- Hocine S, Singer RH, Grünwald D. 2010. RNA processing and export. *Cold Spring Harb Perspect Biol* **2**: a000752.
- Hopkins BD, Fine B, Steinbach N, Dendy M, Rapp Z, Shaw J, Pappas K, Yu JS, Hodakoski C, Mense S et al. 2013. A secreted PTEN phosphatase that enters cells to alter signaling and survival. *Science* **341**: 399-402.
- Howe JG, Shu MD. 1989. Epstein-Barr virus small RNA (EBER) genes: unique transcription units that combine RNA polymerase II and III promoter elements. *Cell* **57**: 825-834.
- Hronová V, Mohammad MP, Wagner S, Pánek J, Gunišová S, Zeman J, Poncová K, Valášek LS. 2017. Does eIF3 promote reinitiation after translation of short upstream ORFs also in mammalian cells? *RNA Biol* **14**: 1660-1667.
- Hu B, Guo H, Zhou P, Shi ZL. 2021. Characteristics of SARS-CoV-2 and COVID-19. *Nat Rev Microbiol* **19**: 141-154.
- Hung M, Patel P, Davis S, Green SR. 1998. Importance of ribosomal frameshifting for human immunodeficiency virus type 1 particle assembly and replication. *J Virol* **72**: 4819-4824.
- Hung V, Udeshi ND, Lam SS, Loh KH, Cox KJ, Pedram K, Carr SA, Ting AY. 2016. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat Protoc* **11**: 456-475.
- Hussain T, Llácer JL, Fernández IS, Munoz A, Martin-Marcos P, Savva CG, Lorsch JR, Hinnebusch AG, Ramakrishnan V. 2014. Structural changes enable start

- codon recognition by the eukaryotic translation initiation complex. *Cell* **159**: 597-607.
- Huston NC, Wan H, Strine MS, de Cesaris Araujo Tavares R, Wilen CB, Pyle AM. 2021. Comprehensive in vivo secondary structure of the SARS-CoV-2 genome reveals novel regulatory motifs and mechanisms. *Mol Cell* **81**: 584-598.e585.
- Ichihashi N, Matsuura T, Kita H, Hosoda K, Sunami T, Tsukada K, Yomo T. 2008. Importance of translation-replication balance for efficient replication by the self-encoded replicase. *ChemBiochem* **9**: 3023-3028.
- leong KW, Indrisiunaite G, Prabhakar A, Puglisi JD, Ehrenberg M. 2021. N 6-Methyladenosines in mRNAs reduce the accuracy of codon reading by transfer RNAs and peptide release factors. *Nucleic Acids Res* **49**: 2684-2699.
- Imataka H, Gradi A, Sonenberg N. 1998. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *Embo j* **17**: 7480-7489.
- Imataka H, Olsen HS, Sonenberg N. 1997. A new translational regulator with homology to eukaryotic translation initiation factor 4G. *Embo j* **16**: 817-825.
- Imataka H, Sonenberg N. 1997. Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol Cell Biol* **17**: 6940-6947.
- Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**: 218-223.
- Ingolia NT, Lareau LF, Weissman JS. 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**: 789-802.
- Ishimaru D, Plant EP, Sims AC, Yount BL, Jr., Roth BM, Eldho NV, Pérez-Alvarado GC, Armbruster DW, Baric RS, Dinman JD et al. 2013. RNA dimerization plays a role in ribosomal frameshifting of the SARS coronavirus. *Nucleic Acids Res* **41**: 2594-2608.
- Ivanov IP, Firth AE, Michel AM, Atkins JF, Baranov PV. 2011. Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res* **39**: 4220-4234.
- Ivanov IP, Loughran G, Atkins JF. 2008. uORFs with unusual translational start codons autoregulate expression of eukaryotic ornithine decarboxylase homologs. *Proc Natl Acad Sci U S A* **105**: 10079-10084.
- Ivanov IP, Loughran G, Sachs MS, Atkins JF. 2010. Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). *Proc Natl Acad Sci U S A* **107**: 18056-18060.
- Ivanov P, Kedersha N, Anderson P. 2019. Stress Granules and Processing Bodies in Translational Control. *Cold Spring Harb Perspect Biol* **11**.
- Iwakiri D, Zhou L, Samanta M, Matsumoto M, Ebihara T, Seya T, Imai S, Fujieda M, Kawa K, Takada K. 2009. Epstein-Barr virus (EBV)-encoded small RNA is released from EBV-infected cells and activates signaling from Toll-like receptor 3. *J Exp Med* **206**: 2091-2099.
- Izidorio MS, Sokabe M, Villa N, Merrick WC, Fraser CS. 2022. Human eukaryotic initiation factor 4E (eIF4E) and the nucleotide-bound state of eIF4A regulate eIF4F binding to RNA. *J Biol Chem* **298**: 102368.

- Jaafar ZA, Kieft JS. 2019. Viral RNA structure-based strategies to manipulate translation. *Nature Reviews Microbiology* **17**: 110-123.
- Jaafar ZA, Oguro A, Nakamura Y, Kieft JS. 2016. Translation initiation by the hepatitis C virus IRES requires eIF1A and ribosomal complex remodeling. *Elife* **5**.
- Jackson CB, Farzan M, Chen B, Choe H. 2022. Mechanisms of SARS-CoV-2 entry into cells. *Nature Reviews Molecular Cell Biology* **23**: 3-20.
- Jackson RJ, Howell MT, Kaminski A. 1990. The novel mechanism of initiation of picornavirus RNA translation. *Trends Biochem Sci* **15**: 477-483.
- Jäger S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, Shales M, Mercenne G, Pache L, Li K et al. 2011. Global landscape of HIV-human protein complexes. *Nature* **481**: 365-370.
- Jagger BW, Wise HM, Kash JC, Walters KA, Wills NM, Xiao YL, Dunfee RL, Schwartzman LM, Ozinsky A, Bell GL et al. 2012. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**: 199-204.
- Jain S, Koziej L, Poulis P, Kaczmarczyk I, Gaik M, Rawski M, Ranjan N, Glatt S, Rodnina MV. 2023. Modulation of translational decoding by m6A modification of mRNA. *Nature Communications* **14**: 4784.
- Jang SK, Kräusslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E. 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol* **62**: 2636-2643.
- Jang SK, Pestova TV, Hellen CU, Witherell GW, Wimmer E. 1990. Cap-independent translation of picornavirus RNAs: structure and function of the internal ribosomal entry site. *Enzyme* **44**: 292-309.
- Jennings MD, Kershaw CJ, White C, Hoyle D, Richardson JP, Costello JL, Donaldson IJ, Zhou Y, Pavitt GD. 2016. eIF2 β is critical for eIF5-mediated GDP-dissociation inhibitor activity and translational control. *Nucleic Acids Res* **44**: 9698-9709.
- Ji Z, Song R, Regev A, Struhl K. 2015. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *Elife* **4**: e08890.
- Jia X, He X, Huang C, Li J, Dong Z, Liu K. 2024. Protein translation: biological processes and therapeutic strategies for human diseases. *Signal Transduction and Targeted Therapy* **9**: 44.
- Jiang X, Liu B, Nie Z, Duan L, Xiong Q, Jin Z, Yang C, Chen Y. 2021. The role of m6A modification in the biological functions and diseases. *Signal Transduction and Targeted Therapy* **6**: 74.
- Johnstone TG, Bazzini AA, Giraldez AJ. 2016. Upstream ORFs are prevalent translational repressors in vertebrates. *Embo j* **35**: 706-723.
- Jones JE, Le Sage V, Lakdawala SS. 2021. Viral and host heterogeneity and their effects on the viral life cycle. *Nature Reviews Microbiology* **19**: 272-282.
- Kahvejian A, Svitkin YV, Sukarieh R, M'Boutchou MN, Sonenberg N. 2005. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev* **19**: 104-113.
- Kaminski A, Hunt SL, Gibbs CL, Jackson RJ. 1994. Internal initiation of mRNA translation in eukaryotes. *Genet Eng (N Y)* **16**: 115-155.
- Kapp LD, Lorsch JR. 2004. GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. *J Mol Biol* **335**: 923-936.

- Katahira J, Ohmae T, Yasugi M, Sasaki R, Itoh Y, Kohda T, Hieda M, Yokota Hirai M, Okamoto T, Miyamoto Y. 2023. Nsp14 of SARS-CoV-2 inhibits mRNA processing and nuclear export by targeting the nuclear cap-binding complex. *Nucleic Acids Res* **51**: 7602-7618.
- Katsafanas GC, Moss B. 2007. Colocalization of transcription and translation within cytoplasmic poxvirus factories coordinates viral expression and subjugates host functions. *Cell Host Microbe* **2**: 221-228.
- Kearse MG, Goldman DH, Choi J, Nwaezeapu C, Liang D, Green KM, Goldstrohm AC, Todd PK, Green R, Wilusz JE. 2019. Ribosome queuing enables non-AUG translation to be resistant to multiple protein synthesis inhibitors. *Genes Dev* **33**: 871-885.
- Kearse MG, Wilusz JE. 2017. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev* **31**: 1717-1731.
- Khaleghpour K, Pyronnet S, Gingras AC, Sonenberg N. 1999. Translational homeostasis: eukaryotic translation initiation factor 4E control of 4E-binding protein 1 and p70 S6 kinase activities. *Mol Cell Biol* **19**: 4302-4310.
- Khan D, Terenzi F, Liu G, Ghosh PK, Ye F, Nguyen K, China A, Ramachandiran I, Chakraborty S, Stefan J et al. 2023. A viral pan-end RNA element and host complex define a SARS-CoV-2 regulon. *Nature Communications* **14**: 3385.
- Khaperskyy DA, Schmaling S, Larkins-Ford J, McCormick C, Gaglia MM. 2016. Selective Degradation of Host RNA Polymerase II Transcripts by Influenza A Virus PA-X Host Shutoff Protein. *PLoS Pathog* **12**: e1005427.
- Khawaja A, Vopalensky V, Pospisek M. 2015. Understanding the potential of hepatitis C virus internal ribosome entry site domains to modulate translation initiation via their structure and function. *Wiley Interdiscip Rev RNA* **6**: 211-224.
- Khong A, Parker R. 2018. mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction. *J Cell Biol* **217**: 4124-4140.
- Kieft JS, Zhou K, Jubin R, Doudna JA. 2001. Mechanism of ribosome recruitment by hepatitis C IRES RNA. *Rna* **7**: 194-206.
- Kim D, Kim S, Park J, Chang HR, Chang J, Ahn J, Park H, Park J, Son N, Kang G et al. 2021. A high-resolution temporal atlas of the SARS-CoV-2 transcriptome and transcriptome. *Nature Communications* **12**: 5120.
- Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. 2020a. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **181**: 914-921.e910.
- Kim JH, Park SM, Park JH, Keum SJ, Jang SK. 2011. eIF2A mediates translation of hepatitis C viral mRNA under stress conditions. *Embo j* **30**: 2454-2464.
- Kim JM, Chung YS, Jo HJ, Lee NJ, Kim MS, Woo SH, Park S, Kim JW, Kim HM, Han MG. 2020b. Identification of Coronavirus Isolated from a Patient in Korea with COVID-19. *Osong Public Health Res Perspect* **11**: 3-7.
- Kim YJ, Moon S, Hwang MY, Han S, Jang H-M, Kong J, Shin DM, Yoon K, Kim SM, Lee J-E et al. 2022. The contribution of common and rare genetic variants to variation in metabolic traits in 288,137 East Asians. *Nature Communications* **13**: 6642.
- Kindler E, Gil-Cruz C, Spanier J, Li Y, Wilhelm J, Rabouw HH, Züst R, Hwang M, V'Kovski P, Stalder H et al. 2017. Early endonuclease-mediated evasion of RNA sensing ensures efficient coronavirus replication. *PLoS Pathog* **13**: e1006195.
- Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, Stanifer ML, Boulant S, Bartenschlager R, Chlanda P. 2020. SARS-CoV-2 structure and

- replication characterized by in situ cryo-electron tomography. *Nature Communications* **11**: 5885.
- Kolitz SE, Lorsch JR. 2010. Eukaryotic initiator tRNA: finely tuned and ready for action. *FEBS Lett* **584**: 396-404.
- Kolupaeva VG, Pestova TV, Hellen CU, Shatsky IN. 1998. Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. *J Biol Chem* **273**: 18599-18604.
- Kopeina GS, Afonina ZA, Gromova KV, Shirokov VA, Vasiliev VD, Spirin AS. 2008. Step-wise formation of eukaryotic double-row polyribosomes and circular translation of polysomal mRNA. *Nucleic Acids Res* **36**: 2476-2488.
- Korneeva NL, First EA, Benoit CA, Rhoads RE. 2005. Interaction between the NH₂-terminal domain of eIF4A and the central domain of eIF4G modulates RNA-stimulated ATPase activity. *J Biol Chem* **280**: 1872-1881.
- Korneeva NL, Lamphear BJ, Hennigan FL, Rhoads RE. 2000. Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. *J Biol Chem* **275**: 41369-41376.
- Koudelka T, Boger J, Henkel A, Schönherr R, Krantz S, Fuchs S, Rodríguez E, Redecke L, Tholey A. 2021. N-Terminomics for the Identification of In Vitro Substrates and Cleavage Site Specificity of the SARS-CoV-2 Main Protease. *Proteomics* **21**: e2000246.
- Kozak M. 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res* **9**: 5233-5252.
- . 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *Proc Natl Acad Sci U S A* **83**: 2850-2854.
 - . 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* **15**: 8125-8148.
 - . 1989a. Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol Cell Biol* **9**: 5073-5080.
 - . 1989b. The scanning model for translation: an update. *J Cell Biol* **108**: 229-241.
 - . 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc Natl Acad Sci U S A* **87**: 8301-8305.
 - . 2001. Constraints on reinitiation of translation in mammals. *Nucleic Acids Res* **29**: 5226-5232.
- Kozak M, Shatkin AJ. 1978. Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. *J Biol Chem* **253**: 6568-6577.
- Krischuns T, Lukarska M, Naffakh N, Cusack S. 2021. Influenza Virus RNA-Dependent RNA Polymerase and the Host Transcriptional Apparatus. *Annu Rev Biochem* **90**: 321-348.
- Kühn R, Luz N, Beck E. 1990. Functional analysis of the internal translation initiation site of foot-and-mouth disease virus. *J Virol* **64**: 4625-4631.
- Kulak NA, Pichler G, Paron I, Nagaraj N, Mann M. 2014. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* **11**: 319-324.
- Kumar P, Hellen CU, Pestova TV. 2016. Toward the mechanism of eIF4F-mediated ribosomal attachment to mammalian capped mRNAs. *Genes Dev* **30**: 1573-1588.

- Kuyumcu-Martinez M, Belliot G, Sosnovtsev SV, Chang KO, Green KY, Lloyd RE. 2004. Calicivirus 3C-like proteinase inhibits cellular translation by cleavage of poly(A)-binding protein. *J Virol* **78**: 8172-8182.
- Kwun HJ, Toptan T, Ramos da Silva S, Atkins JF, Moore PS, Chang Y. 2014. Human DNA tumor viruses generate alternative reading frame proteins through repeat sequence recoding. *Proc Natl Acad Sci U S A* **111**: E4342-4349.
- Lan TCT, Allan MF, Malsick LE, Woo JZ, Zhu C, Zhang F, Khandwala S, Nyeo SSY, Sun Y, Guo JU et al. 2022a. Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. *Nat Commun* **13**: 1128.
- Lan TCT, Allan MF, Malsick LE, Woo JZ, Zhu C, Zhang F, Khandwala S, Nyeo SSY, Sun Y, Guo JU et al. 2022b. Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. *Nature Communications* **13**: 1128.
- Landry DM, Hertz MI, Thompson SR. 2009. RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs. *Genes Dev* **23**: 2753-2764.
- Lapointe CP, Grosely R, Johnson AG, Wang J, Fernández IS, Puglisi JD. 2021. Dynamic competition between SARS-CoV-2 NSP1 and mRNA on the human ribosome inhibits translation initiation. *Proc Natl Acad Sci U S A* **118**.
- Lapointe CP, Grosely R, Sokabe M, Alvarado C, Wang J, Montabana E, Villa N, Shin BS, Dever TE, Fraser CS et al. 2022. eIF5B and eIF1A reorient initiator tRNA to allow ribosomal subunit joining. *Nature* **607**: 185-190.
- Latorre P, Kolakofsky D, Curran J. 1998. Sendai virus Y proteins are initiated by a ribosomal shunt. *Mol Cell Biol* **18**: 5021-5031.
- Launer-Felty K, Wong CJ, Cole JL. 2015. Structural analysis of adenovirus VAI RNA defines the mechanism of inhibition of PKR. *Biophys J* **108**: 748-757.
- Lavie M, Dubuisson J, Belouzard S. 2022. SARS-CoV-2 Spike Furin Cleavage Site and S2' Basic Residues Modulate the Entry Process in a Host Cell-Dependent Manner. *J Virol* **96**: e0047422.
- Lavigne M, Helynck O, Rigolet P, Boudria-Souilah R, Nowakowski M, Baron B, Brûlé S, Hoos S, Raynal B, Guittat L et al. 2021. SARS-CoV-2 Nsp3 unique domain SUD interacts with guanine quadruplexes and G4-ligands inhibit this interaction. *Nucleic Acids Res* **49**: 7695-7712.
- Law GL, Raney A, Heusner C, Morris DR. 2001. Polyamine regulation of ribosome pausing at the upstream open reading frame of S-adenosylmethionine decarboxylase. *J Biol Chem* **276**: 38036-38043.
- Le May N, Dubaele S, Proietti De Santis L, Billecocq A, Bouloy M, Egly JM. 2004. TFIIF transcription factor, a target for the Rift Valley hemorrhagic fever virus. *Cell* **116**: 541-550.
- Le Sage V, Cinti A, Amorim R, Mouland AJ. 2016. Adapting the Stress Response: Viral Subversion of the mTOR Signaling Pathway. *Viruses* **8**.
- Le Sage V, Cinti A, McCarthy S, Amorim R, Rao S, Daino GL, Tramontano E, Branch DR, Mouland AJ. 2017. Ebola virus VP35 blocks stress granule assembly. *Virology* **502**: 73-83.
- Lee DSM, Park J, Kromer A, Baras A, Rader DJ, Ritchie MD, Ghanem LR, Barash Y. 2021. Disrupting upstream translation in mRNAs is associated with human disease. *Nature Communications* **12**: 1515.
- Lee J, Kenward C, Worrall LJ, Vuckovic M, Gentile F, Ton AT, Ng M, Cherkasov A, Strynadka NCJ, Paetzel M. 2022. X-ray crystallographic characterization of the

- SARS-CoV-2 main protease polyprotein cleavage sites essential for viral processing and maturation. *Nat Commun* **13**: 5196.
- Lee KM, Chen CJ, Shih SR. 2017. Regulation Mechanisms of Viral IRES-Driven Translation. *Trends Microbiol* **25**: 546-561.
- Lee S, Liu B, Lee S, Huang SX, Shen B, Qian SB. 2012. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc Natl Acad Sci U S A* **109**: E2424-2432.
- Lee YF, Nomoto A, Detjen BM, Wimmer E. 1977. A protein covalently linked to poliovirus genome RNA. *Proc Natl Acad Sci U S A* **74**: 59-63.
- Lee YY, Cevallos RC, Jan E. 2009. An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2alpha phosphorylation. *J Biol Chem* **284**: 6661-6673.
- LeFebvre AK, Korneeva NL, Trutschl M, Cvek U, Duzan RD, Bradley CA, Hershey JW, Rhoads RE. 2006. Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J Biol Chem* **281**: 22917-22932.
- Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M et al. 2017. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Mol Cell* **66**: 22-37.e29.
- Leppek K, Das R, Barna M. 2018. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* **19**: 158-174.
- Levy-Strumpf N, Deiss LP, Berissi H, Kimchi A. 1997. DAP-5, a novel homolog of eukaryotic translation initiation factor 4G isolated as a putative modulator of gamma interferon-induced programmed cell death. *Mol Cell Biol* **17**: 1615-1625.
- Li A, Chen YS, Ping XL, Yang X, Xiao W, Yang Y, Sun HY, Zhu Q, Baidya P, Wang X et al. 2017. Cytoplasmic m(6)A reader YTHDF3 promotes mRNA translation. *Cell Res* **27**: 444-447.
- Li JJ, Cao C, Fixsen SM, Young JM, Ono C, Bando H, Elde NC, Katsuma S, Dever TE, Sicheri F. 2015. Baculovirus protein PK2 subverts eIF2α kinase function by mimicry of its kinase domain C-lobe. *Proc Natl Acad Sci U S A* **112**: E4364-4373.
- Li K, Kong J, Zhang S, Zhao T, Qian W. 2022. Distance-dependent inhibition of translation initiation by downstream out-of-frame AUGs is consistent with a Brownian ratchet process of ribosome scanning. *Genome Biol* **23**: 254.
- Li L, Kang H, Liu P, Makkinje N, Williamson ST, Leibowitz JL, Giedroc DP. 2008. Structural lability in stem-loop 1 drives a 5' UTR-3' UTR interaction in coronavirus replication. *J Mol Biol* **377**: 790-803.
- Li Q, Imataka H, Morino S, Rogers GW, Jr., Richter-Cook NJ, Merrick WC, Sonenberg N. 1999. Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol Cell Biol* **19**: 7336-7346.
- Li Y, Treffers EE, Naphine S, Tas A, Zhu L, Sun Z, Bell S, Mark BL, van Veelen PA, van Hemert MJ et al. 2014. Transactivation of programmed ribosomal frameshifting by a viral protein. *Proc Natl Acad Sci U S A* **111**: E2172-2181.
- Liang W, Gu M, Zhu L, Yan Z, Schenten D, Herrick S, Li H, Samrat SK, Zhu J, Chen Y. 2023. The main protease of SARS-CoV-2 downregulates innate immunity via a translational repression. *Signal Transduction and Targeted Therapy* **8**: 162.

- Liberman N, Gandin V, Svitkin YV, David M, Virgili G, Jaramillo M, Holcik M, Nagar B, Kimchi A, Sonenberg N. 2015. DAP5 associates with eIF2 β and eIF4AI to promote Internal Ribosome Entry Site driven translation. *Nucleic Acids Res* **43**: 3764-3775.
- Lin Y, May GE, Kready H, Nazzaro L, Mao M, Spealman P, Creeger Y, McManus CJ. 2019. Impacts of uORF codon identity and position on translation regulation. *Nucleic Acids Res* **47**: 9358-9367.
- Linder P, Jankowsky E. 2011. From unwinding to clamping - the DEAD box RNA helicase family. *Nat Rev Mol Cell Biol* **12**: 505-516.
- Linder P, Lasko PF, Ashburner M, Leroy P, Nielsen PJ, Nishi K, Schnier J, Slonimski PP. 1989. Birth of the D-E-A-D box. *Nature* **337**: 121-122.
- Liu B, Qian SB. 2014. Translational reprogramming in cellular stress response. *Wiley Interdiscip Rev RNA* **5**: 301-315.
- Liu Q, Peng X, Shen M, Qian Q, Xing J, Li C, Gregory RI. 2023. Ribo-uORF: a comprehensive data resource of upstream open reading frames (uORFs) based on ribosome profiling. *Nucleic Acids Res* **51**: D248-d261.
- Llácer JL, Hussain T, Saini AK, Nanda JS, Kaur S, Gordiyenko Y, Kumar R, Hinnebusch AG, Lorsch JR, Ramakrishnan V. 2018. Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition. *Elife* **7**.
- Loughran G, Firth AE, Atkins JF. 2011. Ribosomal frameshifting into an overlapping gene in the 2B-encoding region of the cardiovirus genome. *Proc Natl Acad Sci U S A* **108**: E1111-1119.
- Loughran G, Sachs MS, Atkins JF, Ivanov IP. 2012. Stringency of start codon selection modulates autoregulation of translation initiation factor eIF5. *Nucleic Acids Res* **40**: 2898-2906.
- Luukkonen BG, Tan W, Schwartz S. 1995. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J Virol* **69**: 4086-4094.
- Macejak DG, Sarnow P. 1991. Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature* **353**: 90-94.
- Mackowiak SD, Zauber H, Bielow C, Thiel D, Kutz K, Calviello L, Mastrobuoni G, Rajewsky N, Kempa S, Selbach M et al. 2015. Extensive identification and analysis of conserved small ORFs in animals. *Genome Biol* **16**: 179.
- Mader S, Lee H, Pause A, Sonenberg N. 1995. The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol Cell Biol* **15**: 4990-4997.
- Malone B, Urakova N, Snijder EJ, Campbell EA. 2022. Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design. *Nature Reviews Molecular Cell Biology* **23**: 21-39.
- Mangeot PE, Risson V, Fusil F, Marnef A, Laurent E, Blin J, Mournetas V, Massourides E, Sohier TJM, Corbin A et al. 2019. Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nature communications* **10**: 45.
- Mangus DA, Evans MC, Jacobson A. 2003. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* **4**: 223.

- Mao Y, Dong L, Liu X-M, Guo J, Ma H, Shen B, Qian S-B. 2019. m6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nature Communications* **10**: 5332.
- Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. 1997. Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* **89**: 951-961.
- . 1999. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell* **3**: 707-716.
- Marintchev A, Edmonds KA, Marintcheva B, Hendrickson E, Oberer M, Suzuki C, Herdy B, Sonenberg N, Wagner G. 2009. Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. *Cell* **136**: 447-460.
- Martin-Marcos P, Nanda JS, Luna RE, Zhang F, Saini AK, Cherkasova VA, Wagner G, Lorsch JR, Hinnebusch AG. 2014. Enhanced eIF1 binding to the 40S ribosome impedes conformational rearrangements of the preinitiation complex and elevates initiation accuracy. *Rna* **20**: 150-167.
- Masters PS. 2019. Coronavirus genomic RNA packaging. *Virology* **537**: 198-207.
- Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, Wagner G. 1997. Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. *Nat Struct Biol* **4**: 717-724.
- Mayr C. 2019. What Are 3' UTRs Doing? *Cold Spring Harb Perspect Biol* **11**.
- McFadden MJ, McIntyre ABR, Mourelatos H, Abell NS, Gokhale NS, Ipas H, Xhemalçe B, Mason CE, Horner SM. 2021. Post-transcriptional regulation of antiviral gene expression by N6-methyladenosine. *Cell Rep* **34**: 108798.
- McGillivray P, Ault R, Pawashe M, Kitchen R, Balasubramanian S, Gerstein M. 2018. A comprehensive catalog of predicted functional upstream open reading frames in humans. *Nucleic Acids Res* **46**: 3326-3338.
- McKinney C, Zavadil J, Bianco C, Shiflett L, Brown S, Mohr I. 2014. Global reprogramming of the cellular translational landscape facilitates cytomegalovirus replication. *Cell Rep* **6**: 9-17.
- Medina M, Domingo E, Brangwyn JK, Belsham GJ. 1993. The two species of the foot-and-mouth disease virus leader protein, expressed individually, exhibit the same activities. *Virology* **194**: 355-359.
- Melcher SE, Wilson TJ, Lilley DM. 2003. The dynamic nature of the four-way junction of the hepatitis C virus IRES. *Rna* **9**: 809-820.
- Melian EB, Hinzman E, Nagasaki T, Firth AE, Wills NM, Nouwens AS, Blitvich BJ, Leung J, Funk A, Atkins JF et al. 2010. NS1' of flaviviruses in the Japanese encephalitis virus serogroup is a product of ribosomal frameshifting and plays a role in viral neuroinvasiveness. *J Virol* **84**: 1641-1647.
- Mendez AS, Ly M, González-Sánchez AM, Hartenian E, Ingolia NT, Cate JH, Glaunsinger BA. 2021. The N-terminal domain of SARS-CoV-2 nsp1 plays key roles in suppression of cellular gene expression and preservation of viral gene expression. *Cell Rep* **37**: 109841.
- Méthot N, Pause A, Hershey JW, Sonenberg N. 1994. The translation initiation factor eIF-4B contains an RNA-binding region that is distinct and independent from its ribonucleoprotein consensus sequence. *Mol Cell Biol* **14**: 2307-2316.
- Méthot N, Pickett G, Keene JD, Sonenberg N. 1996. In vitro RNA selection identifies RNA ligands that specifically bind to eukaryotic translation initiation factor 4B: the role of the RNA motif. *Rna* **2**: 38-50.

- Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, Pestova TV, Qian SB, Jaffrey SR. 2015. 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* **163**: 999-1010.
- Miao Z, Tidu A, Eriani G, Martin F. 2021. Secondary structure of the SARS-CoV-2 5'-UTR. *RNA Biol* **18**: 447-456.
- Mignone F, Gissi C, Liuni S, Pesole G. 2002. Untranslated regions of mRNAs. *Genome Biol* **3**: Reviews0004.
- Mudhasani R, Tran JP, Retterer C, Kota KP, Whitehouse CA, Bavari S. 2016. Protein Kinase R Degradation Is Essential for Rift Valley Fever Virus Infection and Is Regulated by SKP1-CUL1-F-box (SCF)FBXW11-NSs E3 Ligase. *PLoS Pathog* **12**: e1005437.
- Muhs M, Hilal T, Mielke T, Skabkin MA, Sanbonmatsu KY, Pestova TV, Spahn CM. 2015. Cryo-EM of ribosomal 80S complexes with termination factors reveals the translocated cricket paralysis virus IRES. *Mol Cell* **57**: 422-432.
- Muhs M, Yamamoto H, Ismer J, Takaku H, Nashimoto M, Uchiumi T, Nakashima N, Mielke T, Hildebrand PW, Nierhaus KH et al. 2011. Structural basis for the binding of IRES RNAs to the head of the ribosomal 40S subunit. *Nucleic Acids Res* **39**: 5264-5275.
- Mulvey M, Poppers J, Sternberg D, Mohr I. 2003. Regulation of eIF2alpha phosphorylation by different functions that act during discrete phases in the herpes simplex virus type 1 life cycle. *J Virol* **77**: 10917-10928.
- Munroe D, Jacobson A. 1990. mRNA poly(A) tail, a 3' enhancer of translational initiation. *Mol Cell Biol* **10**: 3441-3455.
- Murray J, Savva CG, Shin B-S, Dever TE, Ramakrishnan V, Fernández IS. 2016. Structural characterization of ribosome recruitment and translocation by type IV IRES. *eLife* **5**: e13567.
- Nachtergaele S, He C. 2018. Chemical Modifications in the Life of an mRNA Transcript. *Annu Rev Genet* **52**: 349-372.
- Nakagawa K, Makino S. 2021. Mechanisms of Coronavirus Nsp1-Mediated Control of Host and Viral Gene Expression. *Cells* **10**.
- Nanda JS, Cheung YN, Takacs JE, Martin-Marcos P, Saini AK, Hinnebusch AG, Lorsch JR. 2009. eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation. *J Mol Biol* **394**: 268-285.
- Naphtine S, Ling R, Finch LK, Jones JD, Bell S, Brierley I, Firth AE. 2017. Protein-directed ribosomal frameshifting temporally regulates gene expression. *Nat Commun* **8**: 15582.
- Naphtine S, Treffers EE, Bell S, Goodfellow I, Fang Y, Firth AE, Snijder EJ, Brierley I. 2016. A novel role for poly(C) binding proteins in programmed ribosomal frameshifting. *Nucleic Acids Res* **44**: 5491-5503.
- Naranda T, Strong WB, Menaya J, Fabbri BJ, Hershey JW. 1994. Two structural domains of initiation factor eIF-4B are involved in binding to RNA. *J Biol Chem* **269**: 14465-14472.
- Ng CS, Jogi M, Yoo JS, Onomoto K, Koike S, Iwasaki T, Yoneyama M, Kato H, Fujita T. 2013. Encephalomyocarditis virus disrupts stress granules, the critical platform for triggering antiviral innate immune responses. *J Virol* **87**: 9511-9522.
- Nguyen TMT, Kim J, Doan TT, Lee MW, Lee M. 2020. APEX Proximity Labeling as a Versatile Tool for Biological Research. *Biochemistry* **59**: 260-269.

- Niedzwiecka A, Marcotrigiano J, Stepinski J, Jankowska-Anyszka M, Wyslouch-Cieszyńska A, Dadlez M, Gingras AC, Mak P, Darzynkiewicz E, Sonenberg N et al. 2002. Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J Mol Biol* **319**: 615-635.
- Nygård O, Nika H. 1982. Identification by RNA-protein cross-linking of ribosomal proteins located at the interface between the small and the large subunits of mammalian ribosomes. *Embo j* **1**: 357-362.
- O'Leary SE, Petrov A, Chen J, Puglisi JD. 2013. Dynamic recognition of the mRNA cap by *Saccharomyces cerevisiae* eIF4E. *Structure* **21**: 2197-2207.
- Ohlmann T, Rau M, Pain VM, Morley SJ. 1996. The C-terminal domain of eukaryotic protein synthesis initiation factor (eIF) 4G is sufficient to support cap-independent translation in the absence of eIF4E. *Embo j* **15**: 1371-1382.
- Othman Z, Sulaiman MK, Willcocks MM, Ulryck N, Blackbourn DJ, Sargueil B, Roberts LO, Locker N. 2014. Functional analysis of Kaposi's sarcoma-associated herpesvirus vFLIP expression reveals a new mode of IRES-mediated translation. *Rna* **20**: 1803-1814.
- Otto GA, Lukavsky PJ, Lancaster AM, Sarnow P, Puglisi JD. 2002. Ribosomal proteins mediate the hepatitis C virus IRES-HeLa 40S interaction. *Rna* **8**: 913-923.
- Otto GA, Puglisi JD. 2004. The pathway of HCV IRES-mediated translation initiation. *Cell* **119**: 369-380.
- Özeş AR, Feoktistova K, Avanzino BC, Fraser CS. 2011. Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B. *J Mol Biol* **412**: 674-687.
- Palam LR, Baird TD, Wek RC. 2011. Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. *J Biol Chem* **286**: 10939-10949.
- Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E et al. 2017. Translation of CircRNAs. *Mol Cell* **66**: 9-21.e27.
- Park GJ, Osinski A, Hernandez G, Eitson JL, Majumdar A, Tonelli M, Henzler-Wildman K, Pawłowski K, Chen Z, Li Y et al. 2022. The mechanism of RNA capping by SARS-CoV-2. *Nature* **609**: 793-800.
- Park HS, Himmelbach A, Browning KS, Hohn T, Ryabova LA. 2001. A plant viral "reinitiation" factor interacts with the host translational machinery. *Cell* **106**: 723-733.
- Park R, El-Guindy A, Heston L, Lin SF, Yu KP, Nagy M, Borah S, Delecluse HJ, Steitz J, Miller G. 2014. Nuclear translocation and regulation of intranuclear distribution of cytoplasmic poly(A)-binding protein are distinct processes mediated by two Epstein Barr virus proteins. *PLoS One* **9**: e92593.
- Parkin NT, Cohen EA, Darveau A, Rosen C, Haseltine W, Sonenberg N. 1988. Mutational analysis of the 5' non-coding region of human immunodeficiency virus type 1: effects of secondary structure on translation. *Embo j* **7**: 2831-2837.
- Parrish S, Moss B. 2007. Characterization of a second vaccinia virus mRNA-decapping enzyme conserved in poxviruses. *J Virol* **81**: 12973-12978.

- Passmore LA, Schmeing TM, Maag D, Applefield DJ, Acker MG, Algire MA, Lorsch JR, Ramakrishnan V. 2007. The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell* **26**: 41-50.
- Pause A, Belsham GJ, Gingras AC, Donzé O, Lin TA, Lawrence JC, Jr., Sonenberg N. 1994a. Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* **371**: 762-767.
- Pause A, Méthot N, Svitkin Y, Merrick WC, Sonenberg N. 1994b. Dominant negative mutants of mammalian translation initiation factor eIF-4A define a critical role for eIF-4F in cap-dependent and cap-independent initiation of translation. *Embo j* **13**: 1205-1215.
- Pavitt GD, Ramaiah KV, Kimball SR, Hinnebusch AG. 1998. eIF2 independently binds two distinct eIF2B subcomplexes that catalyze and regulate guanine-nucleotide exchange. *Genes Dev* **12**: 514-526.
- Peabody DS. 1987. Translation initiation at an ACG triplet in mammalian cells. *J Biol Chem* **262**: 11847-11851.
- . 1989. Translation initiation at non-AUG triplets in mammalian cells. *J Biol Chem* **264**: 5031-5035.
- Pelletier J, Sonenberg N. 1985a. Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell* **40**: 515-526.
- . 1985b. Photochemical cross-linking of cap binding proteins to eucaryotic mRNAs: effect of mRNA 5' secondary structure. *Mol Cell Biol* **5**: 3222-3230.
- . 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**: 320-325.
- Perry RP, Kelley DE, Friderici K, Rottman F. 1975. The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell* **4**: 387-394.
- Pesole G, Mignone F, Gissi C, Grillo G, Licciulli F, Liuni S. 2001. Structural and functional features of eukaryotic mRNA untranslated regions. *Gene* **276**: 73-81.
- Pestova TV, Borukhov SI, Hellen CU. 1998a. Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* **394**: 854-859.
- Pestova TV, de Breyne S, Pisarev AV, Abaeva IS, Hellen CU. 2008. eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: a common role of domain II. *Embo j* **27**: 1060-1072.
- Pestova TV, Hellen CU, Shatsky IN. 1996a. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol Cell Biol* **16**: 6859-6869.
- Pestova TV, Kolupaeva VG. 2002. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev* **16**: 2906-2922.
- Pestova TV, Lomakin IB, Hellen CU. 2004. Position of the CrPV IRES on the 40S subunit and factor dependence of IRES/80S ribosome assembly. *EMBO Rep* **5**: 906-913.
- Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU. 1998b. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev* **12**: 67-83.

- Pestova TV, Shatsky IN, Hellen CU. 1996b. Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol Cell Biol* **16**: 6870-6878.
- Peter D, Igreja C, Weber R, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, Izaurralde E. 2015. Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. *Mol Cell* **57**: 1074-1087.
- Peters JK, Tibble RW, Warminski M, Jemielity J, Gross JD. 2022. Structure of the poxvirus decapping enzyme D9 reveals its mechanism of cap recognition and catalysis. *Structure* **30**: 721-732.e724.
- Pettit SC, Everitt LE, Choudhury S, Dunn BM, Kaplan AH. 2004. Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. *J Virol* **78**: 8477-8485.
- Pfeiffer CT, Paulo JA, Gygi SP, Rockman HA. 2022. Proximity labeling for investigating protein-protein interactions. *Methods Cell Biol* **169**: 237-266.
- Phan L, Schoenfeld LW, Valásek L, Nielsen KH, Hinnebusch AG. 2001. A subcomplex of three eIF3 subunits binds eIF1 and eIF5 and stimulates ribosome binding of mRNA and tRNA(i)Met. *Embo j* **20**: 2954-2965.
- Phan L, Zhang X, Asano K, Anderson J, Vornlocher HP, Greenberg JR, Qin J, Hinnebusch AG. 1998. Identification of a translation initiation factor 3 (eIF3) core complex, conserved in yeast and mammals, that interacts with eIF5. *Mol Cell Biol* **18**: 4935-4946.
- Pisarev AV, Chard LS, Kaku Y, Johns HL, Shatsky IN, Belsham GJ. 2004. Functional and structural similarities between the internal ribosome entry sites of hepatitis C virus and porcine teschovirus, a picornavirus. *J Virol* **78**: 4487-4497.
- Pisareva VP, Pisarev AV, Fernández IS. 2018. Dual tRNA mimicry in the Cricket Paralysis Virus IRES uncovers an unexpected similarity with the Hepatitis C Virus IRES. *eLife* **7**: e34062.
- Pisareva VP, Pisarev AV, Komar AA, Hellen CU, Pestova TV. 2008. Translation initiation on mammalian mRNAs with structured 5'UTRs requires DExH-box protein DHX29. *Cell* **135**: 1237-1250.
- Plant EP, Dinman JD. 2006. Comparative study of the effects of heptameric slippery site composition on -1 frameshifting among different eukaryotic systems. *Rna* **12**: 666-673.
- Poulin F, Gingras AC, Olsen H, Chevalier S, Sonenberg N. 1998. 4E-BP3, a new member of the eukaryotic initiation factor 4E-binding protein family. *J Biol Chem* **273**: 14002-14007.
- Prats AC, De Billy G, Wang P, Darlix JL. 1989. CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *J Mol Biol* **205**: 363-372.
- Preissler S, Ron D. 2019. Early Events in the Endoplasmic Reticulum Unfolded Protein Response. *Cold Spring Harb Perspect Biol* **11**.
- Protter DSW, Parker R. 2016. Principles and Properties of Stress Granules. *Trends Cell Biol* **26**: 668-679.
- Pyle JD, Mandadi KK, Scholthof KBG. 2019. Panicum Mosaic Virus and Its Satellites Acquire RNA Modifications Associated with Host-Mediated Antiviral Degradation. *mBio* **10**.

- Pyronnet S, Imataka H, Gingras AC, Fukunaga R, Hunter T, Sonenberg N. 1999. Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *Embo j* **18**: 270-279.
- Qiu Y, Krug RM. 1994. The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A). *J Virol* **68**: 2425-2432.
- Quade N, Boehringer D, Leibundgut M, van den Heuvel J, Ban N. 2015. Cryo-EM structure of Hepatitis C virus IRES bound to the human ribosome at 3.9-Å resolution. *Nat Commun* **6**: 7646.
- Querol-Audi J, Sun C, Vogan JM, Smith MD, Gu Y, Cate JH, Nogales E. 2013. Architecture of human translation initiation factor 3. *Structure* **21**: 920-928.
- Rabl J, Leibundgut M, Ataide SF, Haag A, Ban N. 2011. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* **331**: 730-736.
- Rabouw HH, Langereis MA, Knaap RC, Dalebout TJ, Canton J, Sola I, Enjuanes L, Bredenbeek PJ, Kikkert M, de Groot RJ et al. 2016. Middle East Respiratory Coronavirus Accessory Protein 4a Inhibits PKR-Mediated Antiviral Stress Responses. *PLoS Pathog* **12**: e1005982.
- Rangan R, Zheludev IN, Hagey RJ, Pham EA, Wayment-Steele HK, Glenn JS, Das R. 2020. RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: a first look. *Rna* **26**: 937-959.
- Ray BK, Lawson TG, Kramer JC, Cladaras MH, Grifo JA, Abramson RD, Merrick WC, Thach RE. 1985. ATP-dependent unwinding of messenger RNA structure by eukaryotic initiation factors. *J Biol Chem* **260**: 7651-7658.
- Rehwinkel J, Letunic I, Raes J, Bork P, Izaurralde E. 2005. Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. *Rna* **11**: 1530-1544.
- Richter C, Hohmann KF, Toews S, Mathieu D, Altincekic N, Bains JK, Binas O, Ceylan B, Duchardt-Ferner E, Ferner J et al. 2021. (1)H, (13)C and (15)N assignment of stem-loop SL1 from the 5'-UTR of SARS-CoV-2. *Biomol NMR Assign* **15**: 467-474.
- Richter-Cook NJ, Dever TE, Hensold JO, Merrick WC. 1998. Purification and characterization of a new eukaryotic protein translation factor. Eukaryotic initiation factor 4H. *J Biol Chem* **273**: 7579-7587.
- Robert F, Kapp LD, Khan SN, Acker MG, Kolitz S, Kazemi S, Kaufman RJ, Merrick WC, Koromilas AE, Lorsch JR et al. 2006. Initiation of protein synthesis by hepatitis C virus is refractory to reduced eIF2.GTP.Met-tRNA(i)(Met) ternary complex availability. *Mol Biol Cell* **17**: 4632-4644.
- Rocak S, Linder P. 2004. DEAD-box proteins: the driving forces behind RNA metabolism. *Nature Reviews Molecular Cell Biology* **5**: 232-241.
- Rogers GW, Jr., Richter NJ, Lima WF, Merrick WC. 2001. Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J Biol Chem* **276**: 30914-30922.
- Rojas M, Vasconcelos G, Dever TE. 2015. An eIF2α-binding motif in protein phosphatase 1 subunit GADD34 and its viral orthologs is required to promote dephosphorylation of eIF2α. *Proc Natl Acad Sci U S A* **112**: E3466-3475.
- Rowlands AG, Panniers R, Henshaw EC. 1988. The catalytic mechanism of guanine nucleotide exchange factor action and competitive inhibition by phosphorylated eukaryotic initiation factor 2. *J Biol Chem* **263**: 5526-5533.

- Royall E, Doyle N, Abdul-Wahab A, Emmott E, Morley SJ, Goodfellow I, Roberts LO, Locker N. 2015. Murine norovirus 1 (MNV1) replication induces translational control of the host by regulating eIF4E activity during infection. *J Biol Chem* **290**: 4748-4758.
- Rozen F, Edery I, Meerovitch K, Dever TE, Merrick WC, Sonenberg N. 1990. Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Mol Cell Biol* **10**: 1134-1144.
- Rozovsky N, Butterworth AC, Moore MJ. 2008. Interactions between eIF4A1 and its accessory factors eIF4B and eIF4H. *Rna* **14**: 2136-2148.
- Rubio RM, Depledge DP, Bianco C, Thompson L, Mohr I. 2018. RNA m(6) A modification enzymes shape innate responses to DNA by regulating interferon β . *Genes Dev* **32**: 1472-1484.
- Russell PJ, Slivka JA, Boyle EP, Burghes AHM, Kearse MG. 2023. Translation reinitiation after uORFs does not fully protect mRNAs from nonsense-mediated decay. *Rna* **29**: 735-744.
- Rutkowski AJ, Erhard F, L'Hernault A, Bonfert T, Schilhabel M, Crump C, Rosenstiel P, Efstathiou S, Zimmer R, Friedel CC et al. 2015. Widespread disruption of host transcription termination in HSV-1 infection. *Nat Commun* **6**: 7126.
- Ryabova LA, Hohn T. 2000. Ribosome shunting in the cauliflower mosaic virus 35S RNA leader is a special case of reinitiation of translation functioning in plant and animal systems. *Genes Dev* **14**: 817-829.
- Sanz MA, González Almela E, Carrasco L. 2017. Translation of Sindbis Subgenomic mRNA is Independent of eIF2, eIF2A and eIF2D. *Scientific Reports* **7**: 43876.
- Schalm SS, Fingar DC, Sabatini DM, Blenis J. 2003. TOS motif-mediated raptor binding regulates 4E-BP1 multisite phosphorylation and function. *Curr Biol* **13**: 797-806.
- Scherer KM, Mascheroni L, Carnell GW, Wunderlich LCS, Makarchuk S, Brockhoff M, Mela I, Fernandez-Villegas A, Barysevich M, Stewart H et al. 2022. SARS-CoV-2 nucleocapsid protein adheres to replication organelles before viral assembly at the Golgi/ERGIC and lysosome-mediated egress. *Sci Adv* **8**: eabl4895.
- Schmitt E, Panvert M, Lazennec-Schurdevin C, Coureux PD, Perez J, Thompson A, Mechulam Y. 2012. Structure of the ternary initiation complex eIF2-GDPNP-methionylated initiator tRNA. *Nat Struct Mol Biol* **19**: 450-454.
- Schubert K, Karousis ED, Ban I, Lapointe CP, Leibundgut M, Bäuml E, Kummerant E, Scaiola A, Schönhut T, Ziegelmüller J et al. 2023. Universal features of Nsp1-mediated translational shutdown by coronaviruses. *Mol Cell* **83**: 3546-3557.e3548.
- Schubert K, Karousis ED, Jomaa A, Scaiola A, Echeverria B, Gurzeler L-A, Leibundgut M, Thiel V, Mühlemann O, Ban N. 2020a. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nature Structural & Molecular Biology* **27**: 959-966.
- Schubert K, Karousis ED, Jomaa A, Scaiola A, Echeverria B, Gurzeler LA, Leibundgut M, Thiel V, Mühlemann O, Ban N. 2020b. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat Struct Mol Biol* **27**: 959-966.
- Schüler M, Connell SR, Lescoute A, Giesebrecht J, Dabrowski M, Schroer B, Mielke T, Penczek PA, Westhof E, Spahn CM. 2006. Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nat Struct Mol Biol* **13**: 1092-1096.

- Schumann U, Zhang H-N, Sibbritt T, Pan A, Horvath A, Gross S, Clark SJ, Yang L, Preiss T. 2020. Multiple links between 5-methylcytosine content of mRNA and translation. *BMC Biology* **18**: 40.
- Sciortino MT, Parisi T, Siracusano G, Mastino A, Taddeo B, Roizman B. 2013. The virion host shutoff RNase plays a key role in blocking the activation of protein kinase R in cells infected with herpes simplex virus 1. *J Virol* **87**: 3271-3276.
- Semper C, Watanabe N, Savchenko A. 2021. Structural characterization of nonstructural protein 1 from SARS-CoV-2. *iScience* **24**: 101903.
- Sen ND, Zhou F, Harris MS, Ingolia NT, Hinnebusch AG. 2016. eIF4B stimulates translation of long mRNAs with structured 5' UTRs and low closed-loop potential but weak dependence on eIF4G. *Proc Natl Acad Sci U S A* **113**: 10464-10472.
- Seo EJ, Liu F, Kawagishi-Kobayashi M, Ung TL, Cao C, Dar AC, Sicheri F, Dever TE. 2008. Protein kinase PKR mutants resistant to the poxvirus pseudosubstrate K3L protein. *Proc Natl Acad Sci U S A* **105**: 16894-16899.
- Shabman RS, Hoenen T, Groseth A, Jabado O, Binning JM, Amarasinghe GK, Feldmann H, Basler CF. 2013. An upstream open reading frame modulates ebola virus polymerase translation and virus replication. *PLoS Pathog* **9**: e1003147.
- Sharma NR, Majerciak V, Kruhlak MJ, Zheng ZM. 2017. KSHV inhibits stress granule formation by viral ORF57 blocking PKR activation. *PLoS Pathog* **13**: e1006677.
- Sharp PM. 2002. Origins of human virus diversity. *Cell* **108**: 305-312.
- Shehata SI, Parker R. 2023. SARS-CoV-2 Nsp1 mediated mRNA degradation requires mRNA interaction with the ribosome. *RNA Biol* **20**: 444-456.
- Shen Z, Zhang G, Yang Y, Li M, Yang S, Peng G. 2021. Lysine 164 is critical for SARS-CoV-2 Nsp1 inhibition of host gene expression. *J Gen Virol* **102**.
- Shi H, Wei J, He C. 2019. Where, When, and How: Context-Dependent Functions of RNA Methylation Writers, Readers, and Erasers. *Mol Cell* **74**: 640-650.
- Shi Y, Vattem KM, Sood R, An J, Liang J, Stramm L, Wek RC. 1998. Identification and characterization of pancreatic eukaryotic initiation factor 2 alpha-subunit kinase, PEK, involved in translational control. *Mol Cell Biol* **18**: 7499-7509.
- Shibuya T, Tange T, Sonenberg N, Moore MJ. 2004. eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nat Struct Mol Biol* **11**: 346-351.
- Shirokikh NE, Spirin AS. 2008. Poly(A) leader of eukaryotic mRNA bypasses the dependence of translation on initiation factors. *Proc Natl Acad Sci U S A* **105**: 10738-10743.
- Sizova DV, Kolupaeva VG, Pestova TV, Shatsky IN, Hellen CU. 1998. Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. *J Virol* **72**: 4775-4782.
- Slavoff SA, Mitchell AJ, Schwaid AG, Cabili MN, Ma J, Levin JZ, Karger AD, Budnik BA, Rinn JL, Saghatelian A. 2013. Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nat Chem Biol* **9**: 59-64.
- Slepenkov SV, Korneeva NL, Rhoads RE. 2008. Kinetic mechanism for assembly of the m7GpppG.eIF4E.eIF4G complex. *J Biol Chem* **283**: 25227-25237.
- Slobodin B, Sehrawat U, Lev A, Hayat D, Zuckerman B, Fraticelli D, Ogran A, Ben-Shmuel A, Bar-David E, Levy H et al. 2022. Cap-independent translation and a

- precisely located RNA sequence enable SARS-CoV-2 to control host translation and escape anti-viral response. *Nucleic Acids Res* **50**: 8080-8092.
- Smith E, Meyerrose TE, Kohler T, Namdar-Attar M, Bab N, Lahat O, Noh T, Li J, Karaman MW, Hacia JG et al. 2005. Leaky ribosomal scanning in mammalian genomes: significance of histone H4 alternative translation in vivo. *Nucleic Acids Res* **33**: 1298-1308.
- Smith MD, Arake-Tacca L, Nitido A, Montabana E, Park A, Cate JH. 2016. Assembly of eIF3 Mediated by Mutually Dependent Subunit Insertion. *Structure* **24**: 886-896.
- Smith RW, Gray NK. 2010. Poly(A)-binding protein (PABP): a common viral target. *Biochem J* **426**: 1-12.
- Sola I, Almazán F, Zúñiga S, Enjuanes L. 2015. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu Rev Virol* **2**: 265-288.
- Sonenberg N, Morgan MA, Merrick WC, Shatkin AJ. 1978. A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proc Natl Acad Sci U S A* **75**: 4843-4847.
- Sood R, Porter AC, Olsen DA, Cavener DR, Wek RC. 2000. A mammalian homologue of GCN2 protein kinase important for translational control by phosphorylation of eukaryotic initiation factor-2alpha. *Genetics* **154**: 787-801.
- Soto Rifo R, Ricci EP, Décimo D, Moncorgé O, Ohlmann T. 2007. Back to basics: the untreated rabbit reticulocyte lysate as a competitive system to recapitulate cap/poly(A) synergy and the selective advantage of IRES-driven translation. *Nucleic Acids Res* **35**: e121.
- Soto-Rifo R, Ohlmann T. 2013. The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism. *Wiley Interdiscip Rev RNA* **4**: 369-385.
- Soto-Rifo R, Rubilar PS, Limousin T, de Breyne S, Décimo D, Ohlmann T. 2012. DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. *Embo j* **31**: 3745-3756.
- Soto-Rifo R, Rubilar PS, Ohlmann T. 2013. The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA. *Nucleic Acids Res* **41**: 6286-6299.
- Spahn CM, Jan E, Mulder A, Grassucci RA, Sarnow P, Frank J. 2004. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* **118**: 465-475.
- Spangle JM, Münger K. 2010. The human papillomavirus type 16 E6 oncoprotein activates mTORC1 signaling and increases protein synthesis. *J Virol* **84**: 9398-9407.
- Spirin AS. 2009. How Does a Scanning Ribosomal Particle Move along the 5'-Untranslated Region of Eukaryotic mRNA? Brownian Ratchet Model. *Biochemistry* **48**: 10688-10692.
- Starck SR, Jiang V, Pavon-Eternod M, Prasad S, McCarthy B, Pan T, Shastri N. 2012. Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science* **336**: 1719-1723.
- Starck SR, Tsai JC, Chen K, Shodiya M, Wang L, Yahiro K, Martins-Green M, Shastri N, Walter P. 2016. Translation from the 5' untranslated region shapes the integrated stress response. *Science* **351**: aad3867.

- Steiner S, Kratzel A, Barut GT, Lang RM, Aguiar Moreira E, Thomann L, Kelly JN, Thiel V. 2024. SARS-CoV-2 biology and host interactions. *Nature Reviews Microbiology* **22**: 206-225.
- Stern-Ginossar N, Thompson SR, Mathews MB, Mohr I. 2019. Translational Control in Virus-Infected Cells. *Cold Spring Harb Perspect Biol* **11**.
- Stewart H, Brown K, Dinan AM, Irigoyen N, Snijder EJ, Firth AE. 2018. Transcriptional and Translational Landscape of Equine Torovirus. *J Virol* **92**.
- Stolboushkina E, Nikonov S, Nikulin A, Bläsi U, Manstein DJ, Fedorov R, Garber M, Nikonov O. 2008. Crystal structure of the intact archaeal translation initiation factor 2 demonstrates very high conformational flexibility in the alpha- and beta-subunits. *J Mol Biol* **382**: 680-691.
- Sukarieh R, Sonenberg N, Pelletier J. 2010. Nuclear assortment of eIF4E coincides with shut-off of host protein synthesis upon poliovirus infection. *J Gen Virol* **91**: 1224-1228.
- Sun C, Querol-Audí J, Mortimer SA, Arias-Palomo E, Doudna JA, Nogales E, Cate JH. 2013. Two RNA-binding motifs in eIF3 direct HCV IRES-dependent translation. *Nucleic Acids Res* **41**: 7512-7521.
- Svitkin YV, Pause A, Haghighat A, Pyronnet S, Witherell G, Belsham GJ, Sonenberg N. 2001. The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. *Rna* **7**: 382-394.
- Sweeney TR, Abaeva IS, Pestova TV, Hellen CU. 2014. The mechanism of translation initiation on Type 1 picornavirus IRESs. *Embo j* **33**: 76-92.
- Takahashi K, Maruyama M, Tokuzawa Y, Murakami M, Oda Y, Yoshikane N, Makabe KW, Ichisaka T, Yamanaka S. 2005. Evolutionarily conserved non-AUG translation initiation in NAT1/p97/DAP5 (EIF4G2). *Genomics* **85**: 360-371.
- Tardivat Y, Sosnowski P, Tidu A, Westhof E, Eriani G, Martin F. 2023. SARS-CoV-2 NSP1 induces mRNA cleavages on the ribosome. *Nucleic Acids Res* **51**: 8677-8690.
- Terenin IM, Dmitriev SE, Andreev DE, Shatsky IN. 2008. Eukaryotic translation initiation machinery can operate in a bacterial-like mode without eIF2. *Nat Struct Mol Biol* **15**: 836-841.
- Thakur A, Marler L, Hinnebusch AG. 2019. A network of eIF2 β interactions with eIF1 and Met-tRNAⁱ promotes accurate start codon selection by the translation preinitiation complex. *Nucleic Acids Res* **47**: 2574-2593.
- Thoms M, Buschauer R, Ameismeier M, Koepke L, Denk T, Hirschenberger M, Kratzat H, Hayn M, Mackens-Kiani T, Cheng J et al. 2020. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* **369**: 1249-1255.
- Tidu A, Janvier A, Schaeffer L, Sosnowski P, Kuhn L, Hammann P, Westhof E, Eriani G, Martin F. 2020. The viral protein NSP1 acts as a ribosome gatekeeper for shutting down host translation and fostering SARS-CoV-2 translation. *Rna* **27**: 253-264.
- Toribio R, Díaz-López I, Boskovic J, Ventoso I. 2016. An RNA trapping mechanism in Alphavirus mRNA promotes ribosome stalling and translation initiation. *Nucleic Acids Res* **44**: 4368-4380.
- Trachsel H, Staehelin T. 1979. Initiation of mammalian protein synthesis. The multiple functions of the initiation factor eIF-3. *Biochim Biophys Acta* **565**: 305-314.

- Trixl L, Lusser A. 2019. The dynamic RNA modification 5-methylcytosine and its emerging role as an epitranscriptomic mark. *Wiley Interdiscip Rev RNA* **10**: e1510.
- Truitt ML, Conn CS, Shi Z, Pang X, Tokuyasu T, Coady AM, Seo Y, Barna M, Ruggero D. 2015. Differential Requirements for eIF4E Dose in Normal Development and Cancer. *Cell* **162**: 59-71.
- Turina M, Desvoyes B, Scholthof KB. 2000. A gene cluster encoded by panicum mosaic virus is associated with virus movement. *Virology* **266**: 120-128.
- Uchiyama-Kadokura N, Murakami K, Takemoto M, Koyanagi N, Murota K, Naito S, Onouchi H. 2014. Polyamine-responsive ribosomal arrest at the stop codon of an upstream open reading frame of the AdoMetDC1 gene triggers nonsense-mediated mRNA decay in *Arabidopsis thaliana*. *Plant Cell Physiol* **55**: 1556-1567.
- V'Kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. 2021. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol* **19**: 155-170.
- V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. 2021. Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology* **19**: 155-170.
- Vachon VK, Conn GL. 2016. Adenovirus VA RNA: An essential pro-viral non-coding RNA. *Virus Res* **212**: 39-52.
- Valásek L, Mathew AA, Shin BS, Nielsen KH, Szamecz B, Hinnebusch AG. 2003. The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S ribosome in vivo. *Genes Dev* **17**: 786-799.
- Valásek L, Nielsen KH, Hinnebusch AG. 2002. Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation in vivo. *Embo j* **21**: 5886-5898.
- Valásek L, Nielsen KH, Zhang F, Fekete CA, Hinnebusch AG. 2004. Interactions of eukaryotic translation initiation factor 3 (eIF3) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. *Mol Cell Biol* **24**: 9437-9455.
- Valente ST, Gilmartin GM, Mott C, Falkard B, Goff SP. 2009. Inhibition of HIV-1 replication by eIF3f. *Proc Natl Acad Sci U S A* **106**: 4071-4078.
- Vanderperre B, Lucier JF, Bissonnette C, Motard J, Tremblay G, Vanderperre S, Wisztorski M, Salzet M, Boisvert FM, Roucou X. 2013. Direct detection of alternative open reading frames translation products in human significantly expands the proteome. *PLoS One* **8**: e70698.
- Vassilenko KS, Alekhina OM, Dmitriev SE, Shatsky IN, Spirin AS. 2011. Unidirectional constant rate motion of the ribosomal scanning particle during eukaryotic translation initiation. *Nucleic Acids Res* **39**: 5555-5567.
- Vattem KM, Wek RC. 2004. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A* **101**: 11269-11274.
- Ventoso I, Blanco R, Perales C, Carrasco L. 2001. HIV-1 protease cleaves eukaryotic initiation factor 4G and inhibits cap-dependent translation. *Proc Natl Acad Sci U S A* **98**: 12966-12971.
- Ventoso I, Sanz MA, Molina S, Berlanga JJ, Carrasco L, Esteban M. 2006. Translational resistance of late alphavirus mRNA to eIF2alpha phosphorylation:

- a strategy to overcome the antiviral effect of protein kinase PKR. *Genes Dev* **20**: 87-100.
- Villa N, Do A, Hershey JW, Fraser CS. 2013. Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. *J Biol Chem* **288**: 32932-32940.
- von Der Haar T, Ball PD, McCarthy JE. 2000. Stabilization of eukaryotic initiation factor 4E binding to the mRNA 5'-Cap by domains of eIF4G. *J Biol Chem* **275**: 30551-30555.
- Vora SM, Fontana P, Mao T, Leger V, Zhang Y, Fu TM, Lieberman J, Gehrke L, Shi M, Wang L et al. 2022. Targeting stem-loop 1 of the SARS-CoV-2 5' UTR to suppress viral translation and Nsp1 evasion. *Proc Natl Acad Sci U S A* **119**.
- Vuyisich M, Spanggord RJ, Beal PA. 2002. The binding site of the RNA-dependent protein kinase (PKR) on EBER1 RNA from Epstein-Barr virus. *EMBO Rep* **3**: 622-627.
- Wacker A, Weigand JE, Akabayov SR, Altincekic N, Bains JK, Banijamali E, Binas O, Castillo-Martinez J, Cetiner E, Ceylan B et al. 2020a. Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy. *Nucleic Acids Research* **48**: 12415-12435.
- Wacker A, Weigand JE, Akabayov SR, Altincekic N, Bains JK, Banijamali E, Binas O, Castillo-Martinez J, Cetiner E, Ceylan B et al. 2020b. Secondary structure determination of conserved SARS-CoV-2 RNA elements by NMR spectroscopy. *Nucleic Acids Res* **48**: 12415-12435.
- Walsh D, Arias C, Perez C, Halladin D, Escandon M, Ueda T, Watanabe-Fukunaga R, Fukunaga R, Mohr I. 2008. Eukaryotic translation initiation factor 4F architectural alterations accompany translation initiation factor redistribution in poxvirus-infected cells. *Mol Cell Biol* **28**: 2648-2658.
- Walsh D, Mohr I. 2004. Phosphorylation of eIF4E by Mnk-1 enhances HSV-1 translation and replication in quiescent cells. *Genes Dev* **18**: 660-672.
- Walsh D, Perez C, Notary J, Mohr I. 2005. Regulation of the translation initiation factor eIF4F by multiple mechanisms in human cytomegalovirus-infected cells. *J Virol* **79**: 8057-8064.
- Wang D, Jiang A, Feng J, Li G, Guo D, Sajid M, Wu K, Zhang Q, Ponty Y, Will S et al. 2021. The SARS-CoV-2 subgenome landscape and its novel regulatory features. *Mol Cell* **81**: 2135-2147.e2135.
- Wang J, Shin BS, Alvarado C, Kim JR, Bohlen J, Dever TE, Puglisi JD. 2022. Rapid 40S scanning and its regulation by mRNA structure during eukaryotic translation initiation. *Cell* **185**: 4474-4487.e4417.
- Wang X, Hennig T, Whisnant AW, Erhard F, Prusty BK, Friedel CC, Forouzmand E, Hu W, Erber L, Chen Y et al. 2020. Herpes simplex virus blocks host transcription termination via the bimodal activities of ICP27. *Nat Commun* **11**: 293.
- Wang Y, Kirkpatrick J, Lage SZ, Carlomagno T. 2023. Structural insights into the activity regulation of full-length non-structural protein 1 from SARS-CoV-2. *Structure* **31**: 128-137.e125.
- Waskiewicz AJ, Flynn A, Proud CG, Cooper JA. 1997. Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. *Embo j* **16**: 1909-1920.

- Waskiewicz AJ, Johnson JC, Penn B, Mahalingam M, Kimball SR, Cooper JA. 1999. Phosphorylation of the cap-binding protein eukaryotic translation initiation factor 4E by protein kinase Mnk1 in vivo. *Mol Cell Biol* **19**: 1871-1880.
- Weingarten-Gabbay S, Elias-Kirma S, Nir R, Gritsenko AA, Stern-Ginossar N, Yakhini Z, Weinberger A, Segal E. 2016. Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science* **351**.
- Weis K, Hondele M. 2022. The Role of DEAD-Box ATPases in Gene Expression and the Regulation of RNA-Protein Condensates. *Annual Review of Biochemistry* **91**: 197-219.
- Weisser M, Voigts-Hoffmann F, Rabl J, Leibundgut M, Ban N. 2013. The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A. *Nat Struct Mol Biol* **20**: 1015-1017.
- Wek RC. 2018. Role of eIF2 α Kinases in Translational Control and Adaptation to Cellular Stress. *Cold Spring Harb Perspect Biol* **10**.
- Wells SE, Hillner PE, Vale RD, Sachs AB. 1998. Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* **2**: 135-140.
- Welnowska E, Sanz MA, Redondo N, Carrasco L. 2011. Translation of viral mRNA without active eIF2: the case of picornaviruses. *PLoS One* **6**: e22230.
- White JP, Cardenas AM, Marissen WE, Lloyd RE. 2007. Inhibition of cytoplasmic mRNA stress granule formation by a viral proteinase. *Cell Host Microbe* **2**: 295-305.
- White SD, Jacobs BL. 2012. The amino terminus of the vaccinia virus E3 protein is necessary to inhibit the interferon response. *J Virol* **86**: 5895-5904.
- Wilson JE, Pestova TV, Hellen CU, Sarnow P. 2000. Initiation of protein synthesis from the A site of the ribosome. *Cell* **102**: 511-520.
- Wolfe AL, Singh K, Zhong Y, Drewe P, Rajasekhar VK, Sanghvi VR, Mavrakis KJ, Jiang M, Roderick JE, Van der Meulen J et al. 2014. RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* **513**: 65-70.
- Wolin SL, Walter P. 1988. Ribosome pausing and stacking during translation of a eukaryotic mRNA. *Embo j* **7**: 3559-3569.
- Wu HY, Guan BJ, Su YP, Fan YH, Brian DA. 2014. Reselection of a genomic upstream open reading frame in mouse hepatitis coronavirus 5'-untranslated-region mutants. *J Virol* **88**: 846-858.
- Xi Q, Cuesta R, Schneider RJ. 2004. Tethering of eIF4G to adenoviral mRNAs by viral 100k protein drives ribosome shunting. *Genes Dev* **18**: 1997-2009.
- . 2005. Regulation of translation by ribosome shunting through phosphotyrosine-dependent coupling of adenovirus protein 100k to viral mRNAs. *J Virol* **79**: 5676-5683.
- Xia X. 2021. Domains and Functions of Spike Protein in Sars-Cov-2 in the Context of Vaccine Design. *Viruses* **13**.
- Yamamoto H, Unbehau A, Loerke J, Behrmann E, Collier M, Bürger J, Mielke T, Spahn CM. 2014. Structure of the mammalian 80S initiation complex with initiation factor 5B on HCV-IRES RNA. *Nat Struct Mol Biol* **21**: 721-727.
- Yan WX, Chong S, Zhang H, Makarova KS, Koonin EV, Cheng DR, Scott DA. 2018. Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein. *Mol Cell* **70**: 327-339.e325.

- Yang D, Wilson JE, Anderson DR, Bohunek L, Cordeiro C, Kandolf R, McManus BM. 1997. In vitro mutational and inhibitory analysis of the cis-acting translational elements within the 5' untranslated region of coxsackievirus B3: potential targets for antiviral action of antisense oligomers. *Virology* **228**: 63-73.
- Yang Q, Sarnow P. 1997. Location of the internal ribosome entry site in the 5' non-coding region of the immunoglobulin heavy-chain binding protein (BiP) mRNA: evidence for specific RNA-protein interactions. *Nucleic Acids Res* **25**: 2800-2807.
- Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen LL, Wang Y et al. 2017. Extensive translation of circular RNAs driven by N(6)-methyladenosine. *Cell Res* **27**: 626-641.
- Yao N, Reichert P, Taremi SS, Prosser WW, Weber PC. 1999. Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure* **7**: 1353-1363.
- Yi SH, Petrychenko V, Schliep JE, Goyal A, Linden A, Chari A, Urlaub H, Stark H, Rodnina MV, Adio S et al. 2022. Conformational rearrangements upon start codon recognition in human 48S translation initiation complex. *Nucleic Acids Res* **50**: 5282-5298.
- Yoder-Hill J, Pause A, Sonenberg N, Merrick WC. 1993. The p46 subunit of eukaryotic initiation factor (eIF)-4F exchanges with eIF-4A. *J Biol Chem* **268**: 5566-5573.
- Young SK, Wek RC. 2016. Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. *J Biol Chem* **291**: 16927-16935.
- Young SK, Willy JA, Wu C, Sachs MS, Wek RC. 2015. Ribosome Reinitiation Directs Gene-specific Translation and Regulates the Integrated Stress Response. *J Biol Chem* **290**: 28257-28271.
- Yuan S, Peng L, Park JJ, Hu Y, Devarkar SC, Dong MB, Shen Q, Wu S, Chen S, Lomakin IB et al. 2020. Nonstructural Protein 1 of SARS-CoV-2 Is a Potent Pathogenicity Factor Redirecting Host Protein Synthesis Machinery toward Viral RNA. *Mol Cell* **80**: 1055-1066.e1056.
- Yueh A, Schneider RJ. 2000. Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to 18S rRNA. *Genes Dev* **14**: 414-421.
- Zaborowska I, Kellner K, Henry M, Meleady P, Walsh D. 2012. Recruitment of host translation initiation factor eIF4G by the Vaccinia Virus ssDNA-binding protein I3. *Virology* **425**: 11-22.
- Zaffagni M, Harris JM, Patop IL, Pamudurti NR, Nguyen S, Kadener S. 2022. SARS-CoV-2 Nsp14 mediates the effects of viral infection on the host cell transcriptome. *Elife* **11**.
- Zerbe CM, Cole JL. 2020. Regulation of Protein Kinase R by Epstein-Barr Virus EBER1 RNA. *Biochemistry* **59**: 1252-1260.
- Zhang B, Morace G, Gauss-Müller V, Kusov Y. 2007. Poly(A) binding protein, C-terminally truncated by the hepatitis A virus proteinase 3C, inhibits viral translation. *Nucleic Acids Res* **35**: 5975-5984.
- Zhang H, Dou S, He F, Luo J, Wei L, Lu J. 2018. Genome-wide maps of ribosomal occupancy provide insights into adaptive evolution and regulatory roles of uORFs during *Drosophila* development. *PLoS Biol* **16**: e2003903.

- Zhang K, Miorin L, Makio T, Dehghan I, Gao S, Xie Y, Zhong H, Esparza M, Kehrer T, Kumar A et al. 2021. Nsp1 protein of SARS-CoV-2 disrupts the mRNA export machinery to inhibit host gene expression. *Sci Adv* **7**.
- Zhang K, Xie Y, Muñoz-Moreno R, Wang J, Zhang L, Esparza M, García-Sastre A, Fontoura BMA, Ren Y. 2019. Structural basis for influenza virus NS1 protein block of mRNA nuclear export. *Nature Microbiology* **4**: 1671-1679.
- Zhang X, Gao X, Coots RA, Conn CS, Liu B, Qian SB. 2015. Translational control of the cytosolic stress response by mitochondrial ribosomal protein L18. *Nat Struct Mol Biol* **22**: 404-410.
- Zhang Y, Dolph PJ, Schneider RJ. 1989. Secondary structure analysis of adenovirus tripartite leader. *J Biol Chem* **264**: 10679-10684.
- Zhao J, Li Y, Wang C, Zhang H, Zhang H, Jiang B, Guo X, Song X. 2020. IRESbase: A Comprehensive Database of Experimentally Validated Internal Ribosome Entry Sites. *Genomics Proteomics Bioinformatics* **18**: 129-139.
- Zhao K, Ke Z, Hu H, Liu Y, Li A, Hua R, Guo F, Xiao J, Zhang Y, Duan L et al. 2021a. Structural Basis and Function of the N Terminus of SARS-CoV-2 Nonstructural Protein 1. *Microbiol Spectr* **9**: e0016921.
- Zhao M-M, Yang W-L, Yang F-Y, Zhang L, Huang W-J, Hou W, Fan C-F, Jin R-H, Feng Y-M, Wang Y-C et al. 2021b. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduction and Targeted Therapy* **6**: 134.
- Zhao Y, Pang TY, Wang Y, Wang S, Kang HX, Ding WB, Yong WW, Bie YH, Cheng XG, Zeng C et al. 2014. LMP1 stimulates the transcription of eIF4E to promote the proliferation, migration and invasion of human nasopharyngeal carcinoma. *Febs j* **281**: 3004-3018.
- Zheng Y, Deng J, Han L, Zhuang MW, Xu Y, Zhang J, Nan ML, Xiao Y, Zhan P, Liu X et al. 2022. SARS-CoV-2 NSP5 and N protein counteract the RIG-I signaling pathway by suppressing the formation of stress granules. *Signal Transduct Target Ther* **7**: 22.
- Zhou J, Wan J, Gao X, Zhang X, Jaffrey SR, Qian SB. 2015. Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* **526**: 591-594.
- Zitomer RS, Walthall DA, Rymond BC, Hollenberg CP. 1984. *Saccharomyces cerevisiae* ribosomes recognize non-AUG initiation codons. *Mol Cell Biol* **4**: 1191-1197.
- Ziv O, Price J, Shalamova L, Kamenova T, Goodfellow I, Weber F, Miska EA. 2020. The Short- and Long-Range RNA-RNA Interactome of SARS-CoV-2. *Mol Cell* **80**: 1067-1077.e1065.

Annexes

1 Annexe I

Revue contrôle traductionnel des coronavirus

Translational control of coronaviruses

Sylvain de Breyne *, Caroline Vindry, Olivia Guillin, Lionel Condé , Fabrice Mure, Henri Gruffat, Laurent Chavatte and Théophile Ohlmann*

SURVEY AND SUMMARY

Translational control of coronaviruses

Sylvain de Breyne¹*, Caroline Vindry, Olivia Guillin, Lionel Condé, Fabrice Mure, Henri Gruffat, Laurent Chavatte and Théophile Ohlmann*

CIRI, Centre International de Recherche en Infectiologie, Univ Lyon, INSERM U1111, Université Claude Bernard Lyon 1, CNRS UMR5308, ENS de Lyon, F-69007, Lyon, France

Received October 02, 2020; Revised October 29, 2020; Editorial Decision October 30, 2020; Accepted November 03, 2020

ABSTRACT

Coronaviruses represent a large family of enveloped RNA viruses that infect a large spectrum of animals. In humans, the severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) is responsible for the current COVID-19 pandemic and is genetically related to SARS-CoV and Middle East respiratory syndrome-related coronavirus (MERS-CoV), which caused outbreaks in 2002 and 2012, respectively. All viruses described to date entirely rely on the protein synthesis machinery of the host cells to produce proteins required for their replication and spread. As such, virus often need to control the cellular translational apparatus to avoid the first line of the cellular defense intended to limit the viral propagation. Thus, coronaviruses have developed remarkable strategies to hijack the host translational machinery in order to favor viral protein production. In this review, we will describe some of these strategies and will highlight the role of viral proteins and RNAs in this process.

INTRODUCTION

Coronaviruses: an overview

Coronaviruses (CoVs) represent a large family of RNA enveloped viruses that contain a positive-sense and single-stranded RNA genome that is one of the largest of all RNA viruses. CoVs are usually associated with enteric or respiratory diseases in their hosts, although hepatic, neurologic and other organ systems may be affected. The current classification of coronaviruses recognizes 46 species in 26 subgenera, 5 genera and 2 subfamilies that belong to the family *Coronaviridae*, suborder *Cornidovirineae*, order *Nidovirales* and realm *Riboviria* (1). The *Orthocoronavirinae* subfamily

can be divided into four genera: α - and β -CoV that only infect mammals, δ - and γ -CoV that mainly target birds. In the late 1960s, the first human coronaviruses (HCoV) were identified (HCoV-229E, HCoV-814 and HCoV-OC43) and also later in the 2000s (HCoV-NL63 and HCoV-HKU1) (2–6); these strains generally cause mild symptoms such as a common cold. In contrast, severe acute respiratory syndrome coronavirus (SARS-CoV) (7–9) and Middle East Respiratory Syndrome coronavirus (MERS-CoV) (10–12) that emerged in 2002 and 2012, respectively, are highly pathogenic coronaviruses. The SARS-CoV-2 is currently at the origin of the COVID-19 pandemic that the world is going through with, as of 10 October 2020, >34 million confirmed cases resulting in >1 million deaths (13). The SARS-CoV-2 genome share 80% and 50% homology with SARS-CoV and MERS-CoV genomes respectively (14–16) and all of them belong to the β -CoV genera. This genus also includes animal strains such as the Mouse Hepatitis Virus (MHV) (17) or the Bovine-CoV (B-CoV) (18,19) but also two human strains (HCoV-OC43 and HCoV-HKU1).

Coronavirus infections begin with interaction between virions and their cellular receptors at the surface of the cells (Figure 1). After release of the nucleocapsid into the cytoplasm, the viral genome serves as mRNA for translation to generate the viral replicase. The genomic viral RNA (gRNA) contains >12 open reading frames (ORFs) and even 14 in the case of SARS-CoV-2 to give rise to about 27 proteins (14). In order to express as many as 27 different proteins, coronaviruses have put in place three different mechanisms. First, when the viral genome enters the cytoplasm, the first ORF (ORF1) is translated into a long polyprotein which is then proteolytically processed to generate several nonstructural proteins (nsp). In addition, this polyprotein is generated in two different forms through the use of a programmed-1 ribosome frameshifting event. This mechanism requires a *cis*-acting RNA element within the coding region that can redirect elongating ribosomes to shift back to the first reading frame by 1 base in the 5' direc-

*To whom correspondence should be addressed. Tel: +33 472 728 893; Fax: +33 472 728 137; Email: sylvain.de.breyne@ens-lyon.fr
Correspondence may also be addressed to Théophile Ohlmann. Tel: +33 472 728 953; Email: theophile.ohlmann@ens-lyon.fr

© The Author(s) 2020. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

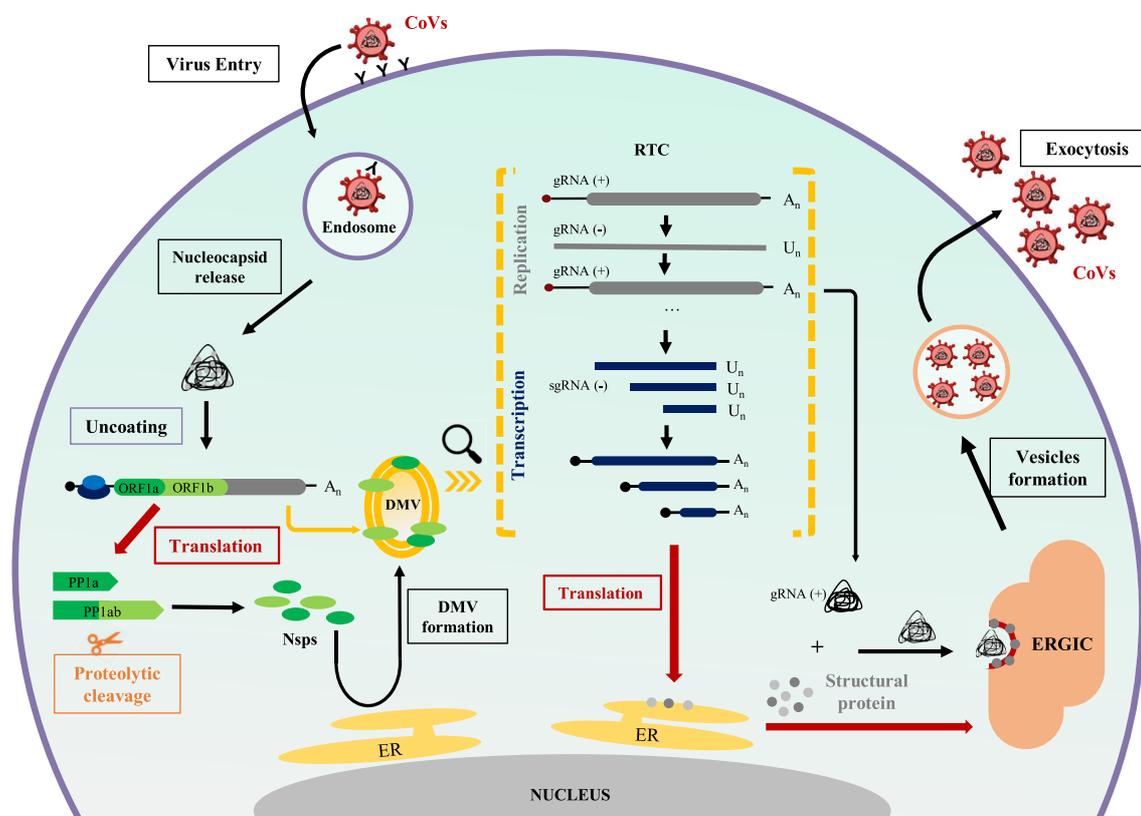


Figure 1. Scheme of the coronaviral replication cycle. After attachment to the cellular receptor, the viral particle is internalized and the gRNA is uncoated and directly translated in the cytoplasm to produce up to 16 nonstructural proteins (nsps) involved in viral proteolytic cleavage, the formation of the viral replication transcription complex (RTC) and in mRNA translational control. Viral genome replication and transcription of the different sgRNAs occur into double membrane vesicles (DMV) that are derived from the endoplasmic reticulum (ER). Some of the sgRNAs are coding for the structural proteins that are mainly expressed through the ER. Viral components (gRNA and viral proteins) assemble at the level of the ERGIC and new virions produced are secreted by exocytosis.

tion. This mechanism was described for the first time in 1985 in the Rous sarcoma virus (20) and 2 years later in the coronavirus Avian Infectious Bronchitis Virus (IBV) (21). All coronaviruses, including SARS-CoV-2, utilize this molecular mechanism (22–27). Therefore, ORF1 is separated into two distinct ORFs: ORF1a and ORF1b. In general, translation of ORF1a allows production of the immediate early proteins involved in the control of the host cellular innate immune response whereas translation of ORF1b generates proteins involved in genome replication and RNA synthesis. The other ORFs, which are expressed later during the viral replication cycle, are translated from distinct subgenomic RNAs (sgRNAs) produced by the synthesis of the minus-strand replicative intermediate, followed by synthesis of new plus-strand sgRNAs that mostly encode structural and accessory proteins. This represents the third strategy used for the translation of the viral ORFs.

Coronavirus RNA synthesis proceeds by a very complex and only partially understood mechanism. The gRNA not only serves as a matrix for the synthesis of the repli-

case polyproteins but also acts as the template for synthesis of negative-sense RNA species (Figure 1). These negative sense subgenomic RNAs (sgRNAs) are used as template for the generation of the positive-sense sgRNAs that are produced by a discontinuous, cotranscriptional process. Then, each positive-sense sgRNA serves as mRNA for translation of each ORF downstream of ORF1. All the sgRNAs are 3' coterminal and contain the same ~70 nucleotides (nt) long 5'-leader. The sites of leader-to-body fusion in the sgRNAs occur at loci in the genome called transcription-regulating sequences (TRSs) that contain a short element that is identical, or nearly identical, to a sequence located at the 5'-end of the gRNA. TRSs contain a conserved 6–7 nt core sequence fairly well conserved within each coronavirus group and surrounded by variable sequences. The viral polymerase, starting from the 3'-end of a genomic RNA, pauses and switches templates at an internal TRS to pursue synthesis at the homologous TRS of the leader gRNA. The resulting negative-strand sgRNA then serves as template for synthesis of multiple copies of the corresponding positive-

strand sgRNAs (28). Transcriptomic analysis of cells infected by coronavirus shows that the majority (>60%) of the reads mapped to viral sequences indicating that viral transcripts become largely predominant over the cellular transcriptome (14). In addition to these viral transcripts, it has been shown that several viral RNAs are produced by different noncanonical recombination events leading to the generation of transcripts harboring different deletions (14,29–31). However, it is not known whether these viral RNAs, which are relatively abundant, can be translated into functional proteins.

One common feature among positive-strand RNA viruses is the assembly of their replication-transcription complexes (RTCs) that is intimately associated with the formation of membranous rearrangements to create virus replication organelle (Figure 1) (32–34). The RTCs are interconnected double membrane vesicles (DMV) that are derived from the endoplasmic reticulum (ER) (35–39). The precise role of these vesicles is not known but they seem to be beneficial to the virus creating a protective microenvironment for the viral genome from attacks by antiviral mechanisms or nucleases present in the cytoplasm; in addition, it provides an environment to concentrate the factors necessary for efficient transcription and replication of the viral genome. As a consequence, formation of infectious progeny virions mainly occurs into the endoplasmic reticulum-golgi intermediate compartment (ERGIC) (Figure 1). After replication, the viral genome is encapsidated by the N protein into the membrane of the ERGIC in which the viral structural proteins are included (40). The viral particles are then released from the cell by a process very similar to exocytosis (Figure 1). Further details on the replication cycle of coronaviruses can be found on these reviews (41–45).

Translation initiation in eukaryotes

Translation is a key step in gene expression in which the messenger RNA (mRNA) is decoded by ribosomes to produce proteins. The eukaryotic mRNA is usually capped at the 5' end and polyadenylated at the 3' end (Figure 2). The ORF codes for the polypeptide and is surrounded by two untranslated regions (UTRs) that control translation initiation by promoting ribosome attachment to the mRNA and the recognition of the initiation codon (AUG codon). The majority of mRNAs uses the cap-dependent mechanism to initiate translation in which eukaryotic initiation factors (eIFs) recognize and bind to the 5' cap structure to bridge the mRNA to the 40S ribosomal subunit (Figure 2). This process requires eIF4F that is a three-subunits complex composed of eIF4E, the cap-binding protein; eIF4A, an ATP dependent RNA helicase which belongs to the DEA(D/H)-box family; and eIF4G, a scaffold protein. EIF4G has binding sites for eIF4A, eIF4E, the poly(A) binding protein (PABP) and eIF3. The eIF4F complex recognizes and binds to the 5' cap structure and unwinds local RNA structures required for adequate positioning of the 40S ribosomal subunit on the mRNA. The 40S ribosomal subunit is recruited to the mRNA as part of the 43S initiation complex that is composed of the 40S ribosome bound to eIF2-GTP/Met-tRNA_i (initiator tRNA), eIF1A, eIF1, eIF3 and eIF5 (46–49). In this conforma-

tion, eIF4G constitutes a real scaffold between the mRNA with the cap structure through eIF4E, and the 40S ribosomal subunit via eIF3. In eukaryotes, the mRNA is pseudocircularized by an interaction between eIF4G at the 5' end and PABP at the 3' end and this contributes to enhance translational efficiency (50–52). Upon its recruitment to the 5' end of the mRNA, the 43S ribosomal complex scans the mRNA in a 5'-3' direction in search of the initiation codon, which is often the first AUG triplet downstream to the 5' cap. During the ribosomal scanning, the GTP bound on eIF2 is hydrolyzed to GDP and phosphate (Pi) in an eIF5-dependent mechanism (53,54). The recognition of the initiation codon by the anticodon of the Met-tRNA_i stops the scanning process and provokes Pi and eIFs release from the ribosomal complex that allows the 60S ribosomal to join and form the 80S ribosome positioned at the initiation codon (Figure 2) (55–57). The guanine nucleotide exchange factor eIF2B promotes the recycling of eIF2-bound GDP into eIF2-bound GTP; as a result, the ternary complex re-assembles for a new round of translation initiation. eIF2 activity is regulated by phosphorylation event that takes place on the serine 51 of the α -subunit (Figure 2). In mammals, four kinases have been well characterized and are activated by stress conditions such as amino acid deprivation (for the GCN2 kinase), ER stress (for the PERK kinase), the presence of dsRNA (for the PKR kinase) or heme deficiency (for the HRI kinase) (58). During viral infection, PKR and PERK are usually activated resulting in a global translation arrest and the formation of stress granules (SGs) (59–61).

Cis-acting elements can also influence initiation codon recognition rate by the 43S complex. Two of them are the nucleotide context of the AUG codon and the presence of short upstream ORFs (uORFs). An optimal consensus sequence around the initiation codon was defined as 5'-A/GccAUGG-3' in which positions -3 and +4 according of the adenosine of the AUG codon are determinant for initiation codon recognition (62–65). When the context is not optimal, a fraction of the scanning ribosomes can bypass the AUG codon to go to the next downstream AUG triplet. This is called leaky scanning and it results in the production of two proteins from two different ORFs or two isoforms of the same protein. However, re-initiation of translation can occur if the length of the first ORF is shorter than 30 amino acids (uORF) and if the main ORF is found approximately 100–200 bp downstream of the stop codon (66). In this case, the different eIFs usually stay in the vicinity of the 40S ribosomal subunit to promote re-initiation. As a consequence, uORFs can regulate the flow of ribosomes on the mRNA (67,68), and this usually results in the reduction of translational efficiency of the main reading frame (69).

Examples of translational control during viral infections

Most of the time, viral infection activates the first host defense mechanism that triggers the production of interferons (IFN) and IFN stimulated genes (70). To counteract this first line of host antiviral innate immunity, viruses have developed strategies to inhibit or escape to this restriction (71) and components of the IFN signal transduction pathway are often inactivated during viral infection (72). As such, some RNA viruses, including *Picornaviruses*, *Orthomyx-*

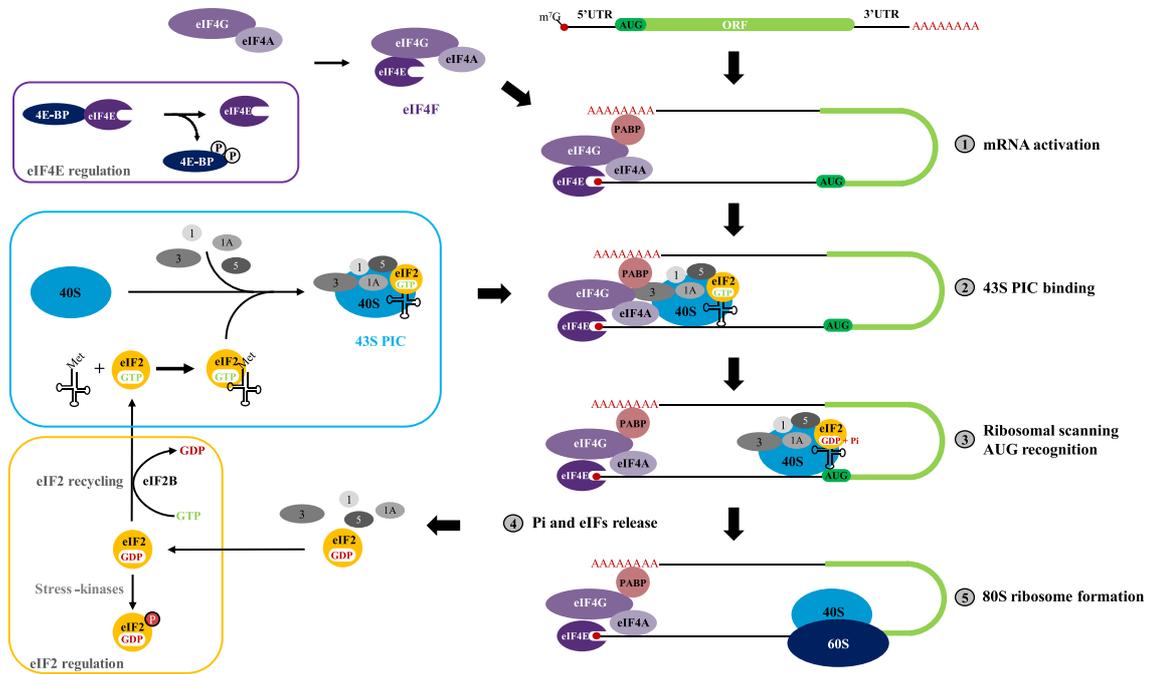


Figure 2. The cap-dependent translation initiation mechanism. (1) Capped and polyadenylated mRNA is activated by the eIF4F complex, composed of eIF4E, eIF4G and eIF4A, to prepare ribosomal landing. 43S ribosomal complex is formed by the association of eIF1, eIF1A, eIF3, eIF5 and the ternary complex eIF2-GTP-Met-tRNA_i with the 40S small ribosomal subunit (see blue square). (2) Attachment of the 43S complex to the mRNA is mediated by multiple protein interactions requiring eIF4E, eIF4G and eIF3. (3) Ribosomal complex scans the 5'UTR with the hydrolysis of GTP into GDP + Pi mediated by eIF5 until it reaches the initiation codon, where its recognition triggers to (4) Pi and eIFs release followed by (5) the association of the 60S and 40S ribosomal subunits to form an 80S ribosome competent for the elongation step. The most known translational controls occur on eIF4E and eIF2. According to the phosphorylation state of 4E-BP, 4E-BP interacts or not with eIF4E and this interaction displaces eIF4E from eIF4F resulting in a decrease of the cap dependent translation efficiency (see violet square). Regeneration of the GDP molecule bound to eIF2 into GTP is catalyzed by the exchange factor eIF2B. Phosphorylation of eIF2 α at a serine in position 61 inhibits this regeneration that also induces a strong reduction of global translation (see orange square).

oviruses, *Coronaviruses* and many others, combine an additional mechanism in which a host translational shutoff is imposed to prevent the infected cells to synthesize new proteins and, amongst those, the IFN-stimulated genes (73). However, such strategy should not affect the expression of viral mRNAs to allow efficient viral replication to spread and continue. Therefore, most of the time, viruses have developed alternative translation initiation mechanisms to ensure continuous viral protein production.

For instance, during *picornaviruses* infection, epitomized by the encephalomyocarditis virus (EMCV) and the poliovirus (PV), a rapid shutoff of the host gene expression is established when viral proteases target and cleave eIF4G, PABP and eIF5B that are required for cap-dependent translation initiation (74–77). This results in a specific inhibition of cap-dependent translation of cellular mRNAs whereas the 5'UTR of the picornaviral mRNAs harbors an internal ribosome entry site (IRES) that enables the recruitment of the ribosomal complex in a cap-independent manner (78). An alternative strategy is used by influenza virus where the viral polymerase complex binds to cellular mRNAs and cleaves few nucleotides downstream to the cap

structure to produce a capped RNA fragment that is subsequently used to prime viral transcription (79,80). The decapped host mRNAs are targeted to degradation that leads to the downregulation of cellular protein production. Finally, translation of the alphavirus mRNAs is resistant to the phosphorylation of eIF2 α that results from PKR activation (81). This is due to a stable RNA hairpin loop structure in the viral genome that stalls the ribosome on the correct AUG and allows translation of the late viral mRNAs (82).

Global translation of cellular mRNAs is not spared during infection with CoVs. Indeed, studies in MHV-infected cells indicate that a strong shutoff of host gene expression is established during the first hours after infection while viral protein expression is not affected (83–87). This inhibition of cellular mRNA expression has been reported in cells infected not only with others coronaviral strains such as in bovines, bats, chickens and pigs (88–92) but also with human CoVs like the HCoV-229E, HCoV-NL43, SARS-CoV and MERS-CoV strains (93–96). Recent studies show that translation repression also occurs in SARS-CoV-2 infected cells (97–99) and analysis of the SARS-CoV-2 tran-

scriptome indicates that host gene expression is almost suppressed in the first hours of infection (14). Most of the viral structural proteins are expressed in the ER that contributes to create an ER stress (100). Furthermore, lipid depletion from the ER which is due to viral replication organelle formation activates the unfolded protein response (UPR) that results to the expression of chaperone proteins, phospholipid biogenesis and translation downregulation (101). Nevertheless, the UPR is usually controlled by viral components to avoid a drastic reduction of global protein synthesis. The first analyses of the SARS-CoV-2 host interacting proteins revealed that many viral proteins target components belonging to the translational machinery but also key elements involved in SG formation and the ER stress response (102).

In this review, we will describe how mRNA translation is controlled in the course of coronaviral infection. Most of the studies retained for this review concern SARS-CoV, MERS-CoV and SARS-CoV-2 strains. In the first part, we will briefly present the different translational pathways that are activated following infection and how viral components can interfere. A large part of this review will focus on viral proteins, especially nsp1, which directly interact with components of the ribosomal complex to promote translation inhibition and mRNA decay. The second part is dedicated to the mechanisms used by viral mRNAs to escape from host protein synthesis inhibition. Finally, the last part will deal with the spectrum of translational mechanisms used by viral mRNAs to be expressed.

TRANSLATIONAL CONTROL BY CORONAVIRUSES

Viral infection generally induces several lines of cellular responses such as innate immunity, apoptosis, activation of different mitogen-activated protein kinases (MAPK) pathway or translational reprogramming. As these different mechanisms during CoV infection have been recently reviewed (103), we will first briefly describe how these different translational control pathways are activated during infection and point out some viral proteins involved in these mechanisms. Next, we will mainly focus on viral proteins expressed from SARS-CoV, MERS-CoV and SARS-CoV-2 strains, which directly alter the integrity of components of the initiation complex. A large part of this section deals with nsp1 and its function in host protein synthesis inhibition as it has been the best studied example.

Interference with the stress activated pathways

CoV replication generates in the cytoplasm a large quantity of double strand RNAs such as positive and negative gRNAs and sgRNAs (Figure 1); this typical signature of RNA virus replication contributes to the activation of PKR. In addition, several aspects of the CoV infection are expected to disrupt ER homeostasis. (i) All CoV structural proteins, except for the N protein, are synthesized, folded and modified in the ER (104–106). (ii) In order to produce new virions, the amount of viral proteins produced largely exceeds the ER folding capacity (100,107). (iii) Viral replication takes place in double membrane vesicles complexes hijacked from the ER and this results in the alteration of its home-

ostasis (Figure 1) (35,108). (iv) The assembly and budding of mature particles of coronaviruses occur into the ERGIC (109) that induces a massive budding accompanied by depletion in the ER membrane (Figure 1). All these different stresses of the ER lead to the activation of the UPR. To restore ER homeostasis, three transmembrane sensors are activated: PERK, inositol-requiring enzyme 1 and activating transcription factor 6. The last two are not directly related to translational control and we invite the reader to consult these reviews (103,110–112), and we will focus on the relationship that may exist between CoVs and both PKR and PERK.

As expected, both PERK and PKR are activated during SARS-CoV infection that in turn phosphorylate the serine 51 of eIF2 α leading to a global translation inhibition (Figure 3A) (113). However, inhibition of PKR alone did not alter the phosphorylation status of eIF2 α meaning that it is essentially controlled by PERK (113). The activation of PKR seems important for other aspects of viral infection, such as virally induced apoptosis that could contribute to the translational inhibition (113). The overexpression of several SARS-CoV proteins alone, including 3a and S, is sufficient to induce an ER stress and the activation of PERK (Figure 3A) (100,114). Expression of the spike protein from MHV or HCoV-HKU1 strains also induces an ER stress (107,115) suggesting a conserved mechanism of PERK activation among CoVs. Interestingly, the phosphorylation of eIF2 α is not such a robust feature for other coronavirus infections as it depends on the cell line used and it can only be a transient phenomenon. For example, infection with MERS-CoV strain induces a low but detectable level of phosphorylated eIF2 α only in Vero cells, but no response in HeLa/CD26 cells (116). In addition, the removal of the gene coding for 4a and 4b proteins in the viral genome of the MERS-CoV (Δ p4 mutant strain) induces the phosphorylation of eIF2 α , activation of PKR and reduces viral replication in HeLa cells suggesting that these proteins are involved in circumventing translational attenuation during infection (Figure 3A). Surprisingly, no difference was observed between MERS-CoV Δ p4 and the wild-type virus in Vero cells, suggesting a cell line specific regulation rather than a solid feature of MERS-CoV infection (116). To the same extent, nsp15 of some CoVs (MHV, HCoV-229E) could also suppress PKR activation (Figure 3A) (117–119). Additionally, during HCoV-OC43 infection of neurons, only an early and transient phosphorylation of eIF2 α was detected (120). There is no doubt that CoV infection is related to the activation of both PKR and PERK that results in the phosphorylation of eIF2 α . Whether the novel SARS-CoV2 also acts on eIF2 α phosphorylation to induce a global translation regulation remains to be investigated.

Translational inhibition induced eIF2 α phosphorylation leads to the formation of SGs and processing bodies (P-bodies). SGs are messenger ribonucleoproteins (mRNPs) composed of mRNAs stalled at the translation initiation step while P-bodies are mRNPs containing untranslated mRNAs associated with proteins involved in mRNA decay and translation repression. These two cytoplasmic granules are dynamic complexes which compact mRNAs for storage, reinitiation of translation or degradation (121–124).

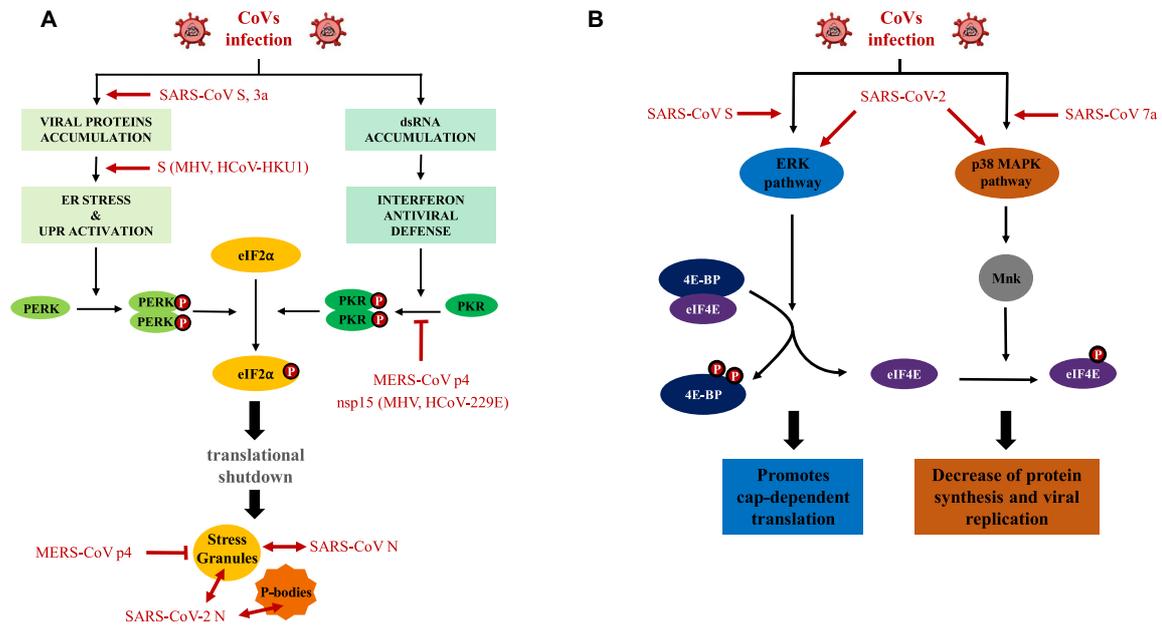


Figure 3. The translational pathways regulation during CoVs infection. (A) Double strand RNA (dsRNA) accumulation produced during the viral replication is one sensor of the interferon antiviral defense and leads to the phosphorylation and activation of PKR. Viral proteins accumulation in the endoplasmic reticulum (ER) constitutes an ER stress and triggers the unfold protein response (UPR). As a result, PERK is phosphorylated and activated. Both PKR and PERK target the α subunit of eIF2 and its phosphorylation induces a global translational shutdown. The stress-activated phosphorylation of eIF2 can lead to the formation of stress granules (SGs) in which mRNAs stalled in translation initiation step are stored. SGs are connected with P-bodies structures as they share RNA and protein components. (B) Modulation of the p38MAPK and ERK pathways. ERK pathway controlled the phosphorylation state of 4E-BP and promotes the dissociation between 4E-BP and eIF4E that enhances the cap-dependent translation. p38 MAPK pathway targets the MAP kinase-interacting serine/threonine-protein kinase (Mnk) that phosphorylates eIF4E. As a consequence and in the context of CoVs infection, activation of this pathway is associated with a decrease of protein synthesis and viral replication. Viral components modulating the pathway are bolded in red.

The assembly of SGs is a cellular response frequently observed upon viral infections but its real contribution to the anti-viral response remains to be established (125). In many cases, viral mRNAs and/or viral genomes are recruited in SGs (125,126). In the particular case of CoV infection, the function of SG and P-bodies has not been fully investigated although an increase in the formation of cellular granules has been observed during the Transmissible GastroEnteritis Virus (TGEV) and MHV infections which correlate with an increase of eIF2 α phosphorylation, translational shutoff and mRNA decay (87,127). During infection with SARS-CoV, the N protein can translocate into SGs (Figure 3A) (128) whereas formation of SGs is not observed in the course of MERS-CoV infection. Nevertheless, mutant viruses lacking either the gene coding for 4a or 4b protein (MERS-CoV Δ p4), become able to induce formation of SGs suggesting that these proteins might be involved in their repression (116,119). Notably, the formation of SGs in cells infected with MERS-CoV Δ p4 is associated with a reduction of viral replication (Figure 3A). Recently, mass spectrometry analysis of the viral-host proteome interaction networks was obtained with overexpressed viral proteins from SARS-CoV-2. Interestingly, the N protein interacts with several components of the SGs (G3BP1 and G3BP2) and P-bodies (Upf1, MOV10, PABPC1 and PABPC2) machineries (Figure 3A) (102). However, the role of these cy-

toplasmic granules as pro- or anti-viral mechanism remains elusive.

Coronavirus infections activates other stress-induced signaling pathways such as the extracellular signal-regulated kinase (ERK) and the p38 MAPK, both of them are known to influence the translational activity of the cell (103). Activation of p38 MAPK is observed during SARS-CoV and MHV infections (Figure 3B) and this leads to the phosphorylation of eIF4E through the MAP kinase-interacting serine/threonine-protein kinase (Mnk) (129–131). The effect of this phosphorylation is quite controversial with some articles showing that phosphorylation of eIF4E is not required for translation (132) whereas in some cases, it can inhibit cap-dependent translation (133). Despite this controversy, most of the data suggest that eIF4E phosphorylation promotes translation initiation (134–137). In the case of CoV, activation of the p38 MAPK pathway is associated with an attenuation of overall cellular protein synthesis. Interestingly, expression of the SARS-CoV accessory protein 7a alone can inhibit cellular protein synthesis and induce p38 MAPK phosphorylation (Figure 3B) (138,139). However, molecules that target the p38 MAPK neither alter SARS-CoV replication nor viral protein synthesis (130). Similarly, infection with HCoV-229E also induces the phosphorylation of eIF4E and the use of a p38 MAPK inhibitor in human fetal lung cells L132 reduces the replication of the

virus (140). To date, the implication of eIF4E phosphorylation during CoV infection is challenging and remains to be explored. Finally, inhibitors that control the phosphorylation of eIF4E, such as tomivosertib, have recently been highlighted to be of interest in the case of treatment during SARS-CoV-2 infection (102) and this confirms the importance of the stress-induced signaling pathways.

In parallel, the ERK signaling cascade is activated during CoV infection (Figure 3B) and seems to counterbalance the action of p38 MAPK in protein synthesis. Indeed, activation of ERK leads to the phosphorylation of eIF4E-binding proteins (4E-BP) that are repressor proteins able to interact directly with the cap-binding protein eIF4E. Phospho-4E-BP cannot bind to eIF4E and sequesters it from being incorporated into the eIF4F complex (Figure 2) (141). The activation of the ERK pathway was observed during infection with SARS-CoV, MERS-CoV and HCoV-229E strains (129,142) but also after overexpression of SARS-CoV S protein (143). In the case of SARS-CoV-2, the investigation of the global phosphorylation landscape of Vero E6 cells reveals the activation of p38 MAPK and ERK pathways, and suggests an impact on cap-dependent translation (144). Clearly, the phosphorylation status of eIF4E and eIF4E-BPs needs further investigation to decipher the role of signaling pathway activation in SARS-CoV2 infection.

The nsp1 protein

The capped and polyadenylated viral genome is directly translated as soon as it is released in the cytoplasm after entry. As a result, two large precursor polyproteins are produced and processed by viral proteases to generate 16 nsps. Nsp1 is the most N-terminal cleavage product and is only expressed in the α -CoVs and β -CoVs genera but it strongly differs in size. Nsp1 of α -CoVs has a size of about 110 amino acids (aa) whereas in β -CoVs the size of nsp1 varies between the four different lineages (A, B, C and D) (145). The lineage A is epitomized by MHV, HCoV-OC43 and HCoV-HKU1, and nsp1 harbors about 245 aa residues. The lineage B includes SARS-CoV and SARS-CoV-2 in which nsp1 has a size of 180 aa. MERS-CoV belongs to the lineage C and encodes a 195 aa long nsp1. Finally, the lineage D regroups mainly bat-CoVs in which nsp1 has a size of 175 aa. Amino acid sequence homology varies between nsp1 from the different β -CoVs but the SARS-CoV-2 nsp1 shares 91% similarities with SARS-CoV nsp1 (102) suggesting a strong conservation of function and the first study on SARS-CoV-2 nsp1 confirmed its role in host protein synthesis inhibition (97–99). In this section, we mainly discuss about the function and mechanism of SARS-CoV nsp1 on host cellular translation during infection and extend these findings to others strains like MERS-CoV and SARS-CoV-2.

SARS-CoV nsp1. In SARS-CoV infected cells, nsp1 can be immunodetected as early as 6 h post infection (hpi) and it has a perinuclear and cytoplasmic localization (94,146). This coincides with the inhibition of host gene expression that starts at 6 hpi and becomes almost complete between 9–12 hpi (95). An ectopic and transient expression of nsp1 in mammalian cells is sufficient to promote a robust and global inhibition of host protein synthesis (94,95,147) with

a strong reduction of the polysome profile shifting to 80S monosomes (148). The deletion of a small region of the carboxy-terminal part of nsp1 (Δ 160–173) is sufficient to reverse the inhibitory effect of the latter and two charged amino acids, K164 and H165, have been identified to be critical for nsp1 function in this translational shutoff (95). Mutation of these two residues into alanine in the SARS-CoV genome to generate the SARS-CoV-mt did not significantly alter viral replication in terms of viral titer, viral RNA transcription or nsp1 expression but it failed to promote robust host endogenous mRNA translation inhibition (95). Partial inhibition of host gene expression was also observed in Vero E6 and 293/ACE2 cells infected by SARS-CoV-mt (95) indicating that either the nsp1-mutant was partially active in the context of infection or that other viral proteins may be involved in this process. Interestingly, SARS-CoV nsp1 targets endogenous cellular mRNAs but also mRNAs resulting from ectopic plasmid expression and transfected mRNAs (94,95). The host cell gene expression shutoff mediated by SARS-CoV nsp1 was not inhibited by actinomycin D treatment that blocked *de novo* RNA synthesis, confirming that nsp1 acts on nascent and pre-existing mRNAs (94,95).

The SARS-CoV nsp1 combines two strategies to inhibit host gene expression: it acts via (i) a direct inhibition of the translation initiation step and (ii) the induction of cellular mRNA degradation (Figure 4). Although translational arrest and cellular degradation are two linked processes, it remains possible to distinguish between them with the use of specific nsp1 mutants (see below). Several evidences indicate that nsp1 interacts with the small 40S ribosomal subunit and this contributes to the reduction of the overall amount of polysomes. This binding is abolished in the case of the specific mutations of the K164 and H165 residues of nsp1 (nsp1-mt) as evidenced by western blot analysis of each fraction in both cellular extracts and rabbit reticulocyte lysate (RRL) (148). In addition, a nsp1 wt-tagged protein, but not the mutant version, co-immunoprecipitates with both S6 ribosomal protein and 18S ribosomal RNA (rRNA) that are key architectural components of the 40S subunit (148). Interestingly, SARS-CoV nsp1 does not appear to be associated with the 80S ribosome suggesting that nsp1 binds exclusively at a pre-initiation stage and is released upon subunit joining. Addition of nsp1 in the RRL affected translation of reporter mRNAs and polysome profile analysis revealed that SARS-CoV nsp1 inhibited 80S ribosome but not 48S initiation complex formation with a critical role for the two charged amino acids in this process (148). As a consequence, translation driven by a cap- and IRES-dependent mechanisms were both inhibited by addition of SARS-CoV nsp1 in RRL. So was the cricket paralysis virus (CrPV) IRES that does not require any of the eIFs and can directly promote the assembly of an 80S ribosome at the AUG codon (149). Therefore, this demonstrates nsp1 targets directly the 40S ribosomal subunit rather than the associated initiation factors.

As mentioned above, SARS-CoV nsp1 also promotes degradation of the host cellular mRNAs in living cells (94,95). Although the mRNAs are also degraded in the RRL, it appears that this does not occur by the same pathway (148). Mutation of two charged and exposed amino acids R124 and K125 of nsp1 (named nsp1-CD for ‘cleav-

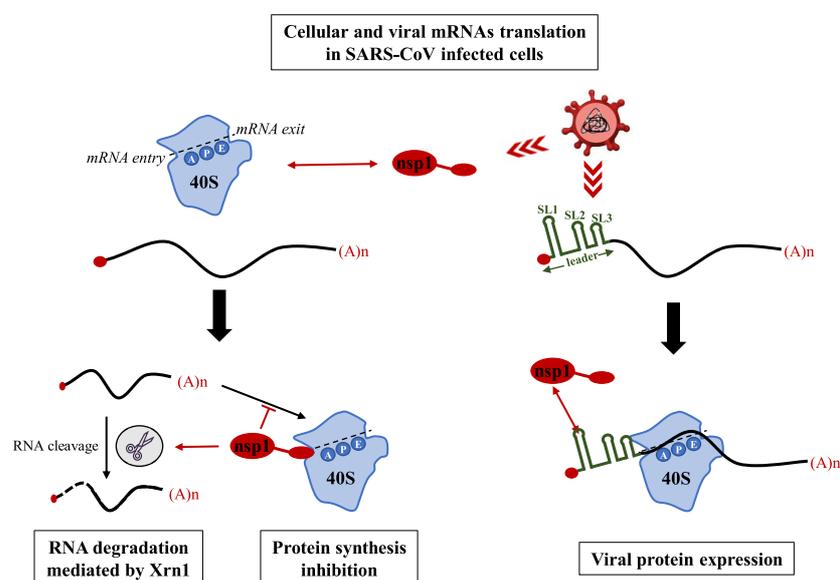


Figure 4. mRNA translation and mRNA decay regulated by SARS-CoV nsp1. A scheme of the 40S ribosomal subunit is shown with the mRNA entry and exit channel and the three tRNA sites: acceptor (A), peptidyl (P) and exit (E). Nsp1 is the first viral protein expressed and is able to interact, through the C-terminal domain, into the mRNA entry channel of the 40S ribosomal subunit. As a result, cellular mRNAs could not be recognized by the ribosome and cellular proteins expression is strongly reduced. In addition, nsp1 promotes endonucleolytic RNA cleavage that conducts the transcript to RNA degradation mediated by Xrn1. In parallel, all SARS-CoV mRNA harbor a common 5' terminal 72 nt leader sequence with three stem-loop (SL) (labeled in green). SL1 interacts with nsp1 and confers a translational resistance of the viral mRNA to the translation repression mediated by nsp1.

age defective') strongly reduces this cleavage without affecting translation repression both in cell culture and in the RRL suggesting that the two mechanisms are distinct (147). Sucrose gradient density and toeprinting assays are both complementary to measure the ratio of ribosomal complexes associated with the mRNA and to map the position of the ribosome on the mRNA at a nucleotide resolution. Both techniques were used to look at the effect of nsp1-CD on the formation of 48S and 80S initiation complexes on mRNAs translated by either a cap- or an IRES-dependent mechanism: formation of both complexes was inhibited in the RRL in the presence of nsp1-CD (147,148). Interestingly, translation of mRNAs in which initiation was driven by the CrPV or Hepatitis C Virus (HCV) IRESes was repressed by nsp1-wt without any evidence of an RNA cleavage (147,148) suggesting that RNA cleavage could also depend on the structure of the targeted RNA. Indeed, these two IRESes are highly structured (150) and directly bind to the 40S ribosome in the absence of any eIFs (149,151). Addition of nsp1-wt and nsp1-CD to the RRL is sufficient to abolish the binary complex formed between the 40S subunit and the CrPV IRES (147) suggesting a mutually exclusive interaction between the 40S, nsp1 and the IRES structure. In contrast, the interaction of 40S subunit with the HCV IRES was not impaired by the presence of nsp1-wt or nsp1-CD but formation of both 48S and 80S ribosome complexes at the AUG codon was inhibited (147). Cryogenic electron microscopy (cryo-EM) analyses revealed that the HCV IRES interacts with a region of the head and platform edge of the 40S subunit where it partially overlaps with

the exit-site (E-site) of the ribosome (152). The CrPV IRES also interacts with the head on the inter-subunit side (153) but, in this case, the CrPV IRES occupies all three tRNA binding sites on the ribosome that are the E-, the peptidyl (P-) and the decoding center (A-) sites (Figure 4, see the 40S ribosome scheme) (154). These difference of occupancy between both IRESes may explain why nsp1 can disrupt the binary complex formed between the 40S subunit and the IRES in the case of CrPV but not with HCV. In agreement with this hypothesis, both Thoms *et al.* and Schubert *et al.* have just shown by cryo-EM studies that nsp1 of a SARS-CoV-2 strain binds to the 40S at the mRNA entry channel (97,98) which is next to the A site (Figure 4). Recent results, obtained by Banerjee *et al.*, show that SARS-CoV-2 nsp1 interacts with the 18S rRNA in a region adjacent to the mRNA entry channel (99). If the structure/function of nsp1 is conserved between the two strains of SARS-CoV, the presence of nsp1 may destabilize the interactions of the CrPV IRES with the ribosome at the A site resulting in the absence of 40S/IRES binary complex.

Endogenous mRNAs, as well as ectopically transfected mRNAs, are both degraded when SARS-CoV nsp1 is expressed in cells (94). However, this is not the case for RNAs that are produced by RNA polymerase type I and III (155). This is consistent with the fact that nsp1 acts through its binding to the 40S subunit and suggests that only mRNAs that are being translated are targeted by nsp1. The Xrn1 protein is a highly conserved 5'-3' exoribonuclease involved in mRNA degradation in the cytoplasm (156-158). Silencing of Xrn1 in the cell abolishes the effect of transfect-

ing SARS-CoV nsp1 and preserves the integrity of cellular mRNAs (155) indicating that mRNA degradation requires Xrn1 (Figure 4). To our knowledge, Xrn1 is not expressed in the RRL and this could explain why mRNAs are not degraded by nsp1 in this system; however, they are cleaved by others, yet unidentified, endonucleases that are activated by the addition of SARS-CoV nsp1 (Figure 4) (147,148,159). The cleavage sites were then mapped by using a toeprinting approach on reporter mRNAs with the 5'UTR of globin or actin that are both expressed by a cap-dependent mechanism. Addition of SARS-CoV nsp1 to the RRL induced about ten endonucleolytic cleavage prints on both sides of the initiation codon with one or two predominant sites but they failed to highlight a consensus sequence or nucleotide preference (159). However, modification of the RNA secondary and tertiary structures by mutagenesis affected the global pattern of cleavage (159) confirming that the 3-D structure of the RNA plays a role in this process. Then, the authors used a comparison between the EMCV and PV IRESes as both IRESes exhibit differences in their site of initiation relative to the binding site of the 40S ribosomal subunit. Indeed, whereas for EMCV, the 40S binds directly at the vicinity of the AUG codon, this is not the case for the PV IRES in which the initiation codon is located 154 nt downstream to the 3' border of the IRES (150). Consequently, in PV, the preinitiation complex needs to reach the initiation codon by linear scanning after its primary attachment to the IRES (160). Addition of the SARS-CoV nsp1 induced several endonucleolytic cleavages at the 3' border of both IRESes but no cleavage could be detected around the authentic AUG codon of the PV mRNA (159) suggesting that the ribosome was stalled after its attachment and could not scan to the initiation codon in the presence of nsp1. This nsp1 mediated RNA cleavage did not occur when both the HCV or CrPV IRESes were used in the RRL (147,159). In the case of CrPV IRES, nsp1 was able to disrupt the binary complex formed between the small ribosomal subunit and the IRES and remained probably bound to the 40S subunit. In such conditions, the mRNA that harbors the CrPV IRES cannot associate with the ribosome and, therefore, was not targeted by nsp1. The situation is even more complex in the case of HCV IRES because a binary complex 40S/IRES was formed in the presence of nsp1 but the RNA remained uncleaved. To date no explanation and experiments have been proposed to understand this resistance to RNA cleavage; however, this IRES is highly structured with pseudoknot structures that could provide a physical barrier against endonucleases. The fact that SARS-CoV nsp1 does not show any intrinsic endonuclease activity means that nsp1 does not act directly on its substrate but rather activates one, or several, unknown cellular endonucleases. As nsp1 blocks ribosomes essentially during the initiation step, these stalled 40S ribosomes on the mRNA may trigger a quality control mechanism that could be reminiscent to the no-go decay which targets mRNAs on which elongation is stalled (161,162). Further studies have to be carried out to understand the molecular mechanism by which nsp1 can induce RNA cleavage.

Nsp1 from others CoV strains. The four different lineages of the β -CoV genera share a relatively low amino acid

sequence similarity: for example, SARS-CoV nsp1 shares only 20.6%, 17.9% and 26.1% sequence similarity with MHV, B-CoV and MERS-CoV nsp1 proteins, respectively (89,163). Despite these differences, the amino acid sequence, LRKxGxKG, is relatively well conserved between the nsp1 of SARS-CoV, MHV, B-CoV and MERS-CoV (163,164) and is localized at positions 123–130 in the SARS-CoV nsp1 as a reference. Importantly, this consensus contains the two residues R124 and K125 that are critical for the endonucleolytic cleavage (96).

As for SARS-CoV, MHV nsp1 is localized exclusively in the cytoplasm of infected cells (165) and has been reported to inhibit the expression of reporter genes (166). Northern blot analyses of MHV infected cellular extracts also indicated a massive degradation of host cell mRNAs (85) with deletion of its carboxy-terminal domain being sufficient to abolish nsp1 function (166). Mutation of the consensus sequence in nsp1 results in a decrease of translation of reporter genes (167). The conserved domain (LRKxGxKG) appears to be needed during *in vivo* viral infection as its mutation leads to a strong attenuation of virulence (167) but the precise molecular mechanism by which MHV nsp1 inhibits gene expression awaits determination.

The MERS-CoV nsp1 exerts a dual function on cellular mRNAs by inducing both translation inhibition and mRNA degradation (96,163). The key charged amino acids identified in SARS-CoV nsp1 are conserved in the MERS-CoV and point mutation of the lysine in position 181, which is the analog of lysine 164 in SARS-CoV nsp1 (95), completely abolishes the function of nsp1 (168). Mutation of the two conserved amino acids R146/K147 in MERS-CoV nsp1, which corresponds to residues R124/K125 in SARS-CoV nsp1, recapitulates the cleavage defective effect and allows to dissociate the two functions of the protein (96,147). Despite these functional similarities, MERS-CoV nsp1 shows many differences with SARS-CoV nsp1 notably at the level of cellular localization, the interaction with the 40S ribosomal subunit and the targeted mRNAs (96). Indeed, only mRNAs transcribed by the RNA polymerase type II in the nucleus of the host infected cell are targeted by MERS-CoV nsp1. As such, mRNAs that are injected or engineered to be produced in the cytoplasm are not sensitive to repression or degradation. Interestingly and in contrast to SARS-CoV, MERS-CoV nsp1 is localized both in the nucleus and the cytoplasm and weakly interacts with the 40S ribosomal subunit raising the question of how MERS-CoV nsp1 can act on actively translated mRNAs. Curiously, the expression of mRNAs produced in the cytoplasm was stimulated by the presence of MERS-CoV nsp1 although the mechanism by which this occurred was not solved. A possible explanation could be attributed to a competitive effect due to the cellular shutoff that increases the availability of the components of the translational machinery for these mRNAs.

Nsp1 is also expressed in strains that belong to the α -CoV genus, such as Porcine Epidemic Diarrhea Virus (PEDV), TGEV, HCoV-229E and HCoV-NL63 but the protein is significantly smaller (110 aa instead of 180 aa for the SARS-CoV) and sequence alignments show only ~21% identity and 23% similarity with the first 128 aa of SARS-CoV nsp1 proteins between both HCoV-229E and HCoV-NL63 (93).

However, bioinformatic analysis revealed that the nsp1 proteins of HCoV-229E, HCoV-NL63 and SARS-CoV share a very similar 3D shaping structure (93). The crystal structures of TGEV and PEDV nsp1 confirmed a high degree of similarity with SARS-CoV nsp1 despite this lack of sequence homology (92,164,169) indicating a conservation of protein function. In agreement with this, nsp1 of TGEV, PEDV, HCoV-229E and HCoV-NL63 strains are able to inhibit global host gene expression at a similar level than the SARS-CoV nsp1 (92,93,159,169). It is also noteworthy that the two K164/H165 residues of SARS-CoV nsp1, which are required for the interaction with the 40S ribosomal subunit (95), are lacking in α -CoV nsp1. However, both HCoV NL63- and HCoV 229E-nsp1 can still immunoprecipitate the S6 ribosomal protein from 293T cell extracts (92,93) suggesting that an interaction between the 40S ribosomal subunit and nsp1 still takes place in α -CoV strains. In addition, TGEV nsp1 cannot induce RNA degradation (159) but HCoV NL63- and 229E-nsp1 are able to do so (93) and several amino acids have been identified in TGEV, PEDV, HCoV-229E and HCoV-NL63 strains to be required to promote efficient translation inhibition (92,169). Thus, nsp1-mediated inhibition of protein synthesis during α -CoV infection appears to be a very conserved process despite many differences in the way it is achieved.

SARS-CoV-2 nsp1. Nsp 1 proteins of SARS-CoV and SARS-CoV-2 share 84.4% sequence identity and 93.9% of sequence similarity suggesting a strong conservation at the level of protein function. Infection of human lung epithelial (Calu3) or monkey kidney (Vero) cells with SARS-CoV-2 strain results in a global translation inhibition (99). Transfection of a plasmid coding for the SARS-CoV-2 nsp1 into 293T cells is sufficient to reduce cellular proteins expression level (99) and induces a shift from the polyribosome fractions to 80S monosomes pointing out that a global inhibition of translation occurs in the presence of SARS-CoV-2 nsp1 (97). SARS-CoV-2 nsp1 is associated with both 40S and 80S but not with the polyribosome fractions (97,98). *In vitro* binding assays have confirmed that recombinant SARS-CoV-2 nsp1 can interact with purified 40S but not with the 60S subunit (97,98) and this binding requires both K164 and H165 residues as it was the case for SARS-CoV nsp1 (97,98,148). Furthermore, addition of recombinant nsp1 protein into a cell-free *in vitro* translation system is also sufficient to inhibit expression of reporter mRNAs (97–99). The interaction between nsp1 and the 40S ribosomal subunit has recently been imaged by cryo-EM at an average resolution of 2.6–2.8 Å, and this revealed that the C-terminus of nsp1 was made of two α -helices and located inside the ribosomal mRNA entry channel at the latch structure formed between rRNA helix h18 of the body and h34 of the head and uS3 of the 40S subunit (97,98,170,171). The C-terminal domain of nsp1 interacts specifically with uS3, uS5 and rRNA helix h18 (97–99). The K164 and H165 residues establish a key interaction with helix h18 and the binding of nsp1 to the mRNA entrance channel blocks mRNA entry (Figure 4) (99). Interestingly, the location of nsp1 could inhibit tRNA recruitment to the 80S ribosome that would block the elongation step (99). This could explain the observation that the mRNA was always missing

from all purified and analyzed nsp1/ribosomal complexes (97,98). As a consequence, addition of SARS-CoV-2 nsp1 to HeLa translation-competent lysates inhibited translation of reporter mRNAs (97–99) except when they are bearing the SARS-CoV-2 genomic 5'UTR (99).

Other CoV viral proteins

Infection with SARS-CoV strains that encodes nsp1-mt (K164A & H165A) results in the inhibition of host mRNA translation (95) suggesting that other viral proteins are involved in the control of cellular mRNA translation.

A two-hybrid screen in yeast has identified interactions between the f subunit of the initiation factor eIF3 and the SARS-CoV S protein that was later confirmed by immunoprecipitation studies in HeLa cells. The N-terminal region of the S protein (20–404 aa) is sufficient to bind to eIF3f and this modifies the cellular localization of eIF3f that become mainly cytoplasmic likewise the S protein. Addition of a S protein truncated in its C-terminal domain (Δ C) into the reticulocyte lysate is able to inhibit the expression of reporter mRNAs and this can be partially restored by the addition of recombinant eIF3f (90). This translation inhibitory effect of spike protein has also been demonstrated in other CoV strains like IBV (90). However, it can be argued that the work was essentially performed with the Δ C mutant in which the peptide signal is deleted that also affects its localization to the cytoplasm compartment; indeed, within the WT S protein, the N-terminal domain that interacts with eIF3f is localized at the extracellular side of the plasma membrane. Therefore, in such conditions, the S protein is not supposed to be able to interact with initiation factors.

The SARS-CoV nucleocapsid N protein is the most abundant amongst all structural proteins and it harbors an RNA-binding domain in its N-terminal region. A two-hybrid yeast screen has shown that the C-terminal domain can interact with the human elongation factor 1- α (EF1 α) (172), which is an essential component for the elongation phase of translation (173,174). The N protein can immunoprecipitate native EF1 α proteins from cell extracts and this property is conserved with the N protein of the HCoV-229E strain (172). Due to its self-association, the viral N protein can form aggregates that include EF1 α (172). Interestingly, HSP70, a chaperone protein that preferentially binds to denatured or aggregated proteins (175), is also found as part of these complexes (172). Thus, it is not surprising that addition of recombinant N protein was able to inhibit translation of reporter mRNAs both in the RRL and in cellular extracts (172). Once again, evidences for inhibition during viral infection are lacking to confirm a function of N protein in the cellular translation shutoff.

Very recently, a global analysis of SARS-CoV-2 host interacting proteins has allowed to identify 332 interactions between viral and human proteins (102). Some of them are involved in translation. For instance, the nsp2 interacts with both GIGYF2 and eIF4E2 (4EHP) (102) that are two components of a complex that represses mRNA translation (176). These interactions were already identified with nsp2 from SARS-CoV (177) but were never investigated further although they suggest a role for nsp2 in regulating mRNA

translation. Similarly, nsp9 interacts also with eIF4H (102), a factor that enhances the ATP dependent helicase activity of eIF4A (178–180). Interestingly, the complex formed between eIF4A and eIF4H can be inactivated using a specific inhibitor that is able to lock eIF4A in an inactive conformation (179). The interaction between nsp9 and eIF4H could modulate eIF4A function and thus, play a role on translational efficiency. Finally, the N protein of SARS-CoV-2 can interact with different isoforms of PABP and the ribosomal protein L36 (102), and this may also participate in the regulation of protein synthesis.

Structure and function of the viral mRNAs

Translation and integrity of cellular mRNAs are significantly altered in cells that are infected by SARS-CoV or MERS-CoV. Despite these conditions, viral mRNAs are not degraded and viral protein synthesis is maintained at high level (95,96,99). In an *in vitro* system, the SARS-CoV nsp1 does not induce the endonucleolytic cleavage on sgRNAs that have been purified from SARS-CoV infected cells (159). These experiments indicate that viral mRNAs could be protected from this host cell shutoff and this may be attributed to RNA *cis*-acting elements that confer this resistance.

In order to better understand this mechanism, it is important to start with a brief description of the architecture of the 5' and 3' termini of the CoV genome. The TRSs that are essential for viral genome transcription (43,181) are located downstream to the leader sequence and upstream to the AUG codon (reviewed in (28)). As such, all coronaviral transcripts (depending on the subtype) share an identical coterminal 3' ends and common 70–90 nt long leader sequences at their 5' terminus (182). These noncoding RNA regions fold into secondary and tertiary structures that are involved in functional RNA–RNA interactions; they also bind to a subset of viral and cellular proteins during replication and translation. In this part, we will not detail the specific role of these RNA structures that are used for genome replication and we advise readers to consult these excellent reviews on the topic (183,184). Rather, we will focus our attention on RNA structures that are directly involved in translational control of the coronavirus genomic and subgenomic RNAs. Thus, for clarity, we will name '5'UTR' the region spanning from the 5' terminus of the transcript and up to the AUG codon. The first 60–95 nt of the 5'UTR refer to the leader sequence, which is common to all coronaviral mRNAs.

The viral 5' UTR. Analysis of the RNA structures of the 5'UTR was essentially performed from the MHV strain. Indeed, by using a combination of phylogenetic analysis and computer-based prediction software, the 140 first nucleotides of the MHV genome have initially been modeled to form three conserved stem-loops (SL) that are named SL1, SL2 and SL4 (185,186). A 3 stem-loop folding is also observed for MERS-CoV whereas SARS-CoV and B-CoV contain an additional SL (SL3) that folds the leader TRS (TRS-L) into a hairpin loop. Interestingly, for all of these four members of the β -CoV family, the RNA structure-based models predict that the AUG codon is embedded

within SL4 (183). Thus, this suggests that the accessibility of the initiator codon may be an issue as it is not positioned in an unstructured region although this has not yet been formerly experimentally determined. In 2015, Selective 2'-Hydroxyl Acylation analyzed by Primer Extension (SHAPE) analysis of the 5'UTR of MHV revealed a conformational structure that was in good agreement with previous predicted models and confirmed the presence of SL1, SL2 and the correct structure of SL4 at position located between nt 80 and 130 (187). Interestingly, the overall structural conformation of these four β -CoV was very similar despite very poor sequence conservation at the level of primary nucleotide sequence suggesting a link between RNA structure and function (143,186,188). Although the function of the 5' UTR in replication has been extensively studied, its role in translation remains, yet, to be determined.

The different CoV RNAs are resistant to the endonucleolytic cleavage induced by nsp1 (Figure 4). This was confirmed by *in vitro* studies with purified sgRNA3 and sgRNA9 (159) indicating that they contain *cis*-acting elements that confer resistance. In order to investigate whether the viral UTRs could be involved in the protection against the action of nsp1, Huang and colleagues engineered different versions of the sgRNA9 coding for the SARS-CoV N protein. They monitored translational efficiency of these constructs in the RRL and they found that the viral 5'UTR was sufficient to protect the transcripts against RNA cleavage (159) but not translation inhibition (147). The 5'UTR of the mRNA encoding the N protein harbors the common 72 nt-long sequence leader with an additional stretch of 8 nt. Mutational analysis of the 5' terminal nucleotides of the leader sequence affected the resistance to the endonucleolytic cleavage indicating that the leader sequence could specifically protect the viral mRNAs. Experiments conducted in cells showed that the leader sequence conferred resistance to both RNA degradation and translation repression induced by nsp1 (189) suggesting that an unidentified cellular partner was required to promote a total resistance to nsp1; this cellular partner would be absent from the RRL. They also showed that nsp1 could specifically interact with SL1 of the leader region and this was sufficient to counteract translational inhibition (Figure 4) (189). Interestingly, the nsp1 amino acid residues needed for 40S binding (K164/H165) and for RNA cleavage (R124/K125) are required to bind SL1 (189). These data confirm that SARS-CoV nsp1 exhibits functional domains involved both in RNA binding and translational regulation, but the molecular mechanism remains unknown. However, as lysine 164 is involved in both viral mRNA and ribosome attachment, we can speculate that these two binding are mutually exclusive and the presence of SL1 on the viral mRNA could displace nsp1 located at the mRNA entry channel (97–99) in order to allow viral protein synthesis (Figure 4). Very recently, Schubert and colleagues have reported that the SARS-CoV-2 nsp1 could also inhibit translation of transcripts harboring the 5'UTR of the gRNA in a HeLa cell free system (98); this was already observed with SARS-CoV nsp1 in a similar experimental setting (159). In contrast, Banerjee and colleagues showed, in a cellular system, that the leader sequence and especially the SL1 structure can protect the mRNA from nsp1-mediated translational inhibition (99).

In addition, the distance between the 5' cap and the SL1 appears to be critical to counteract the function of nsp1 on the viral mRNAs (99). Therefore, the function of nsp1 on mRNA expression may vary from the experimental translational system used (cells, cellular extracts or RRL).

In contrast, MERS-CoV nsp1 only targets cellular mRNAs that are transcribed into the nucleus of the host with no effect on viral mRNAs synthesized from the cytoplasm. However, if the CoV mRNAs are engineered to be artificially synthesized in the nucleus, these viral mRNAs become targets for nsp1. Interestingly, in this case, they can escape translation repression if they possess the leader sequence harboring the SL1 motif (163). However, this resistance appears to be CoV strain specific as SARS-CoV nsp1 represses MERS-CoV mRNAs despite the presence of the MERS-CoV leader sequence (96); the reciprocal has not been studied. In contrast to SARS-CoV nsp1, MERS-CoV nsp1 requires the N-terminal region and more precisely the arginine at position 13 to interact with the SL1 RNA structure (163). Once again and unlike SARS-CoV-nsp1, MERS-CoV-nsp1 can stimulate expression of viral messengers as well as those produced or injected into the cytoplasm. The cleavage defective mutant nsp1 loses this stimulating activity on viral RNAs indicating that residues R146/K147 are required for the efficient expression of viral RNAs (96).

The different functions of nsp1 from SARS-CoV, SARS-CoV-2 and MERS-CoV strains in the cellular mRNAs translation and decay, and in the viral mRNAs protection are recapitulated in Table 1.

Due to the process of viral replication and transcription (190,191), the 5'UTRs of all viral mRNAs start with a cap moiety followed by a strictly conserved ~72 nt sequence leader. The size of the entire 5'UTRs ranges from 72 to 264 nt in the case of SARS-CoV mRNAs (192) and from 75 to 275 nt for SARS-CoV-2 mRNAs (Table 2) (14).

As mentioned above, the impact and the role of the 5' RNA structures on viral protein synthesis remain to be determined. However, it is known that gRNAs and sgRNAs are translated by a canonical cap-dependent translation initiation mechanism and, thus, they require the eIF4F complex: eIF4E, eIF4G and eIF4A. Several studies on different CoVs have studied viral translation in the presence of inhibitors against components of eIF4F. This showed that addition of cap analogue in the RRL strongly reduced translation of reporter mRNAs driven by the 5'UTR derived from the B-CoV strain (193,194), which indicates that eIF4E was required for translation initiation. This result was confirmed by the use of the 4E2RCat compound that inhibited interaction between eIF4E and eIF4G (195) as it almost suppressed HCoV-229E replication with a strong reduction of the expression of the Spike protein (196). Finally, hippuristanol and silvestrol that are two molecules known to specifically inhibit the eIF4A RNA helicase, did also affect HCoV-229E and MERS-CoV replication (196,197). Further analyses showed that this defect of viral replication was due to impairment of viral protein expression (197). Moreover, a comparison of the effects of silvestrol between picornavirus (which are dependent on eIF4A for translation) and HCoV-229E and MERS-CoV showed a stronger sensitivity to silvestrol for coronaviral mRNAs (197).

Interestingly, the initiator AUG codons of most SARS-CoV sgRNAs are surrounded by a poor Kozak context (192) suggesting that a leaky scanning mechanism may occur during viral mRNAs translation initiation. Yang and colleagues have analyzed the ability of these different sgRNA 5'UTR to initiate translation of reporter gene in noninfected cells and they showed that, in such system, leaky scanning was indeed taking place (192). Leaky scanning has been studied in details on the sgRNA7 of both SARS-CoV (198) and B-CoV (193,199) and results in the production of two different proteins from two distinct ORFs. Mutations of the respective AUG codons and/or their nucleotide context impacted their level of expression (193,198,199). In the case of SARS-CoV sgRNA7, the protein ORF7b initiated by leaky scanning was detected in infected cells and incorporated in the viral particle (198). Leaky scanning was also reported from the 5'UTR of the SARS-CoV sgRNA2-1 with recombinant plasmids (192). Analysis of the initiation codons from the different sgRNAs of the SARS-CoV-2 strain shows that several of them are located in a weak Kozak context (namely those initiating for N, S, ORF3a, E and ORF6 proteins) (Table 2). These differences may have an impact on the translation rate and it could also promote expression of unexpected proteins that may be relevant for the virus.

Another layer of complexity in the regulation of translation is provided by the presence of a short uORF located immediately downstream of the genomic leader sequence (Table 2). This uORF is only found on the gRNA, but not on the sgRNAs (200) although this bears some exceptions (201), suggesting that translation of genomic and subgenomic RNAs can be regulated by distinct mechanisms. This uORF potentially encodes a short peptide (3–13 aa) and is found in the vast majority of CoV genomes as established on a sampling of 38 reference strains (200). Based on *in vitro* translation studies, the authors showed that the MHV uORF downregulated translational efficiency from the major ORF1 start codon (200). Ribosomal profiling studies confirmed that this uORF was read by incoming ribosomes during MHV infection despite its weak initiation context (202). Interestingly, MHV mutant viruses lacking this structure showed similar growth properties than the WT but, after 10 passages, all mutants in which the uORF was altered, reverted to the WT sequence or generated a new uORF sequence. This suggests that this uORF plays a beneficial, although nonessential, role in replication. In addition, this effect was exerted at the level of translation as reversion to another uORF resulted in the synthesis of a different short peptide that the one encoded from the WT virus. Mutation of the AUG codon of the uORF in MHV and B-CoV 5'UTR resulted in the enhancement of translation of the main ORF (201,203,204) indicating that it could be a way to control the flux of ribosomes on the mRNA to regulate the level of expression of the nonstructural proteins.

In the case of SARS-CoV, all viral mRNAs start with the same 72 nt leader sequence suggesting a common mechanism for initial ribosome binding. However, the gRNA and the sgRNAs exhibit differences in the length and structure of the untranslated region downstream to nucleotide 72 and up to nucleotide 264 (205). These differences may affect ribosomal scanning and, consequently, translational

Table 1. Comparison of the nsp1 functional characteristics between SARS-CoV, SARS-CoV-2 and MERS-CoV strains; ND, not determined

	Nsp1 functions	SARS-CoV	SARS-CoV-2	MERS-CoV
General characteristics	Sequence identity with SARS-CoV-2	84.4%	100%	~25%
Translation shutoff	Cellular localization Shutoff 40S binding	Cytoplasmic Yes Yes: The C-terminal domain of nsp1 binds at the mRNA entry channel of the 40S ribosome K164/H165		Nuclear and cytoplasmic Yes No
Cellular mRNA degradation	Critical amino acids mRNAs targeted	All mRNAs	ND	K181 Only mRNAs transcribed in the nucleus
Viral mRNA escape	Critical amino acids <i>Cis</i> -acting element required Critical amino acids	R124/K125 SL1 (leader sequence) R124	ND SL1 (leader sequence) ND	R164/K165 SL1 (leader sequence) R13

Table 2. Translational characteristics of the SARS-CoV-2 transcripts

Transcript	ORF	5'UTR length (nt)	AUG initiation codon context	uORF
gRNA	ORF1a and ORF1ab: nsp	265	aag AUGG	ugc AUG C
sgRNA	Spike (S)	76	aca AUG U	no
3a-RNA	ORF3a	77	cuu AUGG	no
E-RNA	Envelope (E)	77	cuu AUG U	no
M-RNA	Membrane (M)	119	gcc AUGG	no
6-RNA	ORF6	230	cag AUG U	no
7a-RNA	ORF7a	75	aac AUG A	no
7b-RNA	ORF7b	151	aga AUG A	no
8-RNA	ORF8	75	aac AUG A	no
N-RNA	Nucleocapsid (N)	83	aaa AUG U	no

efficiency. Although most transcripts have only a few nucleotides extension downstream to the leader sequence, this is not the case for the gRNA and sgRNAs 5, 6 and 8 that have a longer extension of 264, 116, 227 and 155 nt, respectively (192,205). Yang and colleagues have constructed reporter RNAs bearing these 5'UTRs and they compared their translational efficiency. Some 5'UTRs promote a weaker expression of the reporter gene confirming that this region, together with the nucleotide context surrounding the AUG, can impact the translation rate of viral mRNAs (192). Such a difference in expression has also been pointed out in B-CoV in which the length of the 5'UTR of the gRNA and sgRNA 7 are 210 and 77 nt long, respectively. The presence of these additional 133 nt in the 5'UTR of the gRNA is sufficient to significantly decrease the expression of a reporter gene (194). Transcriptomic studies in SARS-CoV-2 infected cells have shown that 10 sgRNA are produced in addition to the gRNA (14). The leader sequence is 72 nt long and the size of the 5'UTR of most sgRNA is similar to the sequence leader (between 75 and 83 nt) but, for three of them (the gRNA, sgRNA5 and sgRNA6), they are 265, 119 and 230 nt long, respectively (Table 2).

The viral 3' UTR. The 3' UTR of coronaviruses has a size that ranges between 300 and 500 nt in length from the stop codon to the beginning of the poly(A) tail. Most of this 3' UTR (436 nt in MHV, 492 nt for TGEV, 338 nt for IBV) is essential for viral replication as it plays a critical role in negative strand RNA synthesis (31,206–208). There has been a lot of studies using biochemical and computer-based algorithms combined with functional studies to modulate the structure and explore the function of the various

motifs that are present in the 3' UTR. This has led to a model initially described for MHV (209) that postulates that the 68 nt just downstream to the stop codon of the N gene folds into a bulged stem loop (BSL), which is followed by a 54 nt hairpin-type pseudoknot (210). These two structures, BSL and pseudoknot, are phylogenetically conserved in length and shape (but, likewise for the RNA structures in the 5'UTR, not at the level of nucleotide sequence) among all coronaviruses studied so far and were proposed to function as a molecular switch during viral RNA synthesis (211). Other secondary and tertiary structures are also found downstream to the pseudoknot but these motifs are less well conserved among strains and appear, for most part, not to be essential for viral replication (for a review see (183)).

The poly(A) tail at the 3' terminus of the transcript also plays a critical role in mRNA stability and translation through its interaction with PABP (212–214). Cytoplasmic PABP interacts both with the deadenylation machinery and most of the translation initiation factors. Thus, the length of the poly(A) tail can modulate expression of mRNAs, and cytoplasmic polyadenylation can occur to regulate translation (215,216). As it is also a target for deadenylases leading to RNA decay and loss of viral genome integrity, the length of the poly(A) tail of viral RNA has to be thoroughly controlled. Nevertheless, the mechanisms implicated in coronaviral RNA polyadenylation remain unclear. It is assumed that the poly(A) tail is probably generated from a short poly(U) stretch template found in the negative-strand genome by using a slippage mechanism involving the viral RNA-dependent RNA polymerase; alternatively, the virus could also use the cytoplasmic poly(A)-polymerase in an AAUAAA-independent manner. Indeed, the canonical

polyadenylation signal (AAUAAA) is not present in CoV mRNAs but it is not required if the mRNA already possesses a short poly(A) stretch. Here, the poly(U) stretch can be used as a template for a short poly(A) synthesis by the viral replicase or the cellular transcriptase (217). It is noteworthy that Peng *et al.* have described the presence of an AGUAAA noncanonical polyadenylation signal that is not required for efficient coronavirus mRNA polyadenylation but may be critical when the length of the tail becomes too short (218).

A dynamic regulation of the control of poly(A) tail length has also been reported in the course of viral infection when the length of the poly(A) tail in the B-CoV strain extends from a 45 (A) at the beginning of infection to a 65 (A) stretch later on during the cycle (219). Interestingly, this length variation was similarly found with IBV infection that suggests a common mechanism among coronaviruses that can occur both in cell culture and in the host organism (220). It is well known that changes in poly(A) tail length are often associated with changes in translational regulation. In the case of coronavirus mRNAs translation, transcripts with longer poly(A) tail are more efficiently translated (219). Moreover, it has been shown that nsp8 from HCoV-229E exhibits an adenylyltransferase activity that indicates for a possible role for the protein in the extension of the viral poly(A) tail (221).

Interactions between viral and host factors with the poly(A) tail have been studied and reports show that cellular PABP can bind to the coronaviral poly(A) tail (222). More recently, data support evidence that the B-CoV N protein can also bind to the poly(A) tail with a high affinity to compete with PABP. As a result, the N protein inhibits mRNA translation in a poly(A)-dependent manner in both the RRL and cultured cells (223). Interestingly, the N protein can interact with both PABP and eIF4G and pull-down assays suggest that the presence of N protein may alter the interaction between eIF4G and eIF4E (223). To summarize, multiple interactions between the poly(A) tail, PABP and N protein occur during CoV infection and contribute to regulate translation of cellular mRNAs. In addition, the N protein interacts with nsp9 (223), which is a replicase protein associated with the viral polymerase (224); Thus, the poly(A) tail may act as a regulator to coordinate the use of the viral genome for translation (binding to PABP) or for replication (binding to N).

Finally, a recent study using transcriptomic analysis has confirmed that SARS-CoV-2 RNAs are polyadenylated with a median length of 47 (A) residues (14). This poly(A) tail is longer for gRNAs than for sgRNAs and among the sgRNAs, two populations are observed: a minor one that harbors a ~30 A residues and a major one that contains a poly(A) tail of about 45 residues. Transcripts with the short poly(A) tail may represent a population of RNAs that have already engaged in the first processes of deadenylation and RNA decay. Interestingly, an epitranscriptomic analysis of the RNA modifications, mainly N6-methyladenosine (m6A) but also 5-methylcytosine methylation (5mC), 2'-O-methylation (Nm), deamination and terminal uridylation shows that all these modified RNAs exhibit shorter poly(A) tail compared to unmodified RNAs. This observation confirms that viral RNA modifications are involved in viral

RNA stability and translational control through the length of the poly(A) tail (14).

To conclude this chapter, the interactions between the two termini were shown to play some role both in viral replication and translation. Indeed, the model of discontinuous transcription of subgenomic RNAs implies that 5' and 3' interactions are needed for efficient replication (225–227). In TGEV, these interactions occur via RNA–RNA, RNA–protein and protein–protein contacts and induce circularization of the coronavirus genome to promote elongation of the minus strand (228). However, more related to the topic of this review, other RNA–protein interactions have been evidenced (222) and suggested that the PABP bound to the 3' poly(A) tail also contributes to genome circularization with its ability to interact with eIF4G (222). Thus, it suggests that molecular interactions mediated by the 5' and 3' UTRs are important for CoV translation and it would be essential, in the future, to define better the nature and function of these interactions between RNAs and potential viral and cellular proteins involved in the process.

CONCLUDING REMARKS

Like in many RNA viruses, protein synthesis is controlled at different levels during CoVs infection. As we have seen, one of the most comprehensive mechanisms is probably translation repression mediated by nsp1. Recent cryo-EM studies have revealed novel insights in how SARS-CoV-2 nsp1 interferes with ribosome function (97,98). However, RNA cleavage and/or degradation induced by nsp1 remains a mystery. In addition, the mechanism by which translational repression is mediated in MERS-CoV nsp1 and α -CoV nsp1 remains to be determined. For clarity and as a summary, we have recapitulated in Table 1 the different functions fulfilled by nsp1 from SARS-CoV, SARS-CoV-2 and MERS-CoV strains.

Based on the NMR structure of SARS-CoV nsp1 (164), Jauregui and colleagues have engineered 38 mutants that target 62 exposed amino acids and analyzed their ability to promote inhibition of host gene expression. From this work, they identified versions of the nsp1 protein that abolished, attenuated or increased the inhibition of host gene expression (229). These data confirm that coronaviruses are able to evolve through mutation of these residues.

Interestingly, SARS-CoV nsp1 can also disrupt localization of the nuclear pore complex protein Nup93 (230), which is consistent with its perinuclear localization (94,231). As a consequence, nsp1 alters the nuclear-cytoplasmic distribution of nucleolin (230). As nucleolin is an RNA binding protein implicated in RNA stability (232), its delocalization in the presence of nsp1 could contribute to the control of host gene expression mediated by nsp1. Interaction between SARS-CoV-2 nsp1 and Nup93 has not been detected from the analysis of SARS-CoV-2 human protein–protein interactome but the authors showed that both nsp9 and ORF6 protein are able to interact with nuclear pore components (102).

Recent data providing from ribosome profiling assays, performed in SARS-CoV-2 infected cells, added an additional layer of complexity in the regulation of translation from the gRNA and sgRNAs. Indeed, the authors have

identified 23 novel ORFs including internal in-frame and out-frame ORFs, and several uORFs that make the translation of these mRNAs more complex and require further functional investigations (27).

To conclude, it is also important to mention recent data on epitranscriptomic of CoV RNAs. To date, 172 RNA modifications can be listed (233) in which the m6A that remains predominant and which is known to play critical roles in mRNA metabolism such as mRNA stability, decay and translation (234). As a result, m6A RNA modification of both viral and cellular mRNAs is a new mean to control post-transcriptional gene expression during viral infection (235). However, the effects of m6A modifications can either promote viral replication as it is the case for enterovirus 71, influenza A virus and the human immunodeficiency virus, but it can also negatively impact viral production as demonstrated during both HCV and ZIKV infection (235). M6A modifications have been studied in cells infected with the α -CoV PEDV strain and the authors showed a decrease in demethylation activity that correlates with an increase of m6A addition in the host cellular RNAs resulting in a decrease of PEDV replication (236). The RNA genome of SARS-CoV-2 contains >50 potential m6A sites based on the presence of specific sequence motifs for m6A modification by the RNA methylase complex METTL3/14 (237). Gain or loss of m6A results in significant functional changes during viral replication (235) and it will be important to determine precisely whether it has a pro- or antiviral effect.

So, the recent data provided by ribosome profiling assay in SARS-CoV-2 infected cells (27), the SARS-CoV-2 protein interactome (102) and the potential mRNA epitranscriptomic regulation confirm the importance of the relationship that exists between the viral proteins and the components of the translational machinery.

FUNDING

Agence nationale de recherche sur le sida et les hépatites virales (ANRS). Funding for open access charge: Institut national de la santé et de la recherche médicale (INSERM).
Conflict of interest statement. None declared.

REFERENCES

- Lefkowitz, E.J., Dempsey, D.M., Hendrickson, R.C., Orton, R.J., Siddell, S.G. and Smith, D.B. (2018) Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res.*, **46**, D708–D717.
- Hamre, D. and Procknow, J.J. (1966) A new virus isolated from the human respiratory tract. *Proc. Soc. Exp. Biol. Med.*, **121**, 190–193.
- McIntosh, K., Becker, W.B. and Chanock, R.M. (1967) Growth in suckling-mouse brain of "IBV-like" viruses from patients with upper respiratory tract disease. *PNAS*, **58**, 2268–2273.
- Tyrell, D.A. and Bynoe, M.L. (1965) Cultivation of a novel type of Common-Cold virus in organ cultures. *Br. Med. J.*, **1**, 1467–1470.
- van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J. and Berkhout, B. (2004) Identification of a new human coronavirus. *Nat. Med.*, **10**, 368–373.
- Woo, P.C., Lau, S.K., Chu, C.M., Chan, K.H., Tsoi, H.W., Huang, Y., Wong, B.H., Poon, R.W., Cai, J.J., Luk, W.K. *et al.* (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.*, **79**, 884–895.
- Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Penaranda, S., Bankamp, B., Maher, K., Chen, M.H. *et al.* (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science*, **300**, 1394–1399.
- Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S., Khattri, J., Asano, J.K., Barber, S.A., Chan, S.Y. *et al.* (2003) The Genome sequence of the SARS-associated coronavirus. *Science*, **300**, 1399–1404.
- Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W. *et al.* (2003) A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.*, **348**, 1953–1966.
- Bermingham, A., Chand, M.A., Brown, C.S., Aarons, E., Tong, C., Langrish, C., Hoschler, K., Brown, K., Galiano, M., Myers, R. *et al.* (2012) Severe respiratory illness caused by a novel coronavirus, in a patient transferred to the United Kingdom from the Middle East, September 2012. *Euro Surveill.*, **17**, 20290.
- Lu, G. and Liu, D. (2012) SARS-like virus in the Middle East: a truly bat-related coronavirus causing human diseases. *Protein Cell*, **3**, 803–805.
- Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D. and Fouchier, R.A. (2012) Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.*, **367**, 1814–1820.
- Dong, E., Du, H. and Gardner, L. (2020) An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.*, **20**, 533–534.
- Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N. and Chang, H. (2020) The architecture of SARS-CoV-2 transcriptome. *Cell*, **181**, 914–921.
- Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L. *et al.* (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, **579**, 270–273.
- Kim, J.M., Chung, Y.S., Jo, H.J., Lee, N.J., Kim, M.S., Woo, S.H., Park, S., Kim, J.W., Kim, H.M. and Han, M.G. (2020) Identification of coronavirus isolated from a patient in Korea with COVID-19. *Osong. Public Health Res. Perspect.*, **11**, 3–7.
- Cheever, F.S. and Daniels, J.B. *et al.* (1949) A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin. *J. Exp. Med.*, **90**, 181–210.
- Mebus, C.A., Stair, E.L., Rhodes, M.B. and Twiehaus, M.J. (1973) Neonatal calf diarrhea: propagation, attenuation, and characteristics of a coronavirus-like agent. *Am. J. Vet. Res.*, **34**, 145–150.
- Mebus, C.A., Stair, E.L., Rhodes, M.B. and Twiehaus, M.J. (1973) Pathology of neonatal calf diarrhea induced by a coronavirus-like agent. *Vet. Pathol.*, **10**, 45–64.
- Jacks, T. and Varmus, H.E. (1985) Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science*, **230**, 1237–1242.
- Brierley, I., Bournsnel, M.E., Binns, M.M., Bilimoria, B., Blok, V.C., Brown, T.D. and Inglis, S.C. (1987) An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.*, **6**, 3779–3785.
- Giedroc, D.P. and Cornish, P.V. (2009) Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res.*, **139**, 193–208.
- Plant, E.P. and Dinman, J.D. (2008) The role of programmed-1 ribosomal frameshifting in coronavirus propagation. *Front. Biosci. J. Virtual Lib.*, **13**, 4873–4881.
- Atkins, J.F. and Baranov, P.V. (2013) Molecular biology: Antibiotic re-frames decoding. *Nature*, **503**, 478–479.
- Dinman, J.D. (2012) Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscipl. Rev. RNA*, **3**, 661–673.
- Kelly, J.A., Olson, A.N., Neupane, K., Munshi, S., San Emeterio, J., Pollack, L., Woodside, M.T. and Dinman, J.D. (2020) Structural and functional conservation of the programmed-1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2). *J. Biol. Chem.*, **295**, 10741–10748.
- Finkel, Y., Mizrahi, O., Nachshon, A., Weingarten-Gabbay, S., Morgenstern, D., Yahalom-Ronen, Y., Tamir, H., Achdout, H., Stein, D., Israeli, O. *et al.* (2020) The coding capacity of SARS-CoV-2. *Nature*, doi:10.1038/s41586-020-2739-1.
- Sola, I., Almazan, F., Zuniga, S. and Enjuanes, L. (2015) Continuous and discontinuous RNA synthesis in coronaviruses. *Ann. Rev. Virol.*, **2**, 265–288.

29. Liao, C.L. and Lai, M.M. (1992) RNA recombination in a coronavirus: recombination between viral genomic RNA and transfected RNA fragments. *J. Virol.*, **66**, 6117–6124.
30. Viehweger, A., Krautwurst, S., Lamkiewicz, K., Madhugiri, R., Ziebuhr, J., Holzer, M. and Marz, M. (2019) Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. *Genome Res.*, **29**, 1545–1554.
31. Luytjes, W., Gerritsma, H. and Spaan, W.J. (1996) Replication of synthetic defective interfering RNAs derived from coronavirus mouse hepatitis virus-A59. *Virology*, **216**, 174–183.
32. den Boon, J.A. and Ahlquist, P. (2010) Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annu. Rev. Microbiol.*, **64**, 241–256.
33. Wolff, G., Limpens, R., Zevenhoven-Dobbe, J.C., Laugks, U., Zheng, S., de Jong, A.W.M., Koning, R.I., Agard, D.A., Grunewald, K., Koster, A.J. *et al.* (2020) A molecular pore spans the double membrane of the coronavirus replication organelle. *Science*, **369**, 1395–1398.
34. Wolff, G., Melia, C.E., Snijder, E.J. and Barcena, M. (2020) Double-Membrane vesicles as platforms for viral replication. *Trends Microbiol.*, **28**, 1022–1033.
35. Knoops, K., Kikkert, M., Worm, S.H., Zevenhoven-Dobbe, J.C., van der Meer, Y., Koster, A.J., Mommaas, A.M. and Snijder, E.J. (2008) SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol.*, **6**, e226.
36. Ulasli, M., Verheije, M.H., de Haan, C.A. and Reggiori, F. (2010) Qualitative and quantitative ultrastructural analysis of the membrane rearrangements induced by coronavirus. *Cell. Microbiol.*, **12**, 844–861.
37. de Wilde, A.H., Raj, V.S., Oudshoorn, D., Bestebroer, T.M., van Nieuwkoop, S., Limpens, R., Posthuma, C.C., van der Meer, Y., Barcena, M., Haagmans, B.L. *et al.* (2013) MERS-coronavirus replication induces severe in vitro cytopathology and is strongly inhibited by cyclosporin A or interferon-alpha treatment. *J. Gen. Virol.*, **94**, 1749–1760.
38. Doyle, N., Hawes, P.C., Simpson, J., Adams, L.H. and Maier, H.J. (2019) The porcine deltacoronavirus replication organelle comprises Double-Membrane vesicles and zippered endoplasmic reticulum with Double-Membrane spherules. *Viruses*, **11**, 1030.
39. Oudshoorn, D., Rijs, K., Limpens, R., Groen, K., Koster, A.J., Snijder, E.J., Kikkert, M. and Barcena, M. (2017) Expression and cleavage of middle east respiratory syndrome coronavirus nsp3-4 Polyprotein induce the formation of Double-Membrane vesicles that mimic those associated with coronaviral RNA replication. *mBio*, **8**, e01658-17.
40. Krijnse-Locker, J., Ericsson, M., Rottier, P.J. and Griffiths, G. (1994) Characterization of the budding compartment of mouse hepatitis virus: evidence that transport from the RER to the Golgi complex requires only one vesicular transport step. *J. Cell Biol.*, **124**, 55–70.
41. Jain, J., Gaur, S., Chaudhary, Y. and Kaul, R. (2020) The molecular biology of intracellular events during Coronavirus infection cycle. *VirusDisease*, **31**, 1–5.
42. Fehr, A.R. and Perlman, S. (2015) Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol. Biol.*, **1282**, 1–23.
43. Masters, P.S. (2006) The molecular biology of coronaviruses. *Adv. Virus Res.*, **66**, 193–292.
44. Weiss, S.R. and Navas-Martin, S. (2005) Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol. Mol. Biol. Rev.*, **69**, 635–664.
45. Hartenian, E., Nandakumar, D., Lari, A., Ly, M., Tucker, J.M. and Glaunsinger, B.A. (2020) The molecular virology of coronaviruses. *J. Biol. Chem.*, **295**, 12910–12934.
46. Asano, K. (2014) Why is start codon selection so precise in eukaryotes? *Translation*, **2**, e28387.
47. Llacer, J.L., Hussain, T., Marler, L., Aitken, C.E., Thakur, A., Lorsch, J.R., Hinnebusch, A.G. and Ramakrishnan, V. (2015) Conformational differences between open and closed states of the Eukaryotic translation initiation complex. *Mol. Cell*, **59**, 399–412.
48. Obayashi, E., Luna, R.E., Nagata, T., Martin-Marcos, P., Hiraishi, H., Singh, C.R., Erzberger, J.P., Zhang, F., Arthanari, H., Morris, J. *et al.* (2017) Molecular landscape of the ribosome Pre-initiation complex during mRNA Scanning: Structural role for eIF3c and its control by eIF5. *Cell Rep.*, **18**, 2651–2663.
49. Herrmannova, A., Prilepskaja, T., Wagner, S., Sikrova, D., Zeman, J., Poncova, K. and Valasek, L.S. (2020) Adapted formaldehyde gradient cross-linking protocol implicates human eIF3d and eIF3c, k and l subunits in the 43S and 48S pre-initiation complex assembly, respectively. *Nucleic Acids Res.*, **48**, 1969–1984.
50. Gallie, D.R. (1991) The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev.*, **5**, 2108–2116.
51. Wells, S.E., Hillner, P.E., Vale, R.D. and Sachs, A.B. (1998) Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell*, **2**, 135–140.
52. Archer, S.K., Shirokikh, N.E., Hallwirth, C.V., Beilharz, T.H. and Preiss, T. (2015) Probing the closed-loop model of mRNA translation in living cells. *RNA Biol.*, **12**, 248–254.
53. Algire, M.A., Maag, D. and Lorsch, J.R. (2005) Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol. Cell*, **20**, 251–262.
54. Majumdar, R. and Maitra, U. (2005) Regulation of GTP hydrolysis prior to ribosomal AUG selection during eukaryotic translation initiation. *EMBO J.*, **24**, 3737–3746.
55. Shirokikh, N.E. and Preiss, T. (2018) Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. *Wiley Interdiscipl. Rev. RNA*, **9**, e1473.
56. Merrick, W.C. and Pavitt, G.D. (2018) Protein synthesis initiation in Eukaryotic cells. *Cold Spring Harb. Perspect. Biol.*, **10**, a033092.
57. Pelletier, J. and Sonenberg, N. (2019) The organizing principles of Eukaryotic ribosome recruitment. *Annu. Rev. Biochem.*, **88**, 307–335.
58. Advani, V.M. and Ivanov, P. (2019) Translational control under Stress: Reshaping the translome. *Bioessays*, **41**, e1900009.
59. Anderson, P. and Kedersha, N. (2008) Stress granules: the Tao of RNA triage. *Trends Biochem. Sci.*, **33**, 141–150.
60. Kedersha, N., Ivanov, P. and Anderson, P. (2013) Stress granules and cell signaling: more than just a passing phase? *Trends Biochem. Sci.*, **38**, 494–506.
61. McCormick, C. and Khapersky, D.A. (2017) Translation inhibition and stress granules in the antiviral immune response. *Nat. Rev. Immunol.*, **17**, 647–660.
62. Kozak, M. (1989) Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol. Cell. Biol.*, **9**, 5073–5080.
63. Kozak, M. (1987) At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.*, **196**, 947–950.
64. Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, **44**, 283–292.
65. Pisarev, A.V., Kolupaeva, V.G., Pisareva, V.P., Merrick, W.C., Hellen, C.U. and Pestova, T.V. (2006) Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev.*, **20**, 624–636.
66. Couso, J.P. and Patraquim, P. (2017) Classification and function of small open reading frames. *Nat. Rev. Mol. Cell Biol.*, **18**, 575–589.
67. Wek, R.C. (2018) Role of eIF2alpha kinases in translational control and adaptation to cellular stress. *Cold Spring Harb. Perspect. Biol.*, **10**, a032870.
68. Chikashige, Y., Kato, H., Thornton, M., Pepper, W., Hilgers, M., Cecil, A., Asano, I., Yamada, H., Mori, C., Brunkow, C. *et al.* (2020) Gen2 eIF2alpha kinase mediates combinatorial translational regulation through nucleotide motifs and uORFs in target mRNAs. *Nucleic Acids Res.*, **48**, 8977–8992.
69. Johnstone, T.G., Bazzini, A.A. and Giraldez, A.J. (2016) Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J.*, **35**, 706–723.
70. Schneider, W.M., Chevillotte, M.D. and Rice, C.M. (2014) Interferon-stimulated genes: a complex web of host defenses. *Annu. Rev. Immunol.*, **32**, 513–545.
71. Acosta, P.L., Byrne, A.B., Hijano, D.R. and Talarico, L.B. (2020) Human Type I Interferon Antiviral Effects in Respiratory and Reemerging Viral Infections. *J. Immunol. Res.*, **2020**, 1372494.
72. Chiang, H.S. and Liu, H.M. (2018) The molecular basis of viral inhibition of IRF- and STAT-Dependent immune responses. *Front. Immunol.*, **9**, 3086.
73. Walsh, D. and Mohr, I. (2011) Viral subversion of the host protein synthesis machinery. *Nat. Rev. Microbiol.*, **9**, 860–875.

74. Gradi, A., Svitkin, Y.V., Imataka, H. and Sonenberg, N. (1998) Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. *PNAS*, **95**, 11089–11094.
75. de Breyne, S., Bonderoff, J.M., Chumakov, K.M., Lloyd, R.E. and Hellen, C.U. (2008) Cleavage of eukaryotic initiation factor eIF5B by enterovirus 3C proteases. *Virology*, **378**, 118–122.
76. Rivera, C.I. and Lloyd, R.E. (2008) Modulation of enteroviral proteinase cleavage of poly(A)-binding protein (PABP) by conformation and PABP-associated factors. *Virology*, **375**, 59–72.
77. Roberts, L.O., Jopling, C.L., Jackson, R.J. and Willis, A.E. (2009) Viral strategies to subvert the mammalian translation machinery. *Prog. Mol. Biol. Transl. Sci.*, **90**, 313–367.
78. Kwan, T. and Thompson, S.R. (2019) Noncanonical translation initiation in Eukaryotes. *Cold Spring Harb. Perspect. Biol.*, **11**, a032672.
79. Walker, A.P. and Fodor, E. (2019) Interplay between influenza virus and the host RNA polymerase II transcriptional machinery. *Trends Microbiol.*, **27**, 398–407.
80. Plotch, S.J., Bouloy, M., Ulmanen, I. and Krug, R.M. (1981) A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell*, **23**, 847–858.
81. Ventoso, I., Sanz, M.A., Molina, S., Berlanga, J.J., Carrasco, L. and Esteban, M. (2006) Translational resistance of late alphavirus mRNA to eIF2alpha phosphorylation: a strategy to overcome the antiviral effect of protein kinase PKR. *Genes Dev.*, **20**, 87–100.
82. Toribio, R., Diaz-Lopez, I., Boskovic, J. and Ventoso, I. (2016) An RNA trapping mechanism in Alphavirus mRNA promotes ribosome stalling and translation initiation. *Nucleic Acids Res.*, **44**, 4368–4380.
83. Siddell, S.G., Wege, H., Barthel, A. and ter Meulen, V. (1980) Coronavirus JHM: cell-free synthesis of structural protein p60. *J. Virol.*, **33**, 10–17.
84. Rottier, P.J., Horzinek, M.C. and van der Zeijst, B.A. (1981) Viral protein synthesis in mouse hepatitis virus strain A59-infected cells: effect of tunicamycin. *J. Virol.*, **40**, 350–357.
85. Hilton, A., Mizzen, L., MacIntyre, G., Cheley, S. and Anderson, R. (1986) Translational control in murine hepatitis virus infection. *J. Gen. Virol.*, **67**, 923–932.
86. Tahara, S.M., Dietlin, T.A., Bergmann, C.C., Nelson, G.W., Kyuwa, S., Anthony, R.P. and Stohlman, S.A. (1994) Coronavirus translational regulation: leader affects mRNA efficiency. *Virology*, **202**, 621–630.
87. Raaben, M., Groot Koerkamp, M.J., Rottier, P.J. and de Haan, C.A. (2007) Mouse hepatitis coronavirus replication induces host translational shutoff and mRNA decay, with concomitant formation of stress granules and processing bodies. *Cell. Microbiol.*, **9**, 2218–2229.
88. Huang, C., Lokugamage, K.G., Rozovics, J.M., Narayanan, K., Semler, B.L. and Makino, S. (2011) Alphacoronavirus transmissible gastroenteritis virus nsp1 protein suppresses protein translation in mammalian cells and in cell-free HeLa cell extracts but not in rabbit reticulocyte lysate. *J. Virol.*, **85**, 638–643.
89. Tohya, Y., Narayanan, K., Kamitani, W., Huang, C., Lokugamage, K. and Makino, S. (2009) Suppression of host gene expression by nsp1 proteins of group 2 bat coronaviruses. *J. Virol.*, **83**, 5282–5288.
90. Xiao, H., Xu, L.H., Yamada, Y. and Liu, D.X. (2008) Coronavirus spike protein inhibits host cell translation by interaction with eIF3f. *PLoS One*, **3**, e1494.
91. Kint, J., Langereis, M.A., Maier, H.J., Britton, P., van Kuppeveld, F.J., Koumans, J., Wiegertjes, G.F. and Forlenza, M. (2016) Infectious bronchitis coronavirus limits interferon production by inducing a host shutoff that requires accessory protein 5b. *J. Virol.*, **90**, 7519–7528.
92. Shen, Z., Ye, G., Deng, F., Wang, G., Cui, M., Fang, L., Xiao, S., Fu, Z.F. and Peng, G. (2018) Structural basis for the inhibition of host gene expression by porcine epidemic diarrhea virus nsp1. *J. Virol.*, **92**, e01896-17.
93. Wang, Y., Shi, H., Rigolet, P., Wu, N., Zhu, L., Xi, X.G., Vabret, A., Wang, X. and Wang, T. (2010) Nsp1 proteins of group 1 and SARS coronaviruses share structural and functional similarities. *Infect. Genet. Evol.*, **10**, 919–924.
94. Kamitani, W., Narayanan, K., Huang, C., Lokugamage, K., Ikegami, T., Ito, N., Kubo, H. and Makino, S. (2006) Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *PNAS*, **103**, 12885–12890.
95. Narayanan, K., Huang, C., Lokugamage, K., Kamitani, W., Ikegami, T., Tseng, C.T. and Makino, S. (2008) Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J. Virol.*, **82**, 4471–4479.
96. Lokugamage, K.G., Narayanan, K., Nakagawa, K., Terasaki, K., Ramirez, S.I., Tseng, C.T. and Makino, S. (2015) Middle east respiratory syndrome coronavirus nsp1 inhibits host gene expression by selectively targeting mRNAs transcribed in the nucleus while sparing mRNAs of cytoplasmic origin. *J. Virol.*, **89**, 10970–10981.
97. Thoms, M., Buschauer, R., Ameismeier, M., Koepke, L., Denk, T., Hirschenberger, M., Kratzat, H., Hayn, M., Mackens-Kiani, T., Cheng, J. *et al.* (2020) Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science*, **369**, 1249–1255.
98. Schubert, K., Karousis, E.D., Jomaa, A., Scaiola, A., Echeverria, B., Gurzeler, L.A., Leibundgut, M., Thiel, V., Muhlemann, O. and Ban, N. (2020) SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat. Struct. Mol. Biol.*, **27**, 959–966.
99. Banerjee, A.K., Blanco, M.R., Bruce, E.A., Honson, D.D., Chen, L.M., Chow, A., Bhat, P., Ollikainen, N., Quinodoz, S.A., Loney, C. *et al.* (2020) SARS-CoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses. *Cell*, **183**, 1325–1339.
100. Chan, C.P., Siu, K.L., Chin, K.T., Yuen, K.Y., Zheng, B. and Jin, D.Y. (2006) Modulation of the unfolded protein response by the severe acute respiratory syndrome coronavirus spike protein. *J. Virol.*, **80**, 9279–9287.
101. Hetz, C., Zhang, K. and Kaufman, R.J. (2020) Mechanisms, regulation and functions of the unfolded protein response. *Nat. Rev. Mol. Cell Biol.*, **21**, 421–438.
102. Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V., Guo, J.Z., Swaney, D.L. *et al.* (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, **583**, 459–468.
103. Fung, T.S. and Liu, D.X. (2019) Human Coronavirus: Host-Pathogen interaction. *Annu. Rev. Microbiol.*, **73**, 529–557.
104. Perrier, A., Bonnin, A., Desmarest, L., Danneels, A., Goffard, A., Rouille, Y., Dubuisson, J. and Belouzard, S. (2019) The C-terminal domain of the MERS coronavirus M protein contains a trans-Golgi network localization signal. *J. Biol. Chem.*, **294**, 14406–14421.
105. Shajahan, A., Supekar, N.T., Gleinich, A.S. and Azadi, P. (2020) Deducing the N- and O- glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology*, doi:10.1093/glycob/cwaa042.
106. Oostra, M., de Haan, C.A., de Groot, R.J. and Rottier, P.J. (2006) Glycosylation of the severe acute respiratory syndrome coronavirus triple-spanning membrane proteins 3a and M. *J. Virol.*, **80**, 2326–2336.
107. Versteeg, G.A., van de Nes, P.S., Bredenbeek, P.J. and Spaan, W.J. (2007) The coronavirus spike protein induces endoplasmic reticulum stress and upregulation of intracellular chemokine mRNA concentrations. *J. Virol.*, **81**, 10981–10990.
108. Reggiori, F., Monastyrska, I., Verheije, M.H., Cali, T., Ulasli, M., Bianchi, S., Bernasconi, R., de Haan, C.A. and Molinari, M. (2010) Coronaviruses Hijack the LC3-I-positive EDEMosomes, ER-derived vesicles exporting short-lived ERAD regulators, for replication. *Cell Host Microbe*, **7**, 500–508.
109. Stertz, S., Reichelt, M., Spiegel, M., Kuri, T., Martinez-Sobrido, L., Garcia-Sastre, A., Weber, F. and Kochs, G. (2007) The intracellular sites of early replication and budding of SARS-coronavirus. *Virology*, **361**, 304–315.
110. Galluzzi, L., Diotallevi, A. and Magnani, M. (2017) Endoplasmic reticulum stress and unfolded protein response in infection by intracellular parasites. *Future Sci. OA*, **3**, FSO198.
111. Smith, J.A. (2014) A new paradigm: innate immune sensing of viruses via the unfolded protein response. *Front. Microbiol.*, **5**, 222.
112. He, B. (2006) Viruses, endoplasmic reticulum stress, and interferon responses. *Cell Death Differ.*, **13**, 393–403.

113. Krahling, V., Stein, D.A., Spiegel, M., Weber, F. and Muhlberger, E. (2009) Severe acute respiratory syndrome coronavirus triggers apoptosis via protein kinase R but is resistant to its antiviral activity. *J. Virol.*, **83**, 2298–2309.
114. Minakshi, R., Padhan, K., Rani, M., Khan, N., Ahmad, F. and Jameel, S. (2009) The SARS Coronavirus 3a protein causes endoplasmic reticulum stress and induces ligand-independent downregulation of the type 1 interferon receptor. *PLoS One*, **4**, e8342.
115. Siu, K.L., Chan, C.P., Kok, K.H., Woo, P.C. and Jin, D.Y. (2014) Comparative analysis of the activation of unfolded protein response by spike proteins of severe acute respiratory syndrome coronavirus and human coronavirus HKU1. *Cell Biosci.*, **4**, 3.
116. Nakagawa, K., Narayanan, K., Wada, M. and Makino, S. (2018) Inhibition of stress granule formation by middle east respiratory syndrome coronavirus 4a accessory protein facilitates viral translation, leading to efficient virus replication. *J. Virol.*, **92**, e00902-18.
117. Deng, X., Hackbart, M., Mettelman, R.C., O'Brien, A., Mielech, A.M., Yi, G., Kao, C.C. and Baker, S.C. (2017) Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *PNAS*, **114**, E4251–E4260.
118. Kindler, E., Gil-Cruz, C., Spanier, J., Li, Y., Wilhelm, J., Rabouw, H.H., Züst, R., Hwang, M., V'Kovski, P., Stalder, H. et al. (2017) Early endonuclease-mediated evasion of RNA sensing ensures efficient coronavirus replication. *PLoS Pathog.*, **13**, e1006195.
119. Rabouw, H.H., Langereis, M.A., Knaap, R.C., Dalebout, T.J., Canton, J., Sola, I., Enjuanes, L., Bredenbeek, P.J., Kikkert, M., de Groot, R.J. et al. (2016) Middle East respiratory coronavirus accessory protein 4a inhibits PKR-Mediated antiviral stress responses. *PLoS Pathog.*, **12**, e1005982.
120. Favreau, D.J., Meessen-Pinard, M., Desforges, M. and Talbot, P.J. (2012) Human coronavirus-induced neuronal programmed cell death is cyclophilin d dependent and potentially caspase dispensable. *J. Virol.*, **86**, 81–93.
121. Ivanov, P., Kedersha, N. and Anderson, P. (2019) Stress granules and processing bodies in translational control. *Cold Spring Harb. Perspect. Biol.*, **11**, a032813.
122. Riggs, C.L., Kedersha, N., Ivanov, P. and Anderson, P. (2020) Mammalian stress granules and P bodies at a glance. *J. Cell Sci.*, **133**, jcs242487.
123. Corbet, G.A. and Parker, R. (2019) RNP Granule Formation: Lessons from P-Bodies and stress granules. *Cold Spring Harb. Symp. Quant. Biol.*, **84**, 203–215.
124. Protter, D.S.W. and Parker, R. (2016) Principles and properties of stress granules. *Trends Cell Biol.*, **26**, 668–679.
125. Eiermann, N., Haneke, K., Sun, Z., Stoecklin, G. and Ruggieri, A. (2020) Dance with the Devil: Stress granules and signaling in antiviral responses. *Viruses*, **12**, 984.
126. Tsai, W.C. and Lloyd, R.E. (2014) Cytoplasmic RNA granules and viral infection. *Ann. Rev. Virol.*, **1**, 147–170.
127. Sola, I., Mateos-Gomez, P.A., Almazan, F., Zuniga, S. and Enjuanes, L. (2011) RNA-RNA and RNA-protein interactions in coronavirus replication and transcription. *RNA Biol.*, **8**, 237–248.
128. Peng, T.Y., Lee, K.R. and Tarn, W.Y. (2008) Phosphorylation of the arginine/serine dipeptide-rich motif of the severe acute respiratory syndrome coronavirus nucleocapsid protein modulates its multimerization, translation inhibitory activity and cellular localization. *FEBS J.*, **275**, 4152–4163.
129. Mizutani, T., Fukushi, S., Murakami, M., Hirano, T., Saijo, M., Kurane, I. and Morikawa, S. (2004) Tyrosine dephosphorylation of STAT3 in SARS coronavirus-infected Vero E6 cells. *FEBS Lett.*, **577**, 187–192.
130. Mizutani, T., Fukushi, S., Saijo, M., Kurane, I. and Morikawa, S. (2004) Phosphorylation of p38 MAPK and its downstream targets in SARS coronavirus-infected cells. *Biochem. Biophys. Res. Commun.*, **319**, 1228–1234.
131. Banerjee, S., Narayanan, K., Mizutani, T. and Makino, S. (2002) Murine coronavirus replication-induced p38 mitogen-activated protein kinase activation promotes interleukin-6 production and virus replication in cultured cells. *J. Virol.*, **76**, 5937–5948.
132. McKendrick, L., Thompson, E., Ferreira, J., Morley, S.J. and Lewis, J.D. (2001) Interaction of eukaryotic translation initiation factor 4G with the nuclear cap-binding complex provides a link between nuclear and cytoplasmic functions of the m(7) guanosine cap. *Mol. Cell. Biol.*, **21**, 3632–3641.
133. Knauf, U., Tschopp, C. and Gram, H. (2001) Negative regulation of protein translation by mitogen-activated protein kinase-interacting kinases 1 and 2. *Mol. Cell. Biol.*, **21**, 5500–5511.
134. Pyronnet, S., Imataka, H., Gingras, A.C., Fukunaga, R., Hunter, T. and Sonenberg, N. (1999) Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *EMBO J.*, **18**, 270–279.
135. Lachance, P.E.D., Miron, M., Raught, B., Sonenberg, N. and Lasko, P. (2002) Phosphorylation of eukaryotic translation initiation factor 4E is critical for growth. *Mol. Cell. Biol.*, **22**, 1656–1663.
136. Panja, D., Kenney, J.W., D'Andrea, L., Zalfa, F., Vedeler, A., Wibrand, K., Fukunaga, R., Bagni, C., Proud, C.G. and Bramham, C.R. (2014) Two-Stage translational control of dentate gyrus LTP consolidation is mediated by sustained BDNF-TrkB signaling to MNK. *Cell Rep.*, **9**, 1430–1445.
137. Bramham, C.R., Jensen, K.B. and Proud, C.G. (2016) Tuning specific translation in cancer metastasis and synaptic Memory: Control at the MNK-eIF4E Axis. *Trends Biochem. Sci.*, **41**, 847–858.
138. Kopecky-Bromberg, S.A., Martinez-Sobrido, L. and Palese, P. (2006) 7a protein of severe acute respiratory syndrome coronavirus inhibits cellular protein synthesis and activates p38 mitogen-activated protein kinase. *J. Virol.*, **80**, 785–793.
139. Padhan, K., Minakshi, R., Bin Towheed, M.A. and Jameel, S. (2008) Severe acute respiratory syndrome coronavirus 3a protein activates the mitochondrial death pathway through p38 MAP kinase activation. *J. Gen. Virol.*, **89**, 1960–1969.
140. Kono, M., Tatsumi, K., Imai, A.M., Saito, K., Kuriyama, T. and Shirasawa, H. (2008) Inhibition of human coronavirus 229E infection in human epithelial lung cells (L132) by chloroquine: Involvement of p38 MAPK and ERK. *Antiviral Res.*, **77**, 150–152.
141. Proud, C.G. (2019) Phosphorylation and signal transduction pathways in translational control. *Cold Spring Harb. Perspect. Biol.*, **11**, a033050.
142. Lim, Y.X., Ng, Y.L., Tam, J.P. and Liu, D.X. (2016) Human Coronaviruses: A review of Virus-Host interactions. *Diseases*, **4**, 26.
143. Chen, S.C. and Olsthoorn, R.C. (2010) Group-specific structural features of the 5'-proximal sequences of coronavirus genomic RNAs. *Virology*, **401**, 29–41.
144. Bouhaddou, M., Memon, D., Meyer, B., White, K.M., Rezelj, V.V., Marrero, M.C., Polacco, B.J., Melnyk, J.E., Ulberts, S., Kaake, R.M. et al. (2020) The global phosphorylation landscape of SARS-CoV-2 infection. *Cell*, **182**, 685–712.
145. Narayanan, K., Ramirez, S.I., Lokugamage, K.G. and Makino, S. (2015) Coronavirus nonstructural protein 1: Common and distinct functions in the regulation of host and viral gene expression. *Virus Res.*, **202**, 89–100.
146. Prentice, E., Jerome, W.G., Yoshimori, T., Mizushima, N. and Denison, M.R. (2004) Coronavirus replication complex formation utilizes components of cellular autophagy. *J. Biol. Chem.*, **279**, 10136–10141.
147. Lokugamage, K.G., Narayanan, K., Huang, C. and Makino, S. (2012) Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. *J. Virol.*, **86**, 13598–13608.
148. Kamitani, W., Huang, C., Narayanan, K., Lokugamage, K.G. and Makino, S. (2009) A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat. Struct. Mol. Biol.*, **16**, 1134–1140.
149. Wilson, J.E., Pestova, T.V., Hellen, C.U. and Sarnow, P. (2000) Initiation of protein synthesis from the A site of the ribosome. *Cell*, **102**, 511–520.
150. Balvay, L., Soto Rifo, R., Ricci, E.P., Decimo, D. and Ohlmann, T. (2009) Structural and functional diversity of viral IRESes. *Biochim. Biophys. Acta*, **1789**, 542–557.
151. Pestova, T.V., Shatsky, I.N., Fletcher, S.P., Jackson, R.J. and Hellen, C.U. (1998) A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev.*, **12**, 67–83.

152. Quade, N., Boehringer, D., Leibundgut, M., van den Heuvel, J. and Ban, N. (2015) Cryo-EM structure of Hepatitis C virus IRES bound to the human ribosome at 3.9-Å resolution. *Nat. Commun.*, **6**, 7646.
153. Spahn, C.M., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P. and Frank, J. (2004) Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell*, **118**, 465–475.
154. Johnson, A.G., Grosely, R., Petrov, A.N. and Puglisi, J.D. (2017) Dynamics of IRES-mediated translation. *Philos. Trans. R. Soc. Lon. Series B, Biol. Sci.*, **372**, 20160177.
155. Gaglia, M.M., Covarrubias, S., Wong, W. and Glaunsinger, B.A. (2012) A common strategy for host RNA degradation by divergent viruses. *Bba-Gene Regul. Mech.*, **1829**, 590–603.
156. Nagarajan, V.K., Jones, C.I., Newbury, S.F. and Green, P.J. (2013) XRN 5' → 3' exoribonucleases: Structure, mechanisms and functions. *Bba-Gene Regul. Mech.*, **1829**, 590–603.
157. Perez-Ortin, J.E., Alepuz, P., Chavez, S. and Choder, M. (2013) Eukaryotic mRNA Decay: Methodologies, pathways, and links to other stages of gene expression. *J. Mol. Biol.*, **425**, 3750–3775.
158. Labno, A., Tomecki, R. and Dziembowski, A. (2016) Cytoplasmic RNA decay pathways - Enzymes and mechanisms. *Bba-Mol. Cell Res.*, **1863**, 3125–3147.
159. Huang, C., Lokugamage, K.G., Rozovics, J.M., Narayanan, K., Semler, B.L. and Makino, S. (2011) SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLoS Pathog.*, **7**, e1002433.
160. Sweeney, T.R., Abaeva, I.S., Pestova, T.V. and Hellen, C.U. (2014) The mechanism of translation initiation on Type I picornavirus IRESs. *EMBO J.*, **33**, 76–92.
161. Powers, K.T., Szeto, J.A. and Schaffitzel, C. (2020) New insights into non-go, non-stop and nonsense-mediated mRNA decay complexes. *Curr. Opin. Struct. Biol.*, **65**, 110–118.
162. Navickas, A., Chamois, S., Saint-Fort, R., Henri, J., Torchet, C. and Benard, L. (2020) No-Go Decay mRNA cleavage in the ribosome exit tunnel produces 5'-OH ends phosphorylated by Trl1. *Nat. Commun.*, **11**, 122.
163. Terada, Y., Kawachi, K., Matsuura, Y. and Kamitani, W. (2017) MERS coronavirus nsp1 participates in an efficient propagation through a specific interaction with viral RNA. *Virology*, **511**, 95–105.
164. Almeida, M.S., Johnson, M.A., Herrmann, T., Geralt, M. and Wuthrich, K. (2007) Novel beta-barrel fold in the nuclear magnetic resonance structure of the replicase nonstructural protein 1 from the severe acute respiratory syndrome coronavirus. *J. Virol.*, **81**, 3151–3161.
165. Brockway, S.M., Lu, X.T., Peters, T.R., Dermody, T.S. and Denison, M.R. (2004) Intracellular localization and protein interactions of the gene 1 protein p28 during mouse hepatitis virus replication. *J. Virol.*, **78**, 11551–11562.
166. Zust, R., Cervantes, L., Weber, F., Kuri, T., Davidson, A., Siddell, S.G., Ludewig, B. and Thiel, V. (2007) Interaction of mouse hepatitis non-structural protein 1 with innate immune effector pathways. *Swiss Med. Wkly.*, **137**, 35s.
167. Lei, L., Ying, S., Baojun, L., Yi, Y. and Xiang, H. (2014) Attenuation of mouse hepatitis virus by deletion of the LLRKxGxKG region of Nsp1 (vol 8, e61166, 2013). *PLoS One*, **9**, e61166.
168. Nakagawa, K., Narayanan, K., Wada, M., Popov, V.L., Cajimat, M., Baric, R.S. and Makino, S. (2018) The Endonucleolytic RNA cleavage function of nsp1 of middle east respiratory syndrome coronavirus promotes the production of infectious virus particles in specific human cell lines. *J. Virol.*, **92**, e01157-18.
169. Shen, Z., Wang, G., Yang, Y.L., Shi, J.L., Fang, L.R., Li, F., Xiao, S.B., Fu, Z.F. and Peng, G.Q. (2019) A conserved region of nonstructural protein 1 from alphacoronaviruses inhibits host gene expression and is critical for viral virulence. *J. Biol. Chem.*, **294**, 13606–13618.
170. Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G. and Yusupov, M. (2011) The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science*, **334**, 1524–1529.
171. Weisser, M. and Ban, N. (2019) Extensions, extra factors, and extreme Complexity: Ribosomal structures provide insights into Eukaryotic translation. *Cold Spring Harb. Perspect. Biol.*, **11**, a032367.
172. Zhou, B., Liu, J., Wang, Q., Liu, X., Li, X., Li, P., Ma, Q. and Cao, C. (2008) The nucleocapsid protein of severe acute respiratory syndrome coronavirus inhibits cell cytokinesis and proliferation by interacting with translation elongation factor 1alpha. *J. Virol.*, **82**, 6962–6971.
173. Riis, B., Rattan, S.I., Clark, B.F. and Merrick, W.C. (1990) Eukaryotic protein elongation factors. *Trends Biochem. Sci.*, **15**, 420–424.
174. Hotokezaka, Y., Tobben, U., Hotokezaka, H., Van Leyen, K., Beatrix, B., Smith, D.H., Nakamura, T. and Wiedmann, M. (2002) Interaction of the eukaryotic elongation factor 1A with newly synthesized polypeptides. *J. Biol. Chem.*, **277**, 18545–18551.
175. Rosenzweig, R., Nillegoda, N.B., Mayer, M.P. and Bukau, B. (2019) The Hsp70 chaperone network. *Nat. Rev. Mol. Cell Biol.*, **20**, 665–680.
176. Morita, M., Ler, L.W., Fabian, M.R., Siddiqui, N., Mullin, M., Henderson, V.C., Alain, T., Fonseca, B.D., Karashchuk, G., Bennett, C.F. et al. (2012) A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. *Mol. Cell Biol.*, **32**, 3585–3593.
177. Cornillez-Ty, C.T., Liao, L.J., Yates, J.R., Kuhn, P. and Buchmeier, M.J. (2009) Severe acute respiratory syndrome coronavirus nonstructural protein 2 interacts with a host protein complex involved in mitochondrial biogenesis and intracellular signaling. *J. Virol.*, **83**, 10314–10318.
178. Marintchev, A. (2013) Roles of helicases in translation initiation: a mechanistic view. *Biochim. Biophys. Acta*, **1829**, 799–809.
179. Sun, Y., Atas, E., Lindqvist, L.M., Sonenberg, N., Pelletier, J. and Meller, A. (2014) Single-molecule kinetics of the eukaryotic initiation factor 4A1 upon RNA unwinding. *Structure*, **22**, 941–948.
180. Rogers, G.W. Jr., Richter, N.J., Lima, W.F. and Merrick, W.C. (2001) Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J. Biol. Chem.*, **276**, 30914–30922.
181. Brian, D.A. and Baric, R.S. (2005) Coronavirus genome structure and replication. *Curr. Top. Microbiol. Immunol.*, **287**, 1–30.
182. Spaan, W., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., van der Zeijst, B.A. and Siddell, S.G. (1983) Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.*, **2**, 1839–1844.
183. Yang, D. and Leibowitz, J.L. (2015) The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res.*, **206**, 120–133.
184. Madhugiri, R., Fricke, M., Marz, M. and Ziebuhr, J. (2016) Coronavirus cis-Acting RNA Elements. *Adv. Virus Res.*, **96**, 127–163.
185. Kang, H., Feng, M., Schroeder, M.E., Giedroc, D.P. and Leibowitz, J.L. (2006) Stem-loop 1 in the 5' UTR of the SARS coronavirus can substitute for its counterpart in mouse hepatitis virus. *Adv. Exp. Med. Biol.*, **581**, 105–108.
186. Liu, P., Li, L., Millership, J.J., Kang, H., Leibowitz, J.L. and Giedroc, D.P. (2007) A U-turn motif-containing stem-loop in the coronavirus 5' untranslated region plays a functional role in replication. *RNA*, **13**, 763–780.
187. Yang, D., Liu, P., Wudeck, E.V., Giedroc, D.P. and Leibowitz, J.L. (2015) SHAPE analysis of the RNA secondary structure of the Mouse Hepatitis Virus 5' untranslated region and N-terminal nsp1 coding sequences. *Virology*, **475**, 15–27.
188. Kang, H., Feng, M., Schroeder, M.E., Giedroc, D.P. and Leibowitz, J.L. (2006) Putative cis-acting stem-loops in the 5' untranslated region of the severe acute respiratory syndrome coronavirus can substitute for their mouse hepatitis virus counterparts. *J. Virol.*, **80**, 10600–10614.
189. Tanaka, T., Kamitani, W., DeDiego, M.L., Enjuanes, L. and Matsuura, Y. (2012) Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. *J. Virol.*, **86**, 11128–11137.
190. Lai, M.M., Brayton, P.R., Armen, R.C., Patton, C.D., Pugh, C. and Stohlman, S.A. (1981) Mouse hepatitis virus A59: mRNA structure and genetic localization of the sequence divergence from hepatotropic strain MHV-3. *J. Virol.*, **39**, 823–834.
191. Lai, M.M., Patton, C.D. and Stohlman, S.A. (1982) Replication of mouse hepatitis virus: negative-stranded RNA and replicative form RNA are of genome length. *J. Virol.*, **44**, 487–492.
192. Yang, Y., Hussain, S., Wang, H., Ke, M. and Guo, D. (2009) Translational control of the subgenomic RNAs of severe acute respiratory syndrome coronavirus. *Virus Genes*, **39**, 10–18.
193. Senanayake, S.D. and Brian, D.A. (1997) Bovine coronavirus I protein synthesis follows ribosomal scanning on the bicistronic N mRNA. *Virus Res.*, **48**, 101–105.

194. Senanayake, S.D. and Brian, D.A. (1999) Translation from the 5' untranslated region (UTR) of mRNA 1 is repressed, but that from the 5' UTR of mRNA 7 is stimulated in coronavirus-infected cells. *J. Virol.*, **73**, 8003–8009.
195. Cencic, R., Hall, D.R., Robert, F., Du, Y., Min, J., Li, L., Qui, M., Lewis, I., Kurtkaya, S., Dingleline, R. *et al.* (2011) Reversing chemoresistance by small molecule inhibition of the translation initiation complex eIF4F. *PNAS*, **108**, 1046–1051.
196. Cencic, R., Desforges, M., Hall, D.R., Kozakov, D., Du, Y., Min, J., Dingleline, R., Fu, H., Vajda, S., Talbot, P.J. *et al.* (2011) Blocking eIF4E-eIF4G interaction as a strategy to impair coronavirus replication. *J. Virol.*, **85**, 6381–6389.
197. Muller, C., Ulyanova, V., Ilinskaya, O., Pleschka, S. and Shah Mahmud, R. (2017) A Novel Antiviral Strategy against MERS-CoV and HCoV-229E Using Binase to Target Viral Genome Replication. *Bionanoscience*, **7**, 294–299.
198. Schaefer, S.R., Mackenzie, J.M. and Pekosz, A. (2007) The ORF7b protein of severe acute respiratory syndrome coronavirus (SARS-CoV) is expressed in virus-infected cells and incorporated into SARS-CoV particles. *J. Virol.*, **81**, 718–731.
199. Senanayake, S.D., Hofmann, M.A., Maki, J.L. and Brian, D.A. (1992) The nucleocapsid protein gene of bovine coronavirus is bicistronic. *J. Virol.*, **66**, 5277–5283.
200. Wu, H.Y., Guan, B.J., Su, Y.P., Fan, Y.H. and Brian, D.A. (2014) Reselection of a genomic upstream open reading frame in mouse hepatitis coronavirus 5'-untranslated-region mutants. *J. Virol.*, **88**, 846–858.
201. Hofmann, M.A., Senanayake, S.D. and Brian, D.A. (1993) An intraleader open reading frame is selected from a hypervariable 5' terminus during persistent infection by the bovine coronavirus. *Adv. Exp. Med. Biol.*, **342**, 105–109.
202. Irigoyen, N., Firth, A.E., Jones, J.D., Chung, B.Y., Siddell, S.G. and Brierley, I. (2016) High-Resolution analysis of coronavirus gene expression by RNA sequencing and ribosome profiling. *PLoS Pathog.*, **12**, e1005473.
203. Chen, W. and Baric, R.S. (1995) Function of a 5'-end genomic RNA mutation that evolves during persistent mouse hepatitis virus infection in vitro. *J. Virol.*, **69**, 7529–7540.
204. Hofmann, M.A., Senanayake, S.D. and Brian, D.A. (1993) A translation-attenuating intraleader open reading frame is selected on coronavirus mRNAs during persistent infection. *PNAS*, **90**, 11733–11737.
205. Hussain, S., Pan, J., Chen, Y., Yang, Y., Xu, J., Peng, Y., Wu, Y., Li, Z., Zhu, Y., Tien, P. *et al.* (2015) Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus. *J. Virol.*, **79**, 5288–5295.
206. Lin, Y.J. and Lai, M.M. (1993) Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontinuous sequence for replication. *J. Virol.*, **67**, 6110–6118.
207. Dalton, K., Casais, R., Shaw, K., Stirrups, K., Evans, S., Britton, P., Brown, T.D. and Cavanagh, D. (2001) cis-acting sequences required for coronavirus infectious bronchitis virus defective-RNA replication and packaging. *J. Virol.*, **75**, 125–133.
208. Mendez, A., Smerdou, C., Izeta, A., Gebauer, F. and Enjuanes, L. (1996) Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. *Virology*, **217**, 495–507.
209. Zust, R., Miller, T.B., Goebel, S.J., Thiel, V. and Masters, P.S. (2008) Genetic interactions between an essential 3' cis-acting RNA pseudoknot, replicase gene products, and the extreme 3' end of the mouse coronavirus genome. *J. Virol.*, **82**, 1214–1228.
210. Hsue, B. and Masters, P.S. (1997) A bulged stem-loop structure in the 3' untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication. *J. Virol.*, **71**, 7567–7578.
211. Goebel, S.J., Hsue, B., Dombrowski, T.F. and Masters, P.S. (2004) Characterization of the RNA components of a putative molecular switch in the 3' untranslated region of the murine coronavirus genome. *J. Virol.*, **78**, 669–682.
212. Deo, R.C., Bonanno, J.B., Sonenberg, N. and Burley, S.K. (1999) Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell*, **98**, 835–845.
213. Ozturk, S. and Uysal, F. (2018) Potential roles of the poly(A)-binding proteins in translational regulation during spermatogenesis. *J. Reprod. Dev.*, **64**, 289–296.
214. Smith, R.W., Blee, T.K. and Gray, N.K. (2014) Poly(A)-binding proteins are required for diverse biological processes in metazoans. *Biochem. Soc. Trans.*, **42**, 1229–1237.
215. Villalba, A., Coll, O. and Gebauer, F. (2011) Cytoplasmic polyadenylation and translational control. *Curr. Opin. Genet. Dev.*, **21**, 452–457.
216. Radford, H.E., Meijer, H.A. and de Moor, C.H. (2008) Translational control by cytoplasmic polyadenylation in *Xenopus* oocytes. *Biochim. Biophys. Acta*, **1779**, 217–229.
217. Hofmann, M.A. and Brian, D.A. (1991) The 5' end of coronavirus minus-strand RNAs contains a short poly(U) tract. *J. Virol.*, **65**, 6331–6333.
218. Peng, Y.H., Lin, C.H., Lin, C.N., Lo, C.Y., Tsai, T.L. and Wu, H.Y. (2016) Characterization of the Role of Hexamer AGUAAA and Poly(A) Tail in Coronavirus Polyadenylation. *PLoS One*, **11**, e0165077.
219. Wu, H.Y., Ke, T.Y., Liao, W.Y. and Chang, N.Y. (2013) Regulation of coronavirus poly(A) tail length during infection. *PLoS One*, **8**, e70548.
220. Shien, J.H., Su, Y.D. and Wu, H.Y. (2014) Regulation of coronavirus poly(A) tail length during infection is not coronavirus species- or host cell-specific. *Virus Genes*, **49**, 383–392.
221. Tvarogova, J., Madhugiri, R., Bylapudi, G., Ferguson, L.J., Karl, N. and Ziebuhr, J. (2019) Identification and characterization of a human coronavirus 229E nonstructural Protein 8-Associated RNA 3'-Terminal Adenylyltransferase activity. *J. Virol.*, **93**, e00291-19.
222. Spagnolo, J.F. and Hogue, B.G. (2000) Host protein interactions with the 3' end of bovine coronavirus RNA and the requirement of the poly(A) tail for coronavirus defective genome replication. *J. Virol.*, **74**, 5053–5065.
223. Tsai, T.L., Lin, C.H., Lin, C.N., Lo, C.Y. and Wu, H.Y. (2018) Interplay between the Poly(A) Tail, Poly(A)-Binding protein, and coronavirus nucleocapsid protein regulates gene expression of coronavirus and the host cell. *J. Virol.*, **92**, e01162-18.
224. Brockway, S.M., Clay, C.T., Lu, X.T. and Denison, M.R. (2003) Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. *J. Virol.*, **77**, 10515–10527.
225. Baric, R.S. and Yount, B. (2000) Subgenomic negative-strand RNA function during mouse hepatitis virus infection. *J. Virol.*, **74**, 4039–4046.
226. Sawicki, S.G. and Sawicki, D.L. (1990) Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J. Virol.*, **64**, 1050–1056.
227. Sethna, P.B., Hung, S.L. and Brian, D.A. (1989) Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *PNAS*, **86**, 5626–5630.
228. Zuniga, S., Sola, I., Alonso, S. and Enjuanes, L. (2004) Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J. Virol.*, **78**, 980–994.
229. Jauregui, A.R., Savalia, D., Lowry, V.K., Farrell, C.M. and Wathelet, M.G. (2013) Identification of residues of SARS-CoV nsp1 that differentially affect inhibition of gene expression and antiviral signaling. *PLoS One*, **8**, e62416.
230. Gomez, G.N., Abrar, F., Dodhia, M.P., Gonzalez, F.G. and Nag, A. (2019) SARS coronavirus protein nsp1 disrupts localization of Nup93 from the nuclear pore complex. *Biochem. Cell. Biol.*, **97**, 758–766.
231. Prentice, E., McAuliffe, J., Lu, X., Subbarao, K. and Denison, M.R. (2004) Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. *J. Virol.*, **78**, 9977–9986.
232. Ghisolfi-Nieto, L., Joseph, G., Puvion-Dutilleul, F., Amalric, F. and Bouvet, P. (1996) Nucleolin is a sequence-specific RNA-binding protein: Characterization of targets on pre-ribosomal RNA. *J. Mol. Biol.*, **260**, 34–53.
233. Boccalletto, P., Machnicka, M.A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T.K., de Crecy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A. *et al.* (2018) MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.*, **46**, D303–D307.

234. Zhao,B.X.S., Roundtree,I.A. and He,C. (2017) Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.*, **18**, 31–42.
235. Williams,G.D., Gokhale,N.S. and Horner,S.M. (2019) Regulation of Viral Infection by the RNA Modification N6-Methyladenosine. *Ann. Rev. Virol.*, **6**, 235–253.
236. Chen,J.N., Jin,L., Wang,Z.M., Wang,L.Y., Chen,Q.B., Cui,Y.R. and Liu,G.L. (2020) N6-methyladenosine regulates PEDV replication and host gene expression. *Virology*, **548**, 59–72.
237. Zhou,Y., Zeng,P., Li,Y.H., Zhang,Z.D. and Cui,Q.H. (2016) SRAMP: prediction of mammalian N-6-methyladenosine (m(6)A) sites based on sequence-derived features. *Nucleic Acids Res.*, **44**, e91.

2 Annexe II

FIGURES SUPPLÉMENTAIRES ARTICLE Condé et al. 2021

Translation of SARS-CoV-2 gRNA Is Extremely Efficient and Competitive despite a High Degree of Secondary Structures and the Presence of an uORF

→ *Renilla Reniformis*
MDTSKVYDPE... - Globin
M**E**TSKVYDPE... - ORF-1ab
M**F**TSKVYDPE... - ORF-S / ORF-6
MDTSKVYDPE... - ORF-3a
M**Y**TSKVYDPE... - ORF-E
M**A**TSKVYDPE... - ORF-M
M**K**TSKVYDPE... - ORF-7a / ORF-8
M**I**TSKVYDPE... - ORF-7b
M**S**TSKVYDPE... - ORF-N

Figure S1. N-terminal modification of the Renilla luciferase coding region. Sequence alignment of the beginning of the Renilla coding region from the SARS-CoV-2 5'UTRs construct used in the figure 3. As the nucleotide context of the AUG codon is conserved; this only induced a modification of the second amino acid (indicated in bold).

Condé et al., revised Figure S1

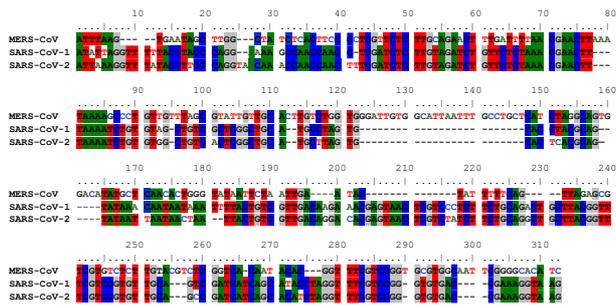


Figure S2. SARS CoV-1/2 and MERS CoV gRNA 5'UTRs conservation. The sequences of the 5'UTR from the SARS-CoV type 1 and 2 and MERS-CoV gRNA have been analyzed with a Multiple Sequence Comparison by Log- Expectation (MUSCLE) using the default parameters. Nucleotide identity has been highlighted in green for adenosine, blue for cytosine, red for tyrosine and grey for guanosine.

Condé et al., revised Figure S2

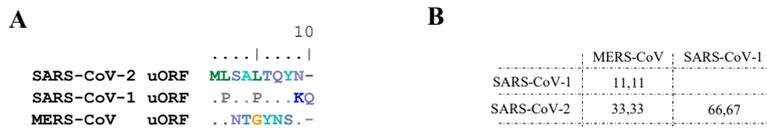


Figure S3. 5'UTR^{ORF1ab} uORFs conservation. (A) The sequence of the 5'UTR^{ORF1ab} uORFs from SARS-CoV-1, SARS-CoV-2 and MERS-CoV were analyzed with a Multiple Sequence Comparison by Log- Expectation (MUSCLE) using the default parameters. (B) Percentage of identity between the uORFs from 5'UTRs of MERS-CoV, SARS-CoV-1 and SARS-CoV-2 gRNAs resulting from a MUSCLE alignment.

Condé et al., revised Figure S3

Constructs	# PCR	5'-3' Primers	Template
pSARS-CoV-2 Leader_Renilla	PCR #1	ATCTCTGTAGATCTGTTCTCTAAACGAACATGGATCTACTCGAAAGTTATG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAA AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pSARS-CoV-2 5'UTR ^{ORF1ab} Renilla	PCR #1	TTCTGCGGTGTCGACCCGATCATCAGCACATCTAGGTTTCGTCGGGGTGTGACCGAAAGGTAAGATGGATCTACTTCGAAAGTTATG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	AATTAATAACTAATTAATCTGCTGTGACAGGACACGAGTAACCTGCTATCTCTGACAGCTGCTTACGGTTTCGTCGGTGTGACGGC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
	PCR #3	GTCTCTAAACGAACCTTAAAACTGTGTGGCTGCTACTGGGCTGATGCTAGTGCACCTACCGAGTAATAAATAACTAATTAATCTGCTGTGACAG AGATGTAAGCTTGCATAAATGTTAT	PCR #2
	PCR #4	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGGCTAGCATTAAAGGTTATACCTCCAGG AGATGTAAGCTTGCATAAATGTTAT	PCR #3
	PCR #5	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGGCTAGCATTAAAGGTTATACCTCCAGG AGATGTAAGCTTGCATAAATGTTAT	PCR #4
pSARS-CoV-1 Leader_Renilla	PCR #1	ATATTAGGTTTTACCTACCCAGGAAAAGCAACCACTCGATCTCTGTAGATCTGTTCTCTAAACGAACATGGATCTACTTCGAAAGTTATG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTACCCAGG AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pSARS-CoV-1 5'UTR ^{ORF1ab} Renilla	PCR #1	TCTTACCGTTTCCTCCGTTGCGATCGATCATCAGCATAGCTAGGTTTCGTCGGGGTGTGACCGAAAGGTAAGATGGATCTACTTCGAAAGTTATG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	ACCTACCCAGTAAACGAATAAATTAATCTGCTGTGACAGAAACGAGTAACCTGCTCTCTGTGACAGCTGCTACGGTTTCGTCGGT AGATGTAAGCTTGCATAAATGTTAT	PCR #1
	PCR #3	TCGATCTCTGTAGATCTGTTCTCTAAACGAACCTTAAATCTGTGTAGCTGCTGCTGCGCTGACCTAGTGCACCTACCGAGTAAACAAATAA AGATGTAAGCTTGCATAAATGTTAT	PCR #2
	PCR #4	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTACCCAGGAAAAGCAACCACTCGATCTCTGTAGATCTGTTCT AGATGTAAGCTTGCATAAATGTTAT	PCR #3
pMERS-CoV Leader_Renilla	PCR #1	ATTTAAGTCGAAAGCTTGGTATCTGCTACTCCCTGTTCTCTGCAACCTTGATTTAACGAACATGGATCTACTTCGAAAGTTATG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGTGAATAGCTGGCTATC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pMERS-CoV 5'UTR ^{ORF1ab} Renilla	PCR #1	TAGAGCTGCTGTTCTGTAGCTGCTCGGTACATACAGGTTTCGTCGGGGTGGCAATTCGGGGACATATGGATCTACTTCGAAAGTTATG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	AATTTGCTGCTCACTACGACGAGGACATATGCTCAACCTGGGTAATACTTAATGAATACTATTTTCAGTTAGAGCGGTGCTGCTCTGTG AGATGTAAGCTTGCATAAATGTTAT	PCR #1
	PCR #3	TAAAGCCCTGTTGTTAGCGTATTGTTGCACTGCTGCTGGGATTTGGCATTAAATTCGCTCATCTAGGC AGATGTAAGCTTGCATAAATGTTAT	PCR #2
	PCR #4	TCACTCCCTGCTCTGCTGCAAGACTTGTGATTTAACGAACCTTAAATGAAGCCCTGTTGTTAGCG AGATGTAAGCTTGCATAAATGTTAT	PCR #3
	PCR #5	TAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGTGAATAGCTGGCTATCTCACTCCCTGTTCTCTG AGATGTAAGCTTGCATAAATGTTAT	PCR #4

Table S1. List of primers used for the cloning of the SARS-CoV-1/2 and MERS-CoV leader and 5'UTR^{ORF1ab}.

Constructs	# PCR	5'-3' Primers	Template
pSARS-CoV-2 5'UTR-ORF-3a Renilla	PCR #1	TTCTGCGGTGTCGACCCGATCATCAGCACATCTAGGTTTCGTCGGGGTGTGACCGAAAGGTAAGATGGATCTACTTCGAAAGTTATG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	AATTAATAACTAATTAATCTGCTGTGACAGGACACGAGTAACCTGCTATCTCTGACAGCTGCTTACGGTTTCGTCGGTGTGACGGC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
	PCR #3	GTCTCTAAACGAACCTTAAAACTGTGTGGCTGCTACTGGGCTGATGCTAGTGCACCTACCGAGTAATAAATAACTAATTAATCTGCTGTGACAG AGATGTAAGCTTGCATAAATGTTAT	PCR #2
	PCR #4	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGGCTAGCATTAAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAA AGATGTAAGCTTGCATAAATGTTAT	PCR #3
	PCR #5	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGGCTAGCATTAAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAA AGATGTAAGCTTGCATAAATGTTAT	PCR #4
pSARS-CoV-2 5'UTR-ORF-3b Renilla	PCR #1	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pSARS-CoV-2 5'UTR-ORF-3c Renilla	PCR #1	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pSARS-CoV-2 5'UTR-ORF-4 Renilla	PCR #1	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pSARS-CoV-2 5'UTR-ORF-4 Renilla	PCR #1	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
	PCR #3	GACCATTCAGTAGAGAGTAAAGTAAATGCTTTCGCTGACAGTAAGTGAACAGAGATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #4	CAGGATCTGAGGTTTCGCTGATACAGTGGTACAGGATGGCACTAATAAATAACAGAGACTTCCAGTAGCAGTGC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pSARS-CoV-2 5'UTR-ORF-5 Renilla	PCR #1	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
	PCR #3	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	PCR #2
	PCR #4	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #3
pSARS-CoV-2 5'UTR-ORF-7a Renilla	PCR #1	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pSARS-CoV-2 5'UTR-ORF-7b Renilla	PCR #1	TATTTGTGGGGAAAGGTTTAAACACTTTCCTACACTCAAAGAAAGCAGAGTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	PCR #1
	PCR #3	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #2
pSARS-CoV-2 5'UTR-ORF-8 Renilla	PCR #1	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #1
pSARS-CoV-2 5'UTR-ORF-9 Renilla	PCR #1	TGTAGATCTGTTCTCTAAACGAACATGTTACTTCGAAAGTTATGATCCAG AGATGTAAGCTTGCATAAATGTTAT	pGlo-Renilla
	PCR #2	TAAGCAGGATCCCTCGAGTAAACGACTACTATAGATTAAGGTTATACCTCCAGGTAACAACCAACCACTTCGATCTCTGTAGATCTGTTCTCTAAACGAAC AGATGTAAGCTTGCATAAATGTTAT	PCR #1

Table S2. List of primers used for the cloning of the different 5'UTRs of the SARS-CoV-2 gRNA and sgRNAs

Constructs	# PCR	5'-3' Primers	Template
pSARS-CoV-2 5'UTR_Renilla Mut-AUG-uORF	PCR #1.1	TAAGCAGGATCCCTCGAGTAATACGACTCACTATAGATTAAGGTTTATACCTTCCAGGT AAGCAGCCGAGTGCAGCCACAGATTTAAAGTTCGTTAGAGAAACAGATCTAC	pSARS-CoV-2_5'UTR ^{ORF1ab} _Renilla
	PCR #1.2	TCTGTGGCTGTCACTCGCTGCTTCTAGTGCCTCAGGAG ACATCAGGATCCATCTTACCTTTCGGTCACACCC	pSARS-CoV-2_5'UTR ^{ORF1ab} _Renilla
	PCR 2	TAAGCAGGATCCCTCGAGTAATACGACTCACTATAGATTAAGGTTTATACCTTCCAGGT ACATCAGGATCCATCTTACCTTTCGGTCACACCC	PCR #1.1 and PCR #1.2
pSARS-CoV-2 5'UTR_Renilla uORF-phase	PCR #1.1	TAAGCAGGATCCCTCGAGTAATACGACTCACTATAGATTAAGGTTTATACCTTCCAGGT GTAATTATATGTAATTATACTCCGTGAGTGCCTAAGCATG	pSARS-CoV-2_5'UTR ^{ORF1ab} _Renilla
	PCR #1.2	AGTATAATTCAATATAAATTAAGTCTGCTGACAGGACAGGAG ACATCAGGATCCATCTTTCCTTTCGGACACACCCGGAC	pSARS-CoV-2_5'UTR ^{ORF1ab} _Renilla
	PCR 2	TAAGCAGGATCCCTCGAGTAATACGACTCACTATAGATTAAGGTTTATACCTTCCAGGT ACATCAGGATCCATCTTTCCTTTCGGACACACCCGGAC	PCR #1.1 and PCR #1.2
pMERS-CoV 5'UTR_Renilla Mut-AUG-uORF	PCR #1	TCACTTCCCCTCGTTCCTTCGAGAAGCTTGTATTTAAGCACTAAATAAAAGCCCTGTGTTAGGG AGATGTAAGCTTGCATAAATGTTAT	pMERS-CoV_5'UTR ^{ORF1ab} _Renilla
	PCR #2	TACCGAAGATCCATAAGCACTCACTATAGATTTAAGTGAATAGCTGGCTATCTCACTTCCCCTCGTTCCTTG AGATGTAAGCTTGCATAAATGTTAT	PCR #1
	PCR #3	TAAAAGCCCTGTGTTAGCGTATGTTGCACTGTCTGGTGGATTGTGGCAATAATTGCTCATCTAGGAGTGGACATTTGCTCAACACTGGG AGATGTAAGCTTGCATAAATGTTAT	PCR #2

Table S3. List of primers used for the cloning of the uORFs mutations

3 Annexe III

Liste des oligonucléotides de clonage

ORF 7a/8	Nom	Sequence oligonucléotides 5' ==> 3'
	1523_aNP0_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP0_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAACAGATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP0x_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTACCAGTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP0x_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGACGATAGAGAACAGATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP1_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP1_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGACCAGGGTTACAAGAACCCGAAAAGAGACCTA
	1523_aNP10_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP10_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAACCGATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP10x_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTACCAGTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP10x_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGACCAGGGTTACAAGAGATCGAAAAGAGACCTA
	1523_aNP11_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP11_R_LC	GTGGTCTCGAAGTTTTCATGTTTTCGTTTAGAGAACCGATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP11x_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTACCAGTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP11x_R_LC	GTGGTCTCGAAGTTTTCATGTTTTCGTTTAGAGAACCGATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP2_F_LC	TAGGTCTCTTCGAAATCTGTATTCTGCTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP2_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAGCAGAAATACAAGATTCGAAAAGAGACCTA
	1523_aNP3_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP3_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGACCAGGGTTACAAGAGATCGAAAAGAGACCTA
	1523_aNP4_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP4_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAGCAGAAATACAAGAGATCGAAAAGAGACCTA
	1523_aNP5_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP5_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAACAGATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP6_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP6_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAACATATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP6x_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTACCAGTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP6x_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAACATATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP7_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP7_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAACATATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP7x_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTACCAGTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP7x_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAACATATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP8_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP8_R_LC	GTGGTCTCGAAGTTTTCATGTTTTCGTTTAGAGAACATATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP8x_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTACCAGTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP8x_R_LC	GTGGTCTCGAAGTTTTCATGTTTTCGTTTAGAGAACATATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP9_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP9_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGAGAACCGATCTACAAGAGATCGAAAAGAGACCTA
	1523_aNP9x_F_LC	TAGGTCTCTTCGATCTCTTGTAGATCTGTTCTCTACCAGTAAACGAACATGAAAACCTCGAGACCAC
	1523_aNP9x_R_LC	GTGGTCTCGAAGTTTTCATGTTTCGTTTAGACGATAGAGAACCGATCTACAAGAGATCGAAAAGAGACCTA
	1523_O7_F_LC	CAAACCAACCACTTTCGAGAGACCGGAGCGCTTCCAGATCAGACAACCTGGTCTCAACTTCGAAAGTTTATGATCCAG
	1523_O7A_R_LC	TCTGTATTCTGAAGTTTTCATGTTTCGTTTAGAGAACAGATCT
	1523_O7B11_R_LC	TCTGTATTCTGAAGTTTTCATGTTTCGTTTAGAGAACAGATCT
	1523_O7B12_R_LC	TCTGTATTCTGAAGTTTTCATGTTTTCGTTTAGAGAACAGATCTACAAG
	1523_O7BX11_R_LC	TCTGTATTCTGAAGTTTTCATGTTTCGTTTAGAGAACAGATCTACAAGAGA
	1523_O7BX12_R_LC	TCTGTATTCTGAAGTTTTCATGTTTTCGTTTAGAGAACAGATCTACAAGAGA
	3321_LD_R_LC	TTCTCTGATATCTTCAGTTTGTGTC

ORF6	Nom	Sequence oligonucléotides 5' ==> 3'
	1523_O6P1_F_LC	TATGCAGGATCCCTAATACGACTCAGTATAGATTAAGGTTTATACCTTCCCAGG
	1523_O6P1_R_LC	TATGCAGGATCCCTAATACGACTCAGTATAGATTAAGGTTTATACCTTCCCAGG
	1523_O6P1_R_LC	TCTGTATTCTGAAGTAAACATCTGTGCACCTGTACAAGCAAAGCAATATTG
	1523_O6P2_R_LC	CTGGATCATAAACTTTCGAAGT
	1523_O6P21_F_LC	CAAACCAACCACTTTCGATCTCTTGTAGATATGTTCTCTAAACGAACGCTTTCCTTA
	1523_O6P21_F_LC	CAAACCAACCACTTTCGATCTCTTGTAGATATGTTCTCTAAACGAACATGAAAAC
	1523_O6P22_F_LC	CAAACCAACCACTTTCGATCTCTTGTAGATCCGTTCTCTAAACGAACGCTTTCCTTA
	1523_O6P23_F_LC	CAAACCAACCACTTTCGATCTCTTGTAAACCTGGTCTCTAAACGAACGCTTTCCTTA
	1523_O6P233_F_LC	CAAACCAACCACTTTCGATCTCTTGTAAACCTGGTCTCTAAACGAACGCTTTCCTTA
	1523_O6P24_F_LC	CAAACCAACCACTTTCGAAATCTTGTATTTCTGCTCTCTAAACGAACGCTTTCCTTA
	1523_O6P244_F_LC	CAAACCAACCACTTTCGATCTCTTGTATTTCTGCTCTCTAAACGAACGCTTTCCTTA
	1523_O6P2B_F_LC	TATGCAGGATCCCTAATACGACTCAGTATAGATTAAGGTTTATACCTTCCCAGGTAACAAACCAACCACTTTCG
	1523_O6P2B_F_LC	TATGCAGGATCCCTAATACGACTCAGTATAGATTAAGGTTTATACCTTCCCAGGTAACAAACCAACCACTTTCG

ORF Spike	Nom	Sequence oligonucléotides 5' ==> 3'
	1523_OS_F_LC	TATGCAGGATCCTAATACGACTCACTATAGATTAAGGTTTATACCTTCCCAGGTAACAAACCAACCACTTTCC
	3321_LD_R_LC	TTCTCTGATATCTTCAGTTGTGTC
	1523_S0_F_LC	CAAACCAACCAACTTTCGATCTCTTGATAGTCTGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_S1_F_LC	CAAACCAACCAACTTTCGATCTCTTGATACCCCTGGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_S2_F_LC	CAAACCAACCAACTTTCGATCTCTTGATACCCCTGGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_S3_F_LC	CAAACCAACCAACTTTCGATCTCTTGATACCCCTGGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_S4_F_LC	CAAACCAACCAACTTTCGATCTCTTGATATTCTGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_S5_F_LC	CAAACCAACCAACTTTCGATCTCTTGATATTCTGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_S6_F_LC	CAAACCAACCAACTTTCGATCTCTTGATATTCTGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_S7_F_LC	CAAACCAACCAACTTTCGATCTCTTGATAGTCTGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_S8_F_LC	CAAACCAACCAACTTTCGATCTCTTGATAGTCTGTCTCTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx0_F_LC	CAAACCAACCAACTTTCGATCTCTTGATAGTCTGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx1_F_LC	CAAACCAACCAACTTTCGATCTCTTGATACCCCTGGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx2_F_LC	CAAACCAACCAACTTTCGATCTCTTGATACCCCTGGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx3_F_LC	CAAACCAACCAACTTTCGATCTCTTGATACCCCTGGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx4_F_LC	CAAACCAACCAACTTTCGATCTCTTGATATTCTGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx5_F_LC	CAAACCAACCAACTTTCGATCTCTTGATATTCTGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx6_F_LC	CAAACCAACCAACTTTCGATCTCTTGATATTCTGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx7_F_LC	CAAACCAACCAACTTTCGATCTCTTGATAGTCTGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_Sx8_F_LC	CAAACCAACCAACTTTCGATCTCTTGATAGTCTGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	1523_aN1_F_LC	CCGGTGTACACATTCGACACCA
	1523_aN1_R_LC	CCGGTGGTGCAGATGTGAAGCA
	1523_Ck1_F_LC	TGCTTACACATTCGACACC
	1523_aN2St_F_LC	CCGGTGTACACATTCGACACCA
	1523_aN2St_R_LC	CCGGTGGTGTAGTGAATGTGAAGCA
	1523_Sx9_F_LC	CAAACCAACCAACTTTCGATCTCTTGATAGTCTGTCTCTAACCGTAAACGAACAATGTTACTTCGAAAGTTTATGATCCAG
	4122_Scag_F_LC	CCAACCTTCGATCTCTTGATAGTCTGTCTCTAACCGTAAACGAACAATGTT
	4122_Saug_F_LC	CCAACCTTCGATCTCTTGATAGTCTGTCTCTAACCGTAAACGAACAATGTT
	4122_Sctg_F_LC	CCAACCTTCGATCTCTTGATCTCTGTCTCTAACCGTAAACGAACAATGTT
	4122_Sstart_F_LC	CCAACCTTCGATCTCTTGATAGTCTGTCTCTAACCGTAAACGAACAATGTT
	4122_S0_F_LC	TAAGCAGGATCCTAATACGACTCACTATAGATTAAGGTTTATACCTTCCCAGGTAACAAACCAACCACTTTCGATCTCTTGTA
	4122_S1_F_LC	TAAGCAGGATCCTAATACGACTCACTATAGATTAAGGTTTATACCTTCCCAGGTAACAAACCAACCACTTTCGATCTCTTGTA

ORF1ab	Nom	Sequence oligonucléotides 5' ==> 3'
	4122_N5_A_F_LC	TGGCCTTTTGCTCACATG
	4122_N5_B_R_LC	GAACATatctacaagagatAgaaggttggttggtttggtaCC
	4122_N5_C_F_LC	cTatctctttagatATGTTCTCTAAACGAACCTTAAAAATCtgtg
	4122_N5_D_R_LC	GTTTACATCTGGCCACCAC
	4122_N6_A_F_LC	TGGCCTTTTGCTCACATG
	4122_N6_B_R_LC	GAACAGaGGtacaagaCctCgaaagttggttggtttggtaCC
	4122_N6_C_F_LC	GaGGtctttagtaCctCTGTCTCTAAACGAACCTTAAAAATCtgtg
	4122_N6_D_R_LC	GTTTACATCTGGCCACCAC
	4122_N7_A_F_LC	TGGCCTTTTGCTCACATG
	4122_N7_B_R_LC	tgctgagtgactaagCGGcagccgagtgacagcca
	4122_N7_C_F_LC	CCGcttagtgactcagcagatAATTAATAACTAATTAAGTCTGCTgtg
	4122_N7_D_R_LC	GTTTACATCTGGCCACCAC
	4122_N8_C_F_LC	CCGcttagtgactcagcagatAATTAaataTAATTAAGTCTGCTgtg
	4122_K0_C_F_LC	GAACCTTAAAAATCtgtgtggctgctact
	4122_K0_B_R_LC	agtgacagccacacaGATTTTAAAGTTC
	4122_L1_A_F_LC	TATGCAGGATCCTAATACGACTCACTATAGATTAAGGTTTATACCTTCCCAGG
	4122_Lwt_D_R_LC	TTACCATTGGAAGTCTCCATCTACCTTTCGGTACACCC
	4122_Lph_R_D_LC	TTACCATTGGAAGTCTCCATCTTCCTTTCGGACACACC
	4122_Mwt_D_R_LC	TTACCATTGGAAGTCTCCAACTTACCTTTCGGTACACCC
	4122_Mph_R_D_LC	TTACCATTGGAAGTCTCCAACTTTCCTTTCGGACACACC
	4122_Pwt_B_R_LC	CCAGGGACAAGGCTCTCCATCTACCTTTCGGTACACCC
	4122_Pph_B_R_LC	CCAGGGACAAGGCTCTCCATCTTCCTTTCGGACACACC
	4122_P42_C_R_LC	ATGGAGAGCCCTTGTCCTCTG

4 Annexe IV

Remerciements techniques

Les figures dans l'introduction ont été créées avec BioRender. Plus précisément, les figures 1, 25 et 26 ont été adaptées à partir des modèles BioRender ("Ribosome Function - Protein Translation," "Mechanism of SARS-CoV-2 Viral Entry," et "Life Cycle of Coronavirus," respectivement).

