



HAL
open science

Le piratage de l'appareil de traduction de la cellule hôte par le virus de la paralysie du criquet (CrPV)

Lauriane Gross

► **To cite this version:**

Lauriane Gross. Le piratage de l'appareil de traduction de la cellule hôte par le virus de la paralysie du criquet (CrPV). Virologie. Université de Strasbourg, 2018. Français. NNT : 2018STRAJ130 . tel-03650061

HAL Id: tel-03650061

<https://theses.hal.science/tel-03650061>

Submitted on 24 Apr 2022

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.



UNIVERSITÉ DE STRASBOURG



ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ
UPR 9002 du CNRS, Architecture et Réactivité de l'ARN (ARN)

THÈSE

présentée par :

Lauriane GROSS

soutenue le : **26 octobre 2018**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline : Sciences de la vie et de la santé

Spécialité : Aspects moléculaires et cellulaires de la biologie

Le piratage de l'appareil de traduction de la cellule hôte par le virus de la paralyse du criquet (CrPV)

THÈSE dirigée par :

Dr. MARTIN Franck

Directeur de recherche, UPR 9002 du CNRS, IBMC

RAPPORTEURS :

Dr. WEIL Dominique

Dr. VAGNER Stéphane

Directeur de recherche, UMR 7622, IBPS

Directeur de recherche, UMR 3348, Institut Curie

EXAMINATEUR :

Pr. DIMITROVA Maria

Professeur, UPR2357 du CNRS, IBMP

REMERCIEMENTS

J'aimerais tout d'abord remercier les Drs. Dominique Weil et Stéphan Vagner d'avoir accepté d'évaluer mon travail de thèse. Je remercie également Pr. Maria Dimitrova d'avoir bien voulu examiner mon travail.

Je remercie Mme Pascale Romby de m'avoir accueilli au sein de son unité de recherche.

J'exprime ma gratitude à Gilbert Eriani pour m'avoir accepté au sein de son équipe. Je le remercie pour son soutien et sa disponibilité, notamment lors de la correction des manuscrits de master et de thèse. Les quatre années passées dans son équipe ont été scientifiquement très enrichissantes.

Mon plus grand merci va à Franck Martin, qui m'a parfaitement préparé pour les épreuves de master, qui a contribué à ma réussite au concours d'obtention des contrats doctoraux et qui m'a encadré jour après jour ces quatre dernières années. Je le remercie pour sa confiance, la liberté de travail qu'il m'a laissé et sa présence dès que j'en ai eu besoin. Il m'a enseigné de nombreuses compétences techniques, la préparation de présentations claires, la rédaction de projets, et la réflexion scientifique... c'est à lui que je dois la majorité de mes acquis.

J'aimerais aussi remercier le reste de l'équipe 333 :

Tout d'abord Christine Allmang, pour sa disponibilité, pour les discussions et réflexions scientifiques, pour la précision de ses explications techniques. Je ne pense pas que j'aurais réussi à faire une qPCR correcte sans elle... Je la remercie aussi pour les nombreux conseils concernant la rédaction de la thèse, les post-doc, les possibilités pour la suite... et surtout pour son écoute et ses conseils plus personnels.

Ensuite Madame Laure... comment dire... une page entière ne suffirait pas à t'exprimer ma gratitude. J'ai appris à me débrouiller dans un labo grâce à toi, à trouver des petits trucs et astuces techniques, à organiser et réussir mes expériences. Mais ton soutien allait bien au-delà d'une aide technique. La piscine, les vendanges, des idées musicales pour les joies à partager ou les stress à évacuer... Tu as fait de ma thèse une belle période de ma vie. Merci.

Je souhaite aussi remercier les libanais. Hassan, une de ces personnes qui a le don d'être toujours agréable et amusante, et Fatima qui est une personne exceptionnelle. J'ai bien ri avec vous deux, et j'ai apprécié les échanges culturels (plus ou moins enflammés).

Pour terminer j'aimerais remercier Aurélie, pour sa compagnie au labo pendant le « miam » et Justine qui m'a aidé à échanger plus facilement sur de sujets scientifiques et à partager mes connaissances.

J'exprime également ma gratitude envers toute l'équipe de la plateforme protéomique : Philippe Hammann, Johana Chicher et Lauriane Kuhn. Merci de tout coeur d'avoir traité mes très nombreux échantillons (Positivons, vous faites désormais parti des quelques rares

personnes qui connaissent le nombre exact de colonnes que contient un classeur excel !!), et merci pour les nombreuses heures que vous m'avez accordé pour analyser ensemble et présenter mes résultats.

Je remercie aussi nos collaborateurs : Jean Luc Imler, Carine Meignin, Evelyne Einhorn, Laurent Daeffler et Estelle Santiago qui m'ont permis d'utiliser les salles de culture cellulaire et de préparation de virus, et qui m'ont fourni le matériel nécessaire à mon étude. Un merci particulier pour Evelyne, Laurent et Estelle qui ont participé à mon travail en m'enseignant toutes les ficelles de la culture cellulaire et les bases de la production de virus.

Je remercie également les « personnes trop cool de la pièce virus »: Vru (merci pour les conseils sur le Kohl indien !), Brice (merci pour les voyages sous le soleil gabonais !) et Nathalie (merci pour ton incroyable gentillesse).

Et bien entendu, je souhaite aussi remercier les copains et amis du 2^{ème}, du 3^{ème} et du 4^{ème}. Antoine, Benji, Floflo, Florian, Raphaël, Alexis (bon on les fait ces vacances Woofing ou pas ?), Lucas (merci pour les moments passés en ta compagnie).

Merci aussi à Carmen et Lili, les correcteurs d'orthographe.

Et les autres ! ceux que j'ai oublié...

Et je terminerai par remercier mes proches, ceux qui me soutiennent depuis quelques mois, depuis 10 ans, depuis 20 ans, depuis 27 ans... Je leur doit tout, ce que j'ai vécu et ce que je suis devenue. Philippe, Anne-chloé, Alex, Emilie, Christophe, Cynthia, Damien, Hélène, Mylène, Stéphanie, Thomas, et surtout Papa, Maman et Lisa.

SOMMAIRE

ABRÉVIATIONS	5
INTRODUCTION	9
I. La synthèse des protéines chez les eucaryotes, différents mécanismes d'initiation de la traduction	9
I.1. Préambule	9
I.2. L'initiation de la traduction canonique ou « coiffe-dépendante »	10
I.2.1. Le recrutement d'un complexe multi-protéique sur la coiffe.....	10
I.2.2. Le recrutement du complexe 43S	10
I.2.3. La formation du complexe 48S sur l'ARNm.....	11
I.2.4. L'assemblage du complexe 80S	11
I.2.5. Le facteur d'initiation eIF3.....	13
I.3. Les mécanismes d'initiation et de régulation non canoniques de la traduction.....	15
I.3.1. L'initiation de la traduction guidée par des structures de l'ARN.....	16
I.3.1.1. L'initiation de la traduction de l'histone H4 par 'tethering' du 43S sur l'AUG.	16
I.3.1.2. L'initiation guidée par les éléments structuraux 3'CITE chez les virus de plante.....	18
I.3.1.3. Le recrutement de la particule 43S par des tiges-boucles qui lient eIF3	20
I.3.1.4. L'initiation de la traduction sur les sites d'entrée interne du ribosome	21
I.3.2. Autres mécanismes d'initiation de la traduction en interne	23
I.3.2.1. L'initiation de la traduction grâce aux modifications des adénosines	23
I.3.2.2. L'initiation de la traduction des ARN circulaires (circARN).....	23
I.3.3. Différents mécanismes de régulation de l'initiation de la traduction.....	24
I.3.3.1. La régulation de la traduction par les facteurs eIF	24
I.3.3.2. La régulation par des éléments structuraux présents en région 5'UTR des ARNm	26
I.3.3.3. Le contrôle de la traduction par les uORF	28
I.3.3.4. Le contrôle de l'initiation de la traduction par le ribosome	29
II. Le piratage de l'appareil de traduction de la cellule hôte par les virus à ARN simple brin grâce aux IRES	30
II.1. La dérégulation de la traduction de l'hôte par les virus à ARN simple brin	30
II.1.1. L'inactivation du complexe eIF4F.....	31
II.1.1.1. Le clivage de eIF4G	31
II.1.1.2. La séquestration de eIF4E	31
II.1.2. La phosphorylation du facteur d'initiation eIF2	32
II.2. Le détournement du ribosome de l'hôte grâce aux IRES viraux.....	33
II.2.1. Les IRES de classe I.....	35
II.2.2. Les IRES de classe II.....	36
II.2.3. Les IRES de classe III.....	39
II.2.3.1. Structure de l'IRES du virus de l'hépatite C.....	39

II.2.3.2. Rôle fonctionnel du domaine I.....	40
II.2.3.3. Formation du complexe de pré-initiation 40S-IRES.....	41
II.2.3.4. Formation du complexe de pré-initiation 48S-IRES.....	42
II.2.3.5. Assemblage du ribosome 80S.....	43
II.2.4. Les IRES de classe IV.....	44
II.2.4.1. Structure adoptée par les IGR des Dicistrovirus.....	44
II.2.4.2. Mécanisme moléculaire de recrutement du ribosome par les IGR.....	46
II.2.5. Les 'IRES Trans-Acting Factors' ou ITAF.....	46
III. L'initiation « IRES-dépendante » de la traduction chez le CrPV.....	48
III.1. Généralités sur les <i>Dicistroviridae</i>	48
III.1.1. Taxonomie et spectre d'hôtes.....	48
III.1.2. Structure du virion des Dicistrovirus.....	49
III.1.3. Organisation génomique et fonction des protéines virales.....	50
III.1.4. Réplication.....	51
III.2. La traduction virale du CrPV, un modèle d'étude de choix.....	52
III.2.1. Les avantages du CrPV.....	52
III.2.2. L'inhibition de la traduction de l'hôte par le CrPV.....	52
III.2.3. L'initiation de la traduction dirigée par les deux IRES du CrPV.....	53
III.2.4. Les points communs entre l'IRES _{5'UTR} du CrPV et l'IRES du HCV.....	54
III.2.5. La régulation différentielle des deux IRES du CrPV.....	55
IV. Objectifs et résumé de mon projet de doctorat.....	56
RÉSULTATS.....	59
I. La 5'UTR du CrPV contient une IRES de classe III – Étude du mode d'action de l'IRES_{5'UTR} de CrPV.....	59
I.1. Résumé de la publication 1.....	59
I.1.1. L'IRES _{5'UTR} est localisé entre les nucléotides 357 et 709 de la région 5'UTR du génome de CrPV.....	59
I.1.2. L'IRES contient un pseudo-nœud nécessaire à son activité traductionnelle.....	60
I.1.3. L'IRES _{5'UTR} dépose le ribosome sur le codon d'initiation AUG ₇₀₉ sans 'scanning'.....	61
I.1.4. Le facteur d'initiation eIF3 est requis pour initier la traduction par l'IRES _{5'UTR}	61
I.1.5. L'IRES _{5'UTR} du CrPV, un IRES de classe III.....	62
I.2. Publication 1.....	65
I.3. Résultats complémentaires non publiés.....	81
I.3.1. Étude de la structure secondaire de la région 5'UTR du CrPV par sondage chimique.....	81
I.3.2. L'extrémité 5'-proximale du génome du CrPV est-elle un bon ARN de référence pour des analyses comparatives par spectrométrie de masse ?.....	82
I.3.3. Identification des partenaires protéiques de l'IRES _{5'UTR} lors de l'initiation de la traduction.....	84
I.3.3.1. Les inhibiteurs de la traduction utilisés dans les analyses.....	84
I.3.3.2. Les partenaires protéiques spécifiques de l'IRES _{5'UTR}	85
I.3.4. Étude du rôle du pseudo-nœud du domaine II dans le recrutement du ribosome.....	88
I.4. Conclusion du chapitre I.....	90

II. Étude de la fonction des domaines II et III de l'IRES_{5'UTR} du CrPV	91
II.1. Objectifs	91
II.2. Résultats	91
II.2.1. Étude du rôle respectif des domaines II et III dans le fonctionnement de l'IRES	91
II.2.2. Identification des partenaires protéiques spécifiques des domaines II et III	93
II.2.3. Étude du rôle de eIF2 dans la traduction guidée par l'IRES _{5'UTR}	94
II.2.4. Étude du rôle du pseudo-nœud dans le recrutement de complexes RNP	95
II.3. Conclusion du chapitre II	98
 DISCUSSION GÉNÉRALE ET PERSPECTIVES	101
 ANNEXES.....	111
 ANNEXE I : MATÉRIEL ET MÉTHODES	111
ANNEXE II : PUBLICATION 2.....	113
ANNEXE III : FIGURE SUPPLÉMENTAIRE	127
 BIBLIOGRAPHIE.....	131

ABRÉVIATIONS

3'CITE	3' Cap-Independent Translation Enhancer
4E-BP	4E-Binding Protein
4E-SE	4E-Sensitive Element
ABPV	Acute Bee Paralysis Virus
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
AMV-RT	Avian Myeloblastosis Virus Reverse-Transcriptase
ARN	Acide ribonucléique
ARNm	ARN messenger
ARNr	ARN ribosomique
ARNsb	ARN simple brin
ARNt	ARN de transfert
ARNt ^{Met} _i	ARNt Methionine initiateur
CAT-1	Cationic Amino acid Transporter 1
CBP	Cap Binding Pocket
circARN	ARN circulaire
CL	Cloverleaf
CMCT	1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate
CrPV	Cricket Paralysis Virus
Cryo-ME	Cryo-Microscopie Electronique
DCV	Drosophila C Virus
DMS	DiMethyl Sulfate
dNTP	désoxyriboNucléotides Tri-Phosphates
eEFs	eukaryotic Elongation Factors
eIF	eukaryotic Initiation Factors
EMCV	Encephalo Myocarditis Virus
eRF	eukaryotic Release Factor
FC	Fold Change
FGF	Fibroblast Growth Factor
FHV	Feline Herpes Virus
FMDV	Foot-and-Mouth Disease Virus
FMRP	Fragile X Mental Retardation Protein
GATA-6	GATA-binding factor 6
GCN2	eIF-2-alpha kinase GCN2
GDP	Guanosine DiPhosphate
GEF	Guanine nucleotide Exchange Factor
GTP	Guanosine TriPhosphate
HCV	Hepatitis C Virus
HEAT-1	<u>H</u> untingtin, e <u>E</u> F3, protein phosphatase 2 <u>A</u> , kinase <u>I</u> OR1
HRI	Heme-Regulated eIF2α kinase
IAPV	Israeli Acute Paralysis Virus

IGR	InterGenic Region IRES
IRES	Internal Ribosome Entry Site
ITAF	IRES <i>Trans</i> -Acting Factor
KBV	Kashmir Bee Virus
L	Loop
La	La autoantigen
m6A	N6-méthyladénosine
mTOR	mammalian Target Of Rapamycin
mTORC1	mammalian Target Of Rapamycin Complex 1
nt	nucléotides
ORF	Open Reading Frame
PABP	Poly(A)-Binding Protein
Patr-1	Protein associated with topo II related-1
PCBP2	Poly(RC) Binding Protein 2
PDGF	Platelet-Derived Growth Factor
PERK	(PKR)-like Endoplasmic Reticulum Kinase
PFU	Plaque Forming Unit
PK	Pseudo-noeud
PKR	Protein Kinase R
PTB	Polypyrimidine tract-Binding Protein
PV	Poliovirus
qPCR	PCR quantitative
RACK1	Receptor for Activated C Kinase 1
RdRp ou RpRd	ARN polymérase ARN dépendante
RLU	Relative Light Units
RNPs	particules ribonuléo-protéiques
RpL	Large subunit Ribosomal Protein
RpS	Small subunit Ribosomal Protein
RT-qPCR	Transcription inverse-qPCR
SHAPE	Selective 2' Hydroxyl acylation Analysed by Primer Extension
SL	Stem-Loop
SQD	Squid
SYBR	Syber Green
TCID50	Tissue Culture Infective Dose
UTR	Untranslated Region
VEGF	Vascular Endothelial Growth Factor
VP0, 1, 2, 3, 4	Viral Protein 0, 1, 2, 3, 4
VPg	genome-linked Viral Protein

INTRODUCTION

INTRODUCTION

I. La synthèse des protéines chez les eucaryotes, différents mécanismes d'initiation de la traduction

I.1. Préambule

La synthèse des protéines est un processus biologique majeur contribuant à la prolifération cellulaire. Sa régulation est importante, car elle permet au métabolisme cellulaire de s'adapter en réponse à différents stimuli. Différents mécanismes de traduction existent, ils permettent à la cellule de répondre efficacement à des changements environnementaux ou à des stress variés. La traduction des ARN messagers chez les eucaryotes est un processus complexe qui se divise en 4 étapes : l'initiation, l'élongation, la terminaison et le recyclage du ribosome. L'initiation est l'étape qui permet d'assembler le ribosome complet 80S sur le codon d'initiation. Les nucléotides du codon s'apparient à ceux de l'anticodon de l'ARNt initiateur (Met-ARNt^{Met} initiateur ou ARNt^{Met}) dans le site P du ribosome (P pour peptidyl-ARNt). Un certain nombre de facteurs d'initiation, nommés eIF (eukaryotic Initiation Factor), sont requis pour cette étape d'assemblage. Lorsque l'initiation se fait de manière conventionnelle les eIF nécessaires sont au nombre de dix-neuf. L'initiation se fait par l'assemblage séquentiel de plusieurs complexes de pré-initiation intermédiaires pour finalement aboutir à la formation d'un ribosome 80S sur le codon d'initiation (Sonenberg & Hinnebusch, 2009; Trachsel et al., 1977). En raison du nombre important de facteurs qui entrent en jeu et de la complexité de son déroulement, l'initiation est l'étape limitante de la synthèse des protéines. Une dérégulation de cette étape peut avoir des conséquences dramatiques sur le métabolisme cellulaire et conduire notamment à la formation de cancers ou au développement de maladies neuro-dégénératives (Bhat et al., 2015; Robichaud & Sonenberg, 2017). Après l'initiation vient l'étape d'élongation qui permet la synthèse de la chaîne polypeptidique. Les ARNt aminoacylés entrent dans le site A du ribosome (A pour Aminoacyl-ARNt), puis, lors d'une étape appelée accommodation, l'ARNt se positionne correctement dans le ribosome de manière à permettre l'étape suivante. L'acide-aminé porté par l'aminoacyl-ARNt entrant est alors attaché de manière covalente au dernier acide aminé de la chaîne peptidique présente dans le site P du ribosome. L'élongation requiert 3 facteurs d'élongation : eEF1A, qui apporte l'aminoacyl-ARNt suivant au ribosome, le facteur d'échange eEF1B, qui recycle le GTP fixé à eEF1 et le facteur eEF2, qui aide à la

translocation du ribosome (Dever & Green, 2012). Lorsque le ribosome atteint la fin de la région codante, c'est l'étape de terminaison. Le codon de terminaison ou codon stop (UAA, UGA, ou UAG) entre dans le site A du ribosome. Ce dernier est alors reconnu par deux facteurs de terminaison appelés 'eukaryotic Release Factor' eRF1 et eRF3. Ces facteurs vont provoquer la dissociation du ribosome et le relargage de la chaîne peptidique (Blanchet et al., 2015; Hatin, et al., 2009). Le ribosome sera dissocié en deux sous-unités, 40S et 60S, puis recyclé pour une nouvelle traduction (Rodnina & Wintermeyer, 2009). Durant ma thèse je me suis focalisée sur l'étape d'initiation de la traduction, et plus précisément sur un mécanisme d'initiation non conventionnel.

1.2. L'initiation de la traduction canonique ou « coiffe-dépendante »

1.2.1. Le recrutement d'un complexe multi-protéique sur la coiffe

Le mécanisme d'initiation de la traduction conventionnel est dépendant de la présence d'une coiffe m⁷Gppp ancrée à l'extrémité 5' de l'ARNm. L'initiation débute avec l'assemblage d'un complexe multi-protéique sur cette coiffe m⁷G. Ce complexe est constitué des facteurs d'initiation eIF4B et eIF4F (Hashem & Frank, 2018). Le facteur eIF4F est lui-même composé d'un assemblage de trois facteurs : le facteur de liaison à la coiffe eIF4E, eIF4A qui est une hélicase à motif 'DEAD' et eIF4G une grande protéine plateforme (Altmann et al., 1985; Sonenberg et al., 1979). L'hélicase eIF4A peut être retrouvée sous deux formes : soit associée à d'autres facteurs au sein du complexe eIF4F, soit isolée sous forme libre dans le cytoplasme. Son rôle est de déplier les structures secondaires de l'extrémité 5' de l'ARNm, avec l'aide d'un facteur auxiliaire eIF4B, de manière à permettre l'attachement du ribosome (Rozen et al., 1990). Le facteur plateforme eIF4G interagit avec les facteurs eIF4E et eIF4A d'une part et les 'polyA binding protein' (PABP) fixés sur la queue poly(A) en 3' de l'ARNm d'autre part (Etchison et al., 1982; Hentze et al., 1997; Lamphear et al., 1995). Ces interactions multiples entraînent le rapprochement des extrémités 5' et 3' de l'ARNm, permettant ainsi sa circularisation (**Figure 1** : étape appelée 'activation de l'ARNm'). L'ARN messenger ainsi circularisé peut recruter le complexe de pré-initiation 43S. Il a été décrit que la configuration de l'ARNm en « boucle fermée » stimule la traduction en augmentant le recyclage et l'assemblage du complexe 48S sur l'ARNm (Jackson et al., 2010; Sonenberg & Hinnebusch, 2009; Svitkin et al., 2009; Topisirovic & Sonenberg, 2011).

1.2.2. Le recrutement du complexe 43S

D'une part, la petite sous-unité du ribosome (40S) issue de l'étape de recyclage s'associe aux facteurs d'initiation eIF3, eIF1, eIF1A et eIF5. D'autre part, le complexe ternaire,

composé de eIF2-GTP-Met-ARNt^{Met}_i, s'associe également avec le 40S. L'ensemble forme le complexe de pré-initiation 43S (Majumdar, et al., 2003; Trachsel & Staehelin, 1978, 1979). Cette particule 43S est ensuite recrutée sur l'ARNm grâce à l'interaction entre eIF3 au sein du 43S, et eIF4G au sein de eIF4F, déjà présent sur l'extrémité 5' de l'ARNm (**Figure 1** : cette étape est appelée « ancrage sur l'ARNm » ou 'attachement to mRNA'). Le facteur d'initiation eIF3 joue un rôle majeur à chaque étape de l'initiation de la traduction. Un paragraphe est dédié à eIF3 ci-après.

1.2.3. La formation du complexe 48S sur l'ARNm

Suite à son recrutement sur l'extrémité 5' du messager, le complexe 43S va glisser le long l'ARNm du 5' vers le 3' jusqu'au codon initiateur AUG selon un mécanisme appelé 'scanning' (Hinnebusch, 2017; Kozak, 1978) (**Figure 1** : '5' to 3' scanning'). Le 'scanning' est assuré par les facteurs d'initiation eIF1/eIF1A et par les hélicases eIF4A/4B, DHX29 et DDX3 qui aident à déplier les structures de la région 5'UTR de l'ARNm (Hashem et al., 2013; Hinnebusch, 2014; Guenther et al., 2018, Martin & Eriani 2018). La reconnaissance du codon d'initiation AUG est validée par le facteur eIF1 (Pestova & Kolupaeva, 2002). Pour être reconnu efficacement, le codon AUG doit se trouver dans un contexte de séquence spécifique, avec une purine en position -3 et un G en position +4 (la numérotation fait référence à la position de ce nucléotide par rapport au A de l'AUG considéré comme le +1). Cette séquence optimale est également appelée séquence de Kozak (Kozak, 1986). Une fois le complexe correctement positionné sur le codon d'initiation dans le site P du ribosome, eIF1 est déplacé pour favoriser l'hydrolyse du GTP lié à eIF2 et l'interaction du Met-ARNt^{Met}_i avec l'AUG (Lorsch, 1999). A ce stade, il est appelé complexe 48S (**Figure 1** : 'initiation codon recognition, hydrolysis of eIF2 bound-GTP').

1.2.4. L'assemblage du complexe 80S

La dernière étape de l'initiation voit la dissociation des facteurs eIF1, eIF1A, eIF3 et eIF2-GDP du complexe de pré-initiation 48S pour permettre l'association de la sous-unité 60S (Jackson et al., 2010) (**Figure 1** : 'subunit joining and factor displacement'). Un autre facteur d'initiation entre en jeu, le facteur eIF5B-GTP. Cette protéine autorise l'assemblage du 60S grâce à son activité GTPasique. L'hydrolyse de son GTP permet de dissocier eIF1A pour que le 80S s'assemble finalement (Lee et al., 2002) (**Figure 1** : 'hydrolysis of eIF5B bound-GTP'). Le ribosome 80S est alors prêt à entrer en phase d'élongation et ce qu'on appelle processus d'initiation canonique est terminé.

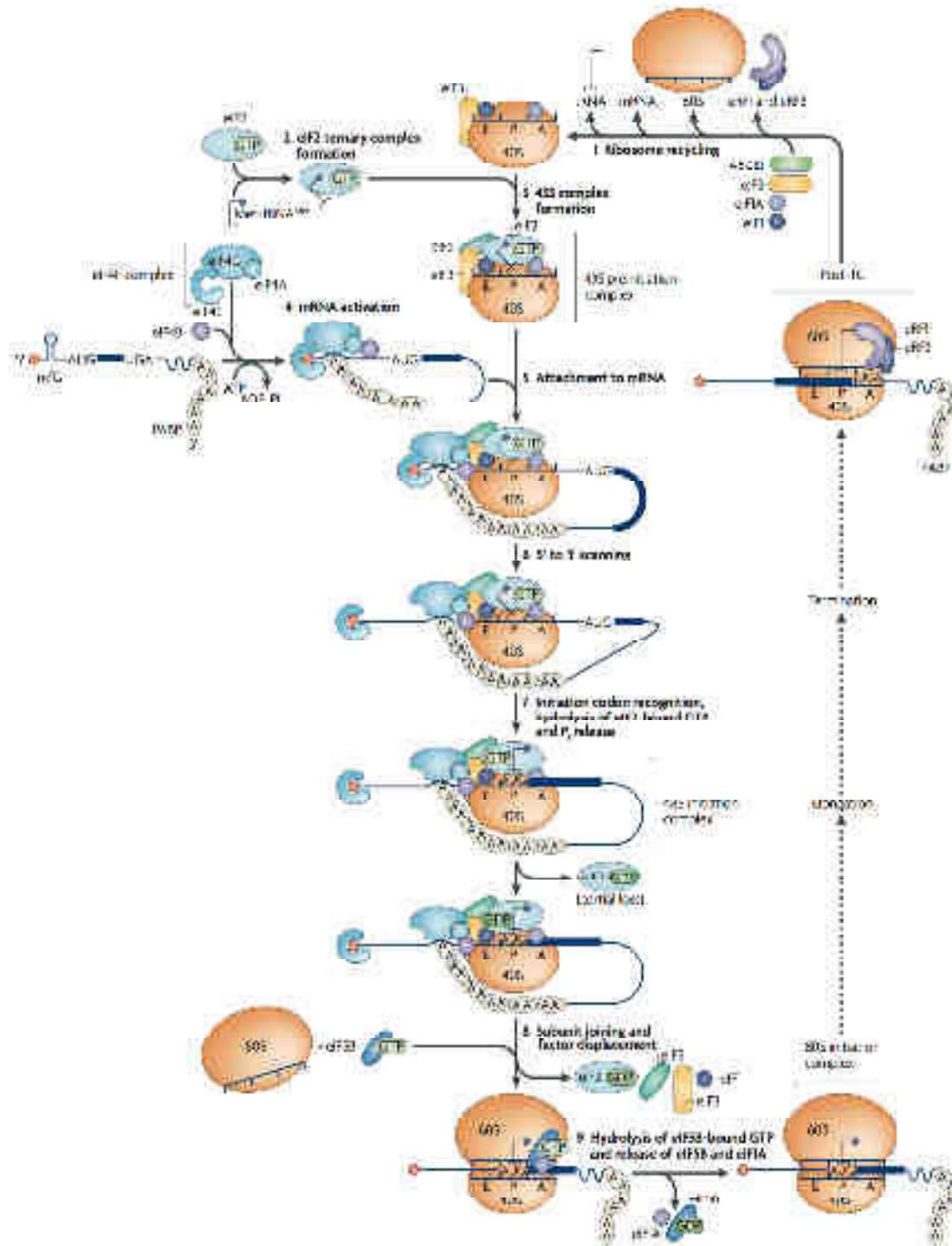


Figure 1 : Le mécanisme d'initiation de la traduction canonique « coiffe-dépendant » chez les eucaryotes. (figure issue de Jackson *et al.*, 2010). Le mécanisme d'initiation canonique dépendant de la coiffe peut être divisé en 8 étapes (étapes 2-9). Il commence après le recyclage (1) des complexes de traduction. Le complexe ternaire eIF2-GTP-Met-ARNt^{Met} (2) est assemblé dans le complexe de pré-initiation 43S (3). L'ARNm est au même moment activé grâce à la liaison du complexe eIF4F à la coiffe, l'interaction de ce dernier avec les PABP en 3' induit une circularisation de l'ARNm (4). Le complexe 43S est alors recruté au niveau de l'extrémité 5' de l'ARNm (5) et va procéder au processus de 'scanning' (6). Ensuite, l'interaction de l'ARNt^{Met} avec l'AUG initiateur induit l'hydrolyse du GTP associé à eIF2 et la formation du complexe 48S (7). L'assemblage de la sous-unité 60S sur le complexe 48S provoque le relargage de la majorité des facteurs d'initiation (8). Puis, le ribosome 80S est prêt à entrer en phase d'élongation.

1.2.5. Le facteur d'initiation eIF3

Le facteur d'initiation eIF3 est le plus gros et le plus complexe des facteurs nécessaires à l'initiation de la traduction (voir la revue Valášek et al., 2017). Chez les mammifères, eIF3 est un complexe multimérique de 800 kDa, composé de 13 sous-unités, nommées eIF3a à eIF3m. Le noyau structural de ce facteur d'initiation est composé de 8 sous-unités (a, c, e, k, l, m, f et h), il est nommé 'central core' ou noyau octamérique de eIF3 (**Figure 2**). Ce noyau octamérique est constitué de six sous-unités protéiques contenant un domaine PCI pour '26S Proteasome, COP9 signalosome, Initiation factor eIF3' et de deux sous-unités contenant un domaine 'Mpr1/Pad1 N-terminal' (MPN). Ce sont les sous-unités a, c, e, k, l, et m qui, grâce à leur domaines PCI, permettent l'assemblage des six sous-unités entre elles. Les sous-unités f et h, quant à elles, possèdent chacune un domaine MPN qui permet la fixation d'autres protéines ou des autres sous-unités sur le noyau octamérique. Les cinq sous-unités restantes (b, d, g, i et j) sont à la périphérie du complexe protéique et sont plus flexibles que les sous-unités du noyau octamérique (**Figure 2**). Les sous-unités b, i, et g forment un module séparé rattaché au noyau octamérique *via* la partie C-terminale de la sous unité eIF3a. Ce complexe est appelé 'Yeast-Like Core' (YLC) car, chez la levure, eIF3a, b, i et g (additionné de eIF3c) constituent le noyau fonctionnel de eIF3 (Hinnebusch, 2006; Valášek, 2012). Les sous-unités b et g possèdent des motifs de liaison à l'ARNm appelés 'RNA Recognition Motif' (RRM). Les sous-unités b et i contiennent des motifs en hélice- β formés par la répétition des acides aminés tryptophane et acide aspartique (motif WD40). Les sous-unités a et c du noyau de eIF3, ainsi que les sous-unités périphériques b et g, sont capables d'interagir directement avec l'ARN messenger. La présence dans le complexe eIF3 de la sous-unité eIF3j est controversée. En effet ce facteur se détache très facilement du complexe eIF3. Chez la levure, eIF3j n'est pas une protéine essentielle à la traduction. Elle s'associe transitoirement au noyau de eIF3 et stimule son activité. eIF3j renforce les interactions entre eIF3 et les autres facteurs eIF. eIF3j permet également l'association du facteur eIF3 à la sous-unité ribosomique 40S lors de la formation de la particule 43S (Block et al., 1998; Valášek et al., 1999). Dans la particule 43S, eIF3j occupe le même emplacement que l'ARNm. Par conséquent, la fixation du 43S sur l'ARNm nécessite le relargage préalable de eIF3j afin de permettre son ancrage sur l'ARNm. Par ailleurs, eIF3j possède d'autres fonctions, notamment dans la biogenèse du 40S. En résumé, certaines publications qualifient eIF3j de sous-unité périphérique de eIF3 alors que d'autres travaux suggèrent qu'il serait plutôt un facteur de traduction qui s'associe transitoirement à eIF3 lors de l'initiation de la traduction (Beznosková et al., 2013; ElAntak et al., 2010; Majzoub et al., 2014).

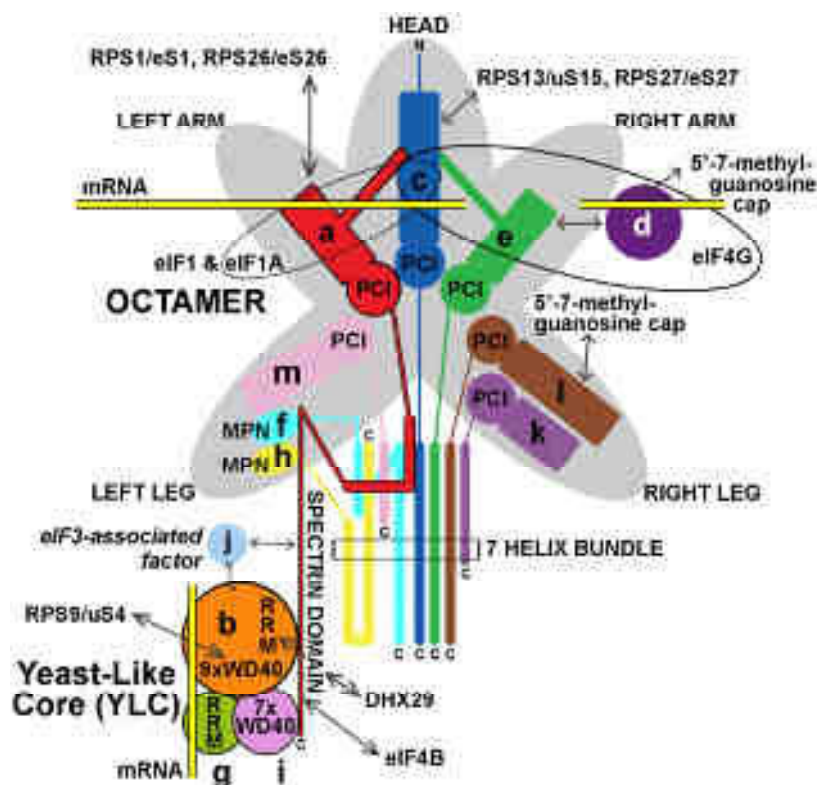


Figure 2 : Représentation schématique du facteur d'initiation eIF3 et des interactions de ses 13 sous-unités. (Figure issue de Valasek *et al.*, 2017). Les sous-unités a, c, e, k, l, m, f et h du facteur d'initiation eIF3 constituent le noyau octamérique (entouré en gris) du facteur indispensable à sa fonction de facteur d'initiation de la traduction. Les sous-unités b, d, g, i et j, quant à elles, sont localisées en périphérie du noyau. Elles participent à la liaison du facteur eIF3 à l'ARN messager grâce à leurs motifs RRM et WD40. Les sous-unités eIF3a, b, i et g sont rattachées au noyau octamérique eucaryote *via* l'extrémité C-terminale de eIF3a. Elles forment le complexe appelé 'Yeast-Like Core' (YLC) qui regroupe quatre des cinq sous-unités constituant le core fonctionnel minimal de eIF3 chez la levure. La sous-unité eIF3j interagit avec eIF3a et b (représenté par des flèches), elle serait associée de manière transitoire à eIF3 pour stimuler son activité. Le facteur eIF3 interagit avec les autres facteurs d'initiation et avec les protéines du ribosome, ces interactions sont représentées par des flèches.

Le facteur d'initiation eIF3 est le seul facteur dont la présence est requise tout au long de l'initiation de la traduction. Dans un premier temps, il empêche l'association des deux sous-unités séparées du ribosome (Hashem *et al.*, 2013). Il assure ensuite le recrutement et la stabilisation du complexe ternaire eIF2-GTP-Met-ARNi^{Met} sur le ribosome (Des Georges *et al.*, 2015). De plus, eIF3 est indispensable lors de l'attachement du complexe de pré-initiation 43S à l'ARNm grâce à son interaction avec eIF4G (Benne & Hershey, 1978; Trachsel & Staehelin, 1979). Enfin, eIF3 assure la fidélité du processus de 'scanning' en garantissant la reconnaissance du codon d'initiation AUG (Hinnebusch, 2006). Les données structurales récentes ont permis de proposer un modèle structural de eIF3 interagissant avec la sous-unité 40S du ribosome eucaryote (**Figure 3**). De manière intéressante, le noyau

octamérique de eIF3 et le complexe YLC sont disposés sur deux faces opposées du ribosome 40S, de telle sorte à l'envelopper presque totalement. Sa localisation aussi étendue sur la sous-unité 40S du ribosome semble expliquer ses nombreuses fonctions lors de l'initiation. Ainsi, l'ensemble de ses caractéristiques fait de eIF3 le facteur central de l'initiation de la traduction ; il est un acteur majeur dans quasiment tous les modes d'initiation de la traduction utilisés par les cellules eucaryotes, canoniques et non-canoniques.

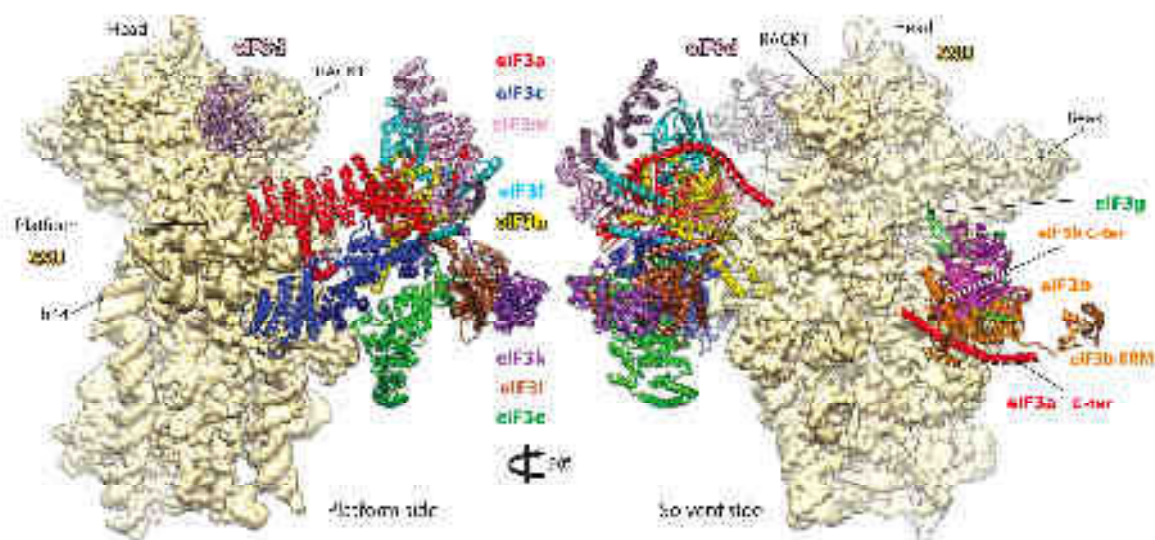


Figure 3 : Modèle structural obtenu par Cryo-Microscopie Électronique de l'organisation des sous-unités du facteur d'initiation eIF3 autour de la sous-unité 40S du ribosome eucaryote. (Figure issue de Des Georges *et al.*, 2015). Les treize différentes sous-unités de eIF3 (représentées avec les mêmes couleurs que dans la **figure 2**) sont localisées de part et d'autre de la petite sous-unité 40S du ribosome (en beige).

I.3. Les mécanismes d'initiation et de régulation non canoniques de la traduction

Le mécanisme d'initiation de la traduction décrit ci-dessus est le mécanisme d'initiation conventionnel des ARN messagers cellulaires coiffés et polyadénylés. Ce mécanisme nécessite la présence de la coiffe en 5' de l'ARNm, c'est pourquoi il est qualifié de « coiffe-dépendant ». Par ailleurs, il nécessite tous les facteurs d'initiation et la localisation du codon initiateur après une étape de 'scanning'. Cependant, la cellule contient différentes populations d'ARNm qui sont traduits par d'autres mécanismes d'initiation de la traduction. Ces mécanismes sont dits « non conventionnels » ou « non canoniques ». Ils permettent généralement la traduction d'ARNm spécifiques dans des conditions particulières. Par exemple, il existe des mécanismes de traduction qui ne nécessitent pas le set complet des facteurs eIF. Ces ARNm spécifiques peuvent également contenir des éléments structuraux

jouant un rôle prépondérant dans le fonctionnement de ces mécanismes d'initiation alternatifs, notamment lors du recrutement du ribosome. De telles structures sont aussi retrouvées dans certains génomes viraux. Ces structures sont plus ou moins sophistiquées et elles sont retrouvées dans toutes les régions de l'ARN (UTR et phase codante). Les modes de recrutement du ribosome utilisés sont également variables. Ces éléments de structure ribonucléique permettent en général d'initier la traduction de l'ARN plus efficacement que le mécanisme canonique. Enfin, ces structures sont généralement utilisées pour la régulation fine de l'expression des ARNm, par exemple à un stade particulier du cycle cellulaire, ou encore en réponse à divers stimuli (stress cellulaire, infection virale, etc.)

1.3.1. L'initiation de la traduction guidée par des structures de l'ARN

1.3.1.1. L'initiation de la traduction de l'histone H4 par 'tethering' du 43S sur l'AUG

Le mécanisme d'initiation de la traduction par 'ribosome tethering' a été décrit dans notre laboratoire pour l'ARNm codant l'histone H4 (Martin et al., 2011) (**Figure 4**). L'ARNm de l'histone H4 de souris (ARNm H4) présente des caractéristiques atypiques. Par exemple, la région 5'UTR ne fait que 9 nucléotides (Meier, Böhni, & Schümperli, 1989). Du côté 3', l'ARNm n'est pas polyadénylé, l'extrémité 3'UTR se termine par une structure en tige-boucle (Jaeger et al., 2005, 2006; Martin et al., 1997). Notre laboratoire a démontré que l'ARNm H4 contient dans la phase codante deux éléments structuraux impliqués dans le processus d'initiation de la traduction. Le premier est une jonction en triple-hélices, localisée 19 nucléotides en aval du codon d'initiation. Dans un premier temps elle séquestre la coiffe en 5' de l'ARNm dans une poche appelée 'Cap Binding Pocket' ou CBP. La coiffe étant alors inaccessible, le facteur eIF4E ne peut pas s'y fixer. En revanche, eIF4E reconnaît un deuxième élément structural au milieu de la phase codante, appelé 'eIF4E Sensitive Element' ou 4E-SE. Cette structure, repliée en double tige-boucle est similaire à l'élément 4E-SE retrouvé dans la région 3'UTR de l'ARNm de la cycline D1 (Culjkovic et al., 2007, 2006). eIF4E interagit directement avec le 4E-SE grâce à son extrémité N-terminale. La poche de liaison à la coiffe de eIF4E utilisée habituellement n'est pas requise dans cette interaction. Dès que le facteur eIF4F (grâce à eIF4E) est recruté sur le 4E-SE, l'hélicase eIF4A déplie les structures autour du codon initiateur AUG. La particule 43S est ensuite directement déposée et positionnée sur le codon AUG, sans étape de 'scanning'. Le positionnement du ribosome se fait par encombrement stérique grâce à la structure en jonction en triple-hélices. Cette structure a donc une fonction double : séquestrer la coiffe dans un premier temps, puis guider le dépôt du ribosome dans un deuxième temps. L'ajustement fin du positionnement du codon AUG dans le site P du ribosome se fait grâce à

une interaction entre un triplet AGG localisé en amont de la jonction triple-hélices et l'hélice h16 de l'ARN ribosomique 18S (Martin et al., 2016) (**Figure 5**). Le complexe 80S s'assemble ensuite avant de s'engager dans l'étape d'élongation. Par conséquent, ce mécanisme d'initiation non conventionnel permet de recruter le complexe de pré-initiation 43S directement sur le codon d'initiation AUG grâce à des structures présentes dans la région codante de l'ARNm. Ce mécanisme de 'ribosome tethering' est plus rapide et plus efficace que la traduction canonique, notamment parce qu'il ne nécessite pas d'étape de 'scanning'. Il est utilisé pendant la phase S du cycle cellulaire. Ce mécanisme permet de produire de façon massive les histones qui seront ensuite utilisées pour former les nucléosomes.

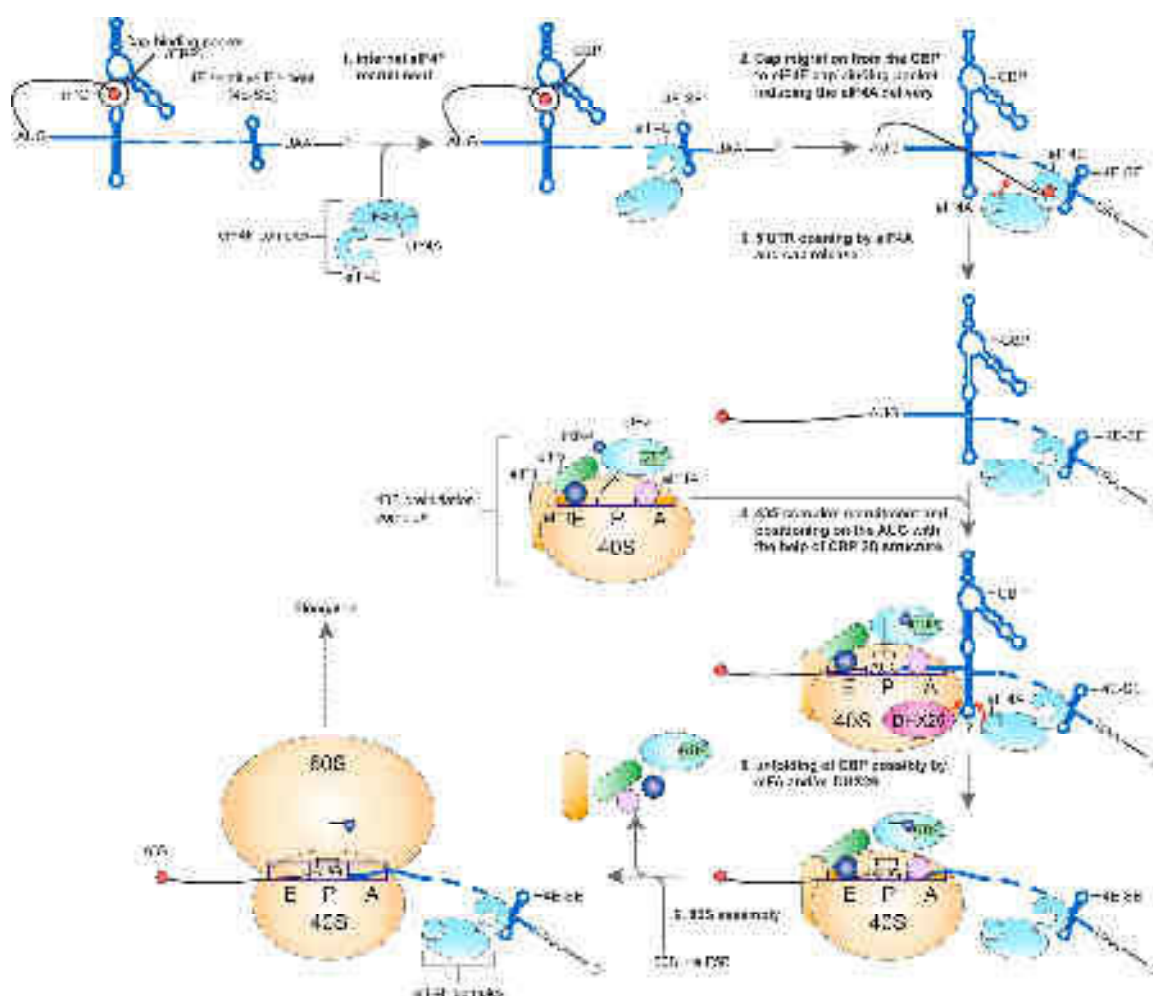


Figure 4 : L'initiation de la traduction « coiffe-dépendante » par 'tethering' du ribosome chez l'ARNm de l'histone H4. La région codante de l'ARNm d'histone H4 contient deux éléments structuraux impliqués dans le processus d'initiation de la traduction : une jonction en triple-hélices contenant une poche de séquestration de la coiffe (CBP) et une double tige-boucle (4E-SE) qui recrute le facteur d'initiation eIF4E de manière interne. La fixation du complexe eIF4E au 4E-SE (1) induit des changements structuraux de la jonction en triple-hélices qui libère la coiffe. Celle-ci vient alors se loger dans la poche de liaison spécifique de la coiffe du facteur eIF4E (2). Le complexe de pré-initiation est ensuite directement déposé sur le codon initiateur AUG du messager (3-4) sans étape de 'scanning' et l'initiation de la traduction se poursuit de manière à assembler le complexe 80S (5-6).

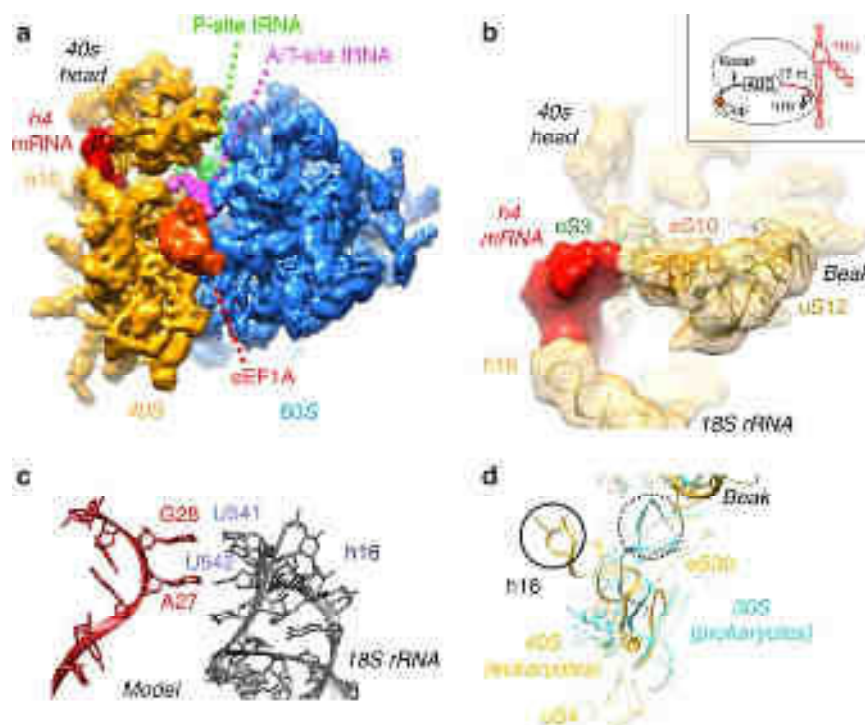


Figure 5 : Structure du complexe d'initiation de la traduction avec l'ARNm H4 et le ribosome 80S. (Figure issue de Martin *et al.*, 2016). L'ajustement du codon initiateur AUG de l'ARN messager de l'histone H4 (ARNm H4) se fait grâce à une interaction entre un triplet AGG localisé en amont de la jonction en triple-hélices de l'ARNm H4 et l'hélice h16 de l'ARN ribosomique 18S. (a) Modèle structural obtenu par cryo-Microscopie Électronique (cryo-ME) du complexe ARNm H4-80S bloqué à l'étape de pré-translocation. L'ARNm H4 est représenté en rouge, eEF1A en orange, l'ARNt entrant dans le site A du ribosome encore associé à eEF1A (A/T-site tRNA) est en magenta, l'ARNt présent dans le site P du ribosome est représenté en vert, la sous-unité ribosomique 60S en bleu et la sous-unité ribosomique 40S en jaune. (b) Une vue rapprochée de la zone de contact entre l'ARNm H4 et le ribosome met en évidence que l'ARNm H4 est en contact avec la boucle apicale de l'hélice h16 (ARNr 18S) et les protéines eS3 et eS10 du ribosome. (c) Modèle des interactions entre l'ARNm H4 (en rouge) et la boucle apicale de l'hélice h16 de l'ARNr 18S (en gris). (d) La superposition des modèles de structure du ribosome eucaryote (en jaune) et du ribosome procaryote (en bleu) met en évidence une différence structurale de l'hélice h16. Chez les eucaryotes, l'orientation de la boucle apicale crée un site de liaison à l'ARNm (cercle plein). Chez les procaryotes l'ARNr 16S n'est pas localisé au même endroit (cercle pointillé), et h16 n'est pas orientée de telle sorte à promouvoir une interaction avec l'ARNm.

1.3.1.2. L'initiation guidée par les éléments structuraux 3'CITE chez les virus de plante

Chez les plantes, moins de 20% des virus à ARN de polarité positive sont coiffés à leur extrémité 5' et polyadénylés à leur extrémité 3'. De ce fait, la grande majorité de ces virus fait appel à des mécanismes de traduction non canoniques pour synthétiser les protéines virales (Lacerda *et al.*, 2017; Simon and Miller, 2011; Miras *et al.*, 2017; Nicholson & White, 2011). Une des stratégies employées par les *Tombusviridae*, *Luteovirus*, et *Umbravirus*, implique un élément structural localisé dans la région 3'UTR du génome viral. Ce type de structures est

appelé '3' Cap-Independent Translation Enhancer' (3'CITE). Les 3'CITE permettent d'assurer une traduction efficace des protéines virales en établissant des interactions à longue distance entre les régions 5'UTR et 3'UTR de l'ARN viral. Ces éléments peuvent prendre différentes structures dites en « I », en « Y », en « T », ou en structure de type pseudo-nœud (Nicholson & White, 2011; Wang et al., 2009). Ils sont capables de recruter soit des facteurs d'initiation, soit la sous-unité 60S du ribosome. L'ARN viral adopte alors une conformation circulaire, ce qui permet l'entrée du complexe 43S sur la région 5'UTR puis le 'scanning' jusqu'au codon d'initiation. La plupart des 3'CITE fonctionnent en recrutant les facteurs d'initiation eIF4E et/ou eIF4G en établissant des interactions ARN-ARN de type 'kissing loop' entre l'élément CITE en 3' et une tige-boucle présente dans la région 5'UTR du génome (Nicholson et al., 2010) (**Figure 6**). D'autres 3'CITE, tel que celui du 'Turnip Crinkle Virus' (TCV), recrutent directement la sous-unité ribosomique 60S au niveau d'une séquence riche en pyrimidines dans la région 5'UTR. Le 60S va ensuite recruter et associer la sous-unité 40S pour terminer l'initiation (McCormack et al., 2008; Simon, 2015; Stupina et al., 2008). Ainsi, les 3'CITE permettent de pirater la machinerie traductionnelle de la cellule hôte pour traduire les ARN viraux au détriment des ARNm cellulaires qui eux ne seront plus traduits.

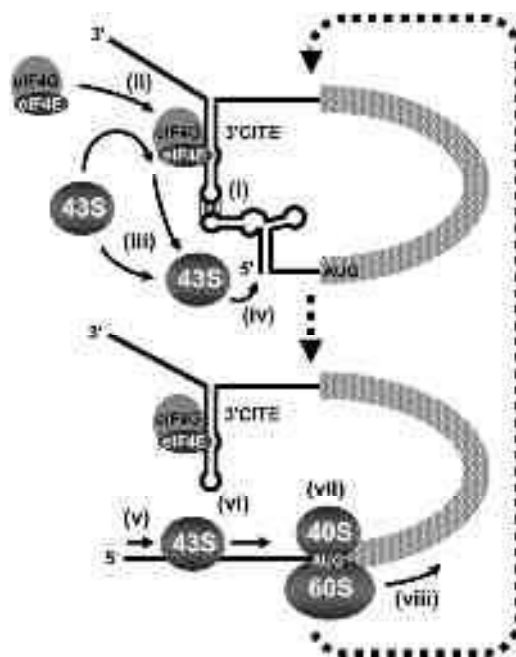


Figure 6 : Modèle du mécanisme d'initiation de la traduction guidé par le 3'CITE des *Tombusvirus*. (Figure issue de Nicholson et al., 2010). Dans ce modèle, l'élément 3'CITE du virus adopte une structure dite en « I ». Des modèles d'initiation similaires ont été proposés pour des virus de plante qui possèdent des éléments 3'CITE en « T », en « Y » ou formant un pseudo-nœud. Dans le modèle ci-dessus, le 3'CITE interagit avec la région 5'UTR via un appariement ARN-ARN de type 'kissing-loop' (i). Dans un premier temps, le complexe eIF4F se fixe au 3'CITE (ii), permettant le recrutement de la particule 43S (iii) sur l'extrémité 5'UTR de l'ARN (iv). La particule 43S effectue ensuite une étape de 'scanning' (v) qui détruit l'interaction ARN-ARN entre

les régions 5'UTR et 3'UTR (vi). Lorsque le ribosome 43S a atteint le codon initiateur, la sous-unité 60S du ribosome s'assemble pour former le complexe d'initiation 80S (vii). L'étape d'élongation de la traduction peut alors commencer (viii). Lorsque l'AUG est à nouveau accessible, un nouveau ribosome peut s'y assembler selon le même mécanisme (flèche pointillée).

1.3.1.3. Le recrutement de la particule 43S par des tiges-boucles qui lient eIF3

Un autre mode d'initiation de la traduction non conventionnel et « coiffe-indépendant » est utilisé pour l'expression des gènes régulateurs de la prolifération cellulaire *c-JUN* et *BTG1*. Les messagers correspondants sont coiffés de manière canonique à leur extrémité 5'. Cependant, ils possèdent également une structure secondaire repliée en une très longue tige-boucle dans leur région 5'UTR. C'est cette structure qui permet le recrutement du ribosome (Blau et al., 2012). Dans ce cas, l'initiation de la traduction ne débute pas par le recrutement du facteur eIF4F sur la coiffe. C'est en fait une interaction directe entre le facteur d'initiation eIF3 et la tige-boucle de la région 5'UTR qui démarre l'assemblage du complexe d'initiation (**Figure 7**). Le recrutement du complexe de pré-initiation 43S se fait ensuite par l'intermédiaire du facteur eIF3 ancré sur la tige boucle. Autre particularité, la sous-unité eIF3d est capable de se lier à la coiffe de ces ARNm. Cette interaction favoriserait également la traduction de ces ARN messagers (Lee et al., 2016). Les bases moléculaires de ce mode de recrutement du ribosome sur ces ARNm sont encore méconnues et controversées. En revanche, le repliement en tige boucle de la région 5'UTR de l'ARNm et son interaction avec eIF3 sont bien caractérisés. Ils ont un rôle crucial dans la régulation traductionnelle, qui peut être une activation ou une répression, de ces messagers cellulaires (Lee et al., 2015).

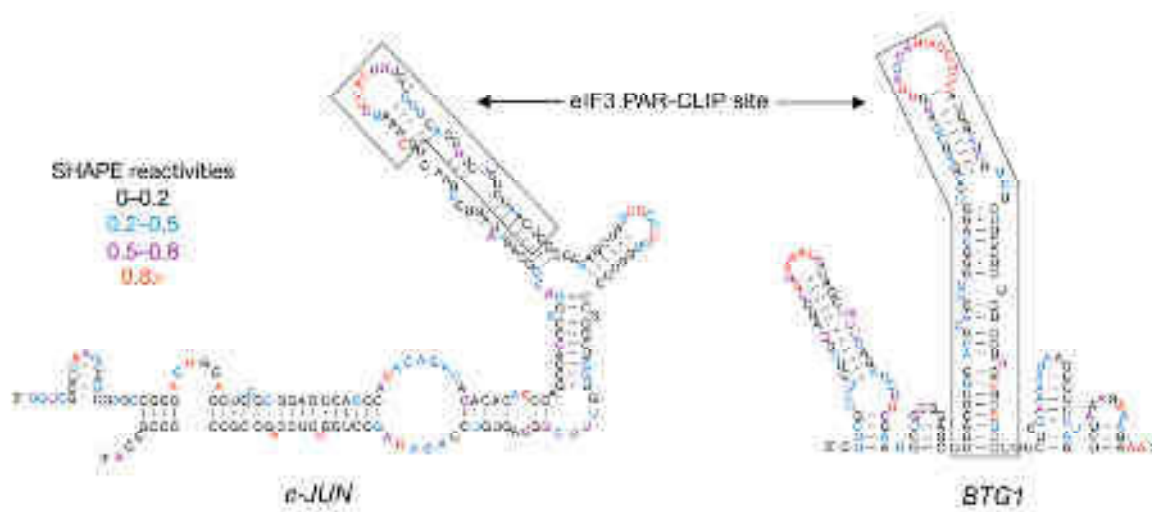


Figure 7 : Modèle de structure secondaire des éléments structuraux retrouvés dans les ARNm *c-JUN* et *BTG1* qui interagissent avec eIF3. (Adaptée de Lee et al., 2015). Récemment, des expériences utilisant la méthode de 'photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation' ou PAR-CLIP,

réalisées sur le transcriptome humain, ont permis de mettre à jour les sites d'interaction entre eIF3 et une sous-classe d'ARNm. Deux exemples sont présentés dans la figure : les ARNm *c-JUN* et *BTG1*, qui sont impliqués dans la prolifération cellulaire. Les sites d'interaction avec eIF3 (site encadré en noir) sont repliés en longues tige-boucles dont la structure a été déterminée par la technique du SHAPE (voir *chapitre 1*). La fixation du facteur eIF3 à la tige-boucle permet le recrutement de la particule 43S pour initier la traduction.

1.3.1.4. L'initiation de la traduction sur les sites d'entrée interne du ribosome (IRES)

Les sites d'entrée internes du ribosome chez les virus :

Les sites d'entrée interne du ribosome ('Internal Ribosome Entry Site' ou IRES) ont été découverts en 1988 dans les génomes de deux virus à ARN de polarité positive : les Poliovirus (PV) et 'Encephalo Myocarditis Virus' (EMCV) (Jang et al., 1988; Pelletier & Sonenberg, 1988; Trono et al., 1988). D'autres IRES ont par la suite été découverts dans d'autres virus à ARN. Les IRES viraux sont des structures ribonucléotidiques complexes présentes dans les régions 5'UTR ou intergéniques des génomes viraux, en amont d'une phase codante ('Open Reading Frame' ou ORF). L'initiation de la traduction guidée par les IRES se fait grâce aux repliements adoptés par l'IRES. Les structures formées par l'IRES sont capables de recruter et d'assembler les composants de la machinerie traductionnelle de la cellule hôte (eIF, facteurs auxiliaires et/ou le ribosome lui-même). La caractéristique commune à tous les IRES est que la coiffe n'est pas requise. La traduction guidée par les IRES est donc qualifiée de « coiffe-indépendante ». Cependant leurs modes d'action sont de plusieurs types. Par exemple, certains IRES nécessitent une étape de 'scanning' pour trouver le codon initiateur. Pour d'autres IRES, le recrutement du ribosome se fait par dépôt direct sur le codon initiateur sans étape de 'scanning' grâce à des interactions entre la petite sous-unité 40S du ribosome et des éléments de structure secondaire de l'IRES. Le nombre de facteurs d'initiation nécessaires pour la formation des complexes d'initiation varie également d'un IRES à l'autre. Comme la coiffe n'est pas requise pour l'initiation, seul le facteur d'initiation eIF4E (qui reconnaît la coiffe) n'est jamais utilisé par les IRES. En revanche tous les autres eIF peuvent être requis pour l'initiation guidée par un IRES. En outre, certains IRES nécessitent également des facteurs auxiliaires, nommés 'IRES trans-acting factor' (ITAF), qui favorisent la traduction. Les IRES viraux permettent donc de diriger la traduction des protéines virales, grâce aux repliements de leur structure, sans l'aide de la coiffe et avec un nombre variable d'eIF (Hellen, 2009; Lozano & Martínez-Salas, 2015; Martínez-Salas et al., 2015). Les IRES permettent aux virus le détournement de la machinerie traductionnelle de la cellule hôte afin de favoriser une traduction rapide et

efficace des protéines virales. Les bases moléculaires du fonctionnement de ces IRES seront détaillées dans la suite de cette introduction (*paragraphe II.2*).

Les sites d'entrée internes du ribosome dans des ARNm cellulaires :

Des IRES ont également été décrits dans les régions 5'UTR de certains ARNm cellulaires. Selon certains travaux, dix à quinze pour cent des ARNm chez les mammifères contiendraient des IRES cellulaires (Spriggs, et al., 2008). À ce jour, environ cent IRES ont été identifiés ; principalement dans des ARNm cellulaires qui codent pour des facteurs de transcription, des facteurs de croissance et des transporteurs (Lacerda et al., 2017; Mokrejš et al., 2010). Les IRES cellulaires les plus étudiés sont localisés dans les ARNm des proto-oncogènes c-myc, de la cycline D1, du facteur de croissance VEGF et de hif-1 α . La traduction de ces ARN messagers est généralement finement régulée durant le cycle cellulaire. Leur expression est également modifiée en réponse à des stress cellulaires. De manière intéressante, ces ARNm particuliers sont généralement traduits par des mécanismes canoniques (« coiffe-dépendants »). Néanmoins, en condition de stress cellulaire leur traduction est prise en charge par les IRES qu'ils contiennent, utilisant alors des mécanismes « coiffe-indépendant » (Komar et al., 2011). Les mécanismes moléculaires de la traduction guidée par les IRES cellulaires sont encore mal compris. En revanche, des études structurales ont été réalisées sur les IRES cellulaires des ARNm c-myc, FGF-2, Apat-1, CAT-1, FGF-1 (pour revues Komar & Hatzoglou, 2005; Komar et al., 2011). Les résultats de ces études démontrent que les IRES cellulaires adoptent des repliements moins complexes que les IRES viraux. Ils contiennent de courtes régions qui se replient généralement en tige-boucle (Baird, et al., 2007). En outre, le fonctionnement des IRES cellulaires nécessite également des ITAF notamment pour stabiliser leurs structures. Certains de ces ITAF sont utilisés par des IRES cellulaires et viraux, comme par exemple l'ITAF 'Polypyrimidine tract-Binding Protein' ou PTB (Komar & Hatzoglou, 2011; Lewis & Holcik, 2008). Signalons cependant que l'activité traductionnelle des IRES cellulaires est généralement bien plus faible que celle observée pour les IRES viraux. De ce fait, le fonctionnement des IRES cellulaires est au cœur d'une polémique au sein de la communauté scientifique depuis plusieurs années et leur existence est quelquefois remise en question (Shatsky et al., 2010).

En résumé de cette partie, les éléments structuraux retrouvés dans les régions 5'UTR, 3'UTR ou dans les phases codantes des ARNm orchestrent des mécanismes d'initiation de la traduction différents. Ces éléments structuraux sont composés de repliements plus ou moins complexes pouvant générer des interactions à longue distance. Nous avons vu que

les tige-boucles qui lient eIF3 sont des structures secondaires simples. En revanche, les éléments 3'CITES et les structures de la phase codantes de l'histone H4 génèrent des interactions beaucoup plus complexes. Dans le cas des IRES viraux, les repliements tridimensionnels peuvent former des modules sophistiqués grâce à des interactions à distance de type pseudo-nœuds par exemple.

1.3.2. Autres mécanismes d'initiation de la traduction en interne

1.3.2.1. L'initiation de la traduction grâce aux modifications des adénosines

Certains ARNm eucaryotes contiennent des nucléotides modifiés dans leurs régions 5'UTR et 3'UTR. Plus précisément, le nucléotide modifié est la N6-méthyladénosine (m^6A). Ces méthylations des adénosines sont réversibles. Il a été démontré qu'elles ont un impact sur la structure et la traduction des ARNm qui les contiennent (Harcourt et al., 2017; Zhao et al., 2017). Des études récentes ont notamment démontré que les modifications m^6A présentes dans les régions 5'UTR de certains ARNm permettent le recrutement du facteur d'initiation eIF3. Ainsi, la présence de m^6A est suffisante pour guider l'assemblage d'un complexe d'initiation (Meyer et al., 2015) (**Figure 8** à gauche). Ce mécanisme a été décrit pour la traduction de l'ARN messager codant la protéine HSP70, en réponse à un stress induit par un choc thermique (Zhou et al., 2015, 2016). Dans ce cas, le recrutement du ribosome puis la reconnaissance du codon d'initiation de la traduction nécessitent respectivement une extrémité 5' libre et une étape de 'scanning'. C'est-à-dire que l'initiation se fait de manière « coiffe-indépendante » et ne nécessite pas la protéine eIF4E (Meyer et al., 2015). Ces travaux suggèrent que les méthylations m^6A dans les régions 5'UTR des ARNm, à l'instar de certains IRES viraux, seraient capable de recruter eIF3 et d'initier la traduction de manière « coiffe-indépendante ». Ces sites de liaison d'eIF3 ont été appelés '(m^6A)-Induced Ribosome Engagement Site' ou MIREs (pour revue voir Lacerda et al., 2017 ; Leppek et al., 2017; Meyer et al., 2015). Les bases moléculaires de ce mécanisme sont encore méconnues.

1.3.2.2. L'initiation de la traduction des ARN circulaires (circARN)

Les ARN circulaires ou circARN sont des produits de l'épissage des introns dans des précurseurs d'ARNm (Salzman, 2016). Des études récentes ont démontré que les circARN agissent comme des éponges à micro-ARN. En outre, ils peuvent servir de sites d'accueil aux protéines qui ont des domaines de liaison à l'ARN (Chekulaeva & Rajewsky, 2018). De manière surprenante, il a été montré que plusieurs circARN cellulaires présents dans le

cytoplasme des cellules eucaryotes sont traduits. Cependant, le mécanisme qu'ils utilisent pour recruter les ribosomes n'est pas encore bien compris (**Figure 8** à droite). Des expériences pionnières d'il y a une vingtaine d'années ainsi que d'autres plus récentes ont démontré que l'insertion d'un IRES viral dans un circARN permet la traduction d'une ORF localisée en aval de celui-ci (Chen & Sarnow, 1998; Wang & Wang, 2015). Ceci suggère que la traduction des circARN se fait de manière « coiffe-indépendante », par exemple sur des sites spécifiques du circARN contenant des nucléotides modifiés m⁶A. Le facteur d'initiation eIF4G y serait recruté, en interne, par l'intermédiaire d'une protéine nommée YTH3 (Chekulaeva & Rajewsky, 2018).

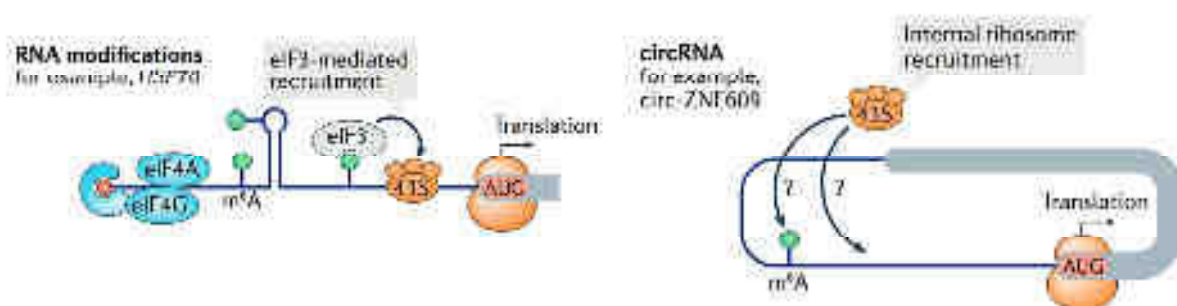


Figure 8 : L'initiation interne de la traduction grâce aux méthylations des adénosines. (Figure issue de Lepek *et al.*, 2017). (**À gauche**) En condition de stress cellulaire, la présence du nucléotide modifié N6-méthyladénosine (m⁶A) dans les régions 5'UTR de certains ARNm eucaryotes favorise le recrutement en interne du facteur d'initiation eIF3. Ce dernier recrute ensuite le complexe de pré-initiation 43S dans la région 5'UTR de l'ARNm. Ce mécanisme d'initiation « coiffe-indépendant » a été observé pour l'ARNm codant pour les protéines de réponses à un choc thermique HSP70. (**À droite**) Un mécanisme de recrutement du complexe 43S similaire serait à l'origine de la traduction de certains ARN circulaires (circRNA ou circARN). Chez les eucaryotes les circARN sont des produits d'épissage des pré-ARNm, leurs rôles et leurs mécanismes moléculaires de traduction sont encore mal compris.

1.3.3. Différents mécanismes de régulation de l'initiation de la traduction

1.3.3.1. La régulation de la traduction par les facteurs eIF

La régulation de l'initiation de la traduction s'effectue à deux niveaux. Premièrement, par la réactivation du facteur d'initiation eIF2, dont le GDP est hydrolysé (eIF2-GDP) après son utilisation. Cette étape est catalysée par le facteur eIF2B. Son rôle est de remplacer la molécule de GDP présente sur la sous-unité γ de eIF2 par une molécule de GTP (Hershey, 2010; Hinnebusch & Lorsch, 2012; Jackson *et al.*, 2010; Sonenberg & Hinnebusch, 2009). Le facteur eIF2-GTP réactivé est alors réintégré dans un nouveau complexe ternaire eIF2-GTP-Met-ARNt^{Met}_i pour l'assemblage d'une nouvelle particule 43S (**Figure 9**). En conditions

de stress cellulaire, comme une infection virale, le recyclage de eIF2-GDP par eIF2B est inhibé suite à l'inactivation du facteur eIF2. Les kinases '(PKR)-like Endoplasmic Reticulum Kinase' (PERK), 'eIF-2-alpha kinase GCN2' (GCN2), 'Heme-Regulated eIF2 α kinase' (HRI) et 'Protein Kinase R' (PKR) phosphorylent de la sous-unité α de eIF2-GDP pour empêcher sa réactivation (**Figure 9** en haut à gauche). L'initiation de la traduction est alors inhibée. Nous décrivons plus en détails les mécanismes de phosphorylation de eIF2 α en réponse à un stress cellulaire dans le *chapitre II.1.2* de ce manuscrit.

Deuxièmement, la régulation de l'initiation de la traduction peut se faire par le facteur eIF4E, le facteur qui se fixe à la coiffe. Le facteur eIF4E contient un motif qui interagit spécifiquement avec la grande protéine plateforme eIF4G. Ce même motif est également reconnu par la protéine '4E-Binding Protein' ou 4E-BP. Chez l'homme il existe trois 4E-BP appelées 4E-BP1, 4E-BP2 et 4E-BP3. La plus étudiée est la protéine 4E-BP1. En se fixant sur eIF4E, 4E-BP empêche l'interaction entre eIF4E et eIF4G lors de l'assemblage du complexe d'initiation eIF4F. De ce fait, la traduction est inhibée. L'hyper-phosphorylation de 4E-BP diminue fortement son affinité pour eIF4E, libérant ce dernier. Le complexe eIF4F se forme alors normalement sur la coiffe et l'initiation de la traduction est possible (Hinnebusch & Lorsch, 2012; Jackson et al., 2010; Sonenberg & Hinnebusch, 2009). L'état de phosphorylation de 4E-BP est contrôlé par la voie 'mammalian Target Of Rapamycin' (mTOR). La voie mTOR intègre quatre signaux cellulaires majeurs, la quantité d'acides aminés, la quantité de facteurs de croissance, la quantité de glucose et la disponibilité de l'hème. En fonction de ces signaux, elle orchestre ensuite les cascades réactionnelles qui en découlent afin de contrôler la croissance cellulaire et le métabolisme cellulaire. Deux complexes multi-protéiques sont au cœur de ces cascades réactionnelles, mTORC1 et mTORC2. Ces deux complexes sont structurellement et fonctionnellement distincts, mais ils contiennent tous deux une protéine kinase nommée mTOR. C'est le complexe mTORC1 qui permet de contrôler l'initiation de la synthèse des protéines en phosphorylant, grâce à la kinase mTOR, l'inhibiteur de traduction 4E-BP (pour le détail de la voie mTOR voir Soulard & Hall, 2007; Roux & Topisirovich 2018). C'est ainsi que l'état de phosphorylation de 4E-BP est une des deux voies classique permettant de contrôler la traduction des ARNm (**Figure 9** en haut à droite).

Cependant, nous avons vu dans les paragraphes précédents que de nombreux ARNm sont en fait traduits par des mécanismes d'initiation « coiffe-indépendants », qui ne nécessitent pas eIF4E. Par conséquent la traduction de cette classe d'ARNm ne peut pas être régulée par 4E-BP. En effet, les mécanismes de régulation de l'initiation de la traduction ne sont pas identiques pour tous les ARNm de la cellule.

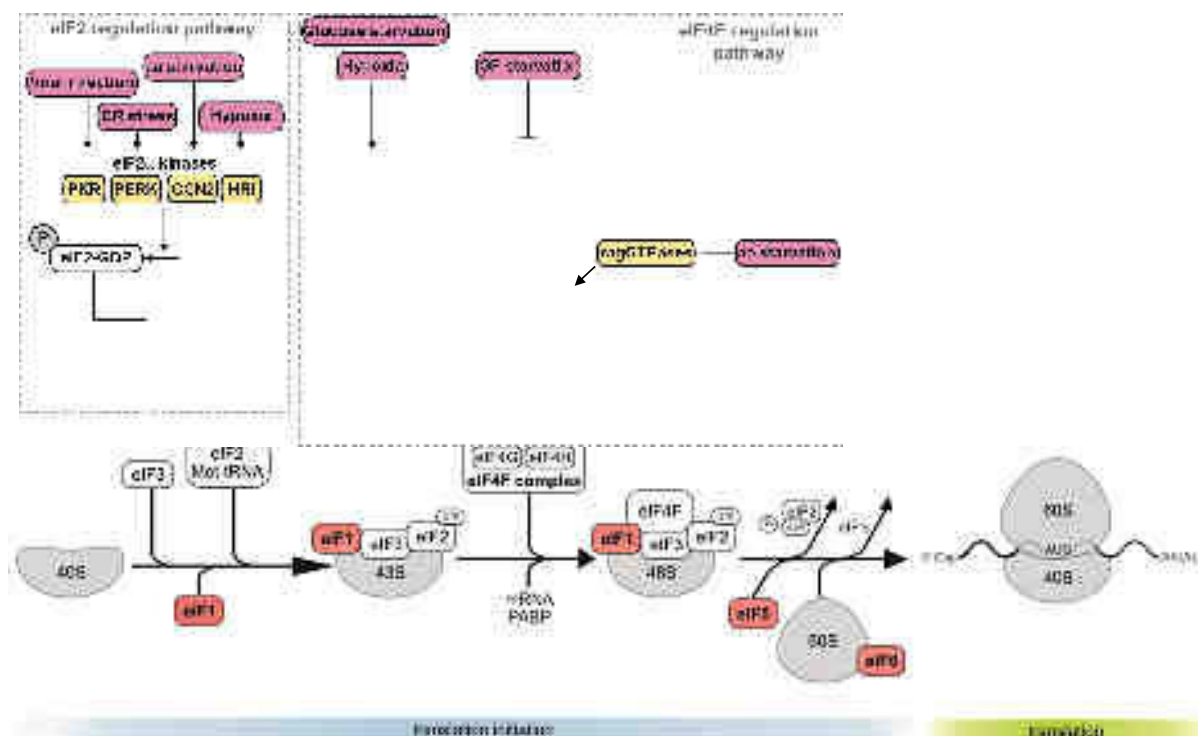


Figure 9 : La régulation de l'initiation de la traduction « coiffe-dépendante » par eIF2 et eIF4E. (Figure adaptée de Golob-Schwarzl *et al.*, 2017). **(À gauche)** Premièrement, l'initiation de la traduction est régulée grâce au recyclage du facteur eIF2-GDP. L'activation de eIF2 dans sa forme GTP est effectuée par le facteur d'initiation eIF2B. Cette étape contrôle la formation du complexe ternaire eIF2-GTP-Met-ARN^{Met}. En conditions de stress cellulaire tel qu'une infection virale, un mauvais repliement tridimensionnel des protéines (ER stress), une carence en hème (hypoxie) ou en acides aminés, la sous-unité α de eIF2-GDP est phosphorylée par les kinases PKR, HRI, PERK ou GCN2. Ceci conduit à l'inhibition du recyclage de eIF2-GTP et donc à l'inhibition de l'initiation de la traduction. **(À droite)** Deuxièmement, la traduction est également régulée par la disponibilité du facteur de liaison de la coiffe eIF4E pour l'assemblage de eIF4F. La disponibilité de eIF4E est régulée par la protéine '4E-Binding Protein' (4E-BP). En conditions cellulaires optimales, 4E-BP est phosphorylée par la protéine kinase mTOR du complexe protéique mTORC1, ainsi eIF4E est libéré. Ce dernier peut alors s'associer avec le facteur d'initiation eIF4G pour former le complexe eIF4F et initier la traduction « coiffe-dépendante ». En réponse à différents signaux cellulaire (carences en glucose, acides aminés, facteurs de croissance, ou en hème) le complexe mTORC1 est inactivé et 4E-BP est dans un état non phosphorylée. De ce fait, 4E-BP séquestre eIF4E et empêche l'initiation de la traduction « coiffe-dépendante ».

1.3.3.2. La régulation par des éléments structuraux présents en région 5'UTR des ARNm

Les pseudo-nœuds :

Certains messagers cellulaires sont capables de former dans la région 5'UTR des interactions à longue distance de type pseudo-nœud avec d'autres parties de l'ARN messager. Par exemple, un pseudo-nœud a été décrit dans la région 5'UTR de l'ARNm codant l'interféron γ chez l'homme (**Figure 10** à gauche). L'initiation de la traduction de

l'interféron γ est « coiffe-dépendante ». Et elle est aussi dépendante de la conformation structurale de la région 5' UTR dans laquelle se trouve le pseudo-nœud. Lorsque le pseudo-nœud est ouvert, la traduction est possible. En revanche, lorsque le pseudo-nœud se forme après le passage du ribosome, des protéines kinases γ sont recrutées pour phosphoryler eIF2 et ainsi inhiber la traduction de l'interféron γ (Ben-Asouli et al., 2002; Cohen-Chalamish et al., 2009). Il s'agit donc ici d'un système de régulation de l'initiation de la traduction « coiffe-dépendante » nécessitant la présence d'une structure tertiaire sur l'ARNm et des protéines régulatrices agissant en *trans*.

Les structures de type G-quartet :

D'autres interactions tertiaires peuvent être présentes en 5'UTR des ARNm comme les structures de type G-quartet (**Figure 10** à droite). Leur rôle dans l'initiation de la traduction n'a été étudié que très récemment. Les G-quartet sont des structures extrêmement stables qui se forment lorsque la séquence ribonucléotidique contient une répétition de résidus G. Les résidus G interagissent entre eux pour former des plateaux de quatre bases G, les plateaux sont espacés de manière régulière. Le canal formé au milieu de G-quartets est occupé par des ions K^+ qui stabilisent les plateaux de bases. La taille de ce canal ne permet pas d'accepter d'autres cations que le K^+ tels que Li^+ et Na^+ . La très grande stabilité de ces structures empêcherait la liaison du complexe de pré-initiation 43S à l'ARN ou bloquerait le 'scanning' du ribosome pendant l'initiation (Halder et al., 2009; Kumari et al., 2008). Les G-quartet sont ainsi qualifiés d'éléments de répression de l'initiation de la traduction agissant en *cis*.

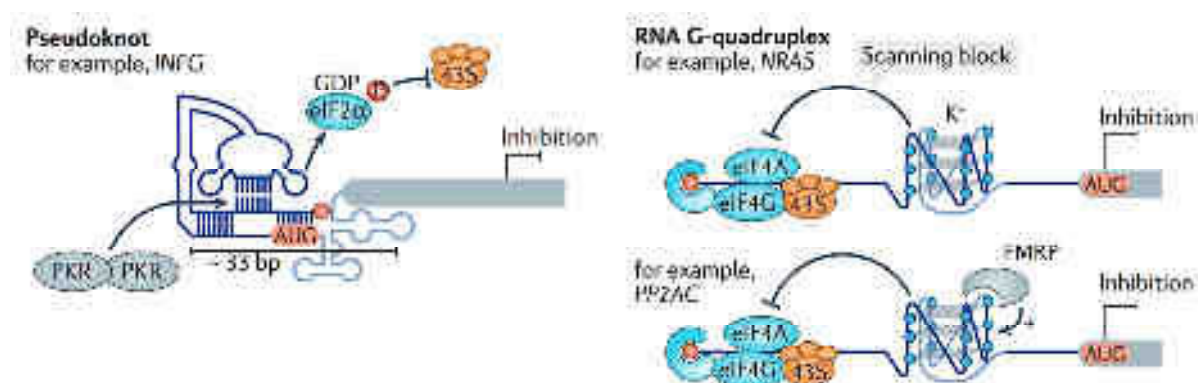


Figure 10 : La régulation de l'initiation de la traduction par des structures présentes dans la région 5'UTR de l'ARNm. (Figure issue de Leppik *et al.*, 2017). (**À gauche**) La région 5'UTR de l'ARN messager codant l'interféron γ contient une structure tertiaire de type pseudo-nœud dont les changements conformationnels induisent le recrutement de kinases PKR. Les kinases PKR phosphorylent eIF2 α , inhibant de ce fait la traduction. (**À droite**) D'autres ARN messagers possèdent des structures tertiaires de type G-quartet dans leurs régions 5'UTR. La formation de ces structures nécessite spécifiquement des ions (K^+). La fixation de la protéine FMRP

sur les G-quartet peut également moduler l'efficacité de la traduction. Le mécanisme moléculaire d'action des G-quartet est encore mal connu, nous savons simplement que ces structures jouent un rôle dans la répression de la traduction.

Les éléments TIE et IRES des Hox :

Un troisième exemple d'élément structural impliqué dans la régulation de la traduction a été décrit. Il s'agit d'éléments structuraux localisés dans les régions 5'UTR des ARNm des gènes homéotiques ou Hox. Ces gènes sont exprimés spécifiquement pendant le développement embryonnaire (pour revue Xue et al., 2014). Les gènes Hox sont exprimés de manière tissu spécifique à un moment précis du développement embryonnaire. Leur fonction est d'activer des cascades d'expression de gènes spécifiques, à des étapes clés du développement embryonnaire. C'est pourquoi il est primordial que leur traduction soit régulée de manière précise. En effet, une perturbation de ces mécanismes de régulation engendre de graves conséquences pendant le développement embryonnaire. La régulation de la traduction des gènes Hox est contrôlée par deux éléments structuraux : d'une part des 'Translation Inhibitory Element' (TIE) qui inhibent la traduction « coiffe-dépendante », et d'autre part des IRES qui guident la traduction de la phase codante des gènes Hox. Ces deux éléments agissent en synergie pour réguler précisément le démarrage et l'arrêt de la traduction pendant le développement embryonnaire. Les mécanismes moléculaires sous-jacents ne sont pas encore élucidés. Des hypothèses suggèrent que les TIE empêcheraient le recrutement du complexe de pré-initiation 43S en région 5'UTR ou bien bloqueraient l'étape de 'scanning', inhibant ainsi l'initiation « coiffe-dépendante » de la traduction. L'inhibition de la traduction canonique grâce aux TIE favoriserait alors l'activation de la traduction « IRES-dépendante » de ces ARNm Hox. La combinaison de ces deux éléments structuraux en *cis* permet une régulation spatio-temporelle fine de l'initiation de la traduction « IRES-dépendante » des gènes Hox.

1.3.3.3. Le contrôle de la traduction par les uORF

De la même manière que les TIE, les 'upstream ORF' (uORF) peuvent agir en synergie avec des IRES présents dans la région 5'UTR du même transcrit dans le but de contrôler l'initiation de la traduction. De tels ARN messagers, possédant à la fois un IRES cellulaire et une uORF, ont été décrits pour les facteurs de croissance et de différenciation cellulaire PDGF, GATA-6, CAT-1, VEGF-A et FGF9. Les modes d'action des uORF sont variés. Par exemple, lors de la synthèse de la protéine FGF9, la traduction de l'uORF permet de réprimer la traduction de l'ORF principale codant pour FGF9 qui est localisée en aval. Lors d'un stress hypoxique la traduction « coiffe-dépendante » est inhibée. La traduction de

l'uORF est donc éteinte. Cet événement conduit à l'activation de la traduction guidée par l'IRES et donc à une surexpression de l'ORF principale FGF9 (Chen et al., 2014). Au contraire, certaines autres uORF peuvent entraîner l'activation de l'ORF en aval. C'est le cas lors de la traduction de l'ARNm codant le transporteur CAT1. La traduction de l'uORF engendre des réarrangements structuraux de la région 5'UTR qui aboutissent au dépliement d'une structure inhibitrice de la traduction. L'IRES localisé en aval n'est alors plus réprimé, il est donc pleinement actif. Ainsi, la traduction de l'ORF principale CAT1 est possible (Fernandez et al., 2002, 2005; Yaman et al., 2003). En conclusion, les uORF sont considérées comme des éléments présents en *cis* sur l'ARNm qui permettent de réguler positivement ou négativement la traduction.

1.3.3.4. Le contrôle de l'initiation de la traduction par le ribosome

Le ribosome lui-même peut jouer un rôle actif et influencer la régulation de l'initiation de la traduction de certaines populations d'ARNm, grâce à sa composition en protéines ribosomiques. Il a été démontré que les ARNm codant pour les protéines Hox sont capables de recruter des « ribosomes spécialisés », dont la composition protéique est différente. Ces ribosomes contiennent notamment la protéine Rpl38 (Topisirovic & Sonenberg, 2011). La présence de Rpl38 dans le ribosome est strictement nécessaire à la traduction d'une sous-classe de onze ARNm Hox du locus A. Sur ces ARNm, la protéine ribosomique Rpl38 contrôle la formation du complexe 80S (Kondrashov et al., 2011). En revanche, la traduction canonique « coiffe-dépendante » des autres ARNm de la cellule ne nécessite pas la présence de Rpl38 sur les ribosomes. D'autres études ont rapporté que la protéine ribosomique 'Receptor for Activated C Kinase1' ou RACK1 joue un rôle identique dans la sélection par le ribosome d'une catégorie d'ARN possédant des IRES. En effet, des ribosomes contenant la protéine RACK1 stimule l'initiation de la traduction « IRES-dépendante » chez les virus 'Hepatitis C Virus' (HCV), 'Drosophila C Virus' (DCV) et 'Cricket Paralysis Virus' (CrPV) (Majzoub et al., 2014). Cependant, à l'instar de Rpl38, la présence de RACK1 dans le ribosome n'est pas nécessaire pour initier la traduction « coiffe-dépendante » d'autres ARNm présents dans la cellule (**Figure 11**). D'autres études encore ont démontré que RACK1 est aussi directement impliquée dans la traduction sélective des ARNm possédant des IRES cellulaires tel que le proto-oncogène c-myc (Ruan et al., 2012). Ces dernières années plusieurs autres protéines du ribosome impliquées dans la sélection de familles d'ARNm ont été caractérisées. Ainsi, le ribosome possède une capacité intrinsèque à choisir les ARNm pour les traduire. Ce phénomène a été baptisé « code ribosome » (pour revue voir Topisirovic & Sonenberg, 2011).

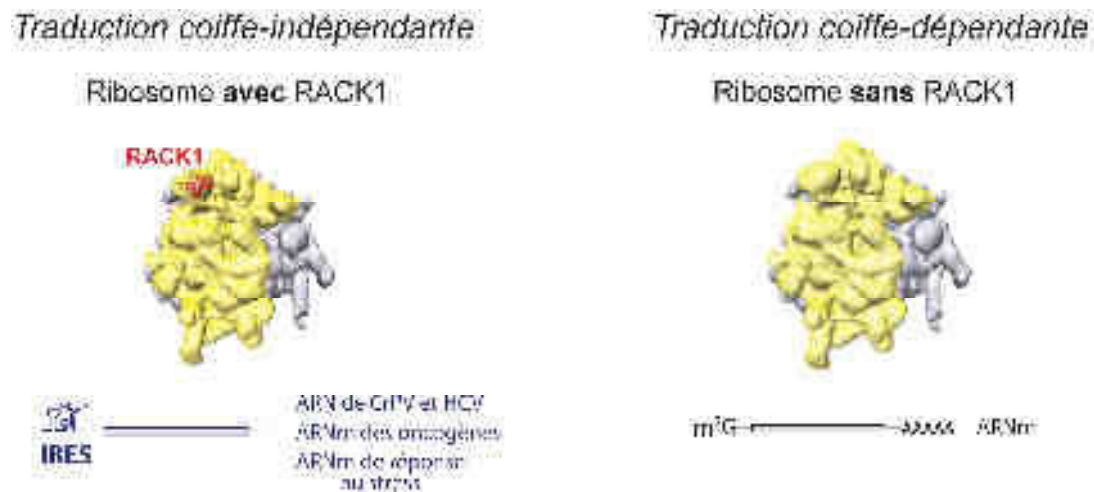


Figure 11 : La protéine du ribosome RACK1 participe à la traduction sélective des ARNm en fonction des conditions physiologiques de la cellule. (À droite) RACK1 n'influence pas la traduction « coiffe-dépendante » des ARNm cellulaires qui sont exprimés de manière constitutive dans la cellule. (À gauche) En revanche RACK1 est indispensable à la traduction spécifique des ARNm possédant des IRES (ARN viraux de CrPV, HCV et ARNm cellulaires codant pour des oncogènes). La traduction de ces ARNm viraux est augmentée lors d'une infection virale. La traduction des ARNm oncogènes est augmentée pendant la propagation des cancers.

II. Le piratage de l'appareil de traduction de la cellule hôte par les virus à ARN simple brin grâce aux IRES

II.1. La dérégulation de la traduction de l'hôte par les virus à ARN simple brin

Chez les virus à ARN simple brin de polarité positive l'ARN viral qui entre dans la cellule lors de l'infection est traduit directement par le ribosome de la cellule hôte. Ces virus déploient des stratégies différentes pour stopper la traduction de l'hôte et pour détourner les composants de la machinerie traductionnelle pour la synthèse exclusive des protéines virales. Dans les stades précoces de l'infection, des protéines virales ciblant les facteurs d'initiation de la traduction de l'hôte sont synthétisées par certains virus. Leur rôle est de stopper la traduction de l'hôte au profit de la traduction virale (Hunnicut, et al., 2006; Martínez-Salas et al., 2015; Pelletier & Sonenberg, 1988). Plus précisément, ces protéines interfèrent avec les mécanismes de régulation traductionnelle de l'hôte. Par exemple, certaines protéines virales clivent ou séquestrent des facteurs d'initiation comme eIF4G. Les virus peuvent également stimuler les mécanismes de phosphorylation du facteur eIF2, entraînant ainsi l'inhibition de la traduction canonique « coiffe-dépendante ». La traduction des protéines virales, quant à elle, est assurée par des IRES localisés dans l'ARN viral. Ne nécessitant pas les eIF ciblés par les protéines virales, la traduction « IRES-dépendante » du

virus est immune à ces mécanismes d'inhibition de la machinerie traductionnelle de l'hôte. Ainsi, la traduction virale est largement favorisée par rapport à la traduction cellulaire. Dans ce *chapitre II* nous présenterons tout d'abord quelques exemples d'inactivation de facteurs eIF de l'hôte par les virus. Ensuite, nous aborderons en détails les bases moléculaires du fonctionnement des IRES viraux.

II.1.1. L'inactivation du complexe eIF4F

II.1.1.1. Le clivage de eIF4G

Les virus utilisent différents mécanismes pour inactiver le complexe eIF4F et inhiber en conséquence la traduction « coiffe-dépendante » de leur hôte. Un premier mécanisme d'inactivation du complexe de liaison à la coiffe eIF4F a été observé dans des cellules infectées par le Poliovirus. La protéase 2A codée par le virus inactive eIF4F en clivant le facteur d'initiation eIF4G (**Figure 12** en haut). La partie C-terminale de eIF4G, qui contient les sites de liaison pour eIF3 et eIF4A, est séparée de la partie N-terminale contenant le site de liaison pour eIF4E et les PABP (Etchison et al., 1982; Lamphear et al., 1993). La traduction « coiffe-dépendante » de l'hôte est de ce fait inactivée. Le complexe eIF4F clivé ne peut plus se fixer sur la coiffe, il contient eIF3 et eIF4A. Ce complexe eIF4F clivé est recruté spécifiquement par le virus pour assurer le 'scanning' du complexe 43S jusqu'au codon d'initiation du génome virale.

II.1.1.2. La séquestration de eIF4E

Un autre mécanisme d'inactivation de eIF4F a été observé chez le Poliovirus, le 'Vesicular Stomatitis Virus' (VSV) et l'EMCV (**Figure 12** en bas). Dans les cellules infectées, les 4E-BP sont déphosphorylées, ce qui entraîne la séquestration de eIF4E et une inhibition spécifique de la traduction cellulaire « coiffe-dépendante » (Gingras et al., 1996; Connor & Lyles, 2002). Le mécanisme utilisé par EMCV et PV pour déphosphoryler 4E-BP n'est pas encore connu. En revanche, dans le cas d'une infection par le VSV, le virus inhibe la voie de régulation mTOR à l'aide d'une kinase virale. Plus précisément, la protéine kinase virale M inhibe la kinase AKT de la voie mTOR. L'inhibition de AKT entraîne une inactivation du complexe mTORC1, et de ce fait une inhibition de la phosphorylation de 4E-BP par la kinase mTOR. Il en résulte une séquestration de eIF4E par 4E-BP (voir la revue Walsh & Mohr, 2011).

Ainsi, les deux mécanismes d'inactivation du complexe eIF4F cités ci-dessus permettent aux virus d'inhiber sélectivement la traduction des ARNm de la cellule hôte dans le but de favoriser la traduction virale guidée par l'IRES.

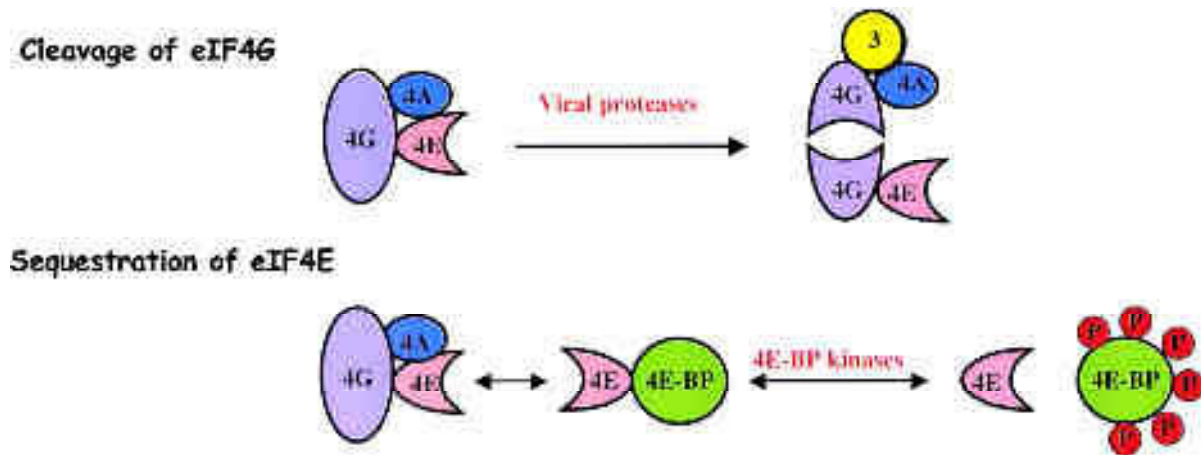


Figure 12 : Dissociation du complexe de liaison de la coiffe eIF4F dans les cellules infectées par les virus à ARN simple brin. (Figure de Bushell et Sarnow 2002). (En haut) Le facteur d'initiation eIF4G est clivé par la protéase virale 2A des Picornavirus. Cette coupure induit l'inactivation du complexe eIF4F et de la traduction des ARN messagers de l'hôte. (En bas) Dans les cellules infectées par les Picornavirus, les '4E-Binding Proteins' (4E-BP) sont déphosphorylées par le virus, ceci induit la séquestration eIF4E et l'inhibition de la formation de eIF4F. La traduction « coiffe-dépendante » de l'hôte est alors impossible.

II.1.2. La phosphorylation du facteur d'initiation eIF2

Lors d'une infection virale, la cellule met en place des mécanismes de défense antivirale. Afin de bloquer la synthèse des protéines virales, elle procède notamment à l'inactivation du recyclage du facteur d'initiation eIF2 en phosphorylant sa sous-unité régulatrice eIF2 α (voir **figure 9**). La phosphorylation de eIF2 α réduit ainsi le pool de complexe ternaire eIF2-GTP-Met-ARNt^{Met}. Chez les mammifères, il existe quatre kinases capables de phosphoryler eIF2 α : Les kinases PERK et GCN2, induites par le stress du réticulum endoplasmique (RE) qui signale l'accumulation d'ARNt déacylés résultant de carences en acides aminés. Les autres kinases sont la kinase HRI, activée lors de privations en hème (stress hypoxique) et la kinase PKR, induite lors de la traduction de l'interféron γ . Cette dernière est capable de se fixer sur des motifs spécifiques d'ARN double brin (voir *paragraphe 1.3.3.2* et **figure 10**), tels que des pseudo-nœuds ou des tige-boucles de plus de 30 paires de base. En condition de stress cellulaire déclenché par la présence d'un virus dans le cytoplasme de la cellule, la kinase PKR est activée suite à sa liaison à l'ARN viral double brin qui est en cours de répllication. En réponse à la détection de l'ARN viral, PKR phosphoryle eIF2 α . Elle est alors

responsable de l'inhibition de traduction du virus mais aussi de la traduction de l'hôte. Les kinases PERK et GCN2 sont également activées lors d'infections virales, notamment lors de l'infection par le VSV et les membres de la famille des *Alphavirus* (Berlanga et al., 2006; Connor & Lyles, 2005; McInerney et al., 2005; Ryman et al., 2005; Ventoso et al., 2006). Ces virus dépendent du RE pour la synthèse et le repliement correct des protéines virales. Outre l'accumulation d'ARNt non chargés, la surcharge de protéines sur le RE conduit également à l'activation de PERK. Celle-ci phosphoryle eIF2 α qui va inhiber la synthèse protéique globale et de ce fait résorber l'excès de protéines dans le RE (Harding et al., 1999). Certains virus comme l'Herpes Simplex Virus' (HSV), le HCV ou le 'Kaposi's Sarcoma-associated Herpes Virus' (KSHV) contrebalancent cette inactivation en neutralisant, en amont de la cascade réactionnelle, les kinases PKR ou PERK responsables de la phosphorylation de eIF2 α . Ainsi, une partie des facteurs d'initiation eIF2 de la cellule n'est pas phosphorylée. Ce pool d'eIF2 actif est utilisé par le virus pour sa traduction (Mulvey et al., 2007, 2003). D'autres virus, tels que HCV ou le 'Sindbis Virus' (SV) sont capables d'utiliser eIF2 dans son état phosphorylé (Dmitriev et al., 2010; Garaigorta & Chisari, 2009; Lancaster et al., 2006; Ventoso et al., 2006). Dans ce cas, le mécanisme d'inactivation de eIF2 éteint la traduction « coiffe-dépendante » des messagers de l'hôte. En revanche, la traduction virale n'est pas affectée.

II.2. Le détournement du ribosome de l'hôte grâce aux IRES viraux

Certains virus à ARN détournent la machinerie traductionnelle de leur hôte grâce à des éléments structuraux dans leur ARN génomique appelés IRES. Les IRES leur permettent également de contourner l'inhibition traductionnelle déclenchée par leur présence dans les cellules (*paragraphe ci-dessus*). La majorité des IRES viraux connus à ce jour ont été découverts dans les génomes des virus appartenant à l'ordre des *Picornavirales* (**Figure 13**) et plus précisément dans les virus des familles des *Dicistroviridae* et des *Picornaviridae*. En outre, des IRES ont également été décrits chez les *Flaviviridae* et les *Retroviridae*. Malgré quelques caractéristiques communes, les IRES sont à première vue très différents. Ils se distinguent par leur mode d'action pour initier la traduction et par la complexité de leurs structures secondaires et tertiaires. Ces deux critères ont permis de classer les IRES viraux en quatre classes que nous allons maintenant détailler (pour revue voir Mailliot & Martin, 2018).

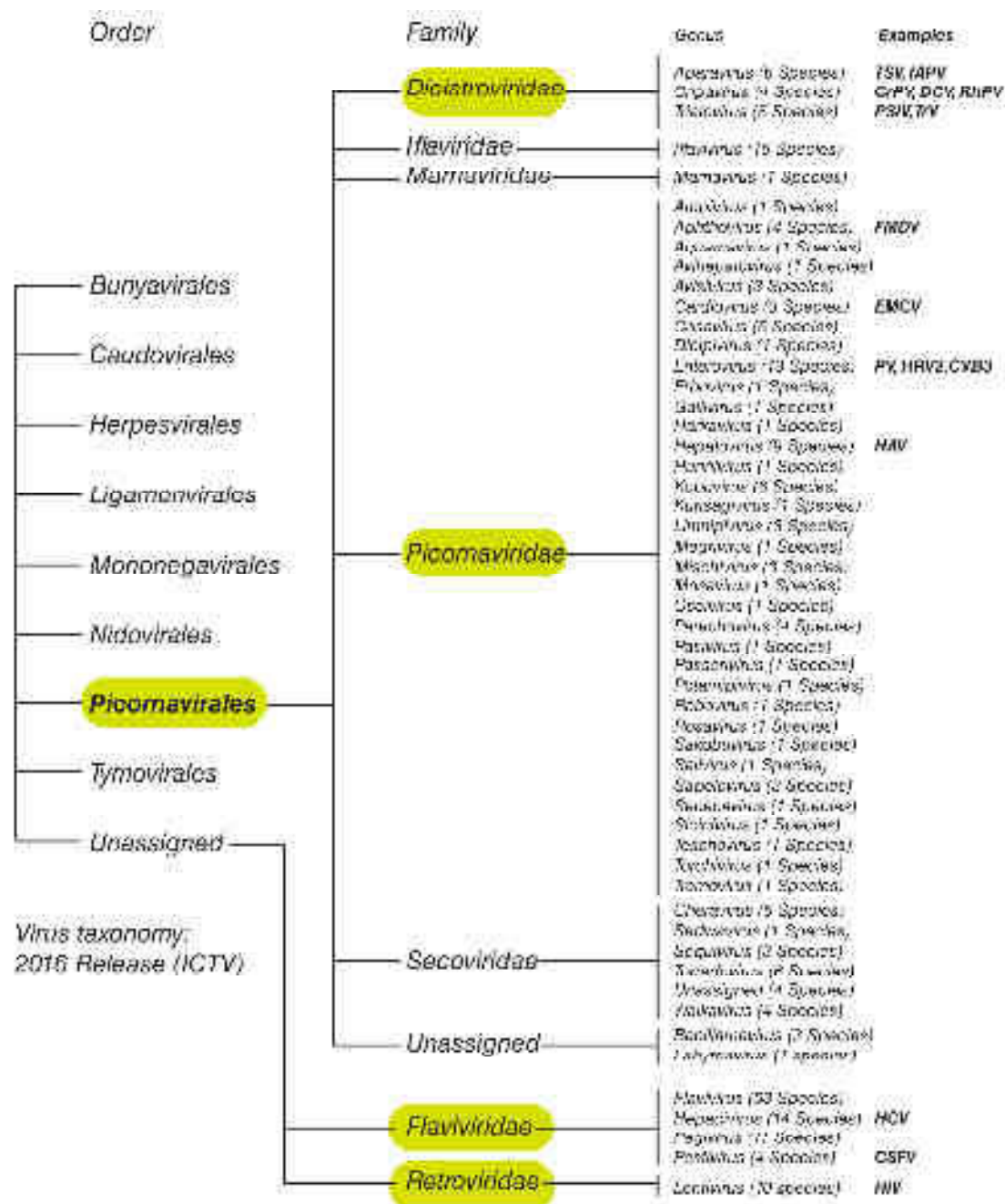


Figure 13 : La classification des virus. (Figure issue de Martin et Maillot 2018). Les ordres et les familles de virus qui contiennent des 'Internal Ribosome Entry Site' (IRES) sont surlignés en vert. Des exemples d'espèces de virus contenant les IRES les mieux décrits sont indiqués à droite.

II.2.1. Les IRES de classe I

Les IRES de classe I sont majoritairement présents dans les génomes des *Picornaviridae*, comme par exemple les Enterovirus A71 (EV-A71), Coxsackievirus (CVB3), Poliovirus (PV), et Rhinovirus Humain (HRV). Parmi ces IRES de classe I, le mieux caractérisé est l'IRES présent dans la région 5'UTR du Poliovirus (voir la revue Mailliot & Martin, 2018). L'extrémité 5'UTR du PV comprend 743 nucléotides et se replie de manière à former des éléments structuraux qui sont regroupés en six domaines (domaines structuraux I à VI). Les 100 premiers nucléotides forment le domaine I, ce dernier adopte une structure en feuille de trèfle ('cloverleaf'). Cette structure est requise lors de l'étape de réplication des brins d'ARN viraux positifs et négatifs (**Figure 14** surlignée en vert clair) (Andino et al., 1990, 1993). La partie IRES proprement dite comprend les domaines II à VI (des nucléotides 100 à 743). L'IRES du poliovirus utilise tous les facteurs d'initiation pour recruter le ribosome, à l'exception de eIF4E. Le facteur eIF4F est dissocié après coupure de eIF4G par la protéase virale 2A. Après coupure, le fragment eIF4G restant peut encore s'associer avec eIF4A, mais plus avec eIF4E. Ce complexe eIF4G clivé-eIF4A est recruté sur le domaine V de l'IRES (**Figure 14** surligné en vert foncé). Il guide le recrutement du complexe de pré-initiation 43S et permet son 'scanning' jusqu'au codon d'initiation AUG (Gradi et al., 1998; Lamphear et al., 1995). De manière intéressante, un premier codon AUG en position 586-588 permet l'entrée du ribosome mais pas l'initiation de la traduction. Le ribosome passe outre ce premier AUG et poursuit le 'scanning' jusqu'au deuxième AUG en position 743. La traduction ne démarre que sur ce codon AUG en position 743 (Pilipenko et al., 1992). L'IRES du poliovirus contient également des motifs essentiels pour le recrutement de facteurs auxiliaires, les ITAF (**Figure 14**). Par exemple, le domaine IV contient un motif riche en C impliqué dans le recrutement de la 'Poly(RC)-Binding Protein 2' (PCBP2) qui est nécessaire à la traduction chez le PV (Perera et al., 2007; Toyoda et al., 2007; Sweeney et al., 2014). Ce même domaine présente également une tige-boucle contenant la séquence consensus GNRA qui est impliquée dans une interaction à longue distance avec une autre région du génome viral (Malnou et al., 2002; Fiore et al., 2002). Sa fonction précise ainsi que l'autre région du génome impliquée n'ont pas encore été clairement identifiées. Cependant, il a été démontré que le motif GNRA est essentiel pour l'activité de l'IRES. En outre, cet IRES possède également un motif riche en pyrimidines qui est le site de liaison de la protéine PTB (Kafasla et al., 2010).

Ainsi, les IRES de classe I nécessitent tous les facteurs d'initiation excepté eIF4E pour initier efficacement la traduction. Ils contiennent des structures secondaires simples, généralement constituées de tiges-boucles plus ou moins courtes. Ils recrutent le ribosome en amont du codon d'initiation AUG. Puis, une étape classique de 'scanning' du 5' vers le 3' permet de localiser le codon de démarrage AUG.

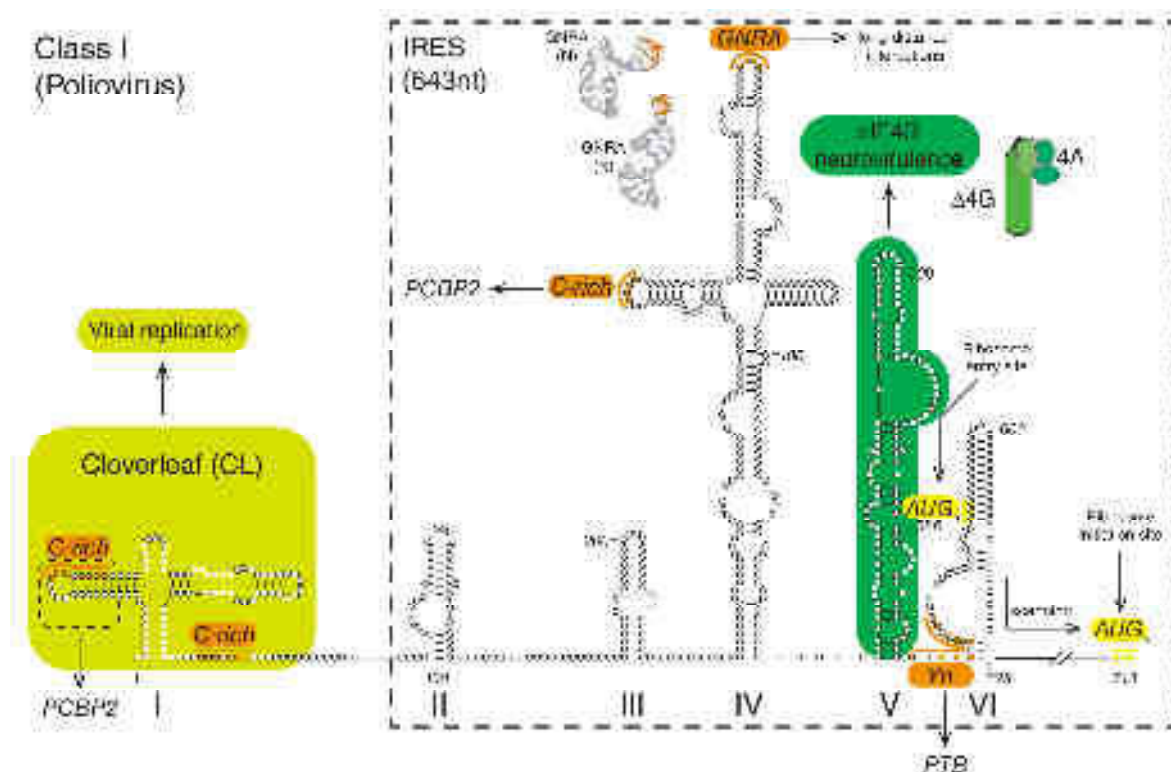


Figure 14 : Les IRES de classe I : exemple du Poliovirus (PV). (Figure issue de Martin et Mailliot, 2018). La région 5'UTR du PV adopte une structure secondaire comprenant six domaines distincts. Le domaine I se replie en une structure en feuille de trèfle ('Cloverleaf' ou CL) qui est impliquée dans la réplication virale (vert clair). Les domaines II à VI constituent l'IRES proprement dit du PV (encadré en pointillés). Certains nucléotides de ces domaines sont impliqués dans le recrutement des 'IRES trans-acting factors' (ITAF) ou sont requis pour la formation d'interactions à distances qui contribuent à l'intégrité structurale de l'IRES (ces nucléotides sont indiqués en orange). Des données structurales de l'interaction à distance de type tétra-boucle GNRA ont été obtenues par RMN (N), les nucléotides sont représentés selon le même code couleur. Le domaine V de l'IRES interagit spécifiquement avec le facteur d'initiation de la traduction eIF4G clivé par la protéase 2A (en vert foncé). Les deux codons AUG du PV sont surlignés en jaune.

II.2.2. Les IRES de classe II

Des IRES appartenant à la classe II sont également retrouvés chez certains virus de la famille des *Picornaviridae*. Dans cette catégorie, les IRES les plus étudiés sont ceux du 'Encephalo Myocarditis Virus' EMCV du genre des *Cardiovirus* et du 'Foot and Mouth Disease Virus' (FMDV) du genre des *Aphthovirus*. Ces deux IRES possèdent de nombreuses caractéristiques structurales communes, bien que leur séquence nucléotidique diverge. Ils sont d'une taille similaire (439 nucléotides pour l'EMCV et 445 nucléotides pour le FMDV) et contiennent cinq domaines structuraux organisés de la même manière. Le domaine I est impliqué dans la réplication virale comme chez le poliovirus présenté plus haut. Les

domaines II à V contiennent l'IRES. On y retrouve des sites d'interaction pour des ITAF dont la présence est indispensable au recrutement du ribosome (**Figure 15** : représentation de l'IRES de l'EMCV). Par exemple, le domaine II contient une séquence riche en pyrimidine qui recrute PTB (Andreev et al., 2007; Jang & Wimmer, 1990; Luz & Beck, 1991). Lorsque la protéine PTB est recrutée par un IRES elle contribue à la formation du complexe de pré-initiation 48S sur l'IRES (Andreev et al., 2007). Le domaine III se replie indépendamment (Fernández & Martínez-Salas, 2003; Serrano et al., 2007), il adopte une structure secondaire en longue tige-boucle avec deux jonctions à quatre hélices. Dans les boucles apicales du domaine III se trouve un motif RAAA ainsi qu'un motif GNRA capable de former une interaction à distance avec une boucle riche en C de ce même domaine III (Dupont & Snoussi, 2009; Martínez-Salas et al., 2015; Phelan, 2004). Bien que la fonction du motif GNRA ne soit pas encore claire, il est essentiel pour l'activité de l'IRES de EMCV (Robertson et al., 1999) et du FMDV (López de Quinto & Martínez-Salas, 1997) comme décrit précédemment chez le poliovirus. Le domaine IV adopte un repliement en Y, également conservé chez les deux virus (Hoffman & Palmenberg, 1995). Ce domaine constitue le site de liaison de eIF4G à l'IRES (Kolupaeva et al., 1998; López de Quinto & Martínez-Salas, 2000). Les deux tige-boucles qui surplombent la jonction en triple-hélices s'associent au domaine HEAT-1 de eIF4G (Imai et al., 2016). Le domaine V contient un autre motif riche en pyrimidines auquel se lient plusieurs facteurs, notamment le facteur d'initiation eIF4B, PTB ainsi que d'autres protéines qui possèdent des sites de fixation à l'ARN (Fernandez-Chamorro et al., 2014; Jang & Wimmer, 1990; Kafasla et al., 2009; López de Quinto et al., 2001; Luz & Beck, 1991). Les IRES de FMDV et EMCV sont donc structurellement très similaires. En revanche, ils fonctionnent différemment lors de la reconnaissance du codon d'initiation AUG. Les deux IRES possèdent deux codons AUG qui sont en phase. Dans le cas de EMCV seul le deuxième codon AUG est utilisé pour l'initiation de la traduction. En revanche, les deux codons AUG sont utilisés lors de la traduction virale du FMDV (Jackson, 2005; López de Quinto & Martínez-Salas, 1999). Les deux codons sont séparés de 84 nucléotides, mais grâce à un repliement tridimensionnel ils seraient spatialement proches, permettant un démarrage de la traduction sur les deux codons AUG (Andreev et al., 2007). Contrairement aux IRES de classe I, la reconnaissance des codons d'initiation chez le FMDV et l'EMCV ne nécessite pas d'étape de 'scanning'. En effet le complexe de pré-initiation 43S est déposé directement sur le codon AUG. De plus, les deux virus utilisent des stratégies différentes pour bloquer la traduction cellulaire. Dans le cas du virus FMDV, le facteur d'initiation eIF4G est clivé grâce à la protéase virale 3C-pro (Mosenkis et al., 1985). Lors de l'infection par EMCV, ce sont les 4E-BP déphosphorylées qui séquestrent eIF4E (Gingras et al., 1996). Dans les deux cas, l'initiation de la traduction « coiffe-dépendante » de l'hôte est inhibée et la traduction « IRES-dépendante » de ces virus favorisée.

En résumé, les IRES de classe II adoptent donc des repliements simples, nécessitent tous les facteurs d'initiation excepté eIF4E et procèdent à un dépôt direct du ribosome sur le codon d'initiation AUG sans étape de 'scanning'. Ils sont fonctionnellement similaires aux IRES de classe I mais ils s'affranchissent de l'étape de 'scanning' du ribosome, lente et coûteuse en énergie.

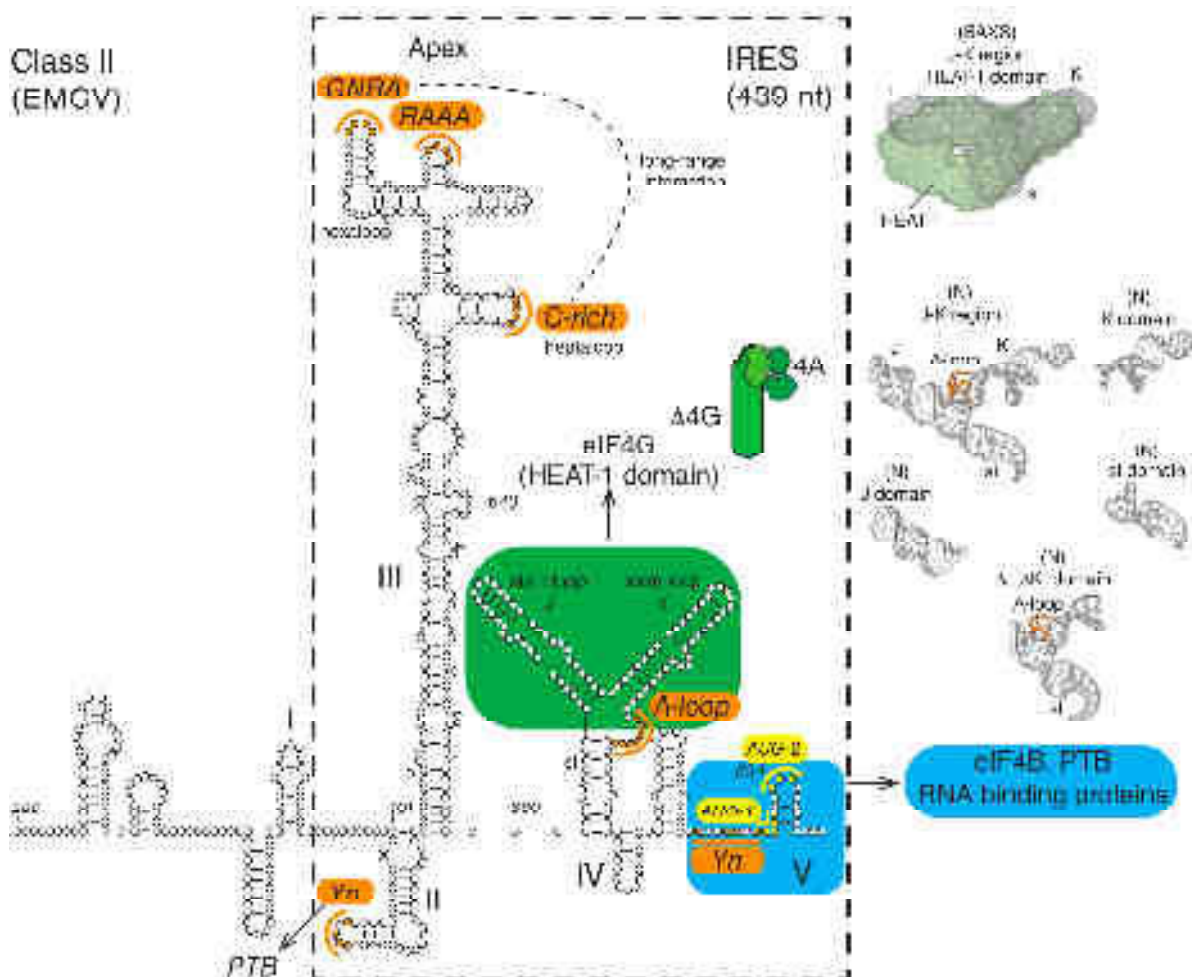


Figure 15 : Le virus de l'Encéphalomyocardite (EMCV) possède un IRES de classe II. La figure, issue de Martin et Maillot, 2018, représente la structure secondaire de la région 5'UTR de l'ARN viral de l'EMCV. Les nucléotides essentiels pour l'activité de l'IRES ou pour le recrutement des ITAF sont indiqués en orange. Le domaine minimal nécessaire à l'activité IRES de l'EMCV est encadré en pointillés. Le domaine IV recrute spécifiquement le facteur eIF4G (en vert foncé) et le domaine V interagit avec le facteur d'initiation eIF4B et l'ITAF PTB, ainsi qu'avec d'autres protéines liant l'ARN (en bleu). Les deux codons initiateurs AUG de l'EMCV sont surlignés en jaune. Des données structurales obtenues par RMN (N) sur le domaine IV isolé ou sur le domaine IV en complexe avec le domaine HEAT-1 de eIF4G (par SAXS) sont représentées à droite.

II.2.3. Les IRES de classe III

II.2.3.1. Structure de l'IRES du virus de l'hépatite C

Les IRES de classe III ont été décrits chez les virus de la famille des *Flaviviridae* et chez certains membres des *Picornaviridae*, notamment chez le 'Simian Virus 2' (SV2) et le 'Porcine Enterovirus 8' (PEV-8). L'IRES de cette classe le plus étudié est celui du virus de l'hépatite C (HCV) qui appartient à la famille des *Flaviviridae*. Ses caractéristiques servent ainsi de référence pour les IRES de classe III (**Figure 16**). L'IRES du HCV comprend 298 nucléotides, il est subdivisé en quatre domaines structuraux (nommés domaines I à IV). Les repliements de ces domaines sont globalement plus complexes que dans les IRES des classes I et II. Des études structurales par cristallographie aux rayons X, par Résonance Magnétique Nucléaire (RMN), par Cryo-Microscopie Électronique (Cryo-ME) ou encore par 'Small Angle X-ray Scattering' (SAXS) ont démontré que les différents domaines de l'IRES de HCV sont rigides et peuvent se replier indépendamment les uns des autres (Kieft et al., 2002; Lukavsky, 2009; Lukavsky et al., 2003; Pérard et al., 2013). De plus, ces domaines sont séparés par des régions charnières non structurées permettant une flexibilité inter-domaines (Kieft et al., 1999; Kieft, 2008; Pérard et al., 2013). La plupart de ces domaines sont conservés chez les Flavivirus ainsi que chez certains Picornavirus (Asnani et al., 2016). Le domaine II adopte une structure en longue tige-boucle (Lukavsky et al., 2003). Ce domaine est lui-même composé de deux sous-domaines nommés module IIa et module IIb (**Figure 16**). Le module IIa contient une boucle charnière qui autorise la flexibilité de cette longue tige-boucle. Le module IIb abrite également une boucle interne et une tige-boucle apicale. Le domaine III se replie en une grande structure comprenant six sous-domaines, nommés respectivement modules IIIa à IIIf. Les modules IIIabc et IIIef forment des jonctions à quatre hélices (Laing & Schlick, 2009), alors que le module IIIcd forme une jonction en triple-hélices. Le module IIIf établit une interaction à distance de type pseudo-nœud. Enfin, le domaine IV se replie en une tige-boucle dans laquelle se trouve le codon d'initiation AUG. De manière intéressante, lorsque le ribosome est en interaction avec l'IRES, la structure du domaine IV change de conformation et se déplie pour libérer le codon d'initiation.

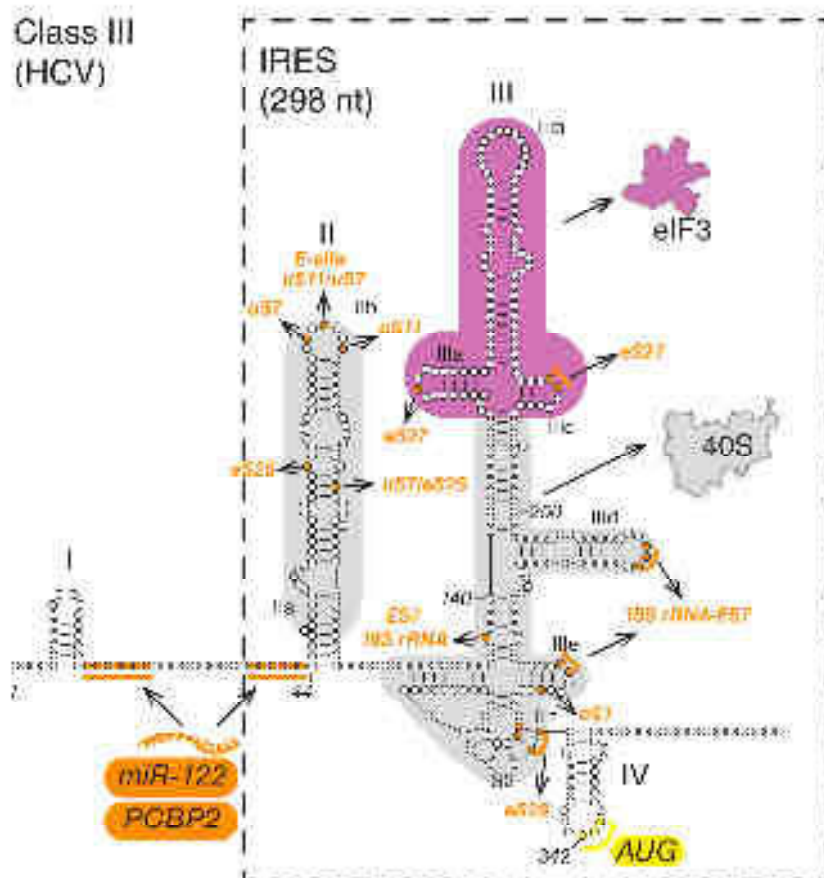


Figure 16 : Les IRES de classe III : exemple de l'IRES du virus de l'hépatite C (HCV). La figure, issue de Martin et Maillot, 2018, représente la structure secondaire de la région 5'UTR de l'ARN génomique du HCV. Les nucléotides essentiels pour l'activité de l'IRES ou pour le recrutement des ITAF sont indiqués en orange. Le domaine minimal de l'IRES du HCV est encadré en pointillés. Les domaines II, IIIc, IIIe et IIIf interagissent directement avec la petite sous-unité du ribosome 40S (en gris). La partie apicale du domaine III (surlignée en rose) interagit spécifiquement avec eIF3. Les contacts entre l'IRES et les protéines du ribosome sont indiqués par des flèches. Le codon initiateur AUG dans la boucle apicale du domaine IV est surligné en jaune.

II.2.3.2. Rôle fonctionnel du domaine I

Comme pour les IRES de classe I et II, le domaine qui précède l'IRES est également impliqué dans la réplication virale. De manière intéressante, le domaine I du HCV contient deux sites de liaison pour le micro-ARN miR-122 (**Figure 16**). Ce micro-ARN joue un rôle dans l'activation de la réplication du HCV, il stimule également la traduction virale (Henke et al., 2008; Jopling et al., 2005; Roberts et al., 2011; Sagan et al., 2015). Le mécanisme de stimulation de la traduction par le miR-122 reste largement méconnu. Des études suggèrent que la liaison de miR-122 et de la protéine Ago2 sur l'ARN pourrait favoriser le repliement de l'IRES dans une conformation optimale pour initier la traduction. D'autres études récentes

montrent que la protéine PCBP2 et miR-122 se fixent au même site de liaison sur le génome de HCV. Par conséquent, ces deux molécules sont en compétition. La fixation de l'une ou de l'autre ferait pencher la balance soit vers la traduction virale soit vers la réplication de l'ARN viral (Masaki et al., 2015).

II.2.3.3. Formation du complexe de pré-initiation 40S-IRES

Les domaines II à IV de la région 5'UTR du HCV constituent l'IRES proprement-dit du virus. De manière intéressante, les domaines II et III permettent d'assurer un recrutement direct de la petite sous-unité du ribosome 40S, grâce à de multiples interactions entre des nucléotides de ces domaines et des composants de la sous-unité ribosomique 40S. (**Figure 17**). Par exemple, des nucléotides des sous-domaines IIIa et IIIc interagissent avec les protéines ribosomiques eS7 et eS27. Le sous-domaine IIIe reconnaît spécifiquement la protéine eS1 (Laing & Schlick, 2009; Yamamoto et al., 2015). Le domaine II interagit avec les protéines uS7 (RpS5) (**Figure 17** en bleu foncé) et uS11 (en rouge) qui sont localisées dans le site E du ribosome (H. Yamamoto et al., 2015). Ainsi, le ribosome 40S interagit avec une grande affinité pour l'IRES ($K_d = 2-4$ nM) et le complexe HCV-40S formé est très stable ($k_{off} = 0,002$ s⁻¹ ; $k_{on} = 1,1$ μMs⁻¹) (Fuchs et al., 2015). Des études structurales ont également mis en évidence une interaction spécifique entre la boucle apicale du module IIIId (₂₆₆GGG₂₆₈) de l'IRES et le segment d'expansion ES7 de l'ARNr 18S (₁₁₁₆CCC₁₁₁₈) (Malygin et al., 2013; Matsuda & Mauro, 2014). Un autre contact entre le module IIIe de l'IRES et les nucléotides ₁₁₁₄UU₁₁₁₅ de ce même segment d'expansion ES7 de l'ARNr 18S a également été mis en évidence. En outre, les interactions avec la boucle IIIId sont essentielles, car elles provoquent des changements de conformation de l'IRES indispensables pour la traduction (Angulo et al., 2016). Enfin, le module IIIf, formant le pseudo-noeud, permet d'orienter l'IRES afin de positionner correctement le codon AUG dans le site P du ribosome (Berry et al., 2010; Quade et al., 2015; Yamamoto et al., 2015). Cette interaction permet au ribosome d'être directement déposé sur le codon d'initiation de la traduction, s'affranchissant ainsi de l'étape de 'scanning'.

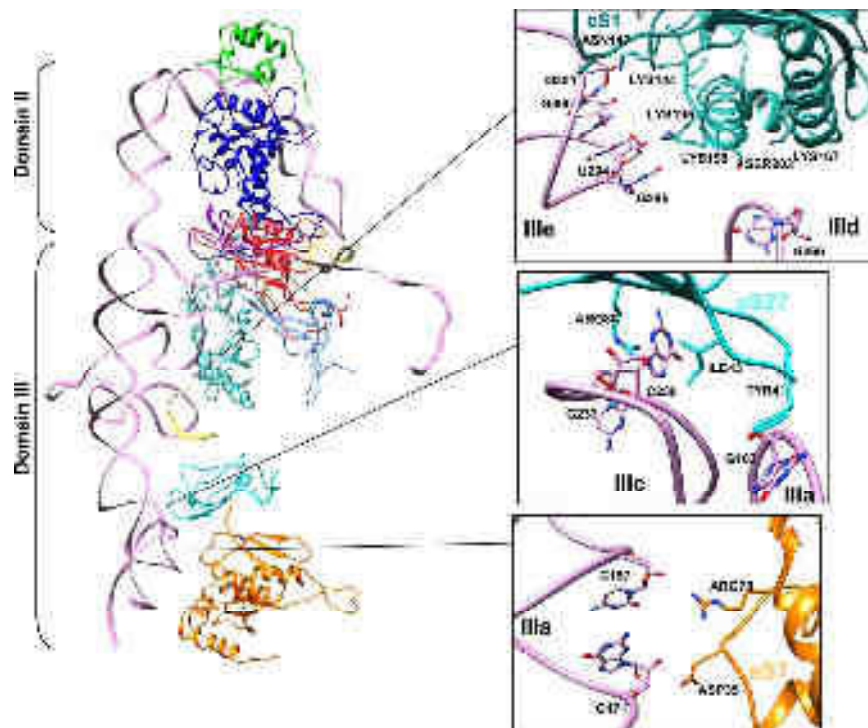


Figure 17 : La petite sous-unité du ribosome 40S interagit directement avec les domaines II et III de l'IRES du HCV. (Figure issue de Yamamoto *et al.*, 2015). Le modèle structural de l'IRES de HCV dans le ribosome 40S, obtenu par Cryo-Microscopie Électronique (Cryo-ME), met en évidence les interactions entre les domaines II et III de l'IRES et les protéines de la petite sous-unité du ribosome 40S (à gauche). L'IRES est représenté en rose et les protéines du ribosome sont indiquées selon les couleurs suivantes : de haut en bas : eS25 en vert, uS7 en bleu foncé, uS11 rouge, eS28 en violet, eS26 en cyan, eS1 en bleu canard, eS27 en turquoise et eS7 en orange. (À droite) Vues rapprochées sur 3 sites spécifiques d'interaction de l'IRES du HCV et des protéines ribosomales (eS1, eS27 et eS7) du ribosome 40S.

II.2.3.4. Formation du complexe de pré-initiation 48S-IRES

Après la fixation du ribosome 40S, l'IRES de HCV recrute spécifiquement le facteur d'initiation eIF3 et le complexe ternaire eIF2-GTP-Met-ARNt^{Met}_i pour former le complexe de pré-initiation HCV-48S. Des études ont démontré que le noyau octamérique du facteur eIF3 (sous-unités a, c, e, f, h, k, l et m) interagit avec le module IIIabc de l'IRES HCV avec une grande affinité (Sun *et al.*, 2011). Cette interaction entre l'IRES et eIF3 se fait essentiellement par des contacts entre l'IRES et les sous-unités eIF3a et eIF3c (Erzberger *et al.*, 2014; Hashem *et al.*, 2013; Sun *et al.*, 2013). Les sous-unités b, d, g, i et j de eIF3 ne sont pas requises pour l'interaction avec l'IRES, bien qu'elles soient essentielles pour initier la traduction de manière efficace. La sous-unité eIF3j joue également un rôle important dans la formation du complexe HCV-48S. Elle est initialement positionnée, sur eIF3, dans le canal qui va accueillir l'ARNm au sein de la particule 48S. Par conséquent eIF3j doit être éjectée

pour laisser la place à l'ARN viral et permettre une mise en place correcte du codon d'initiation AUG dans le site P du ribosome (Fraser et al., 2007). De plus, il a été démontré que eIF3j est absolument nécessaire à l'infection par HCV (Majzoub et al., 2014), cependant, les bases moléculaires de son mode d'action ne sont pas connues. Il en est de même pour la protéine ribosomale RACK1 qui est localisée à proximité de la sous-unité d du facteur eIF3 au sein de la particule 48S. Son mode d'action dans la formation du HCV-48S n'est pas bien défini, mais sa présence est elle aussi requise pour l'infection par ce virus (Majzoub et al., 2014).

II.2.3.5. Assemblage du ribosome 80S

L'initiation de la traduction s'achève par la formation du complexe d'initiation 80S sur l'IRES HCV. Ceci s'accompagne de réarrangements structuraux importants de l'IRES. Le domaine II, notamment, subit plusieurs changements conformationnels qui sont probablement importants pour le positionnement précis du Met-ARNt^{Met} dans le site P du ribosome (**Figure 18**). Ces réarrangements structuraux garantissent ainsi la transition du complexe HCV-48S vers le complexe HCV-80S (Yamamoto et al., 2015). D'autres études suggèrent que le domaine II est également requis pour l'hydrolyse du GTP présent sur facteur eIF2 au sein du complexe ternaire, étape conduisant à la formation du 48S. Il semble également que le domaine II soit nécessaire pour la dissociation des eIF du complexe 48S afin de permettre l'association de la sous-unité 60S du ribosome, étape induite par le facteur eIF5B (Locker et al., 2007; Otto & Puglisi, 2004).

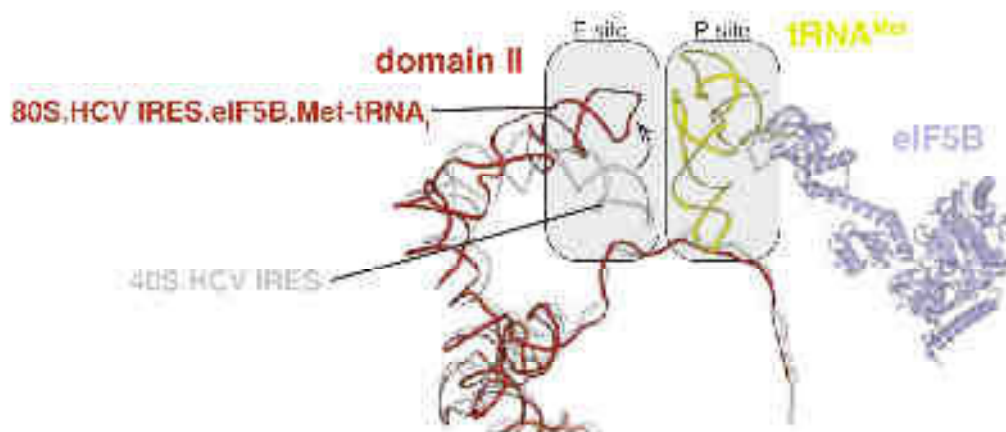


Figure 18 : Le mouvement du domaine II dans le site E du ribosome permet de positionner précisément le Met-ARNt^{Met} dans le site P du ribosome. Des études structurales par cryo-ME ont mis en évidence les changements conformationnels importants (indiqué par la flèche) du domaine II de l'IRES du HCV entre les stades 40S-HCV (en gris) et 80S-HCV (en rouge). Les mouvements du domaine II assurent le bon positionnement l'ARNt initiateur dans le site P du ribosome. Dans le complexe d'initiation 80S-HCV, le facteur d'initiation eIF5B (en bleu) et l'ARNt^{Met} initiateur (en jaune) sont également visibles.

En résumé, les IRES de classe III adoptent des repliements structuraux complexes. Certaines de leurs structures sont capables d'interagir directement avec la sous-unité 40S du ribosome. Cette classe d'IRES n'utilise qu'un nombre limité de facteurs d'initiation. En effet, les IRES de classe III interagissent directement avec eIF3 et utilisent eIF2, mais ils peuvent s'affranchir des autres eIF. De plus, l'initiation guidée par les IRES de classe III ne nécessite pas d'étape de 'scanning' pour former la particule 48S sur le codon d'initiation.

II.2.4. Les IRES de classe IV

II.2.4.1. Structure adoptée par les IGR des Dicistrovirus

Les IRES de la quatrième classe sont retrouvés chez les virus de l'ordre des *Picornavirales*. À ce jour, ce type d'IRES n'a été identifié que chez les virus de la famille des *Dicistroviridae*. Les IRES de classe IV sont également appelés 'Intergenic Region' (IGR) en raison de leur localisation intergénique entre les deux ORF du génome des Dicistrovirus. Les IGR sont les plus petits IRES connus à ce jour, ils comprennent environ 200 nucléotides. L'IRES IGR le plus étudié et le mieux caractérisé dans cette classe est celui du 'Cricket Paralysis Virus' (CrPV). Cependant, un certain nombre d'autres IGR ont également été étudiés, comme par exemple ceux de 'Israeli Acute Paralysis Virus' (IAPV), du 'Platia Stali Intestine Virus' (PSIV), du 'Taura Syndrome Virus' (TSV) et d'autres encore. La structure et la topologie de ces IRES IGR sont très conservées (Kerr & Jan, 2016). Ces IRES se replient en une structure compacte et très sophistiquée qui implique des interactions à distance et plusieurs pseudo-nœuds (Costantino & Kieft, 2005; Jan & Sarnow, 2002; Kanamori & Nakashima, 2001; Pflugsten et al., 2007). L'IRES IGR du CrPV contient trois domaines (domaines 1 à 3), contenant chacun un pseudo-nœud nommé PKI, PKII ou PKIII (Jan & Sarnow, 2002; Kanamori & Nakashima, 2001) (**Figure 19a**). Les domaines 1 et 2 sont essentiels pour la liaison du ribosome (**Figure 19b**) grâce à leur interaction directe entre les structures 'Stem-Loop IV', 'Stem-Loop V' (SLIV et SLV), 'Loop 1.1' (L1.1), et des protéines des sous-unités 40S et 60S du ribosome (Costantino & Kieft, 2005; Jan & Sarnow, 2002; Nishiyama et al., 2003; Pflugsten et al., 2006; Schüller et al., 2006). Le domaine 3, chez le CrPV, contient le pseudo-nœud PKI ainsi qu'une boucle variable riche en A, nommée 'Variable Region Loop' (VRL). Il existe deux types de domaines 3 dans cette classe d'IRES : des domaines 3 de type I avec une simple hélice en amont de PKI, comme dans le CrPV ; ou des domaines 3 de type II qui contiennent une tige-boucle supplémentaire SLIII, comme celui retrouvé dans l'IGR du TSV (voir l'encadré vert dans la **figure 19a**). Le pseudo-nœud PKI du domaine 3 adopte dans tous les IGR un repliement qui mime structuralement la structure du duplexe

codon-anticodon. Cette structure formée par le PKI est appelée motif 'Anticodon Stem-Loop-like' ou 'ASL-like'.

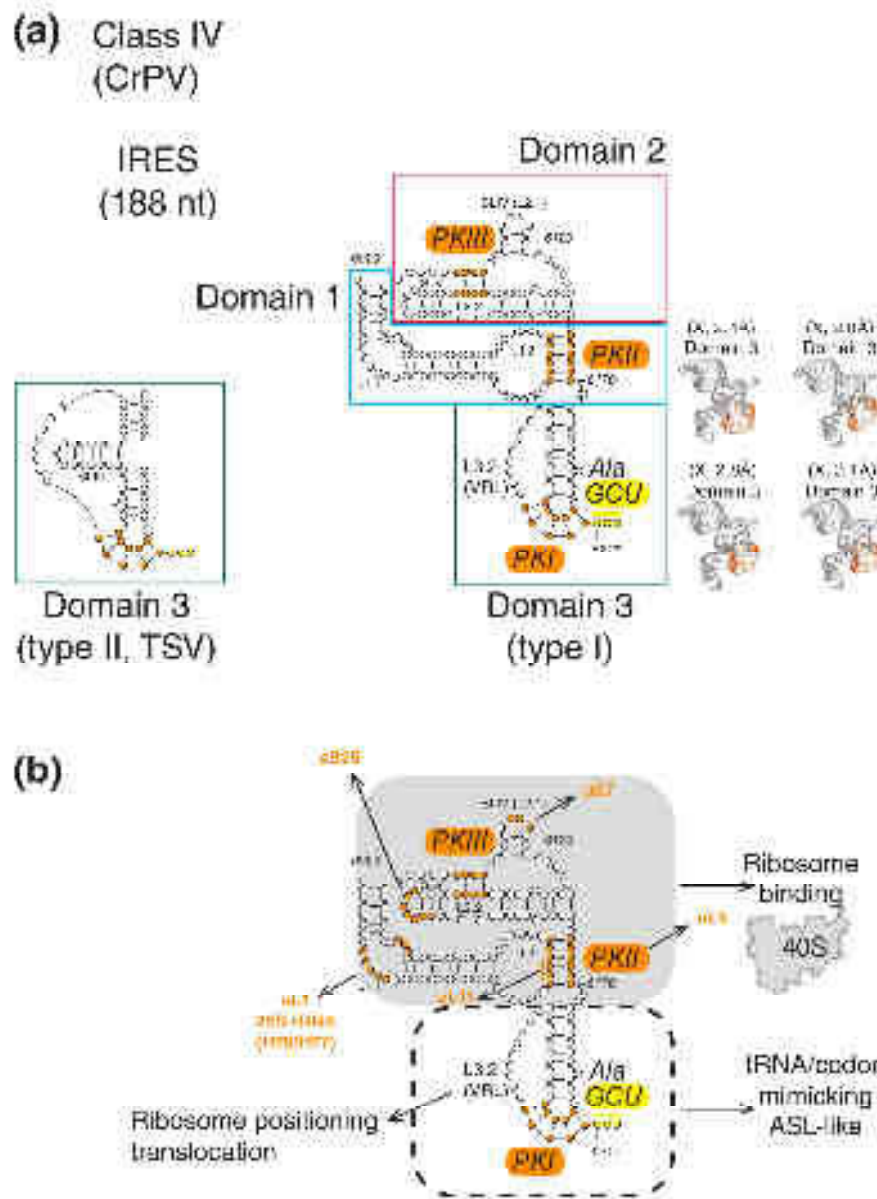


Figure 19 : L'IGR du CrPV est un IRES de classe IV. (Figure issue de Mailliot et Martin, 2018). (a) Structure secondaire de l'IRES de la région intergénique (IGR) du 'Cricket Paralysis Virus' (CrPV). L'IGR du CrPV initie la traduction sur un codon initiateur non AUG (surligné en jaune). Il possède trois domaines structuraux : le domaine 1 (encadré en bleu), le domaine 2 (encadré en rouge) et le domaine 3 (encadré en vert). Chaque domaine contient une interaction à distance de type pseudo-nœud (indiqués en orange). Parmi les IGR étudiés à ce jour, deux types de domaine 3 ont été observés : des domaines 3 de type I comme dans CrPV et des domaines de type II comme dans TSV. Les structures obtenues par cristallographie (X) des domaines et sous-domaines de l'IGR sont également présentées. (b) L'IGR initie la traduction en recrutant directement et uniquement la petite sous-unité ribosomique 40S. Les contacts spécifiques entre l'IGR du CrPV et les sous-unités ribosomales 40S et 60S sont indiquées par des flèches.

II.2.4.2. Mécanisme moléculaire de recrutement du ribosome par les IGR

L'initiation de la traduction sur l'IGR débute avec le recrutement de la petite sous-unité du ribosome 40S grâce aux domaines 1 et 2 de l'IRES (**Figure 19b**). Lorsque la sous-unité 40S du ribosome est fixée, le PKI du domaine 3, qui mime la structure du duplexe codon-anticodon, est placé dans le site A du ribosome. Puis, la sous-unité 60S du ribosome est assemblée. Ensuite, une étape de pseudo-translocation, guidée par le facteur d'élongation eEF2, permet de déplacer le PKI du site A vers le site P du ribosome (Hussain et al., 2014; Koh et al., 2014; Yamamoto et al., 2007). La flexibilité de la boucle variable VRL du domaine 3 est importante pour le bon positionnement du PKI et la pseudo-translocation (Au et al., 2015; Au & Jan, 2012; Ruehle et al., 2015) (**Figure 18b**). L'étape de pseudo-translocation permet de placer le codon d'initiation GCU dans le site A du ribosome et d'assurer, avec l'aide des facteurs d'élongation eEF1 et eEF2, l'entrée du premier aminoacyl-ARNt (l'alanyl-ARNt^{Ala}) dans le site A du ribosome (Hussain et al., 2014; Jan et al., 2003; Pestova & Hellen, 2003; Pestova et al., 2004; Yamamoto et al., 2007). Finalement, une seconde étape de translocation permet de déplacer le premier aminoacyl-ARNt du site A vers le site P, et le PKI du site P vers le site E. Ensuite, l'étape d'élongation se poursuit de manière conventionnelle.

Pour résumer, le mode d'action des IGR est très simple. Ces IRES sont capables d'initier la traduction sur un codon de démarrage qui n'est pas un AUG, sans aucun facteur d'initiation et sans le Met-ARNt^{Met}_i (Jan & Sarnow, 2002; Pestova & Hellen, 2003; Wilson et al., 2000; Pestova et al., 2004). Seules les sous-unités 40S et 60S du ribosome sont requises pour initier la traduction sur les IRES de classe IV. La structure adoptée par les IRES de cette classe est très compacte et très sophistiquée.

II.2.5. Les 'IRES Trans-Acting Factors' ou ITAF

Outre les facteurs d'initiation eIF, certains facteurs auxiliaires peuvent influencer l'initiation de la traduction guidée par les IRES. Ces facteurs cellulaires sont nommés ITAF pour 'IRES Trans-Acting Factor'. Leur action sur les IRES permet de moduler l'initiation de la traduction en la stimulant ou en la réprimant. De nombreux ITAF ont été identifiés lors des traductions virales. Ce sont des protéines généralement impliquées dans des mécanismes tels que la régulation de la transcription, l'épissage, le transport des ARN ou le contrôle de la traduction. Parmi ces protéines multifonctionnelles, les plus couramment recrutées par les IRES viraux sont les protéines PTB, PCBP2, le facteur d'épissage SRp20, l'auto-antigène La, la nucléoline ou encore la gémine 5 (pour revue Lee et al., 2017; Martínez-Salas et al., 2015).

Le **tableau 1** présente, à titre d'exemple, la liste des ITAF identifiés chez les *Picornavirales*. D'autres ITAF sont utilisés par les virus de la famille des *Flaviviridae* et des *Rétroviridae*. Les premiers ITAF qui ont été découverts sont PTB, utilisé par EMCV, et la protéine 'La autoantigène' (La), requise par le Poliovirus (pour revue Flather & Semler, 2015). La protéine PTB se fixe sur des séquences de l'ARNm riches en pyrimidines, elle est impliquée dans les mécanismes d'épissage et de transport des ARNm (Ghetti et al., 1992; Oh et al., 1998; Sawicka et al., 2008). PTB favorise la traduction guidée par les IRES viraux de FMDV et de EMCV. La protéine PTB stimule également la traduction guidée par certains IRES cellulaires comme celui du proto-oncogène c-myc (Cobbold et al., 2008). La protéine La, quant à elle, est un antigène reconnu par les anticorps de patients atteints de la maladie auto-immune appelé lupus. Elle stimule le recrutement du ribosome 40S sur les IRES du Poliovirus et de HCV (Ali & Siddiqui, 1997; Costa-mattioli et al., 2004; Meerovitch et al., 1993; Meerovitch et al., 1989). Les ITAF sont souvent capables de s'associer en plusieurs exemplaires, de manière à former des poymères d'ITAF. Ils favorisent ainsi la maturation des ARNm et les protègent contre les ribonucléases. De manière analogue, la fonction des ITAF sur les IRES est de stabiliser l'interaction entre l'IRES et le ribosome. Ainsi, les ITAF sont des facteurs auxiliaires qui agissent en coopération avec les eIF. Dans le cas de HCV, les ITAF exercent une fonction supplémentaire ; ils contribuent aux changements conformationnels du ribosome et de l'IRES, changements qui sont indispensables lors de l'initiation de la traduction virale (Niepmann, 2013).

Protéine	IRES	Effet	Référence
PTB	EMCV, EMCV, FMDV, PV, HCV, EMCV, HAV	Stimulation	Cohen et al. (2001), Jung and Wimmer (1990), Sawicka et al. (2008), Liu and Neck (2001), Verheijen et al. (2001), Chern et al. (2004)
ITAF1	PV, HCV	Stimulation	
ITAF2	PV, HCV, CBV3, EV71, BEV, HAV, EMCV, FMDV#	Stimulation#	Gomonov et al. (2002), Sawicka (2008), Wimmer et al. (1990), Gruff et al. (1988)
ITAF41	EV71, HCV	Stimulation	Wong et al. (2011)
SR200	PV	Stimulation	Bednarek et al. (2007)
ITAF42	EMCV, EMCV#	Stimulation#	Wang et al. (2007), Chapman et al. (1991)
Lim	PV, HCV	Stimulation	Massella et al. (2007), Smith et al. (1991), Wang and Semler (1988)
ITAF45	PV	Stimulation	Smith and Jung (2009), Ray and Das (2002), Wang et al. (2011)
La	PV, CBV3, BEV, EMCV, HAV	Stimulation	Ali et al. (1997)
ITF1	EV71	Stimulation	Wang et al. (2011)
ADP1	EV71	Repression	Chapman et al. (2001)
ITAF46	EV71	Repression	Chen et al. (2011)
ITAF47	EMCV	Inverse régulation	Chen et al. (2011)
ITAF48/ITAF49	HCV	Repression	Ali et al. (1997)
GARS	PV	Stimulation	Adrian et al. (2012)

Tableau 1: Liste des 'IRES Trans-Acting Factor' (ITAF) associés aux IRES des *Picornavirales*. (Tableau issu de Martinez-Salas *et al.*, 2015). Les colonnes du tableau indiquent respectivement : le nom de l'ITAF, l'IRES viral sur lequel il a été identifié, l'effet induit par l'ITAF sur la traduction guidée par l'IRES et l'article dans lequel a été décrit l'ITAF. Le signe # indique les ITAF n'entraînant aucun effet sur la stimulation ou la répression de la traduction de l'IRES sur lequel ils se fixent. Par exemple, l'ITAF₄₅ stimule l'IRES du FMDV mais ne stimule pas la traduction guidée par l'IRES EMCV.

III. L'initiation « IRES-dépendante » de la traduction chez le CrPV

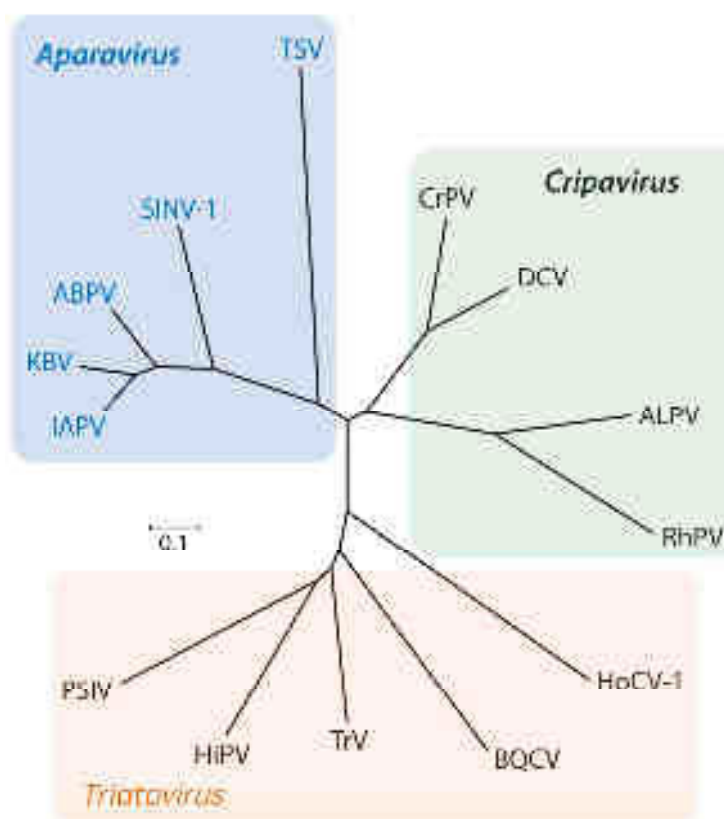
III.1. Généralités sur les *Dicistroviridae*

La famille des *Dicistroviridae* est constituée de petits virus possédant un génome à ARN de polarité positive qui infectent les arthropodes parmi lesquels les abeilles, les criquets ou les crevettes. Actuellement, notre compréhension de la biologie de ces virus est encore incomplète. De nombreuses études ont permis de comprendre la structure des particules virales ainsi que les mécanismes d'initiation de la traduction guidée par l'IRES IGR. En revanche, les mécanismes de réplication de l'ARN viral, la fonction des gènes viraux ou l'entrée du virus dans les cellules hôtes représentent encore de nombreux champs d'investigation. Ces virus sont également d'excellents modèles utilisés pour l'étude du système immunitaire des insectes.

III.1.1. Taxonomie et spectre d'hôtes

La famille des *Dicistroviridae* est classée dans l'ordre des *Picornavirales*. En effet, l'organisation du génome et la structure des virions sont homologues à celles des *Picornaviridae*, une autre famille bien caractérisée des *Picornavirales* (Bonning & Miller, 2010; Le Gall et al., 2008). C'est pourquoi les Dicistrovirus sont également connus sous le nom de 'picorna-like viruses'. À ce jour, les *Dicistroviridae* se décomposent en trois genres de virus : les *Aparavirus*, les *Cripavirus* et les *Triatovirus* (Valles et al., 2017) (**Figure 20**). Ces trois genres se distinguent par des divergences phylogénétiques et par la structure du domaine 3 de l'IRES intergénique (voir *paragraphe II.2.4.1*). Les quinze représentants aujourd'hui identifiés de la famille des *Dicistroviridae* ont été isolés d'insectes tels que les Diptères, les Hémiptères, les Hyménoptères, les Lépidoptères, les Orthoptères, mais aussi des crustacés notamment des décapodes (Bonning & Miller, 2010).

Figure 20 ci-contre: Arbre phylogénétique de la famille des *Dicistroviridae*. (Adaptée de Bonning and Miller 2010). L'arbre phylogénétique a été généré à partir de l'alignement des séquences complètes des génomes des *Dicistroviridae*. La famille des *Dicistroviridae* comprend trois ordres : l'ordre des *Aparavirus* (en bleu) qui contient 5 espèces dont trois infectant les abeilles, l'ordre des *Cripavirus* (en vert) contenant 4 espèces dont le CrPV et l'ordre des *Triatovirus* (en orange) qui regroupe 5 espèces dont le 'Triatoma virus' servant de référence à cet ordre. Les virus représentés sont: 'Acute Bee Paralysis Virus' (ABPV), 'Taura Syndrome Virus' (TSV), 'Kashmir Bee Virus' (KBV), 'Solenopsis Invicta Virus-1' (SINV-1), 'Israeli Acute Paralysis Virus' (IAPV), 'Cricket Paralysis Virus' (CrPV), 'Drosophila C Virus' (DCV), 'Aphid Lethal Paralysis Virus' (ALPV), 'Rhopalosiphum Padi Virus' (RhPV), 'Black Queen Cell Virus' BQCV, 'Himetobi P Virus' (HiPV), 'Plautia Stali Intestine Virus' (PSIV), 'Triatoma Virus' (TrV), 'Homalodisca Coagulata Virus-1' (HoCV-1). La longueur des branches est inversement proportionnelle au degré d'homologie génétique entre les différentes espèces virales.



III.1.2. Structure du virion des *Dicistrovirus*

Les virions des *Dicistroviridae* sont sphériques, non enveloppés et d'un diamètre d'environ 30 nm. La capside virale présente une symétrie icosaédrique de type pseudo-T=3. Les capsides icosaédriques contiennent un minimum de 60 sous-unités protéiques, les plus grandes en contiennent un multiple de 60 afin de respecter la symétrie de l'icosaèdre. Le nombre de triangulation « T » permet de calculer le nombre de sous-unités protéiques constituant la capside. Par exemple, les virus possédant une capside icosaédrique de type T=3, sont constitués de $60 \times 3 = 180$ sous-unités protéiques (nombre minimal de sous-unités protéiques \times T). La capside des *Dicistrovirus* est qualifiée de « pseudo »-T=3 parce que, contrairement aux capsides virales de type T=3 qui sont constitués de 180 copies d'une seule protéine virale, elle contient 60 copies (protomères) d'un assemblage des trois protéines virales de capside. Les protéines de capside sont nommées 'Viral Protein 1, 2, 3' (VP1, VP2, VP3) (**Figure 21**). Le virion contient une quatrième protéine de capside, la protéine VP4, qui est plus petite et qui est en contact avec le génome au sein de la particule virale (Mullapudi et al., 2016; Sánchez-Eugenia et al., 2015; Valles et al., 2017).

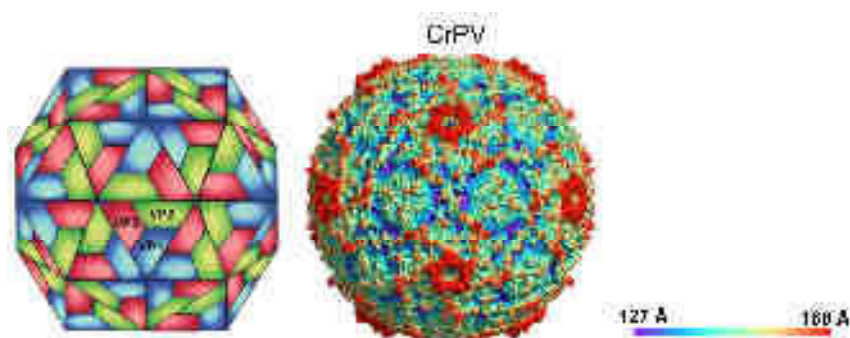


Figure 21 : La structure des virions des *Dicistroviridae*. (Figure issue de Valles *et al.*, 2017). **(À gauche)** Représentation schématique d'une particule virale de Dicistrovirus qui illustre l'organisation des 60 protomères constitués chacun d'une copie de chaque protéine de capsid virale (VP1, VP2 et VP3). **(À droite)** Structure par Cryo-ME du virion du CrPV. L'échelle de couleur indique la distance entre les protéines de capsides et le centre du virion.

III.1.3. Organisation génomique et fonction des protéines virales

Le génome des *Dicistroviridae* est une molécule d'ARN simple-brin de polarité positive, linéaire et monopartite. Sa particularité est qu'il contient deux phases codantes principales (ORF1 et ORF2), non chevauchantes. Chaque ORF est précédée par un IRES qui contrôle leur traduction, l'IRES_{5'UTR} pour l'ORF1 et l'IGR pour l'ORF2 (**Figure 22**). L'ORF1 code pour une polyprotéine précurseur des protéines non structurales. Les protéines non structurales sont : une hélicase, une protéase et une ARN-polymérase ARN-dépendante ou RpRd. En outre, l'ORF1 code pour la 'genome-linked Viral Protein' (VPg) qui se lie de manière covalente à l'extrémité 5' du génome viral. Cette dernière protège l'extrémité 5' du génome viral dans les virions. L'ORF1 code également pour une protéine suppresseur de l'ARN interférence chez le DCV (Van Rij *et al.*, 2006) et le CrPV (Besnard-Guérin *et al.*, 2015; Garrey *et al.*, 2010; Wang *et al.*, 2006). L'ORF2 code pour une polyprotéine qui est le précurseur des protéines structurales de capsid VP0, VP1 et VP2. La protéine VP0 est clivée ensuite en protéines de capsid VP3 et VP4. L'extrémité 3' du génome viral est protégée par une queue poly(A).

Selon la nomenclature officielle, les polyprotéines des *Picornaviridae* et des *Dicistroviridae* sont classées en trois régions protéiques nommées P1, P2 et P3 (**Figure 22**, pour les *Dicistroviridae*). La région P1 correspond à la polyprotéine de l'ORF2, et les régions protéiques P2 et P3 combinées correspondent à la polyprotéine de l'ORF1. La région P2 code pour des protéines non structurales nommées 2A, 2B et 2C (2C étant l'hélicase virale).

Et la région P3 code les protéines non structurales 3A, 3B, 3C (3C étant la protéase) et 3D (pour la réplicase RpRd).

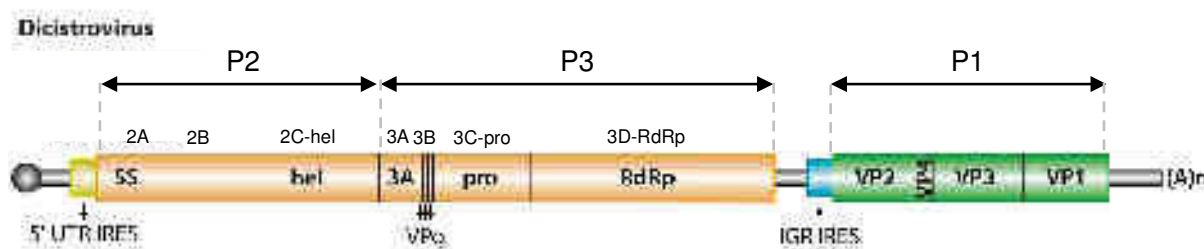


Figure 22 : Organisation du génome des Dicistrovirus. (Adaptée de Bonning and Miller, 2010). Le génome des Dicistrovirus est un ARN simple brin bicistronique. La phase codante 1 (ORF1) représentée en orange code pour le précurseur des protéines non structurales du virus. On y retrouve dans l'ordre les protéines suivantes : 'Silencing Suppressor', SS (spécifique au CrPV et au DCV) ; une hélicase, hel ; la protéine 3A de la région P3 des protéines non structurales, 3A ; la protéine VPg, 'genome-linked protein' répété entre 1 et 6 copies selon l'espèce ; une protéase, pro ; une ARN polymérase ARN dépendante, RdRp. Les sites de clivage protéolytique du précurseur sont indiqués par des lignes verticales. L'ORF2 (en vert) code pour les précurseurs des protéines structurales de capsid VP1, 2, 3 et 4. Chacune des ORF est précédée d'un IRES (IRES_{5'UTR} et IGR) permettant d'initier la traduction des polyprotéines virales. À l'extrémité 5' du génome est associée une VPg (sphère à gauche) et l'extrémité 3' du génome est polyadénylée. Les polyprotéines codées par le génome sont nommées P1 (polyprotéine de l'ORF2) et P2-P3 (polyprotéine de l'ORF1). Les régions P1, P2 et P3 sont subdivisées en sous-régions (2A, 2B, 2C-hélicase, 3A, 3B, 3C-protéase, 3D-RdRp) correspondant aux produits de maturation de la polyprotéine.

III.1.4. Réplication

À ce jour, l'étape de réplication chez les *Dicistroviridae* n'est pas encore bien connue. En raison de leur ressemblance avec les *Picornaviridae*, il a été spéculé que les *Dicistroviridae* utiliseraient les mêmes mécanismes de réplication. Chez les *Picornavirus*, la synthèse de l'ARN débute par la liaison d'un complexe protéique appelé 3CD constituée par la protéase et la RpRd (voir **figure 22**) sur la structure en feuille de trèfle ou 'cloverleaf' (CL) localisée du côté proximal de l'extrémité 5' du génome (voir *paragraphes II.2.1 et II.2.2*). Le complexe 3CD ainsi positionné en 5' du génome interagit avec les PABP, permettant ainsi de circulariser l'ARN génomique. C'est alors que la synthèse des brins d'ARN complémentaires négatifs commence. En outre, il a également été démontré que l'initiation de la transcription du génome du Poliovirus est amorcée grâce à la protéine VPg liée à l'extrémité 5' de l'ARN (Van Boom et al., 1998). Les mécanismes de réplication des *Dicistroviridae* ne sont pas encore bien caractérisés.

III.2. La traduction virale du CrPV, un modèle d'étude de choix

III.2.1. Les avantages du CrPV

Le CrPV appartient au genre des *Cripavirus*, il est le virus de ce genre qui a été le plus étudié. Il sert actuellement de référence pour l'étude des virus de la même famille. Le CrPV possède le plus grand spectre d'hôtes observé chez les virus qui infectent les invertébrés. En effet, le CrPV est capable d'infecter au moins vingt-quatre espèces différentes d'insectes (Bonning & Miller, 2010). Cette caractéristique fait du CrPV un modèle d'étude de choix. Autre particularité intéressante, le CrPV se réplique également de façon efficace dans des cellules d'insectes en culture (Scotti, 1975; Scotti, Hoefakker, & Dearing, 1996), ce qui facilite considérablement son étude. Plusieurs outils précieux, dont un clone infectieux du CrPV ainsi qu'un système rapporteur bicistronique qui mime le génome du virus (Kerr et al., 2015), ont été construits il y a peu. Ces nouveaux outils offrent des perspectives expérimentales de choix pour le décryptage des mécanismes d'action du CrPV à l'échelle moléculaire et cellulaire.

III.2.2. L'inhibition de la traduction de l'hôte par le CrPV

L'infection des cellules d'insectes par le CrPV conduit à une inhibition de la traduction « coiffe-dépendante ». Les deux étapes clef de la régulation de l'initiation de la traduction canonique sont ciblées par le virus pendant l'infection. Premièrement, au tout début de l'infection, la formation du complexe eIF4F est altérée suite à la dissociation des facteurs eIF4G et eIF4E (Garrey et al., 2010). Le mécanisme de dissociation des facteurs eIF4G et eIF4E est encore mal compris. Il ne s'agit cependant pas d'un clivage protéolytique du facteur eIF4G par une protéase virale. Cette étape précoce permet au virus d'inhiber la traduction de l'hôte dès son entrée dans la cellule. Cette inhibition initiale de la traduction « coiffe-dépendante » permet surtout au virus de réquisitionner les ribosomes et les eIF, à présent disponibles, pour la traduction virale guidée par ses IRES. Deuxièmement, après trois heures d'infection, la sous-unité α du facteur eIF2 est phosphorylée (Garrey et al., 2010) de telle sorte à empêcher le recyclage du GDP associé au facteur. Ainsi, eIF2-GDP est dans un état inactif. La formation de nouveaux complexes d'initiation 43S est de ce fait réduite dans les cellules et la quantité des sous-unités 40S et 60S libres augmente. La kinase responsable de la phosphorylation de eIF2 α n'a pas encore été identifiée, mais nous savons qu'il ne s'agit pas des kinases PERK ou GCN2 induites par le stress du réticulum endoplasmique. Ainsi, la traduction canonique « coiffe-dépendante » de l'hôte est inhibée en phase tardive d'infection, alors que la production des protéines virales se poursuit.

Le virus CrPV utilise ainsi de multiples mécanismes pour inhiber la traduction de l'hôte et pirater sa machinerie traductionnelle pour la traduction exclusive des protéines non structurales et structurales. Il utilise deux mécanismes d'inhibition de la traduction « coiffe-dépendante » différents, qui sont activés à des moments distincts du cycle viral. Ces phénomènes reflètent vraisemblablement les modes d'action différents des deux IRES du CrPV vis-à-vis de certains facteurs spécifiques nécessaires à leur traduction.

III.2.3. L'initiation de la traduction dirigée par les deux IRES du CrPV

Le CrPV constitue un excellent modèle d'étude de la traduction « IRES-dépendante » en raison de la présence de ses deux IRES (**Figure 23**). L'IRES IGR localisé dans la région intergénique du virus, qui contrôle la synthèse des protéines structurales, a fait l'objet de nombreuses recherches ces deux dernières décennies. Il est actuellement très bien caractérisé, tant au niveau structural que fonctionnel. Il a notamment permis de comprendre les bases moléculaires du mode d'action des IRES de la classe IV (décrit dans le *paragraphe II.2.4*).

En revanche, l'IRES localisé dans la région 5'UTR, en amont de l'ORF1, n'était pas du tout caractérisé au début de mon doctorat. Cet IRES a été appelé IRES_{5'UTR}. Contrairement aux séquences nucléotidiques et aux structures secondaires et tertiaires des IGR qui sont conservées au sein des *Dicistroviridae*, les IRES_{5'UTR} au sein de la famille sont très hétérogènes. La diversité des IRES_{5'UTR} chez les Dicistrovirus pourrait ainsi refléter l'évolution des différentes stratégies d'adaptation des virus à leurs hôtes respectifs pour détourner leurs ribosomes.

L'existence d'un IRES dans la région 5'UTR du CrPV a été proposée en 2000 par Wilson et collaborateurs grâce à l'utilisation de systèmes rapporteurs « bicistroniques » ; c'est-à-dire contenant deux ORF codant pour des protéines rapotrices sur le même ARN messager (Wilson et al., 2000). Plus récemment, des études ont démontré que l'initiation de la traduction de l'ORF1 du CrPV est dépendante de la présence sur le ribosome de la protéine RACK1 (Majzoub et al., 2014). Au contraire des autres ARNm de la cellule qui ne nécessitent pas la protéine RACK1 sur le ribosome pour initier efficacement leur traduction « coiffe-dépendante ». D'autres travaux ont montré que la sous-unité j du facteur d'initiation eIF3 est importante pour la traduction guidée par l'IRES_{5'UTR} du CrPV (Majzoub et al., 2014). Enfin, le facteur d'initiation de la traduction eIF2-GTP semble également important pour la traduction guidée par l'IRES_{5'UTR}. Les auteurs Garrey *et al.*, ont prouvé qu'une phosphorylation de eIF2 α en phase précoce d'infection réduisait la synthèse des protéines virales et la réplication du CrPV.

Génome du Virus de la Paralyse du Criquet (CrPV)

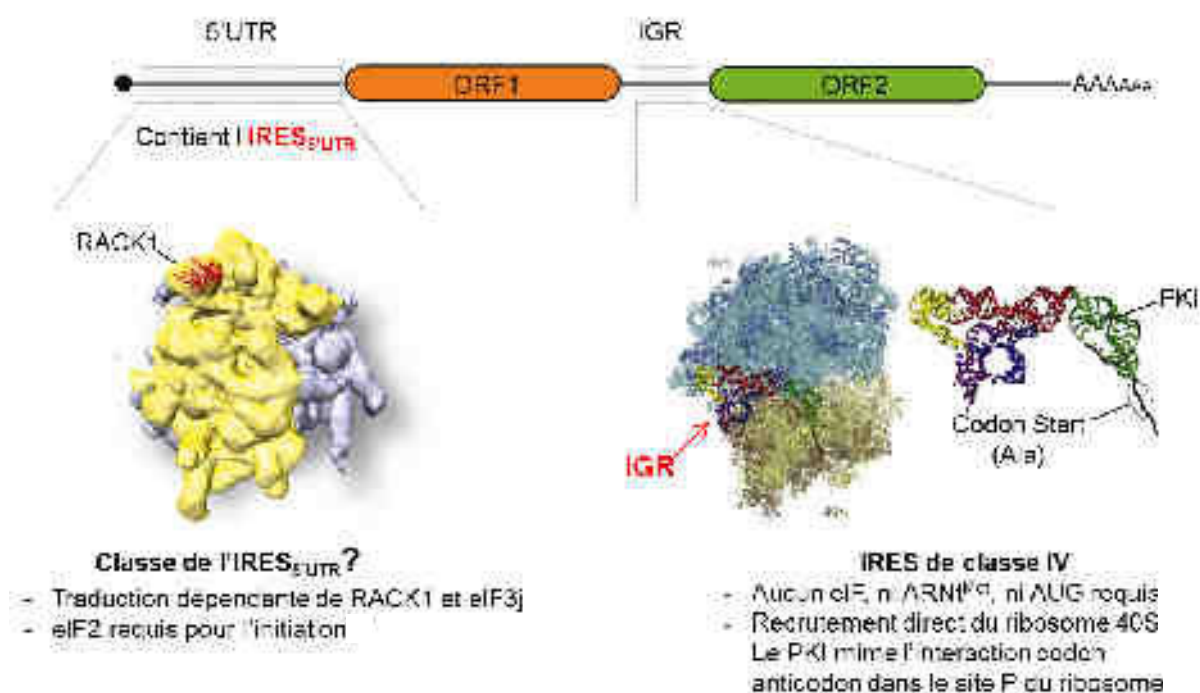


Figure 23 : Résumé des caractéristiques des deux IRES CrPV connus avant mon doctorat. (La figure du ribosome 80S avec la protéine RACK1 en rouge est issue de Majzoub *et al.*, 2014 et le modèle structural de l'IGR dans le ribosome 80S est issu de Fernandez *et al.*, 2015). L'IRES de la région intergénique du CrPV (IGR) était très bien caractérisé, il appartient à la classe IV des IRES (à droite). En revanche, l'IRES de la région 5'UTR du CrPV, nommé IRES_{5'UTR} était très peu connu. La traduction guidée par cet IRES nécessite la présence de la protéine ribosomique RACK1 dans le ribosome ainsi que le sous-unité eIF3j au sein du facteur multifactoriel eIF3 (à gauche).

III.2.4. Les points communs entre l'IRES_{5'UTR} du CrPV et l'IRES du HCV

Le HCV est un membre de la famille des *Flaviviridae*, il appartient au genre des *Hepacivirus*. Son génome à ARN simple brin de polarité positive est constitué d'une seule longue ORF qui code pour un précurseur polyprotéique unique. Ce précurseur protéique est ensuite maturée pour former les protéines structurales et non structurales nécessaires à la propagation du virus. L'ORF du HCV est flanquée de régions 5'UTR et 3'UTR formant des structures secondaires et tertiaires importantes lors de la réplication et de la traduction virale. La région 5'UTR du HCV contient l'IRES du HCV, décrite dans le *paragraphe II.2.3* de ce manuscrit. Cet IRES appartient à la classe III des IRES. De manière intéressante, l'IRES du HCV nécessite lui aussi les protéines RACK et eIF3j pour guider la synthèse des protéines virales. Ainsi, les IRES du CrPV et du HCV semblent apparentés, leur mode de fonctionnement partage notamment ces deux particularités fonctionnelles.

III.2.5. La régulation différentielle des deux IRES du CrPV

De manière intéressante, des études récentes ont montré que les deux IRES du CrPV ne sont pas utilisés par le virus de la même manière pendant le cycle infectieux. L'IRES_{5'UTR} est exprimé de manière constitutive pendant tout le cycle infectieux jusqu'à la mort de la cellule hôte. En revanche, l'IGR, est éteint au début de l'infection, puis, est activé durant la phase tardive de l'infection afin de produire de grandes quantités de protéines structurales à partir de l'ORF2 (Khong et al., 2016). Les bases fonctionnelles et structurales de cette expression différentielle des deux IRES ne sont pas encore connues. Deux hypothèses ont été avancées par les auteurs de ces études. Ils ont proposés que l'IGR serait bloqué en début de cycle par une protéine cellulaire ou qu'il serait activé en phase tardive par une protéine virale produite pendant la phase précoce du processus infectieux. Une autre possibilité serait que la machinerie traductionnelle de la cellule hôte, endommagée ou modifiée par les protéases virales durant l'infection précoce, soit désormais adaptée à l'IGR et donc utilisée par celui-ci uniquement pendant la phase tardive. L'ensemble des données énoncées dans ce *chapitre III.2* indique que les deux IRES du CrPV fonctionnent différemment. Ces IRES agissent néanmoins en synergie lors de l'infection virale pour assurer une traduction efficace des protéines virales.

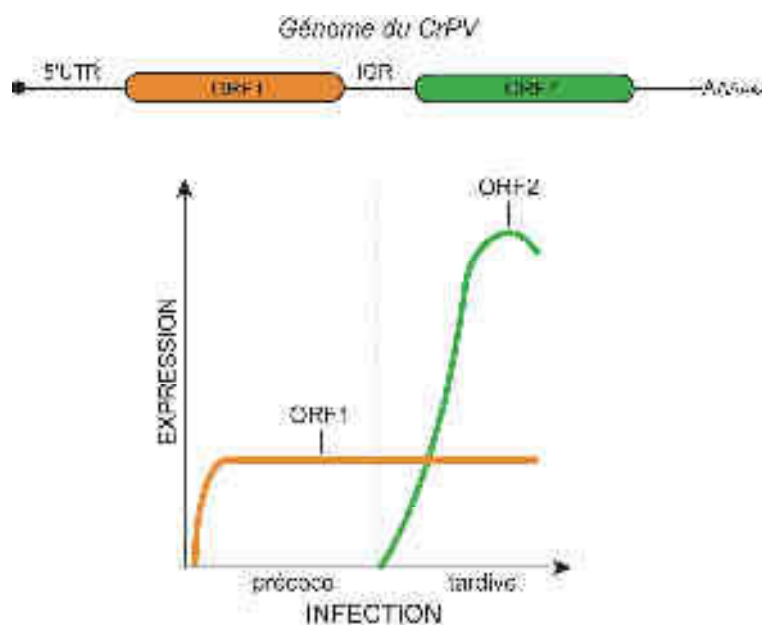


Figure 24 : La régulation différentielle des deux ORF du CrPV lors de l'infection. L'ORF1 (en orange) dont l'initiation de la traduction est guidée par l'IRES_{5'UTR} est exprimée tout au long du processus d'infection virale de manière constante. L'ORF2 (en vert) sous la dépendance de l'IGR, en revanche, ne s'exprime qu'en phase tardive du cycle viral. Les protéines structurales sont alors produites en grandes quantités à partir de l'ORF2.

IV. Objectifs et résumé de mon projet de doctorat

L'objectif principal de mon doctorat était de caractériser d'un point de vue fonctionnel et structural l'IRES_{5'UTR}. Dans un premier temps, je me suis attachée à établir un modèle de la structure secondaire de la région 5'UTR du CrPV. Par la suite, j'ai localisé précisément les bornes en 5' et en 3' de l'IRES_{5'UTR}. Par des expériences à l'aide de système de traduction *in vitro* acellulaires, nous avons étudié le mode d'action de l'IRES_{5'UTR}. Afin d'aller plus loin dans la compréhension de ce mode d'action, j'ai entrepris d'identifier les facteurs agissant *in trans* nécessaires à l'activité traductionnelle de cet IRES. De cette manière, nous avons démontré que l'IRES_{5'UTR} recrute spécifiquement le facteur eIF3. Nous avons également caractérisé le mode d'initiation de la traduction guidée par l'IRES_{5'UTR}, le ribosome est déposé directement sur le codon AUG initiateur sans passer par une étape de 'scanning'. Les données ainsi collectées sur l'IRES_{5'UTR} du CrPV ont permis de le ranger dans la classe III des IRES. Ces résultats sont présentés dans le *chapitre I*. De manière intéressante, nos résultats montrent que l'IRES_{5'UTR} partage des caractéristiques structurales et fonctionnelles avec l'IRES de HCV. Ainsi, nos travaux viennent corroborer les études antérieures suggérant que ces deux IRES possèdent de fortes similarités fonctionnelles, notamment concernant leur dépendance à la protéine ribosomale RACK1 lors de la traduction. Pour aller plus loin dans la caractérisation du mécanisme moléculaire d'action de l'IRES_{5'UTR}, j'ai étudié ses domaines structuraux indépendamment. L'objectif était de déterminer quels sont les domaines responsables du recrutement des facteurs d'initiation et du ribosome. Nous avons constaté que le domaine III recrute spécifiquement le facteur eIF3 et que le domaine II est impliqué dans la formation du complexe de pré-initiation 48S. Ces deux domaines semblent donc avoir des fonctions distinctes mais ils doivent être liés physiquement pour garantir l'assemblage efficace du complexe d'initiation. Ainsi, nos études ont montré que les différents domaines structuraux de l'IRES_{5'UTR} agissent en synergie pour recruter le ribosome. L'ensemble de ces résultats est consigné dans le *chapitre II* de ce manuscrit.

Le deuxième objectif principal de mon projet de doctorat a été de comprendre la synergie des deux IRES de CrPV pendant le cycle infectieux. Nous avons voulu comprendre les expressions différentielles des deux ORF de CrPV lors de l'infection grâce à l'action différée des deux IRES. L'objectif était de déterminer les bases moléculaires de ce phénomène afin de mieux comprendre l'évolution de la traduction virale guidée par l'IRES_{5'UTR} et l'IGR au cours du cycle infectieux. Dans le *chapitre III* de ce manuscrit nous détaillons les moyens mis en œuvre ainsi que les résultats obtenus lors de l'identification des partenaires moléculaires cellulaires et viraux des IRES_{5'UTR} et IGR, et ce à plusieurs étapes du processus d'infection des cellules hôtes par le CrPV.

RÉSULTATS

RÉSULTATS

I. La 5'UTR du CrPV contient une IRES de classe III – Étude du mode d'action de l'IRES_{5'UTR} de CrPV

I.1. Résumé de la publication 1

L'objectif premier de mon doctorat était de caractériser d'un point de vue structural et fonctionnel l'IRES localisé en 5'UTR du génome de CrPV (IRES_{5'UTR}) qui guide la traduction de l'ORF1. L'IRES_{5'UTR} a été peu étudié et son mode de fonctionnement était encore mal connu il y a peu. C'est pourquoi je me suis attelée à la caractérisation fine de cet IRES_{5'UTR}. Les contours de l'IRES_{5'UTR} ont été cartographiés de manière précise, sa structure secondaire a ensuite été déterminée, puis ses partenaires protéiques nécessaires pour l'initiation de la traduction des protéines non structurales du CrPV ont été caractérisés. Les résultats de ces travaux ont été compilés dans un article publié dans le journal *Nucleic Acids Research* en 2017 (Gross et al., 2017).

I.1.1. L'IRES_{5'UTR} est localisé entre les nucléotides 357 et 709 de la région 5'UTR du génome de CrPV

J'ai tout d'abord entrepris de cartographier l'extrémité 5'UTR de l'ARN génomique du CrPV afin d'y localiser précisément l'IRES_{5'UTR}. Dans un premier temps, ce travail a nécessité la mise au point d'un système de traduction acellulaire *in vitro*. Pour ce faire, des extraits de cellules embryonnaires de drosophile S2 ont été préparés. Ces extraits ont permis de suivre la traduction d'ARN messagers rapporteurs lors de tests de traduction *in vitro*. Afin de localiser précisément les contours de l'IRES_{5'UTR}, la totalité de l'extrémité 5'UTR du CrPV, ou bien des sous-fragments de celle-ci, ont été insérés en amont de la phase codante de la luciférase Rénilla. Pour chacune de ces constructions, l'activité luciférase guidée par l'IRES_{5'UTR} a été mesurée après traduction *in vitro* dans les extraits de cellules de drosophile S2. A l'aide de délétions grossières, puis grâce à des délétions de plus en plus précises, j'ai réussi à déterminer les bornes de l'IRES_{5'UTR}. L'IRES_{5'UTR} est précisément localisé entre les nucléotides 357 et 709, en amont du codon d'initiation AUG en position 709.

1.1.2. L'IRES contient un pseudo-nœud nécessaire à son activité traductionnelle

Nous avons ensuite déterminé la structure secondaire de la région 5'UTR de CrPV (1-709) en utilisant la méthode de 'Selective 2' Hydroxyl acylation Analysed by Primer Extension' (SHAPE). Cette technique consiste à modifier les riboses de chaque nucléotide de l'ARN non apparié, à l'aide d'un réactif chimique (Mortimer & Weeks, 2007; Mortimer & Weeks, 2009). La position de la modification est ensuite détectée par extension d'une amorce fluorescente. La réactivité du ribose avec le réactif renseigne sur l'implication de la base nucléotidique dans des éléments de structure secondaire et/ou tertiaire. Ainsi les nucléotides dans les régions simple brin sont réactifs, alors que les nucléotides appariés ne le sont pas, ou peu. Ces expériences ont permis d'établir un modèle de la structure secondaire de la région 5'UTR de l'ARN génomique du CrPV. Puis, j'ai affiné ces résultats par la détermination de la structure secondaire de l'IRES_{5'UTR} (357-709) par sondage chimique avec d'autres réactifs chimiques. Le DMS (dimethyl sulfate) et le CMCT (1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate). Contrairement aux sondes SHAPE qui modifient la position 2' du ribose des nucléotides, les sondes DMS et CMCT présentent l'avantage de réagir spécifiquement avec les bases. En effet, le DMS modifie uniquement et spécifiquement les bases C et A et le CMCT modifie les bases G et U. Les résultats de ces expériences de sondage chimique ont permis de construire un modèle fiable de l'IRES_{5'UTR} de CrPV.

Le modèle global de la région 5'UTR du CrPV se décompose en trois domaines, appelés respectivement du 5' vers le 3', domaine I, II et III. Le domaine I (1-263) n'est pas nécessaire à l'activité de l'IRES_{5'UTR}. En revanche, il contient des éléments structuraux intéressants. Des régions structurées en amont d'un IRES ont déjà été décrites dans les régions 5'UTR d'autres virus. Ces structures seraient impliquées dans la réplication de ces virus (Andino et al, 1993; Andino et al., 1990). Les deux domaines suivant, les domaines II (357-466) et III (505-709) recouvrent l'IRES_{5'UTR}. Ils sont séparés par une région flexible peu structurée contenant simplement une petite tige-boucle. L'analyse de ce modèle a permis de mettre en évidence des séquences complémentaires, distantes l'une de l'autre, qui semblent évoquer une structure de type pseudo-nœud. Ces séquences GGGA (382-385) et UCCC (436-439) sont localisées dans le domaine II. Par la suite, nous avons validé l'existence de ce pseudo-nœud par l'étude de mutations qui abolissent ou qui restaurent cette interaction à distance. La structure de ces mutants a également été analysée par SHAPE. Le sondage a confirmé que les mutations qui détruisent le pseudo-nœud déplient la structure de l'IRES_{5'UTR}, alors que des mutations compensatoires permettent de reformer le pseudo-nœud dans un repliement quasiment identique à celui de l'IRES_{5'UTR} sauvage. D'un point de vue fonctionnel des tests de traduction *in vitro* avec ces mêmes mutants ont permis de démontrer

que ce pseudo-nœud est indispensable à l'activité traductionnelle de l'IRES. Ceci a également été confirmé, dans le cadre d'une collaboration avec le laboratoire de Jean-Luc Imler (UPR 9022), par des expériences de traduction *in vivo* dans des cellules S2.

1.1.3. L'IRES_{5'UTR} dépose le ribosome sur le codon d'initiation AUG₇₀₉ sans 'scanning'

Afin de mieux comprendre le mode de fonctionnement de cet IRES_{5'UTR} j'ai entrepris des études fonctionnelles, toujours à l'aide des extraits de traduction acellulaires de cellules S2 et de constructions ARN possédant le rapporteur Rénilla luciférase en aval de l'IRES_{5'UTR}. Le codon AUG en position 709-711 est le codon de démarrage de la traduction de l'ORF1 contrôlée par l'IRES_{5'UTR}. Lorsque ce codon est muté en ACG, l'IRES_{5'UTR} est alors totalement inactive. Ensuite, d'autres codons AUG ont été introduits à des distances plus ou moins éloignées du codon d'initiation afin de tester si le ribosome recruté sur l'IRES_{5'UTR} est capable de 'scanner' pour trouver un autre codon de démarrage en aval de celui localisé en position 709-711. Les résultats de ces expériences ont très clairement démontré que l'IRES_{5'UTR} du CrPV recrute le ribosome sur le codon AUG en position 709-711 et ne permet pas le 'scanning' jusqu'à un codon de démarrage alternatif localisé un peu plus en aval. L'IRES_{5'UTR} dépose donc directement et précisément le ribosome sur le codon AUG localisé en position 709-711 du génome de CrPV. Le ribosome n'est pas capable d'entamer une étape de 'scanning' à la recherche d'un autre codon AUG en aval de celui-ci.

1.1.4. Le facteur d'initiation eIF3 est requis pour initier la traduction par l'IRES_{5'UTR}

Pour aller plus loin dans la compréhension des mécanismes moléculaires mis en jeu lors de l'initiation de la traduction guidée par l'IRES_{5'UTR}, j'ai entrepris l'identification des facteurs d'initiation de la traduction (eIF) requis pour l'activité de l'IRES_{5'UTR}. Une technique novatrice mise au point par mon laboratoire d'accueil a été utilisée. Elle consiste à utiliser un ARN messager biotinylé comme appât pour purifier des complexes d'initiation assemblés dans des extraits de traduction *in vitro*. Les complexes sont ensuite purifiés à l'aide de billes magnétiques couplées à la streptavidine (Chicher et al., 2015; Prongidi-Fix et al., 2013). La composition en protéines des complexes d'initiation est alors déterminée par spectrométrie de masse. Grâce à cette approche, j'ai pu démontrer que l'IRES_{5'UTR} recrute spécifiquement le complexe multifactoriel eIF3 qui est composé de 13 sous-unités. Nous avons identifié sur l'IRES_{5'UTR} 12 des 13 sous-unités d'eIF3. La sous-unité manquante est eIF3j, elle est néanmoins nécessaire à l'activité de l'IRES_{5'UTR} (Majzoub et al., 2014). Cependant, il a été montré que eIF3j est la sous-unité la plus labile du complexe et qu'elle se décroche facilement pendant la purification. Une explication probable concernant son absence serait

que eIF3j se soit dissociée des complexes d'initiation pendant la purification. Une autre possibilité serait que eIF3j se décroche pendant l'assemblage du complexe d'initiation avant la purification (voir section *discussion et perspectives*). Ensuite, grâce aux mutants décrits plus haut, nous avons pu montrer que la fixation d'eIF3 se fait également lorsque le pseudo-nœud est détruit. Ceci suggère que l'interaction avec eIF3 se fait plutôt *via* le domaine III dans l'IRES_{5'UTR}. Une autre possibilité serait qu'eIF3 est recruté par des structures du domaine II qui sont conservées lorsque le pseudo-nœud est détruit (discuté dans la section *discussion et perspectives*).

1.1.5. L'IRES_{5'UTR} du CrPV, un IRES de classe III

L'ensemble de ces travaux a permis de collecter des données importantes sur l'IRES localisé dans la région 5'UTR du CrPV. En résumé, nous avons localisé précisément l'IRES_{5'UTR} dans la région 5'UTR du CrPV. Nous avons déterminé la structure secondaire de la totalité de la région 5'UTR et montré que l'IRES_{5'UTR} contient une interaction à distance qui forme un pseudo-nœud dont la présence est nécessaire à l'activité de l'IRES_{5'UTR}. L'IRES_{5'UTR} est capable de recruter spécifiquement le facteur d'initiation eIF3, recrutement qui ne nécessite pas la présence du pseudo-nœud. Enfin l'IRES_{5'UTR} recrute le ribosome directement sur le codon AUG de démarrage de la traduction sans passer par une étape de 'scanning'. L'ensemble de ces résultats permet de conclure que l'IRES_{5'UTR} du CrPV appartient à la classe III des IRES (**Figure 25**). De manière intéressante, l'IRES du HCV appartient également à cette classe III des IRES et il partage les mêmes caractéristiques que ceux énoncés dans ce paragraphe pour l'IRES_{5'UTR} du CrPV (Mailliot & Martin, 2018). Par conséquent nos travaux ont confirmés le fait que l'IRES_{5'UTR} de CrPV est très proche de l'IRES de HCV d'un point de vue structural et fonctionnel. Ceci avait déjà été suggéré par des études fonctionnelles sur la protéine ribosomique RACK1 qui est nécessaire à la traduction des deux IRES (voir *paragraphe III.2.3* de l'introduction).

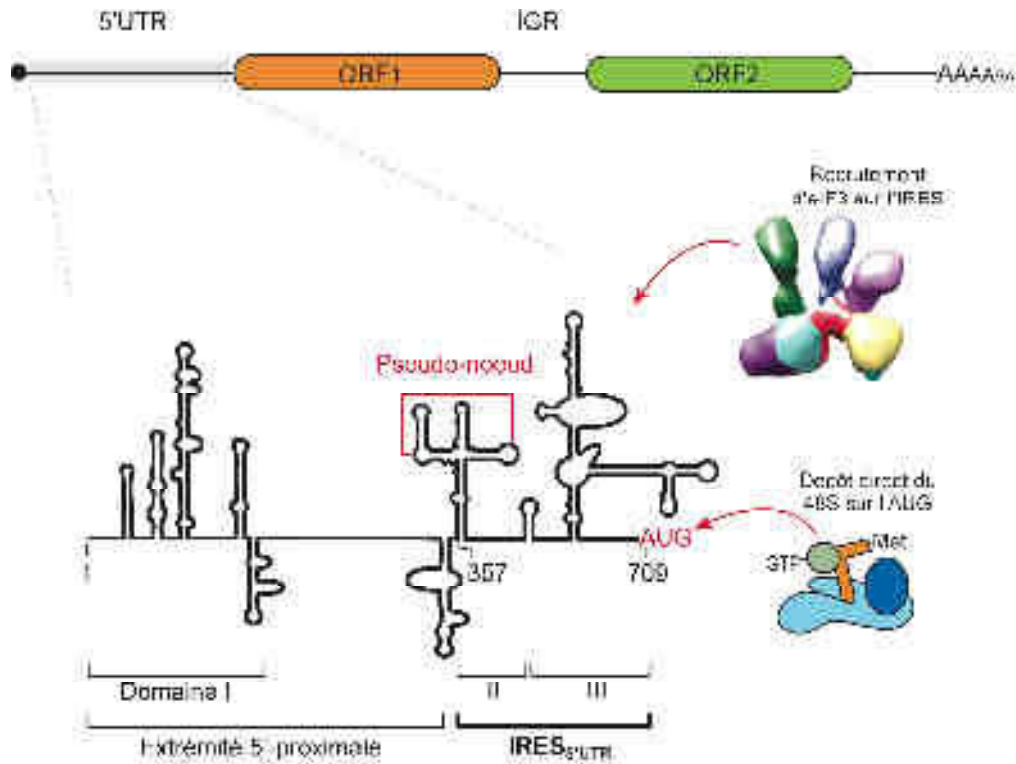


Figure 25 : L'IRES_{5'UTR} du CrPV est un IRES de classe III. Le génome du CrPV (schéma en haut) possède deux ORF, chacune d'elle est précédée d'un IRES. L'IRES de la région 5'UTR (IRES_{5'UTR}) dirige l'expression des protéines non structurales (ORF1) et l'IGR contrôle la traduction des protéines structurales (ORF2). La région 5'UTR (rectangle gris) se découpe en trois grands domaines structuraux nommés, I, II et III. Les domaines II et III constituent l'IRES_{5'UTR}. Le domaine II contient une interaction à distance de type pseudo-nœud, essentiel à l'activité traductionnelle de l'IRES_{5'UTR}. Lors de l'initiation de la traduction le facteur d'initiation eIF3 est recruté sur l'IRES_{5'UTR} et le complexe d'initiation 48S est directement déposé sur le codon d'initiation AUG, sans étape de 'scanning'.

I.2. Publication 1

The IRES_{5'UTR} of the dicistrovirus cricket paralysis virus is a type III IRES containing an essential pseudoknot structure.

Lauriane Gross, Quentin Vicens, Evelyne Einhorn, Audrey Noireterre, Laure Schaeffer, Lauriane Kuhn, Jean-Luc Imler, Gilbert Eriani, Carine Megnin et Franck Martin

Nucleic Acids Research, 2017, Vol. 45, No. 15, 8993-9004

The IRES_{5'UTR} of the dicistrovirus cricket paralysis virus is a type III IRES containing an essential pseudoknot structure

Lauriane Gross¹, Quentin Vicens¹, Evelyne Einhorn², Audrey Noireterre¹, Laure Schaeffer¹, Lauriane Kuhn³, Jean-Luc Imler², Gilbert Eriani¹, Carine Meignin² and Franck Martin^{1,*}

¹Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, F-67000 Strasbourg, France,

²Université de Strasbourg, CNRS, Réponse Immunitaire et Développement chez les Insectes, UPR 9022, F-67000 Strasbourg, France and ³Université de Strasbourg, CNRS, Plateforme Protéomique Strasbourg—Esplanade, F-67000 Strasbourg, France

Received November 03, 2016; Revised June 30, 2017; Editorial Decision July 06, 2017; Accepted July 07, 2017

ABSTRACT

Cricket paralysis virus (CrPV) is a dicistrovirus. Its positive-sense single-stranded RNA genome contains two internal ribosomal entry sites (IRESs). The 5' untranslated region (5'UTR) IRES_{5'UTR} mediates translation of non-structural proteins encoded by ORF1 whereas the well-known intergenic region (IGR) IRES_{IGR} is required for translation of structural proteins from open reading frame 2 in the late phase of infection. Concerted action of both IRES is essential for host translation shut-off and viral translation. IRES_{IGR} has been extensively studied, in contrast the IRES_{5'UTR} remains largely unexplored. Here, we define the minimal IRES element required for efficient translation initiation in drosophila S2 cell-free extracts. We show that IRES_{5'UTR} promotes direct recruitment of the ribosome on the cognate viral AUG start codon without any scanning step, using a Hepatitis-C virus-related translation initiation mechanism. Mass spectrometry analysis revealed that IRES_{5'UTR} recruits eukaryotic initiation factor 3, confirming that it belongs to type III class of IRES elements. Using Selective 2'-hydroxyl acylation analyzed by primer extension and DMS probing, we established a secondary structure model of 5'UTR and of the minimal IRES_{5'UTR}. The IRES_{5'UTR} contains a pseudoknot structure that is essential for proper folding and ribosome recruitment. Overall, our results pave the way for studies addressing the synergy and interplay between the two IRES from CrPV.

INTRODUCTION

Viruses use various strategies to hijack the host cellular translational machinery in order to produce their viral proteins. Among these, positive-stranded RNA viruses down-regulate host translation while increasing viral translation (1). For example, during poliovirus infection, an RNA structural element on the viral genome, also named internal ribosome entry site (IRES), is able to recruit the host ribosome while cap-dependent cellular translation is shut-off upon cleavage of essential canonical translation factors such as eukaryotic initiation factor eIF4G and PolyA Binding Protein (2–4). IRES has been classified into four main types according to their structural organization and their eIF requirement (5,6). Type I and II are large IRES that need most of the eIF except cap-binding protein eIF4E. Type I IRES recruits the ribosome upstream of the AUG start codon and then undergoes a scanning step to localize the AUG. In contrast, type II IRES loads the ribosome directly on the start codon without any scanning step. Type III IRES needs only eIF2 and eIF3 in order to bind directly to the 40S ribosomal subunit and to load the ribosome on the start codon without scanning. Finally, type IV IRES is the most compact: they usually contain pseudoknots, do not need any eIFs at all and can initiate translation on a non-AUG start codon.

The genome of *Dicistroviridae* consists of an ~9 kb monopartite positive-stranded RNA containing two open reading frames, ORF1 and ORF2, which encode respectively non-structural and structural proteins. Cricket paralysis virus (CrPV) is a prototype member of this family that has been thoroughly investigated. Translation of viral proteins is exclusively driven by two IRES: the 5'UTR contains IRES_{5'UTR} for ORF1 translation, while a type IV IRES is located in the intergenic region (IGR) IRES_{IGR} between the two ORFs and controls expression of structural proteins from ORF2 (7–10). Both IRES act in synergy to pro-

*To whom correspondence should be addressed. Tel: +33 388417042; Fax: +33 388602218; Email: f.martin@ibmc-cnrs.unistra.fr

duce the viral proteins required for rapid shut-off host protein translation to favor preferential viral protein synthesis. Whereas expression of non-structural proteins driven by IRES_{5'UTR} is constant during the whole infectious process, the expression of structural proteins from the IRES_{IGR} begins during the late phase of infection, with concentrations gradually increasing until reaching supramolar concentration at the end of infection (8–10). The dramatic increase in structural protein expression from the IRES_{IGR} is directly dependent on the expression of non-structural proteins from IRES_{5'UTR}, although the details of the underlying mechanisms remain elusive (11).

Translation initiation mediated by IRES_{IGR} has been extensively studied in the last two decades (12–17). Briefly, a pseudoknot structure named PKI mimics a codon–anticodon interaction in the P-site of the ribosome, therefore allowing direct ribosome recruitment without any translation factors and translation initiation from a non-AUG codon (12,13,18–20). It was recently observed by Cryo-electron microscopy (Cryo-EM) that PKI enters the ribosome in the A-site of the ribosome and is then further translocated into the P-site by eEF2, leaving the A-site free to accept the first aminoacyl-tRNA in order to proceed to elongation (18,19,21). In contrast to IRES_{IGR}, translation mediated by IRES_{5'UTR} has been much less studied. The IRES_{5'UTR} of the related dicistrovirus *Rhaphosilum padi* virus requires the scanning factors eIF1, eIF2 and eIF3 for efficient translation in a reconstituted cell-free translation extract (22). However, while the IRES_{IGR} sequences of dicistroviruses are highly conserved (23), IRES_{5'UTR} elements are largely variable and do not share any common consensus sequence suggesting the existence of different organizations among the *Dicistroviridae* family (24). Although it was discovered in early 2000, the CrPV IRES_{5'UTR} remains largely uncharacterized (10). Recently, it was demonstrated that IRES_{5'UTR}-driven translation requires the ribosomal protein RACK1 while the IRES_{IGR} can still promote efficient translation initiation with RACK1-depleted ribosomes (25). Therefore, CrPV IRES_{5'UTR} and IRES_{IGR} are using fundamentally different strategies to initiate translation. In order to better understand the translational events leading to viral propagation, a better understanding of the structure and function of the CrPV IRES_{5'UTR} is needed.

Here, we have characterized structurally and functionally the CrPV IRES_{5'UTR}. We have mapped the IRES to the 760 nt-long 5'UTR and, using selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) and dimethyl sulfate (DMS) probing, established a secondary structure model of the entire 5'UTR. Then, we have determined the minimal IRES element, and shown that it contains a pseudoknot structure. The existence of the pseudoknot was validated by mutational analysis. Finally, we have shown that this structure is essential for proper folding and activity of the IRES_{5'UTR} *in vitro* in *drosophila* cell-free translation extracts and *in vivo* in S2 cells.

MATERIALS AND METHODS

RNA transcription

Renilla reporter mRNAs were synthesized from DNA template by *in vitro* transcription using recombinant T7

RNA polymerase. After transcription, unincorporated nucleotides were trapped on a G-25 column and RNA transcripts were phenol-extracted and precipitated. RNA pellets were resuspended in water and their concentration was determined by absorbance measurements.

RNA translation in cell-free translation extract

In vitro translation competent extracts from *Drosophila melanogaster* S2 cells were prepared as previously described (25,26). Cells were lysed in 40 mM HEPES–KOH [pH8], 100 mM potassium acetate, 1 mM magnesium acetate, 1 mM DTT at a density of 10⁹ ml⁻¹ using a Cell Disruption Bomb (Parr Instrument Company). The lysate was then cleared by centrifugations at 4°C and supplemented with creatine kinase at 0, 24°C, aliquoted and stored at –80°C. *In vitro* translation experiments were performed as previously described under subsaturating conditions to avoid substrates titration (25,26). Translation efficiency was determined by Renilla Luciferase assay. RNA integrity of translated reporter mRNAs was checked by polyacrylamide gel electrophoresis (PAGE) on denaturing 4% acrylamide gels.

RNA translation in S2 cells

The pACT5C–IRES_{5'CrPV}–Renilla Luciferase was mutated by site-directed mutagenesis to obtain the m1–m4 plasmids (25). *Drosophila* S2 cells were transfected with reporter plasmid DNAs by the CaPO₄ precipitation method (adapted from (27)). Twenty-four hours later, medium was changed and copper was added to the culture medium (0.5 mM) to induce the expression of the capped Firefly construct (pMT–Firefly). Forty-eight hours later, cells were lysed and luciferase activity was measured with the Promega Dual-Luciferase assay, using a Berthold luminometer.

SHAPE analysis

Full-length RNA was designed to comprise the following elements: a leader sequence (19 nt), the entire CrPV 5'UTR (709 nt), the beginning of the ORF1 (47 nt), a 3' linker and a generic reverse transcription primer-binding site (28). Mutants comprising the 357–754 sequence did not possess any leader or 3' linker and generic primer-binding site. RNA was transcribed and purified on 4% acrylamide denaturing gels.

RNA folding assays. For the full-length wt RNA, 18 pmoles RNA in 11 µl of a buffer containing 25 mM NaCl, 15 mM MgCl₂, 25 mM Na cacodylate pH 6.5, were incubated at 60°C for 2 min, except for the no-refolding (NR) control, which was not heated. Then, 'R_a' samples underwent 1-h long cool down to 37°C. A total of 10 pmoles of NR RNA were used for SHAPE modification (see below). A total of 8 pmoles of NR and refolded RNA were added to an equivalent volume (5 µl) of loading buffer (50% glycerol, 2.5 mM Na cacodylate pH 6.5). Native PAGE was performed at 4°C on 4% acrylamide, 10 mM glycine, 10 mM Tris base. RNA were detected with toluidine blue.

For each wt or mutant RNA (m1–m4), 15 pmoles RNA in 10 µl of translation buffer and incubated at 50°C for 5

min, except for NR controls. Then, 'R_b' samples were kept at 25°C for 10 min, while 'R_c' samples underwent a 15 min-long cool down to 25°C. A total of 4–8 pmoles of R_b RNA were used for SHAPE modification.

SHAPE. SHAPE was performed in 10 µl containing 2 pmoles RNA (0.5 µM final concentration), 6.5 mM 1-methyl-7-nitroisatoic anhydride (1M7) or 80 mM benzoyl cyanide (BzCN; for full length wt RNA dataset only), 10% dimethylsulfoxide (DMSO), 90 mM Na HEPES pH 8.0. Modification is complete after 10 min at RT (~23°C). The modified RNA was precipitated, washed, dried and resuspended in 0.5X TE.

Reverse transcription was performed in 20 µl containing 2 pmoles RNA (except for BzCN dataset: 3 pmoles RNA), 0.9 pmoles of a fluorescently labeled primer (except for BzCN dataset: 6 pmoles; primers used: cr-rev421: 5'-GACCACGCGAGTCGTAATC-3'; cr-rev529: 5'-CAAGGGCTAACTAATCAGGTGTAC-3'; cr-rev769: 5'-GAGTTGATGTTGTTGTTGCGTTG-3'; 3'-gen: 5'-GAACCGGACCGAAGCCCG-3'), 160 U SuperScript III reverse transcriptase, 83 mM KCl, 56 mM Tris-HCl (pH 8.3), 0.56 mM each deoxynucleotides (dNTP), 5.6 mM DTT and 3 mM MgCl₂. Denaturing occurred at 95°C for 2 min, followed by annealing at 65°C for 5 min and incubation on ice for 2 min. RT extension parameters were: 42°C for 2 min, 50°C for 30 min and 65°C for 5 min. Sequencing reactions were performed in parallel in similar conditions, but containing 0.5 mM dideoxythymidine triphosphate (ddTTP). Reactions were stopped by the addition of 4 µl 50 mM EDTA pH 8.0, phenol-chloroform extracted, precipitated, washed, dried and resuspended in 10 µl deionized formamide. Samples were loaded on a 96-well plate for sequencing on an Applied Biosystems 3130xl genetic analyzer.

The resulting electropherograms were analyzed using QuSHAPE (29), which aligns signal within and across capillaries, as well as to the dideoxy references and corrects for signal decay. Normalized reactivities range from 0 to ~2, with 1.0–1.2 being the average reactivity for highly reactive positions. For the sake of simplicity, the final *.shape file contains reactivities for the full length RNA (752 nt) that were combined from reactivities for nucleotides 1–353 from primer cr-rev421, reactivities for nucleotides 354–711 from primer cr-rev769 (three datasets with 1M7 modification) and reactivities for nucleotides 712–752 from primer 3'-gen (two datasets with BzCN modification, for which 75% of the residues have reactivities either >1.2 or undetermined, which suggest that the region is highly unstructured).

Structure prediction. Secondary structure prediction was obtained using the Fold and ShapeKnots algorithms available in RNAstructure v. 5.7 (30). The recommended parameters for pseudoknot and probing-based prediction were used, using the *.shape file as input. A value of –500 was given to nucleotides for which reactivities were not determined. The output 2D models were rendered using VARNA (31). The secondary structure of the region 357–466 region was used as input for automated three-dimensional (3D) modeling using the web-based RNAComposer server (32), using default parameters.

DMS probing

Like the SHAPE experiments, the DMS probing reactions were performed on 2 pmoles of the minimal IRES_{5'UTR}. Briefly, the RNA is incubated for 10 min in dimethylsulfate (DMS) buffer (50 mM Na cacodylate, pH 7.5, 5 mM MgCl₂ and 100 mM KCl). The RNA was modified in the presence of 1.25% DMS. The reaction is performed for 10 min at 20°C and terminated on ice. Then, the modified RNA are precipitated in ethanol. The modification sites were detected by primer extension with two fluorescent primers complementary to nucleotides 732–755 and 517–707, respectively. The resulting electropherograms were analyzed by the same method as previously described for SHAPE analysis.

CMCT probing

The 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-p-toluene sulfonate (CMCT) experiments were performed on 4 pmoles RNA. The RNA is incubated for 10 min in CMCT buffer (Na borate 50 mM pH 8, 5; MgCl₂ 5 mM; KCl 100 mM added with 1 µg of total tRNA). The modifications were done with 10, 5 g/l; final CMCT. The reaction is completed in 30 min at 20°C, then the RNA is precipitated. The primer extension and capillary electrophoresis steps are the same as for the DMS probing.

Sucrose gradients

Pre-initiation complexes were assembled on 3' 32P-labeled RNA transcript by incubation in S2 cell-free extracts in the presence of 2 mM of a non hydrolysable GTP analog (GMP-PNP). The complexes are then resolved on sucrose gradients 7–47% and centrifuged in a SW41 rotor for 2 h 30 min at 37K at 4°. After centrifugation, the gradient is collected into 45 distinct fractions, then the positions of pre-initiation complexes were monitored by detection of radio-labeled RNA transcripts by Cerenkov counting of each fraction.

Mass spectrometry analysis and data processing.

Protein extracts were digested with sequencing-grade trypsin (Promega) as described previously (33). Resulting peptides were analyzed by nanoLC-MS/MS and MS data were searched by the Mascot algorithm against the UniProtKB *D. melanogaster* database. Identifications were validated with a protein False Discovery Rate (FDR) < 1% using a decoy database strategy. The total number of MS/MS fragmentation spectra was used to quantify each protein from three independent biological replicates. This spectral count was submitted to a negative-binomial test using an edgeR GLM regression through the R-package. For each identified protein, an adjusted *P*-value corrected by Benjamini–Hochberg was calculated, as well as a protein fold-change (FC = average spectral count in IRES/average spectral count in Domain I). The results are presented in a Volcano plot using protein log₂ fold changes and their corresponding adjusted log₁₀ *P*-values highlighted proteins up-regulated in each condition (Domain I and IRES).

RNA immunoprecipitation

The reporter plasmids containing the Wt IRES_{5'UTR}, mut2, mut4 upstream of Renilla coding sequence were transfected into drosophila S2 cells (Invitrogen) by the Electroporation method following manufacturer's instructions (Qiagen). Forty-eight hours later, cells were lysed and immunoprecipitated with an eIF3b antibody (gift from M. Hentze) coupled to Dynabeads Protein G (Life technologies). After overnight incubation, RNA and proteins were extracted with TriZol reagent (MRC). For reverse transcriptase-quantitative polymerase chain reaction (RT-qPCR), 500 ng RNA were used to perform reverse transcription (iScript cDNA Synthesis Kit from BioRad) followed by qPCR (BioRad SYBR-Green). Primers qPCR Renilla: Fw 5'-GGATGATAACTGGTCCGCAG-3', Rev 5'-TTGCCTGATTTGCCCATACC-3'.

RESULTS AND DISCUSSION

The 712-nt long 5'UTR of CrPV contains an IRES that mediates translation of ORF1 polyprotein (7). In order to characterize the minimal IRES element, we inserted nucleotides 1–761 from the 5'UTR or truncated versions upstream of the reporter gene Renilla luciferase which leads to the production of a fusion protein containing the first 17 viral N-terminal amino acids fused to Renilla luciferase peptide sequence. These constructs were *in vitro* transcribed and used for *in vitro* translation assays using S2 cell-free translation extracts. Renilla luciferase activity was used to measure the IRES translational activity driven by each RNA construct.

First, we demonstrated that capped and uncapped RNA constructs are stable in S2-cell extracts and have identical translational activity (Supplementary Figure S1). We only used uncapped RNA in subsequent experiments. We then proceeded to map the location of the IRES. Truncations from the 5' end revealed that the 5' distal half of the 5'UTR is not required for IRES activity, but that the 5' proximal half is essential (Figure 1A). More precise 5' and 3' truncations revealed that the minimal IRES element is located between nucleotides 357 and 761 (Figure 1B).

Translation efficiency was not modified when the N-terminal viral sequence coding for the first 17 amino acids was deleted, indicating that the minimal IRES_{5'UTR} is located precisely between nucleotides 357 and 709 (see construct 0 aa in Figure 2A). We next investigated whether the IRES_{5'UTR} is able to recruit the ribosome and promote scanning for a downstream AUG start codon. We mutated the viral AUG start codon at position 709 to ACG and observed a dramatic decrease of translation activity. This indicates that IRES_{5'UTR} is not able to promote scanning to reach the Renilla luciferase AUG codon further downstream and that it needs a genuine AUG start codon in its immediate vicinity (Figure 2A). To further analyze the ability of IRES_{5'UTR} to drive efficient ribosomal scanning, we re-introduced in-frame AUG start codons (at codon number 5 and 8) downstream of the cognate AUG start codon mutated to ACG. In order to avoid AUG context effects, we kept the wt AUG flanking sequence, namely A at position –3 and U at position +4 for these AUGs. (Figure 2B). When the start codon was placed at codon 5, the translation

efficiency was already dramatically reduced (~4-fold reduction), and when it was placed at codon 8 the IRES_{5'UTR} was almost totally inactive.

Interestingly, when a stop codon is introduced between the two in-frame AUGs, translation is less severely affected. This suggests that the ribosome cannot undergo a shunting mechanism to bypass the stop codon but rather terminates at the UAA codon and then proceed to partial re-initiation on the next Renilla AUG codon. This probably explains the better translation efficiency with the variant containing the UAA codon (Figure 2A). These experiments confirmed that IRES_{5'UTR} drives efficient ribosome assembly on the genuine viral start codon but is not able to promote scanning for another AUG start codon further downstream. The observation that IRES_{5'UTR} was able to drive translation initiation with the same efficiency for three other constructs containing frame shifting of the coding sequence by inserting 1, 2 and 3 nts upstream of the first AUG supports this interpretation (Figure 2C).

Altogether, our data show that IRES_{5'UTR} recruits ribosomes to its cognate viral AUG start codon, and is not able to scan for an alternative AUG start codon further downstream. This is evocative of type III IRES (5), and contrasts with the model proposed for the related discistovirus *Rhaphidopsilum padi* virus (RhPV). Reconstitution experiments revealed that the IRES_{5'UTR} from RhPV requires the scanning factor eIF1 in addition to eIF2 and eIF3 for translation and belongs to the type I category of IRES (22). Similarly, the IRES_{5'UTR} from another positive-strand unclassified virus, Halastavi arva virus (34), behaves like a type I IRES and allows the recruited ribosome to perform retrograde scanning (35). The cadicivirus IRES_{5'UTR} from the picornavirus family, which shares a similar dicistronic genome structure (36), is also a type I IRES that promotes normal 5' scanning (37). These differences in the translation mechanism are consistent with the large differences both in size and sequence between the IRES_{5'UTR} from members of this virus family, which contrast with the high conservation of the IRES_{IGR} (10,23,24). They suggest that IRES_{5'UTR} from discistoviruses might use different strategies to recruit the host translational machinery, reflecting evolution of distinct host adaptation strategies among members of this virus family.

We next determined the secondary structure of the whole 5'UTR using the SHAPE method (28), after checking that the 709-nt long RNA was homogeneously folded (Supplementary Figure S2A). The resulting predicted 2D model shows that the overall 5'UTR contains three highly structured domains that are separated by flexible linkers (Figure 3). Domain I encompasses nucleotides 1–263 and contains five stem-loops. Domain II (302–466) and III (505–689) present a more sophisticated secondary structure with hairpins and three- and four-way junctions. Based on our deletion experiments (Figures 1 and 2), we mapped the minimal IRES_{5'UTR} to nucleotides 357–709, which is in good agreement with our predicted secondary structure model. Indeed, construct 365 in which the first-half of P1 in domain II is deleted has a 4-fold reduction of its translation activity compared with constructs 360 and 357. Therefore, the minimal IRES_{5'UTR} requires domains II and III to be fully active.

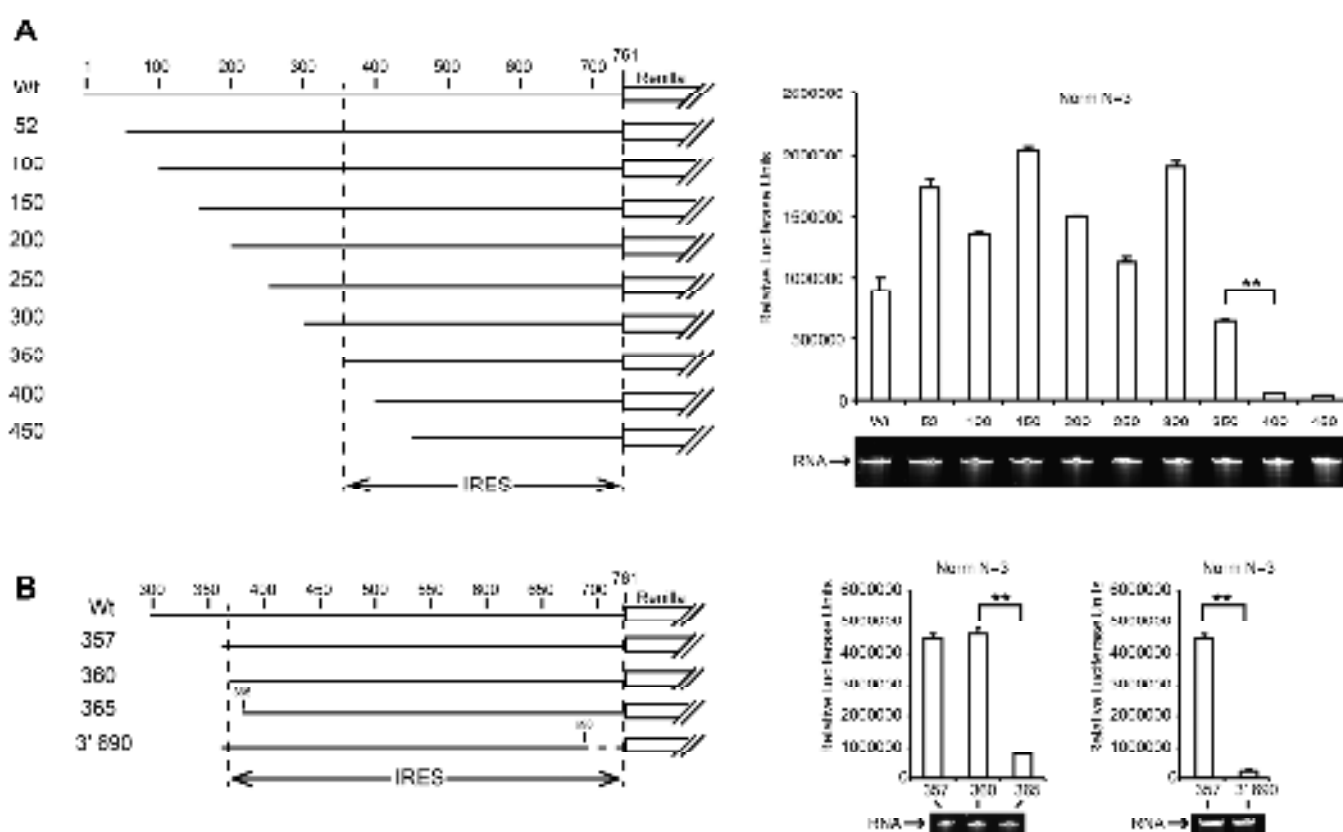


Figure 1. Mapping of the minimal IRES sequence in the 5' UTR of CrPV by 5' and 3' truncation. On the left panels, a cartoon representation of *Musca domestica* reporter transcripts used in S2 cell-free translation extracts is shown. On the right panels, translation efficiencies are represented as raw bioluminescence activity (Relative Light Units or RLU) for each transcript. Standard deviations of translational activity for each transcript are shown and corrected from their relative reporter activity (1.17 \times 10⁰⁵ based on Bradford's assay). RNA integrity was controlled by agarose-forming polyacrylamide gel electrophoresis (PAGE) and Etanidic Bromide staining. (A) IRES mapping within the entire CrPV 5' UTR and (B) precise mapping of 5' and 3' ends of the IRES_{CrPV} from CrPV.

Domains II and III from IRES_{CrPV} are separated by a flexible region, which contains a 7-nt loop and 6 bp stem. Since domains II and III are sufficient for IRES_{CrPV} activity, we also determined the secondary structure of the minimal IRES_{CrPV} construct isolated from the full-length 5' UTR. The SHAPE analysis revealed highly similar reactivities between the full length and minimal IRES construct (Supplementary Figure S2) indicating that domain II and III retain their structure in an RNA fragment containing residues 357–754 (Pearson correlation coefficient $R_{Pearson} = 0.81$). Therefore, domains II and III can fold independently from domain I.

A closer look at the SHAPE reactivity in domain II revealed that loops J3/J3a (382GGGA385) and L5 (436A/C/C/C/139) are completely inaccessible to the SHAPE reagent (Figure 4A; Supplementary Figure S2B and C). Moreover, the J3/J3a and L5 loops contain complementary sequences suggesting long-range base pairs mediated by a pseudoknot structure, as the ShapeKnots algorithm helped us pinpoint (30). To strengthen this observation, we performed DMS and CMCT probing. The pattern of DMS/CMCT reactivities supports the 2D structure predicted on the basis of the SHAPE data (Figure 3, see insert). The absence of DMS/CMCT reactivity for the nucleotides involved in the putative pseudoknot structure confirms the

SHAPE analysis and further supports the existence of a long-range interaction between these residues.

To demonstrate the presence of the pseudoknot structure, we constructed minimal IRES_{J3/J3a} fragments containing mutations in J3/J3a and L5. Mutants m1 and m2 contained non-complementary sequences in the two loops, whereas mutants m3 and m4 contained compensatory mutations that are complementary but different from the wt sequences (Figure 4A). In compensatory mutant m3, 3 nts in the loops of the putative pseudoknot have been swapped whereas in m4, the 4 nts have been swapped. To assess the impact of these mutations on the IRES_{J3/J3a} secondary structure, we determined the SHAPE reactivity profiles of the four mutants in the buffer used for translation profiles of the four mutants (see 'Materials and Methods' section, Supplementary Figure S4) and compared it with the wt IRES_{J3/J3a}. As expected, the loops J3/J3a and L5 are more reactive to the SHAPE reagent in the non-complementary mutants m1 and m2 than in the wt sequence indicating an increased flexibility (Figure 4B, green boxes). On the contrary, compensatory mutants m3 and m4 have reduced accessibility to the SHAPE reagent in loops J3/J3a and L5, as observed for the wt IRES_{J3/J3a}, suggesting that these loops are indeed involved in a pseudoknot interaction. Moreover, mutants m1 and m2 not only destabilize the long-range interaction but also affect

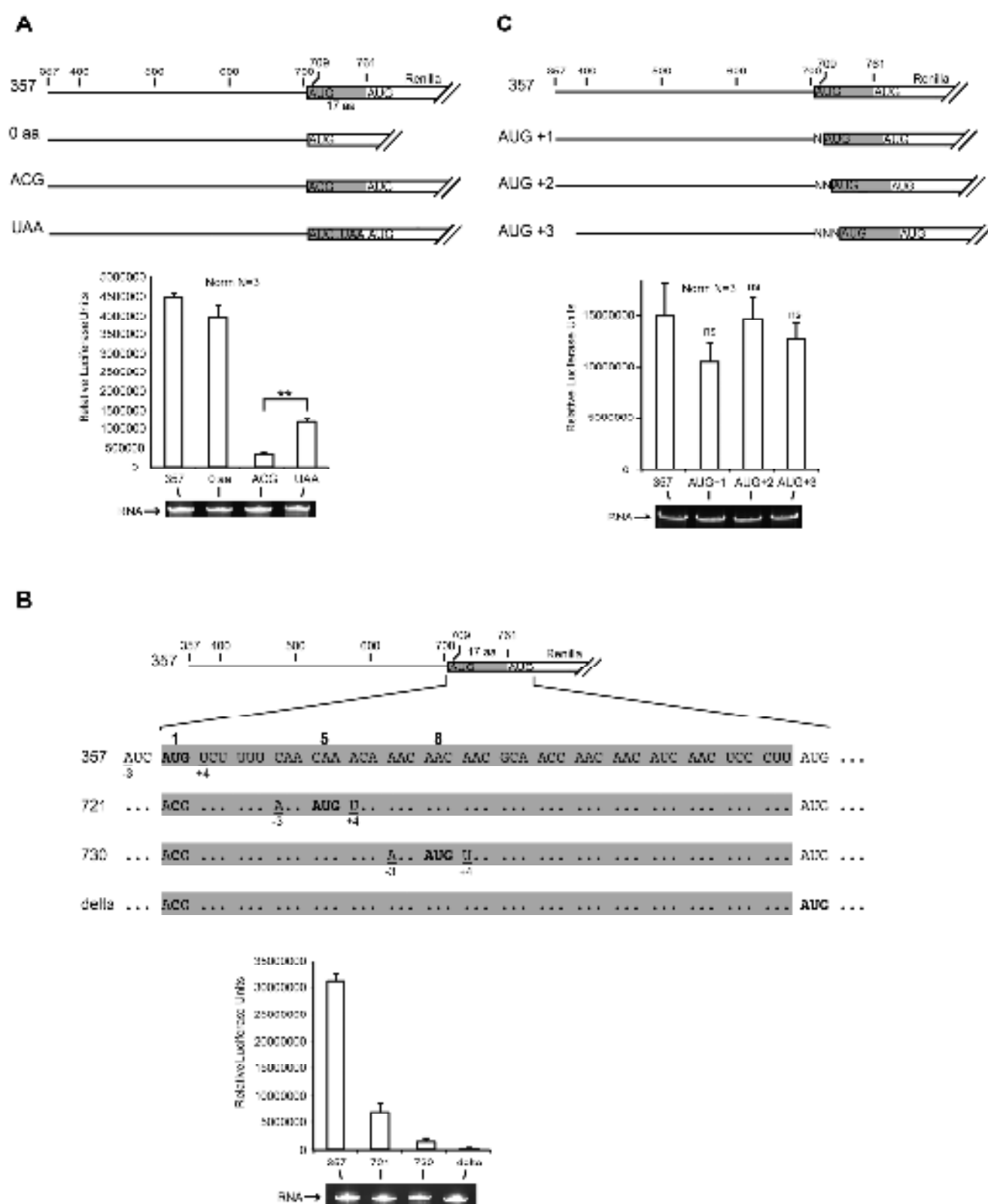


Figure 2. AUG start codon recognition during CrPV IRES₁₋₁₀ driven translation initiation. Ren1la luciferase reporter transcripts used in S2 cell free translation are represented as bar histograms (relative luciferase activity (RLU) for each translation; standard deviations of translational activity for each transcript are shown and calculated from three independent experiments, $P < 0.005$ based on Student's *t* test); non-functional RNA integrity was controlled by 4% denaturing PAGE and ethanol-*in situ* denaturing A-*in situ* roding approaches are shown in gray and fused to Ren1la luciferase coding sequence. (A) IRES₁₋₁₀ shows translation initiation at final cognate AUG start codon, but does not promote scanning for the downstream to find the Ren1la AUG when the 5' AUG is near to ACG. (B) In basal transcripts, the cognate AUG start codon is near to ACG. In addition, in 4 cases ACGs were inserted at codon position 5 and 8 with the same context. In all cases, cognate AUG. The 3' proximal AUG codons are shown in bold. (C) In these transcripts the cognate AUG is shifted by +2 or +3nt.

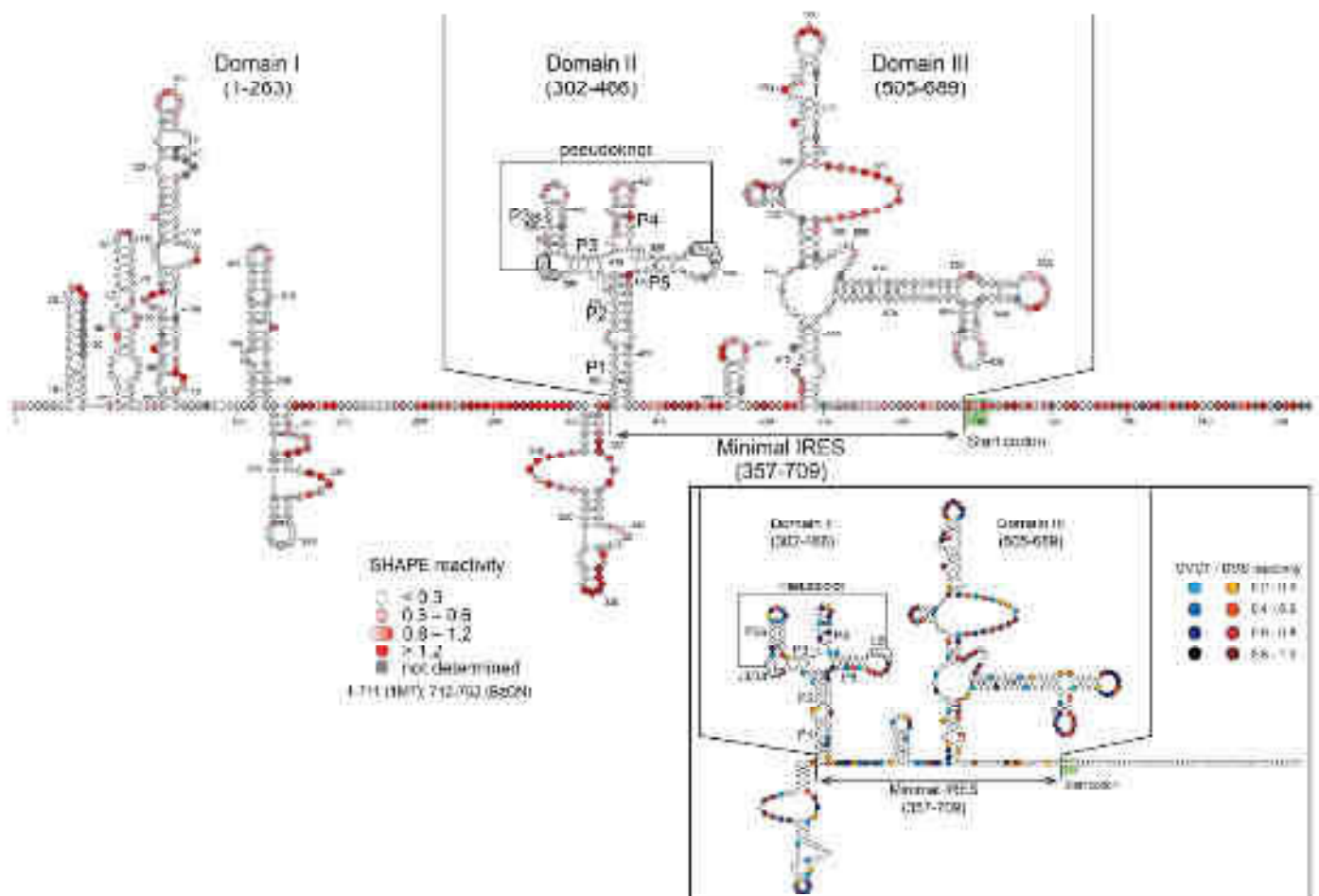


Figure 3. The predicted secondary structure of the IRES_{5177R} of CPV reveals two highly structured domains named I, II and III, separated by flexible linkers. Selected 2' hydroxyl acylation analyzed by primer extension (SHAPE) data are overlaid on the structure prediction for the full length 5177R RNA. Reactivities are shown as averages for a three independent experiments (except for region 711-754, see text for two experiments). Position of the minimal IRES is indicated (357-709). DMS and CMCT reactivity data for the minimal IRES domain (357-710). Reactivities are represented as averages from three independent experiments in the box of the figure part of the figure. Nucleotides from 1-7/36 and 1-5/36 are involved in long-distance interactions to form the pseudoknot structure. Reactivity values with standard deviations are listed in Supplementary Table S1-1.

the overall folding of domain II when compared with the wt IRES_{5177R} as shown by the Pearson correlation coefficients for both m1 and m2 mutants with the wt structure ($R_{\text{Pearson}} = 0.2$) (Alternative folds for m1 and m2 are presented in Supplementary Figure S5). Both compensatory mutants m3 and m4 present an overall folding of domains II and III very close to that of the wt IRES_{5177R} ($R_{\text{Pearson}} = 0.9$) (Figure 4B). Altogether, these data demonstrate that loops 13/5a and 1.5 are involved in a pseudoknot structure and that this long-range interaction is essential to reach the native folding of domain II.

We next addressed the requirement of the pseudoknot structure for the IRES_{5177R} translation activity. Mutants m1-m4 were inserted upstream of Renilla luciferase coding sequences in reporter constructs and the IRES activity was tested in *in vitro* translation assays with S2 cell-free extracts. The mutants m1 and m2 were inactive and unable to drive efficient translation, whereas mutants m3 and m4 showed partial or even fully restored IRES_{5177R} activity (60 and 90% respectively) (Figure 4C). We conclude that the pseudoknot structure is essential for the IRES_{5177R} activity,

indicating that these long-range distances are required for proper folding and translation initiation. In addition, a 3 bp inversion in mutant m3 is less efficient than a full 4 bp inversion as shown by the near wt IRES activity driven by mutant m4.

In order to determine the precise role of the pseudoknot structure, we performed translation initiation complex assembly and analyzed these complexes by sucrose gradient analysis. We first checked that the mutations did not affect RNA stability in S2-cell extracts (Supplementary Figure S6). Then, in the presence of GMP-PNP (a non-hydrolysable GTP-analog), the wt IRES showed accumulation of the 48S pre-initiation complexes (Figure 4D and Supplementary Figure S7). As expected the null mutants m1 and m2 showed a significant decrease in the 48S complex formation. The compensatory mutant m3 and m4 restored 48S complex assembly although with different efficiencies. The most efficient mutant was m4 as already observed in previous *in vitro* translation experiments with S2 cell-free extracts (Figure 4C). Taken together, these experiments indicate that the pseudoknot structure in domain II

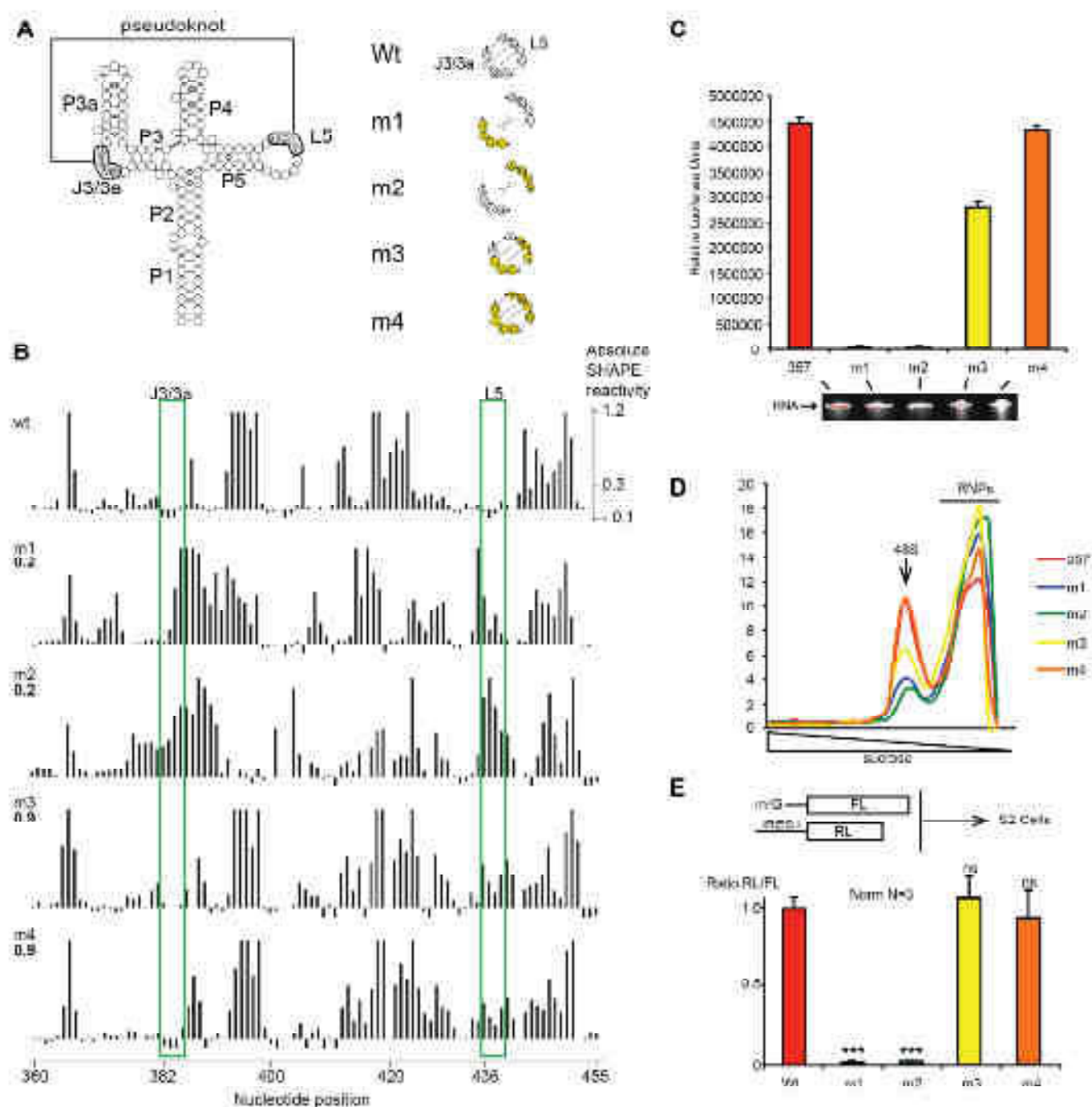


Figure 4. A pseudoknot is required for correct folding and efficient IRES₇₀₇ (773) with or without translation. (A) Mutants at the J3/3a and/or L5 regions used in this study; mutated nucleotides are shown in yellow. (B) Absolute SHAPE reactivities for 357–754 RNA transcripts with either wt or m1–m4 mutant sequence. Reactivities are shown as averages from two independent experiments (reactivity values with standard deviation are given in Supplementary Table S2). The Pearson correlation coefficients (R_{wt}) with the 357–754 wt dataset are shown on the left of each histogram. J3/3a and L5 are boxed in green. (C) Translation activity of IRES 357–754 wt and mutants m1–m4 when placed upstream of Renilla luciferase coding region. RNA integrity control by densitometry (420 nm) on ethylnaphthylsulfonamide-PAGE gel image under the same conditions as in (C). (D) Secondary structure prediction (3–420) of wt and mutant m1–m4 IRESs. (E) *In vivo* translation using the reporter in reporter-transfected HeLa cells. The mutants m1 and m2 accomplish with \sim IRES₇₀₇ activity. * $P < 0.05$ and ** $P < 0.005$, *** $P < 0.0005$ based on Student's *t*-test.

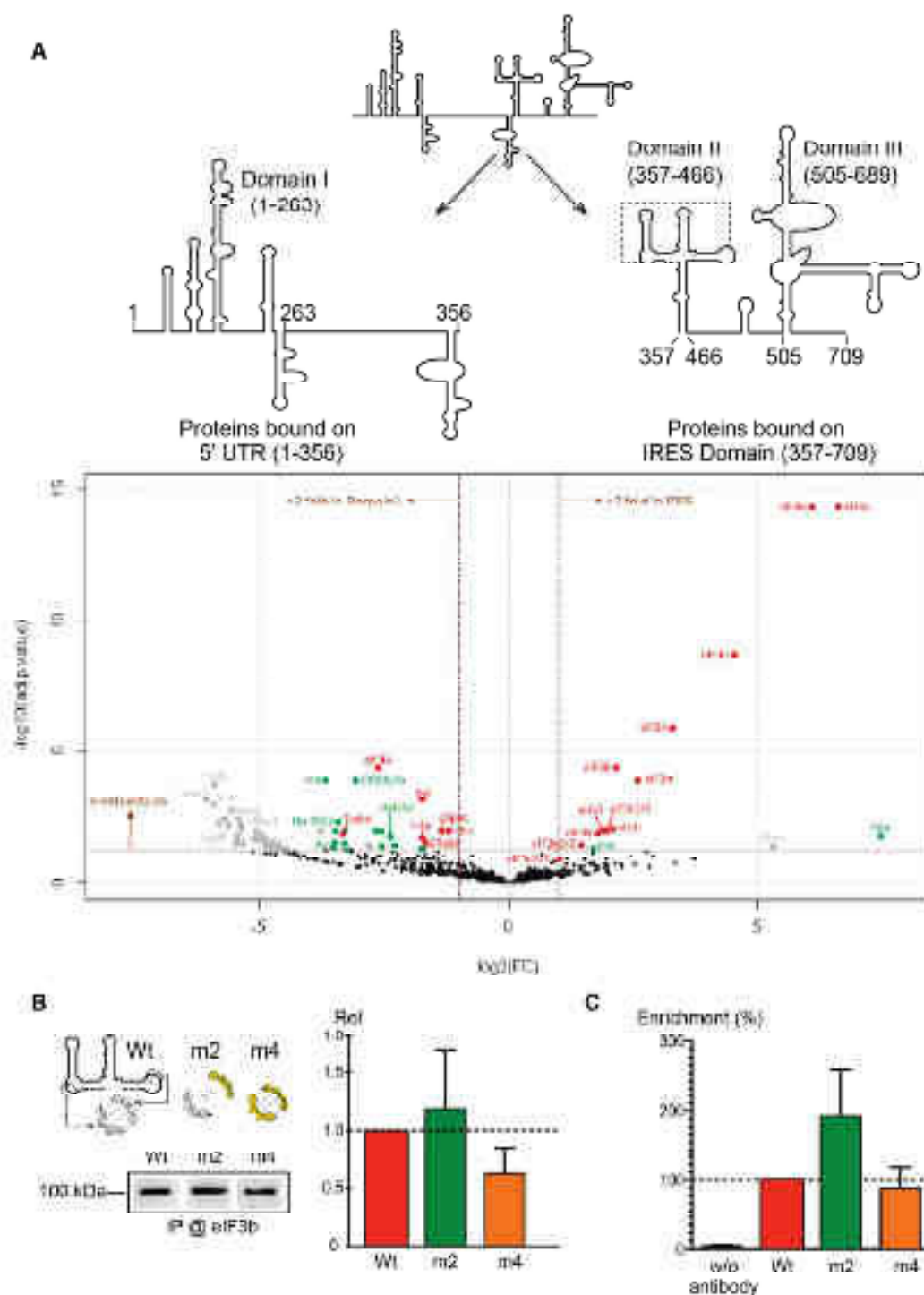


Figure 5. Eukaryotic initiation factor 4E1a is recruited specifically by the IRES_{5'UTR} in a protein-knot-independent manner. **(A)** Mass spectrometry analysis of adenine-blocked translation initiation complexes on the 5' proximal segment (1-356) and the minimal IRES (357-709). Graphical representation of parametric data (protein log₂ system) count fold changes (on the x-axis) and the corresponding adjusted log₁₀ P-values (on the y-axis) are plotted in a parametric volcano plot. The significance thresholds are represented by a horizontal dashed line (P -value = 0.05, two-tailed test with Benjamini-Hochberg adjustment) and two vertical dashed lines (± 2.0 fold on the left and ± 2.0 fold on the right). Data points in the upper left and upper right quadrants indicate significant negative and positive changes in protein abundance. Protein names are labeled near the off-scale red spots and they are depicted according to the following color code: proteins represented by a red spot are identified by more than 30 MS/MS spectra, by a green spot, when identified by ≥ 10 spectra and by a gray spot when identified by ≥ 0 spectra. Data points are plotted on the basis of average spectral counts from triplicate analysis. For an exhaustive list of data, see Supplementary Table S5. **(B)** Western blot analysis of eIF3b recruitment by WT, m2 and m4 mutants. Blocked translation initiation complexes programmed with WT, m2 and m4 mutants. The complexes were assembled *in vitro* in S2 cell extracts. The histogram represents the quantification of three independent experiments. **(C)** *In vivo* RNA immunoprecipitation of IRES_{5'UTR}-Rubi1a reporter mRNA in S2 cells. WT, m2 and m4 mRNAs were immunoprecipitated using anti-eIF3b antibody and quantified by RT-qPCR. The histogram represents the quantification of three independent experiments. The enrichment fold was calculated by the ratio between immunoprecipitated and the input and was set to 100% for the WT.

is essential for 48S pre-initiation complex formation and is therefore essential at an early step of translation initiation.

Finally, mutants m1–m4 were tested for IRES activity *in vivo* using monocistronic reporters transfected in S2 cells (Figure 4E). Here again, m1 and m2 mutants were totally inactive, whereas the IRES activity was restored to the same extent for m3 and m4 mutants. Taken together, this demonstrates that the pseudoknot structure is essential for IRES_{5'UTR} activity *in vivo* as well.

To further characterize the CrPV IRES_{5'UTR}, we performed mass spectrometry analysis on translation initiation complexes in order to determine the factors that bind specifically to the IRES. To do so, translation initiation complexes were assembled on a biotinylated CrPV IRES_{5'UTR} in S2-cell extracts. Translation initiation complexes were then first immobilized on magnetic streptavidin beads and subsequently eluted by DNase treatment as previously described (33,38). The composition of the purified translation initiation complex was then determined by mass spectrometry. To determine the factors required at the initial step of the translational process, translation initiation was blocked by addition of edeine, which prevents initiator tRNA^{Met} codon-anticodon interaction in the P-site of the ribosome (39–41). As a negative control, we performed the same experiments with the 5' proximal part of the 5'UTR that is not essential for IRES activity. Interestingly, 12 of the 13 subunits of eIF3 are specifically found on the IRES fragment, indicating that IRES_{5'UTR} recruits this factor (Figure 5A and Supplementary Table S5), the only missing factor being eIF3j. However, we have previously shown that eIF3j is required for the activity of IRES_{5'UTR} (25). In yeast, eIF3j/hcr1 has been shown to participate in efficient AUG recognition during the scanning process by directly interacting with eIF1A for canonical cap-dependent translation (42). A similar function has also been proposed for mammalian eIF3j (43,44). Since we have shown that IRES_{5'UTR} does not promote scanning for initiation on the viral AUG, our data show that eIF3j has a different function for IRES_{5'UTR}. We propose that eIF3j is essential for ribosome recruitment to the IRES_{5'UTR} and that it dissociates after complex formation prior to AUG recognition.

In order to determine whether the pseudoknot is necessary for eIF3 recruitment, pre-initiation complexes were programmed *in vitro* with S2-cell extracts with Wt, m2 and m4 and were purified as previously described. Western blot analysis with an antibody specific for drosophila eIF3b subunit showed that mutant m2 (which has no pseudoknot) is also able to recruit eIF3 suggesting that eIF3 interaction does not require the pseudoknot (Figure 5B). To further corroborate this statement *in vivo*, Renilla reporter mRNAs containing the Wt IRES_{5'UTR} and mutants m2 and m4 were transfected in S2 cells. Then we performed RNA immunoprecipitation with an antibody raised against drosophila eIF3b subunit, we quantified the Renilla reporter mRNA by RT-qPCR. Indeed, mutant m2 is also able to interact with eIF3 *in vivo* confirming that the pseudoknot is not required for eIF3 recruitment (Figure 5C). The interaction between the IRES_{5'UTR} and eIF3 is transient during the translation initiation process. In both experiments, mutant m2 seems to interact more efficiently than Wt and m4, this might be ex-

plained by the fact that mutant m2 is inactive and therefore accumulates more complexes with eIF3 than Wt and m4.

The IRES_{5'UTR} is preceded by domain I which is dispensable for translational activity. A few other proteins bind specifically to this domain (Figure 5A and Supplementary Table S5). These secondary structures might be involved in other steps required for viral propagation such as initiation of replication or encapsidation. Our data are in good agreement with previous observations made while constructing an infectious molecular clone of CrPV (45). The authors isolated a clone containing a duplication of fragment 75–271, which inhibits viral infectivity. In our secondary structure model, this duplication is located in domain I, away from the minimal IRES sequence. The duplication however has no effect on viral translation and RNA accumulation during CrPV infection suggesting it impacts viral entry and/or viral packaging (46).

In addition, the structured arrangements of domains II and III with several helices connected by three- and four-way junctions and loops is reminiscent of those in the classical swine fever virus (CSFV), the Hepatitis-C virus (HCV), the foot-and-mouth disease virus (FMDV) and other IRESes (6). These structured domains were shown to be important for IRES function, through binding to eIF3 (47), promoting long-range interactions (48,49), or binding to IRES-*trans* acting factors (50,51). The secondary structure of the apical domain II in CrPV IRES_{5'UTR} that harbors a pseudoknot bears some resemblance to that of residues 153–255 in CSFV and residues 134–249 in FMDV (Supplementary Figure S8). Automatic 2D-based 3D-RNA modeling suggests that, because of the pseudoknot, the four-way junction would adopt a topology closer to that of the L-shaped CSFV (52), than to the topology modeled for FMDV (49). Future studies using X-ray crystallography or cryo-EM will be needed to pinpoint the similarities and differences in how type III IRES domains fold, and what role they play in IRES function.

In summary, the IRES_{5'UTR} contains a three-way junction structure with a pseudoknot that recruits the ribosome on the cognate viral AUG start codon without any scanning step. Moreover IRES_{5'UTR} specifically recruits eIF3. This is reminiscent of translation driven by the HCV-IRES, the prototypic type III IRES (5). This similarity between CrPV IRES_{5'UTR} and HCV-IRES is in agreement with the fact that both IRES are strictly dependent on the ribosomal protein RACK1. On the contrary IRES_{IGR}-driven translation efficiency does not require the presence of RACK1 on the ribosome, suggesting fundamentally distinct molecular mechanisms for ribosome recruitment (53). The structural characterization of the IRES_{5'UTR} from CrPV will facilitate the investigation of the role played by RACK1 in cap-independent translation. During CrPV infection, the IRES_{5'UTR} promotes translation from ORF1 in a constitutive manner during the whole infectious process, whereas IRES_{IGR}-driven translation starts in the second half of infection and is boosted in the late phase of infection (11). The molecular basis for this differential expression pattern, which is essential for the progression of the viral infection, remains unexplored. Our data on the IRES_{5'UTR} represent a first step toward a better understanding of the concerted ac-

tion of the two structurally and functionally different IRES active in dicistroviruses.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Claire Batisse and Adeline Renaud for S2-cell extracts preparation. We are thankful to Matthias Hentze for *drosophila* eIF3 antibodies. We also thank Redmond P. Smyth for critical reading of the manuscript.

FUNDING

Centre National de la Recherche Scientifique (CNRS); Université de Strasbourg; Investissement d'Avenir program [NetRNA ANR-10-LABX-36, ANR ANR-11-SVSE802501]. Funding for open access charge: Centre National de la Recherche Scientifique.

Conflict of interest statement. None declared.

REFERENCES

- Walsh,D. and Mohr,I. (2011) Viral subversion of the host protein synthesis machinery. *Nat. Rev. Microbiol.*, **9**, 860–875.
- Joachims,M., Van Breugel,P.C. and Lloyd,R.E. (1999) Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J. Virol.*, **73**, 718–727.
- Etchison,D., Milburn,S.C., Ederly,I., Sonenberg,N. and Hershey,J.W. (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.
- Kuyumcu-Martinez,N.M., Joachims,M. and Lloyd,R.E. (2002) Efficient cleavage of ribosome-associated poly(A)-binding protein by enterovirus 3C protease. *J. Virol.*, **76**, 2062–2074.
- Jackson,R.J., Hellen,C.U.T. and Pestova,T. V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.*, **11**, 113–127.
- Filbin,M.E. and Kieft,J.S. (2009) Toward a structural understanding of IRES RNA function. *Curr. Opin. Struct. Biol.*, **19**, 267–276.
- Bonning,B.C. and Miller,W.A. (2010) Dicistroviruses. *Annu. Rev. Entomol.*, **55**, 129–150.
- Garrey,J.L., Lee,Y.-Y., Au,H.H.T., Bushell,M. and Jan,E. (2010) Host and viral translational mechanisms during cricket paralysis virus infection. *J. Virol.*, **84**, 1124–1138.
- Moore,N.F., Kearns,A. and Pullin,J.S. (1980) Characterization of cricket paralysis virus-induced polypeptides in *Drosophila* cells. *J. Virol.*, **33**, 1–9.
- Wilson,J.E., Powell,M.J., Hoover,S.E. and Sarnow,P. (2000) Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol. Cell Biol.*, **20**, 4990–4999.
- Khong,A., Bonderoff,J.M., Spriggs,R. V., Tammper,E., Kerr,C.H., Jackson,T.J., Willis,A.E. and Jan,E. (2016) Temporal regulation of distinct internal ribosome entry sites of the dicistroviridae cricket paralysis virus. *Viruses*, **8**, doi:10.3390/v8010025.
- 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.
- Costantino,D.A., Pflugsten,J.S., Rambo,R.P. and Kieft,J.S. (2008) tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat. Struct. Mol. Biol.*, **15**, 57–64.
- Spahn,C.M.T., 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.
- Pestova,T. V. and Hellen,C.U.T. (2003) Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev.*, **17**, 181–186.
- Jan,E., Kinzy,T.G. and Sarnow,P. (2003) Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 15410–15415.
- Jan,E. and Sarnow,P. (2002) Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J. Mol. Biol.*, **324**, 889–902.
- Fernández,I.S., Bai,X.C., Murshudov,G., Scheres,S.H.W. and Ramakrishnan,V. (2014) Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell*, **157**, 823–831.
- Muhs,M., Hilal,T., Mielke,T., Skabkin,M.A., Sanbonmatsu,K.Y., Pestova,T. V. and Spahn,C.M.T. (2015) Cryo-EM of ribosomal 80s complexes with termination factors reveals the translocated cricket paralysis virus IRES. *Mol. Cell*, **57**, 422–433.
- Pflugsten,J.S., Costantino,D.A. and Kieft,J.S. (2006) Structural basis for ribosome recruitment and manipulation by a viral IRES RNA. *Science*, **314**, 1450–1454.
- Murray,J., Savva,C.G., Shin,B.-S., Dever,T.E., Ramakrishnan,V. and Fernández,I.S. (2016) Structural characterization of ribosome recruitment and translocation by type IV IRES. *Elife*, **5**, doi:10.7554/eLife.13567.
- Terenin,I.M., Dmitriev,S.E., Andreev,D.E., Royall,E., Belsham,G.J., Roberts,L.O. and Shatsky,I.N. (2005) A cross-kingdom internal ribosome entry site reveals a simplified mode of internal ribosome entry. *Mol. Cell Biol.*, **25**, 7879–7888.
- Kanamori,Y. and Nakashima,N. (2001) A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. *RNA*, **7**, 266–274.
- Roberts,L.O. and Groppe,E. (2009) An atypical IRES within the 5' UTR of a dicistrovirus genome. *Virus Res.*, **139**, 157–165.
- Majzoub,K., Hafirassou,M.L., Meignin,C., Goto,A., Marzi,S., Fedorova,A., Verdier,Y., Vinh,J., Hoffmann,J.A., Martin,F. et al. (2014) RACK1 controls IRES-mediated translation of viruses. *Cell*, **159**, 1086–1095.
- Wakiyama,M., Kaitsu,Y. and Yokoyama,S. (2006) Cell-free translation system from *Drosophila* S2 cells that recapitulates RNAi. *Biochem. Biophys. Res. Commun.*, **343**, 1067–1071.
- Green,M.R. and Sambrook,J. (2012) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, NY.
- Wilkinson,K.A., Merino,E.J. and Weeks,K.M. (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.*, **1**, 1610–1616.
- Karabiber,F., McGinnis,J.L., Favorov,O.V. and Weeks,K.M. (2013) QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA*, **19**, 63–73.
- Hajdin,C.E., Bellaousov,S., Huggins,W., Leonard,C.W., Mathews,D.H. and Weeks,K.M. (2013) Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 5498–5503.
- Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
- Popenda,M., Szachniuk,M., Antczak,M., Purzycka,K.J., Lukasiak,P., Bartol,N., Blazewicz,J. and Adamiak,R.W. (2012) Automated 3D structure composition for large RNAs. *Nucleic Acids Res.*, **40**, e112.
- Chicher,J., Simonetti,A., Kuhn,L., Schaeffer,L., Hammann,P., Eriani,G. and Martin,F. (2015) Purification of mRNA-programmed translation initiation complexes suitable for mass spectrometry analysis. *Proteomics*, **15**, 2417–2425.
- Boros,Á., Pankovics,P., Simmonds,P. and Reuter,G. (2011) Novel positive-sense, single-stranded RNA (+ssRNA) virus with di-cistronic genome from intestinal content of freshwater carp (*Cyprinus carpio*). *PLoS One*, **6**, e29145.
- Abaeva,I.S., Pestova,T.V. and Hellen,C.U.T. (2016) Attachment of ribosomal complexes and retrograde scanning during initiation on the Halastavi árva virus IRES. *Nucleic Acids Res.*, **44**, 2362–2377.

36. Woo, P.C.Y., Lau, S.K.P., Choi, G.K.Y., Huang, Y., Teng, J.L.L., Tsoi, H.-W., Tse, H., Yeung, M.L., Chan, K.-H., Jin, D.-Y. *et al.* (2012) Natural occurrence and characterization of two internal ribosome entry site elements in a novel virus, Canine Picodistrovirus, in the Picornavirus-like superfamily. *J. Virol.*, **86**, 2797–2808.
37. Asnani, M., Pestova, T.V and Hellen, C.U.T. (2016) Initiation on the divergent Type I dicistrovirus IRES: factor requirements and interactions with the translation apparatus. *Nucleic Acids Res.*, **44**, 3390–3407.
38. Prongidi-Fix, L., Schaeffer, L., Simonetti, A., Barends, S., Ménétret, J.-F., Klaholz, B.P., Eriani, G. and Martin, F. (2013) Rapid purification of ribosomal particles assembled on histone H4 mRNA: a new method based on mRNA-DNA chimaeras. *Biochem. J.*, **449**, 719–728.
39. Odon, O.W., Kramer, G., Henderson, A.B., Pinphanichakarn, P. and Hardesty, B. (1978) GTP hydrolysis during methionyl-tRNA^f binding to 40 S ribosomal subunits and the site of edeine inhibition. *J. Biol. Chem.*, **253**, 1807–1816.
40. Garreau de Loubresse, N., Prokhorova, I., Holtkamp, W., Rodnina, M.V., Yusupova, G. and Yusupov, M. (2014) Structural basis for the inhibition of the eukaryotic ribosome. *Nature*, **513**, 517–522.
41. Kozak, M. and Shatkin, A.J. (1978) Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. *J. Biol. Chem.*, **253**, 6568–6577.
42. Aylett, C.H.S., Boehringer, D., Erzberger, J.P., Schaefer, T. and Ban, N. (2015) Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. *Nat. Struct. Mol. Biol.*, **22**, 269–271.
43. ElAntak, L., Wagner, S., Herrmannova, A., Karaskova, M., Rutkai, E., Lukavsky, P.J. and Valasek, L. (2010) The indispensable n-terminal half of eIF3j/hcr1 cooperates with its structurally conserved binding partner eIF3b/prtl1-rrm and with eIF1A in stringent aug selection. *J. Mol. Biol.*, **396**, 1097–1116.
44. Sokabe, M. and Fraser, C.S. (2014) Human eukaryotic initiation factor 2 (eIF2)-GTP-Met-tRNAⁱ ternary complex and eIF3 stabilize the 43 S preinitiation complex. *J. Biol. Chem.*, **289**, 31827–31836.
45. Kerr, C.H., Wang, Q.S., Keatings, K., Khong, A., Allan, D., Yip, C.K., Foster, L.J. and Jan, E. (2015) The 5' untranslated region of a novel infectious molecular clone of the dicistrovirus cricket paralysis virus modulates infection. *J. Virol.*, **89**, 5919–5934.
46. Kerr, C.H., Wang, Q.S., Keatings, K., Khong, A., Allan, D., Yip, C.K., Foster, L.J. and Jan, E. (2015) The 5' untranslated region of a novel infectious molecular clone of the dicistrovirus cricket paralysis virus modulates infection. *J. Virol.*, **89**, 5919–5934.
47. Hashem, Y., des Georges, A., Fu, J., Buss, S.N., Jossinet, F., Jobe, A., Zhang, Q., Liao, H.Y., Grassucci, R.A., Bajaj, C. *et al.* (2013) High-resolution cryo-electron microscopy structure of the *Trypanosoma brucei* ribosome. *Nature*, **494**, 385–389.
48. Fernández-Miragall, O. and Martínez-Salas, E. (2003) Structural organization of a viral IRES depends on the integrity of the GNRA motif. *RNA*, **9**, 1333–1344.
49. Jung, S. and Schlick, T. (2013) Candidate RNA structures for domain 3 of the foot-and-mouth-disease virus internal ribosome entry site. *Nucleic Acids Res.*, **41**, 1483–1495.
50. Pacheco, A. and Martínez-Salas, E. (2010) Insights into the biology of IRES Elements through riboproteomic approaches. *J. Biomed. Biotechnol.*, **2010**, 1–12.
51. Lozano, G. and Martínez-Salas, E. (2015) Structural insights into viral IRES-dependent translation mechanisms. *Curr. Opin. Virol.*, **12**, 113–120.
52. Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., Pestova, T. V., Hellen, C.U.T. and Frank, J. (2013) Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature*, **503**, 539–543.
53. Majzoub, K., Hafirassou, M.L., Meignin, C., Goto, A., Marzi, S., Fedorova, A., Verdier, Y., Vinh, J., Hoffmann, J.A., Martin, F. *et al.* (2014) Ribosomal protein RACK1 is a specific host factor required for IRES-mediated translation of fly and human viruses. *Cell*, **159**, 1086–1095.

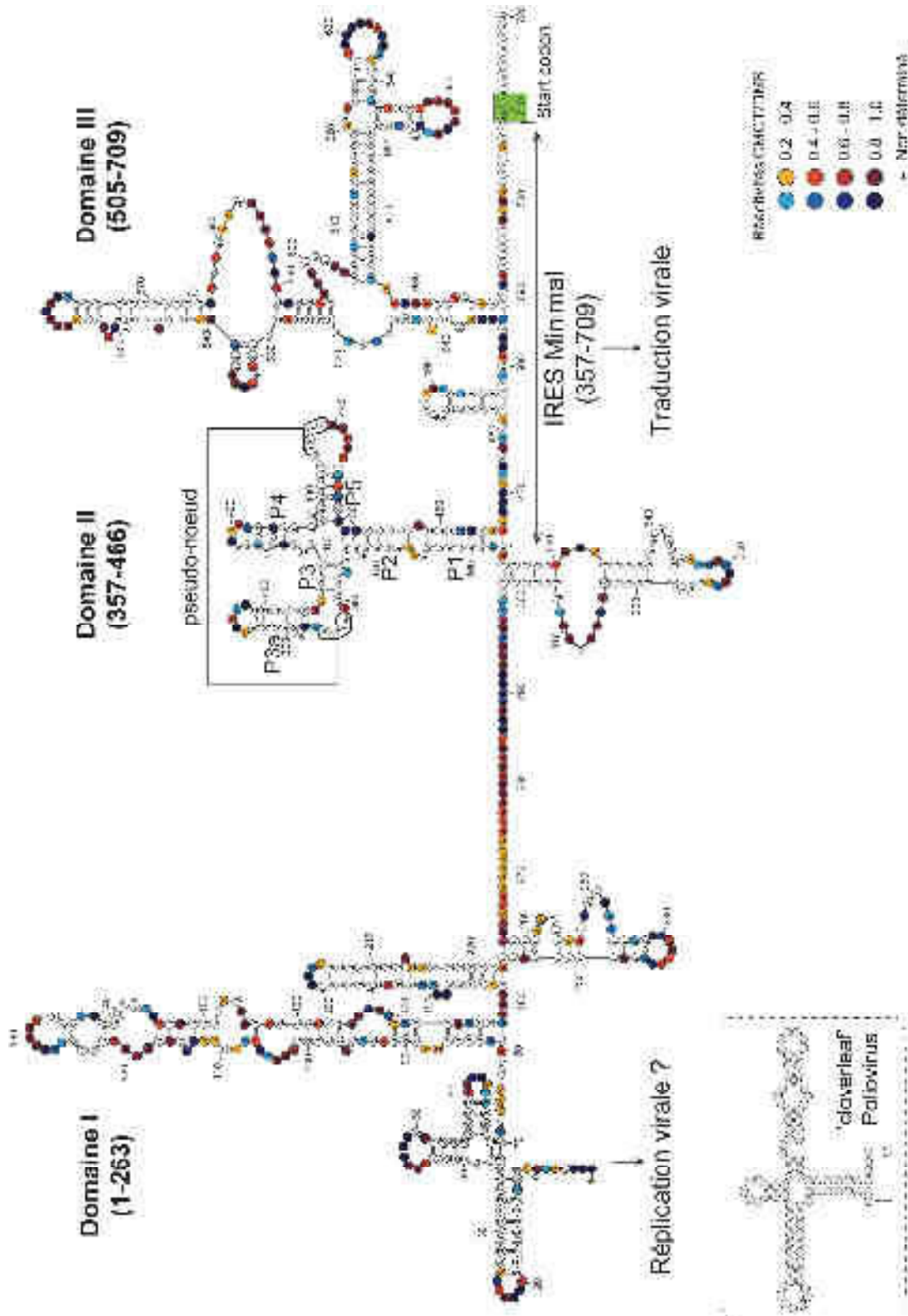


Figure 26 : Modèle de structure secondaire de la région 5'UTR complète du CrPV. Le modèle de structure secondaire de départ obtenu par SHAPE a été affiné par sondage chimique au DMS (pour les bases A et C) et au CMCT (pour les bases G et U). La région 5'UTR du CrPV se subdivise en trois domaines. Le domaine I (1-263) possède une structure similaire à la feuille de trèfle ou 'cloverleaf' du poliovirus (encadré en pointillés gris), il pourrait être impliqué dans la réplication virale. Les domaines II (357-466) et III (505-709) recouvrent l'IRES_{5'UTR} minimal (357-709) nécessaire à la traduction de l'ORF1 du CrPV. Le domaine II contient un appariement à distance de type pseudo-nœud. Les réactivités des nucléotides au DMS et CMCT sont représentées selon le code couleur indiqué dans la légende de la figure. La réactivité de chaque nucléotide est calculée en moyennant les réactivités de trois expériences de sondage indépendantes pour chaque réactif.

I.3. Résultats complémentaires non publiés

Par la suite, j'ai réalisé différentes analyses sur la totalité de la région 5'UTR du CrPV ainsi que sur le fragment minimal contenant l'IRES_{5'UTR}. Les données collectées ont permis de décortiquer avec plus de précision les bases moléculaires de l'initiation de la traduction guidée par l'IRES_{5'UTR}. Ainsi, j'ai affiné la structure secondaire de la région 5'UTR complète du CrPV. Les éléments structuraux de l'IRES_{5'UTR} ont également été décortiqués et leurs rôles dans l'initiation de la traduction caractérisés.

1.3.1. Étude de la structure secondaire de la région 5'UTR du CrPV par sondage chimique

La structure secondaire des 709 nucléotides de la région 5'UTR ont été sondés en utilisant la méthode du SHAPE. Cette méthode génère souvent des données ambiguës qui sont difficiles à interpréter pour certaines zones d'un ARN donné. En effet, les ARN contiennent presque toujours des nucléotides qui, bien que présents dans des régions structurées, demeurent réactifs dans une expérience de SHAPE. Il est donc difficile d'établir un modèle de structure secondaire fiable à partir des seules données de SHAPE. Afin d'affiner le modèle préliminaire de structure secondaire obtenu grâce au SHAPE, j'ai mis au point une méthode couplant les techniques de modification des bases nucléotidiques au DMS et CMCT avec la technique d'analyse par fluorescence utilisée dans le SHAPE. Ces deux sondes chimiques avaient déjà permis d'affiner le modèle de structure secondaire du fragment contenant l'IRES_{5'UTR} (résultats présentés dans la *publication 1*). Par la suite, j'ai également utilisé ces deux sondes DMS et CMCT pour étudier la structure secondaire de l'extrémité 5'-proximale (1-356) de la région 5'UTR du CrPV.

La région 5'-proximale possède des éléments structurés (1-263) ainsi qu'une région flexible simple brin (264-301) et enfin une tige-boucle (302-356) en amont de l'IRES_{5'UTR}. Le modèle de repliement des nucléotides 1 à 180 a pu être affiné grâce à ces nouvelles données de sondage (**Figure 26**). Il apparaît que les 80 premiers nucléotides en 5' du domaine I adoptent une structure de type feuille de trèfle ou 'cloverleaf'. De manière intéressante, des éléments structuraux similaires ont déjà été décrits dans la région 5'UTR d'autres virus comme par exemple le poliovirus (voir encadré). Chez ce virus, cette structure est essentielle pendant l'infection virale car elle serait impliquée dans l'étape de réplication du virus (Andino et al., 1993, 1990). En aval de cette structure 'cloverleaf' du CrPV, les sondages DMS/CMCT ont mis en évidence trois longues tige-boucles. Les données obtenues par sondage au DMS/CMCT pour les nucléotides 263 à 356 ont permis de confirmer le modèle initial établi

pour cette région lors des études préliminaires par SHAPE. Cette région comporte un domaine « charnière » quasiment dénué de structure (et donc très réactif aux sondes chimiques) et la tige-boucle 302-356. Le domaine charnière sépare le domaine I, vraisemblablement impliqué dans la réplication virale, des domaines II et III qui composent l'IRES_{5'UTR} responsable de la traduction de l'ORF1.

En outre, nous avons établi par SHAPE que les réactivités des nucléotides de l'IRES_{5'UTR} dans le contexte de la 5'UTR complète corrélaient à 80% avec les réactivités de l'IRES_{5'UTR} analysé séparément. Nous en concluons que, lorsqu'ils sont isolés, le domaine I et l'IRES_{5'UTR} adoptent un repliement identique à celui qu'ils adoptent dans la molécule complète. Il n'y a donc aucune interaction entre le domaine I et l'IRES_{5'UTR}, ils se replient indépendamment l'un de l'autre.

1.3.2. L'extrémité 5'-proximale du génome du CrPV est-elle un bon ARN de référence pour des analyses comparatives par spectrométrie de masse ?

L'objectif suivant de mon étude a été d'identifier les partenaires protéiques spécifiques de l'IRES_{5'UTR} du CrPV. Pour cela, nous avons purifié des complexes d'initiation assemblés sur des ARN messagers en présence d'inhibiteurs du ribosome. Leur composition en protéine a ensuite été déterminée par spectrométrie de masse. Nous avons réalisé ces expériences avec le domaine minimal IRES_{5'UTR} d'une part, et avec l'autre moitié de la région 5'UTR du CrPV (1-356), appelée extrémité 5'-proximale, d'autre part. L'extrémité 5'-proximale a été utilisée comme ARN de référence afin d'identifier par comparaison les protéines fixées spécifiquement à l'IRES_{5'UTR}.

L'extrémité 5'-proximale a servi d'ARN de référence pour plusieurs raisons. Premièrement, parce que la région 5'-proximale et l'IRES_{5'UTR} sont deux fragments d'ARN de taille comparable (352 et 357 nucléotides respectivement) qui sont issus de la même molécule de départ. Ensuite, parce que ces deux fragments forment deux domaines structurellement indépendants qui se replient sans interactions inter-domaines. De plus, nous avons établi précédemment que la région 5'-proximale (1-356) n'est pas nécessaire à l'activité IRES_{5'UTR}. Elle peut être totalement enlevée sans que l'activité de l'IRES ne soit affectée (*publication 1* et **Figure 27A**). Enfin nous avons validé que le domaine 5'-proximal ne permet pas l'assemblage de complexe d'initiation de la traduction (**Figure 27B**). Ceci a été vérifié en incubant le domaine 5'-proximal et l'IRES_{5'UTR} radio-marqués dans les extraits acellulaires de traduction de cellules S2 de drosophile. L'assemblage des complexes ainsi formés a été analysé sur gradient de sucrose. Nous constatons que l'IRES_{5'UTR} permet la formation efficace de complexes d'initiation 48S et 80S. La région 5'-proximale, quant à elle, recrute un

complexe d'une taille intermédiaire (Y). Ce domaine étant inactif en traduction (**Figure 27A**), le complexe 'Y' n'est donc probablement pas un complexe d'initiation de la traduction.

L'ensemble des éléments énoncés dans ce paragraphe a conduit à l'utilisation de la région 5'-proximale de l'extrémité 5'UTR du CrPV comme ARN de référence dans la caractérisation des partenaires protéiques spécifiques à chaque sous-fragment de la région 5'UTR du CrPV.

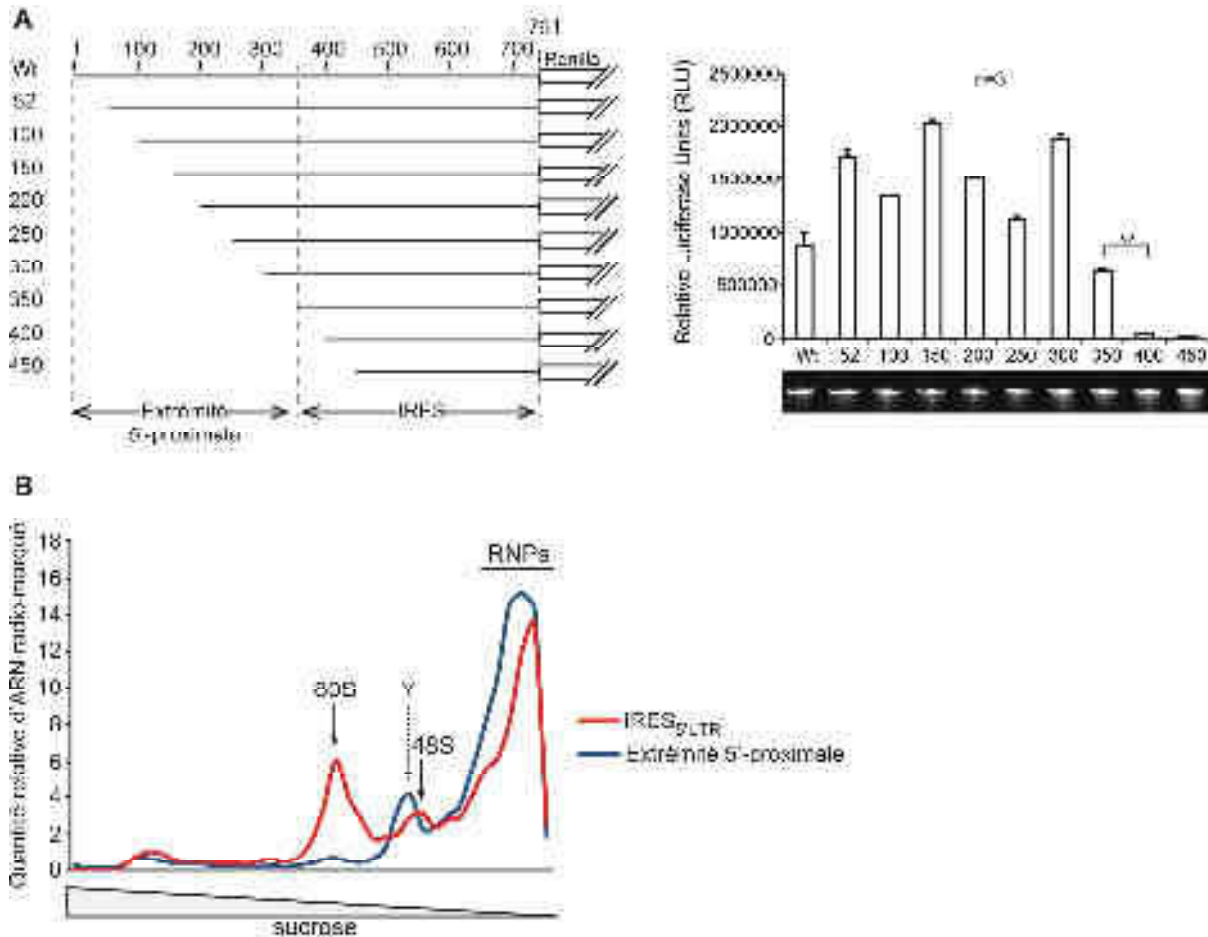


Figure 27 : Le domaine 5'-proximal et l'IRES_{5'UTR} ont des fonctions distinctes. (A) Seule la partie IRES permet l'initiation de la traduction (adapté de la *publication 1*). Des tests de traduction *in vitro* dans des extraits de cellules S2 de drosophile ont été réalisés avec des fragments d'ARN tronqués progressivement du côté 5' (représentées à gauche). L'expression du gène rapporteur est représentée sur l'histogramme à droite. Les tests de traduction *in vitro* sont réalisés en triplicatas (n=3). Pour chaque ARN testé l'intensité de l'activité luciférase (RLU) est normalisée par rapport à son intensité dans un test sans ARN. Les écarts types sont calculés à partir des valeurs de RLU des triplicatas (**valeur statistique p < 0,005 basée sur un test de Student). (B) Seule la partie IRES permet de recruter des complexes d'initiation de la traduction 48S et 80S. Les deux ARN, IRES (en rouge) et domaine 5'-proximal (en bleu) ont été marqués au ³²P à leur extrémité 3'. Les complexes ARN-protéines ont été assemblés dans des extraits acellulaires, puis séparés sur gradient de sucrose. Aucun inhibiteur de la traduction n'a été utilisé. La quantité d'ARN présente dans chaque fraction a été déterminée par comptage radioactif Cerenkov. Le graphique représente le pourcentage d'ARN radioactif retrouvé dans chaque fraction par rapport à l'ARN total introduit au départ de l'expérience. Les fractions contenant les complexes d'initiation 80S, 48S et Y sont indiquées par des flèches.

1.3.3. Identification des partenaires protéiques de l'IRES_{5'UTR} lors de l'initiation de la traduction

1.3.3.1. Les inhibiteurs de la traduction utilisés dans les analyses

Nous avons réalisé des expériences différentielles d'assemblages de complexes d'initiation de la traduction dans le but de comparer de manière systématique les partenaires protéiques spécifiques de la partie 5'-proximale à ceux spécifiques de l'IRES_{5'UTR}. Cela a été réalisé en présence de trois inhibiteurs de la traduction différents qui permettent de bloquer les complexes d'initiation aux stades 43S, 48S ou 80S.

L'édéine :

L'édéine, bloque la formation de la particule 48S en empêchant l'interaction entre l'ARNt^{Met} initiateur et le codon AUG de démarrage de la traduction (**Figure 28**). L'édéine se loge entre les sites P et E du ribosome et induit des changements conformationnels de l'ARN ribosomique. Par conséquent, l'édéine empêche le positionnement de l'ARNm dans le centre de décodage. Il en résulte un blocage de l'interaction codon-anticodon de l'ARNt^{Met} initiateur sur l'AUG. Ainsi, l'édéine bloque efficacement la traduction des messagers initiant avec l'ARNt^{Met}_i dans le site P du ribosome. En revanche, l'édéine n'inhibe pas les IRES de type IV car ils ne nécessitent pas l'interaction entre l'ARNt^{Met}_i et le codon d'initiation. Leur PKI mime l'interaction codon-anticodon et se loge dans le site A du ribosome (Dinos et al., 2004; Garreau de Loubresse et al., 2014; Wilson et al., 2000). En résumé, l'édéine permet d'accumuler les complexes d'initiation qui se forment en amont de l'interaction codon-anticodon, c'est à dire les particules 43S.

Le GMP-PNP :

Le GMP-PNP, est un analogue non hydrolysable du GTP. Le complexe formé par le facteur d'initiation eIF5B et cet inhibiteur ne peut plus se détacher du 48S. Par conséquent, le GMP-PNP empêche l'assemblage de la grande sous-unité (60S) du ribosome (**Figure 28**) et il entraîne donc l'accumulation de particules 48S bloquées sur le codon AUG de démarrage.

La cycloheximide :

La cycloheximide nous sert à étudier la formation des 80S sur l'IRES. Cet antibiotique inhibe l'étape de translocation du ribosome en se liant à l'ARNt déacylé présent dans le site E du ribosome (Schneider-Poetsch et al., 2010; Garreau de Loubresse et al., 2014). Lorsque la cycloheximide est rajoutée avant l'ARNm dans un milieu réactionnel, la première étape de translocation du ribosome est bloquée. Par conséquent, l'ajout de cycloheximide dans les

expériences d'assemblage de complexes d'initiation conduit à l'accumulation de complexes 80S bloqués sur le codon d'initiation AUG. La cycloheximide empêche l'entrée en élongation du ribosome (**Figure 28**).

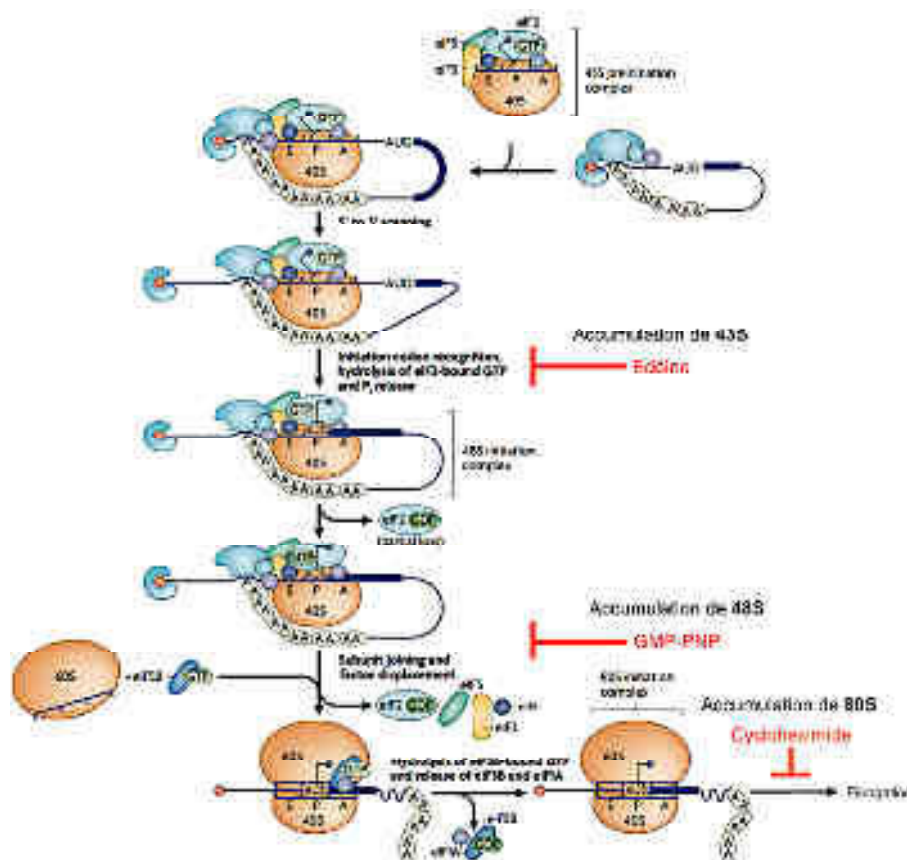


Figure 28 : Les inhibiteurs de la traduction utilisés pour la formation de complexes 43S, 48S et 80S. (Figure adaptée de Jackson *et al.*, 2010). L'édicine est un inhibiteur qui empêche la reconnaissance entre le codon AUG initiateur et l'anticodon de l'ARN^{tMet}, ainsi l'initiation de la traduction est bloquée au stade où la particule 43S est présente sur l'ARN messager. Le GMP-PNP empêche le relargage des eIF et l'association de la grande sous-unité ribosomique 60S. Le complexe de pré-initiation 48S s'accumule sur l'ARNm. La cycloheximide est un antibiotique qui bloque la translocation du ribosome et bloque l'initiation de la traduction au stade 80S.

1.3.3.2. Les partenaires protéiques spécifiques de l'IRES_{5'UTR}

La méthode de purification des complexes et d'analyse par spectrométrie de masse utilisée est décrite dans les articles Gross *et al.*, 2017 ; Prongidi-Fix *et al.*, 2013 et Chicher *et al.*, 2015. La composition de complexes ARN-protéines purifiés a été déterminée par spectrométrie de masse. La **figure 29** présente les résultats de spectrométrie de masse. En comparant les protéines qui interagissent spécifiquement avec chaque ARN lors des blocages avec les trois inhibiteurs nous déduisons trois conclusions principales :

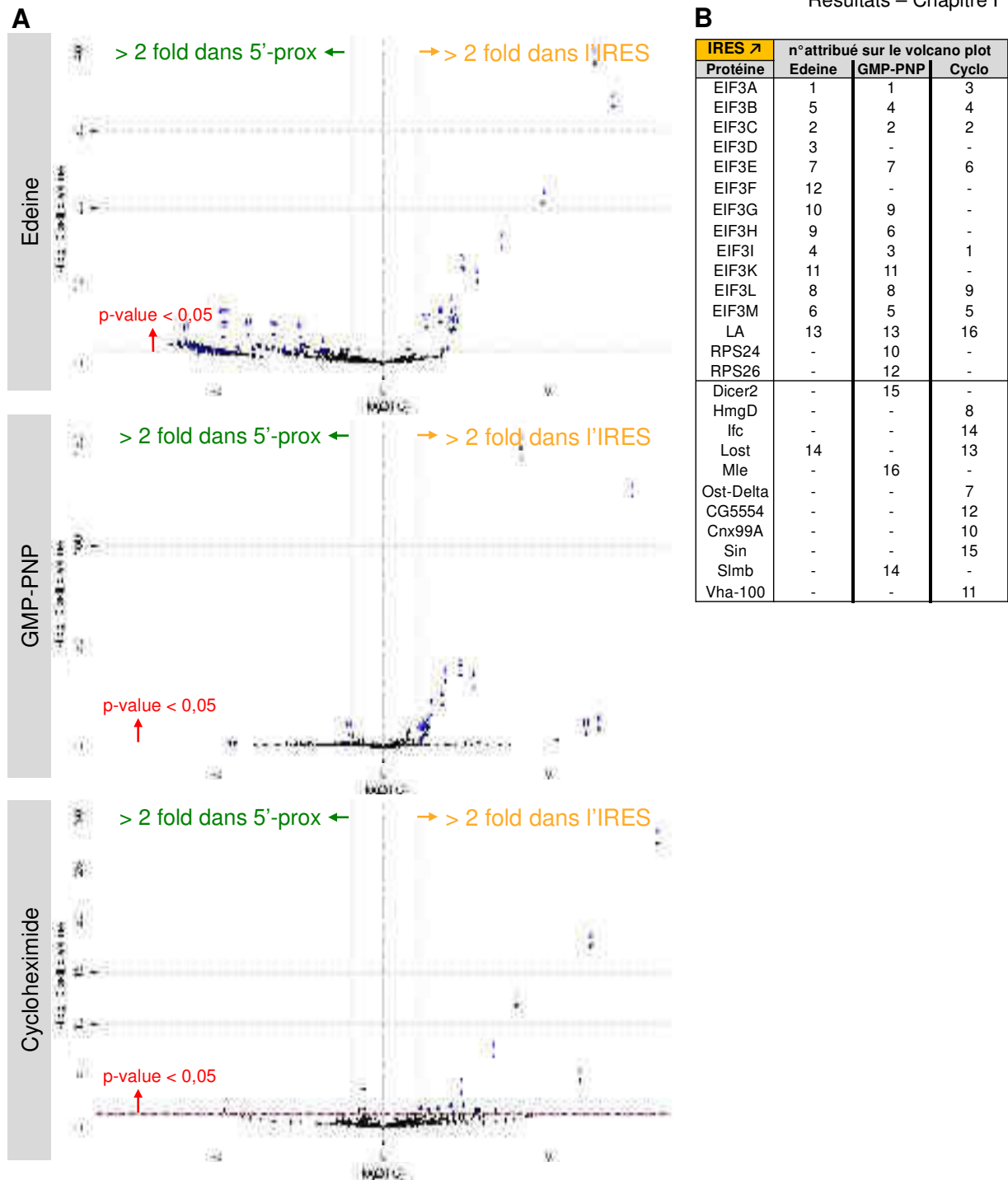


Figure 29 : Le facteur d'initiation eIF3 est recruté spécifiquement sur l'IRES_{5'UTR}. (A) Représentation graphique des résultats de spectrométrie de masse : protéines présentes sur le domaine 5'-proximal (5'-prox à gauche sur les 'volcano plots') et sur l'IRES_{5'UTR} (IRES à droite sur les 'volcano plots') aux stades 43S (édeine), 48S (GMP-PNP) et 80S (Cycloheximide) de l'initiation. Sur les 'volcano plots' le Fold Change (FC) d'une protéine indique si elle est plus présente spécifiquement (enrichie) sur l'un ou l'autre des ARN testés (5'prox ou IRES). Les points bleus signalés par des chiffres indiquent les protéines présentes significativement (valeur statistique p ou 'p-value' < 0,05) et enrichies plus de 2 fois (> 2 fold) sur chacun des ARN. Les expériences de spectrométrie de masse sont réalisées en triplicatas pour chaque échantillon, les données présentées sont la moyenne des triplicatas. (B) Le tableau donne le détail des protéines enrichies sur l'IRES_{5'UTR} qui sont indiquées par les points bleus et les chiffres sur les 'volcano plots'. Les protéines ayant une fonction dans la traduction sont regroupées dans la première moitié du tableau. Le détail des protéines enrichies sur le domaine 5'-proximal ainsi que leur nombre de spectres se trouvent dans l'annexe III.

Premièrement, nous observons sur les ‘volcano plots’ que, quel que soit l’inhibiteur testé, seul l’IRES_{5’UTR} recrute spécifiquement des protéines impliquées dans la formation des complexes d’initiation de la traduction. Le facteur d’initiation de la traduction eIF3 et l’ITAF La ont notamment été observés (**Figure 29A**, **figure 29B** et **figure supplémentaire** en annexe). Ceci vient corroborer les résultats du paragraphe précédent indiquant que seule la partie IRES de la région 5’UTR du CrPV est en mesure de recruter des complexes d’initiation de la traduction et d’initier la traduction du gène en aval. Les protéines du ribosome ont elles aussi été détectées sur l’IRES_{5’UTR}. Cependant, elles ont aussi été détectées en quantités équivalentes sur la région 5’-proximale. C’est pourquoi aucune protéine ribosomale n’est visible sur les ‘volcano plot’. Ceci semble indiquer que le ribosome se fixe de manière aspécifique à tous les ARN de ces analyses.

Deuxièmement, en nous focalisant sur les partenaires protéiques de l’IRES_{5’UTR}, nous constatons la présence du facteur d’initiation eIF3. L’élément nouveau ici est que le nombre de sous-unités d’eIF3 ainsi que leur nombre de spectre diminue fortement lorsque l’on va vers le 80S. En présence d’édéine, douze des treize sous-unités d’eIF3 sont significativement enrichies sur l’IRES_{5’UTR}. Par exemple, la sous-unité eIF3f, qui est la moins abondante, présente 32 spectres et la sous-unité eIF3a, qui est la plus abondante, en possède 207 (**Figure 30**, Édéine). En présence de GMP-PNP ou de cycloheximide le nombre de spectre de chaque sous-unité de eIF3 diminue progressivement jusqu’à disparaître totalement pour les sous-unités eIF3d, f, g, h et k (**Figure 30**, GMP-PNP et cycloheximide ‘Cyclo’). Les sous-unités restantes, a, b, c, e, i, l, et m sont présentes avec un nombre de spectres très faible (entre 14 et 50 spectres), mais il s’agit certainement de résidus d’une contamination par eIF3 du 48S. Nous observons en effet dans les analyses en gradient (**Figure 31**, courbe IRES_{5’UTR}) que le blocage à la cycloheximide n’est pas total et qu’il reste une quantité détectable de complexes 48S sur l’IRES_{5’UTR} lorsqu’il est bloqué au stade 80S. L’ensemble de ces résultats suggère ainsi que le facteur d’initiation eIF3 sert au recrutement du ribosome 43S sur l’IRES_{5’UTR} pour former le 48S puis se décroche du complexe au stade 80S.

Le troisième et dernier point important est que la sous-unité eIF3j n’a pas été détectée dans les complexes bloqués en édéine, GMP-PNP ou en cycloheximide. Cette sous-unité est toujours absente des complexes d’initiation fixés à l’IRES_{5’UTR}, alors que les 12 autres sous-unités de eIF3 sont présentes au début de l’initiation. L’absence d’eIF3j peut s’expliquer car cette sous-unité est en périphérie du complexe multifactoriel eIF3, elle est labile et se dissocie facilement (cette hypothèse sera discutée plus en détails dans la partie *discussion*).

IRES ↗	nb de spectres (moyenne)		
	Edeine	GMP-PNP	Cyclo
EIF3A	207,0	225,7	42,3
EIF3B	106,0	97,3	28,7
EIF3C	141,5	146,0	43,3
EIF3D	122,5	-	-
EIF3E	84,0	67,3	22,0
EIF3F	32,5	-	-
EIF3G	60,5	44,7	-
EIF3H	64,0	80,7	-
EIF3I	112,5	116,3	50,3
EIF3K	40,5	34,7	-
EIF3L	74,0	58,0	14,0
EIF3M	93,5	85,7	24,7

Figure 30 : Le facteur d'initiation eIF3 est recruté spécifiquement sur l'IRES_{5'UTR} au début de l'initiation puis se décroche lorsque le ribosome complet 80S est formé. Liste des sous-unités de eIF3 enrichies sur l'IRES_{5'UTR} pour chaque condition expérimentale testée (Édeine, GMP-PNP et cycloheximide 'Cyclo'). L'histogramme jaune superposé ainsi que le chiffre attribué à chaque protéine représente la moyenne du nombre de spectres (nb de spectres (moyenne)) des trois expériences de spectrométrie de masse réalisées pour chaque échantillon. En spectrométrie de masse le nombre de spectres d'une protéine est relatif à son abondance dans l'échantillon.

1.3.4. Étude du rôle du pseudo-nœud du domaine II dans le recrutement du ribosome

Nous avons montré précédemment (*publication 1*) la présence d'une interaction à distance de type pseudo-nœud dans le domaine II de l'IRES_{5'UTR}. De plus, les mutations qui détruisent le pseudo-nœud (m1 et m2) déplient la structure de l'IRES_{5'UTR}, empêchent le recrutement du 48S sur l'IRES_{5'UTR} et sont délétères pour la traduction. Au contraire, les mutations compensatoires (m3 et m4) qui reforment le pseudo-nœud dans un repliement identique à celui de l'IRES_{5'UTR} sauvage (WT) permettent d'assurer le recrutement du 48S ainsi qu'une traduction efficace (pour les constructions m1 à m4 voir **figure 31A**). Ce pseudo-nœud a donc un rôle essentiel dans le recrutement du complexe d'initiation 48S et dans l'activité traductionnelle de l'IRES_{5'UTR}.

Afin de déterminer précisément à quel stade de l'initiation est impliqué le pseudo-nœud, j'ai également étudié le comportement des mutants du pseudo-nœud lors de la formation de complexes d'initiation 80S. Comme précédemment, les quatre mutants ainsi que l'IRES_{5'UTR} sauvage ont été incubés dans des extraits acellulaires de cellules S2. Les complexes d'initiation ont alors été bloqués au stade 80S sur l'AUG à l'aide de la cycloheximide. Les complexes protéines-IRES_{5'UTR} ainsi bloqués ont été séparés sur gradient de sucrose. Nous observons (**Figure 31B**) qu'aucun des 4 mutants du pseudo-nœud ne permet l'assemblage de 80S de manière aussi efficace l'IRES_{5'UTR} sauvage. Ceci est particulièrement intrigant

pour le mutant m4, sur qui le 48S s'assemble aussi efficacement que sur l'IRES_{5'UTR} sauvage (*publication 1*). Notons que les séquences du pseudo-nœud ne sont pas les mêmes dans l'IRES_{5'UTR} sauvage et dans le mutant m4. Ainsi, ces résultats suggèrent que non seulement la structure du pseudo-nœud, mais aussi sa séquence, sont requises pour la formation du 48S ou pour une étape ultérieure conduisant à la formation efficace du 80S. Cette conclusion est en accord avec le fait que le pseudo-nœud n'est pas nécessaire au recrutement du facteur d'initiation eIF3, qui se fait lors d'une étape antérieure à la formation du 48S (*publication 1*). Lors des tests fonctionnels *in vitro* et *in vivo*, nous avons cependant observé que les mutants m3 et m4 sont presque aussi actifs que l'IRES_{5'UTR} sauvage (*publication 1*). La différence d'efficacité d'assemblage de particules 80S observée sur les gradients pourrait résulter d'une instabilité des complexes 80S bloqués par la cycloheximide qui se dissocieraient des ARN mutants m3 et m4. En revanche, lors des tests fonctionnels, les particules 80S assemblées ne sont pas bloquées, elles pourraient donc immédiatement s'engager en phase d'élongation. Ceci expliquerait pourquoi nous observons une traduction correcte des mutants m3 et m4 (*publication 1*).

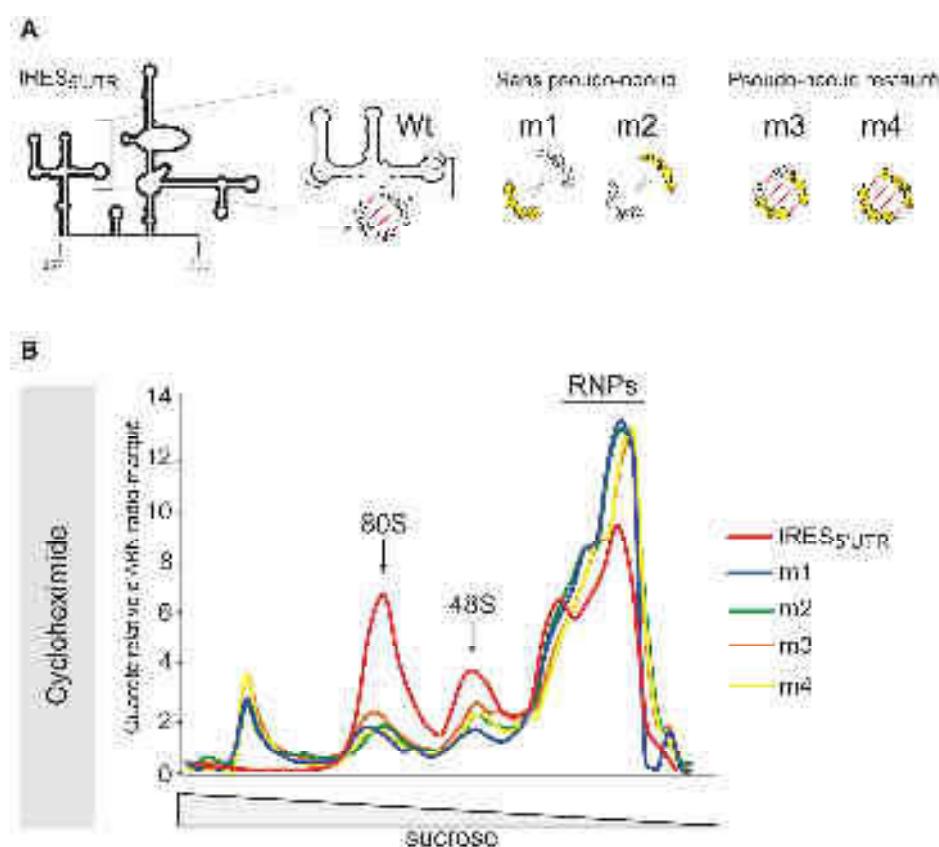


Figure 31 : Étude du rôle du pseudo-nœud dans la formation du complexe 80S sur l'IRES_{5'UTR}. (A) Représentation schématique de l'IRES_{5'UTR} sauvage et des mutants du pseudo-nœud. (B) Analyse des complexes d'initiation 80S sur les ARN marqués au ³²P par gradient de sucrose. L'emplacement du 80S et du 48S est indiqué par des flèches. Les résultats des gradients sont une moyenne de 3 expériences indépendantes d'assemblage et de séparation des complexes d'initiation. Remarque : les pics présents dans les fractions

lourdes des gradients (à gauche) sont liés au fait que lors de la collecte des fractions, un capillaire perce le gradient par le haut et passe au travers des fractions légères avant d'arriver au fond du tube. Par conséquent, ces pics au fond des tubes correspondent à une légère contamination due à notre méthode de collecte des fractions. Il ne faut donc pas tenir compte de ces pics dans l'interprétation des résultats.

En résumé, ces études semblent suggérer que la structure du pseudo-nœud est importante pour la formation du 48S et que sa séquence est également importante pour la formation du 80S. Cette interprétation est notamment compatible avec une ouverture du pseudo-nœud après la formation du complexe 48S rendant ainsi accessible ces séquences qui seraient importantes pour la formation du 80S (Cette hypothèse sera discutée plus en détail dans la section *discussion et perspectives*). À ce stade, ces conclusions devront être confirmées par des expériences complémentaires, notamment par des analyses structurales de l'évolution des appariements du pseudo-nœud au cours du processus d'initiation de la traduction.

I.4. Conclusion du chapitre I

L'IRES_{5'UTR} du CrPV s'inscrit dans un contexte structural complexe. Il se situe au sein d'une région 5'UTR constituée de 3 domaines distincts, possédant chacun leur fonction. Le premier domaine est probablement impliqué dans la réplication virale alors que les deux autres domaines constituent l'IRES. Il reste à savoir si ces deux domaines distincts qui composent l'IRES_{5'UTR} sont fonctionnellement indépendants ou non. Nous savons d'ores et déjà que le pseudo-nœud présent dans le domaine II est primordial pour la formation des particules 48S sur l'IRES_{5'UTR}. Son ouverture définitive par l'introduction de mutations cause non seulement l'impossibilité de former des 48S mais inactive aussi l'activité de traduction de l'IRES. De plus, la séquence nucléotidique sauvage favorise une structure globale de l'IRES_{5'UTR} intègre et active qui assure un assemblage efficace du ribosome 80S sur l'AUG et son passage en phase d'élongation. En conclusion de ce chapitre nous proposons que l'IRES_{5'UTR} du CrPV est un IRES de classe III, qui recrute le facteur d'initiation eIF3. Ce dernier est nécessaire en début d'initiation pour recruter le ribosome puis est éjecté du complexe d'initiation lorsque le ribosome complet 80S est prêt à entrer en élongation.

II. Étude de la fonction des domaines II et III de l'IRES_{5'UTR} du CrPV

II.1. Objectifs

Pour aller plus loin dans la caractérisation de l'IRES_{5'UTR} du CrPV et préciser son mécanisme moléculaire d'action, je me suis intéressée d'avantage au rôle des domaines II et III. Pour ce faire, les domaines II et III de l'IRES ont été étudiés séparément. J'ai notamment utilisé des techniques de traduction *in vitro* et d'assemblages de complexes d'initiation sur les deux domaines dans les systèmes acellulaires de cellules S2 de drosophile décrits précédemment.

II.2. Résultats

II.2.1. Étude du rôle respectif des domaines II et III dans le fonctionnement de l'IRES

Pour étudier le rôle du domaine III de l'IRES_{5'UTR} du CrPV lors de la traduction, nous avons testé dans un premier temps l'activité traductionnelle *in vitro* du gène rapporteur Rénilla luciférase placé sous la dépendance de l'IRES_{5'UTR} tronqué contenant uniquement le domaine II. La traduction *in vitro* réalisée avec ce mutant montre une diminution importante de l'activité de l'IRES lorsque le domaine III (501-709) est enlevé (**Figure 32**). Le domaine II à lui seul n'est donc pas suffisant à l'activité traductionnelle de l'IRES_{5'UTR}, le domaine III y joue lui aussi un rôle déterminant.

Dans un second temps, afin de tester si le domaine II et le domaine III isolés pouvaient assembler le ribosome 80S, nous les avons incubés dans des extraits de traduction bloqués à la cycloheximide. Les complexes d'initiation de la traduction formés ont été fractionnés sur gradient de sucrose et comparés à ceux obtenus avec l'IRES_{5'UTR} sauvage (**figure 33B**). Nous avons constaté qu'aucun des domaines n'assemble de complexe d'initiation 80S lorsqu'ils sont isolés, pas même le domaine III qui contient le codon initiateur AUG. Ensuite, afin de déterminer si le domaine II et III pouvaient agir en *trans*, nous les avons mélangé en présence des mêmes extraits de traduction bloqués à la cycloheximide. Encore une fois, aucune particule 80S ne se forme sur l'AUG du domaine III (**figure 33C**). Ces résultats démontrent ainsi que les domaines II et III doivent être reliés pour être fonctionnels et qu'ils ne peuvent en aucun cas être considéré comme deux entités fonctionnelles indépendantes capables d'agir à distance sur un AUG initiateur.

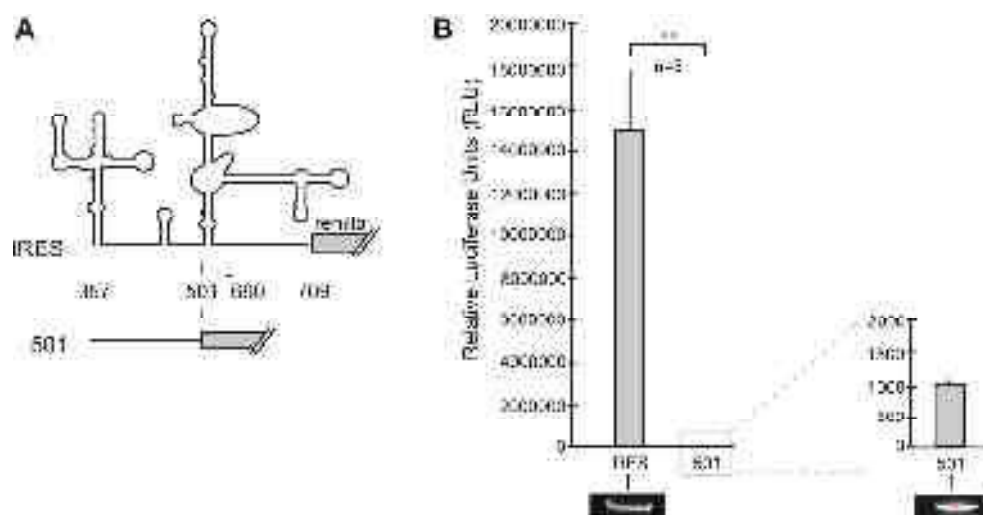


Figure 32 : Le domaine III est indispensable à l'activité traductionnelle de l'IRES_{5'UTR}. (A) Les constructions ARN appelées 'IRES' pour l'IRES_{5'UTR} complet (357-709) et '501' pour l'IRES tronqué, dont les nucléotides 501 à 709 ont été supprimés, sont utilisées lors (B) des tests de traduction *in vitro* dans des extraits de cellules de drosophile. L'activité luciférase du gène rapporteur dans chaque échantillon (IRES et 501) est normalisée par rapport à l'activité luciférase d'un échantillon sans ARN (bruit de fond). L'analyse a été réalisée en triplicatas (n=3). La valeur statistique p indique si les différences d'expression du gène rapporteur sont significatives (** p < 0,005 selon un test de Student).

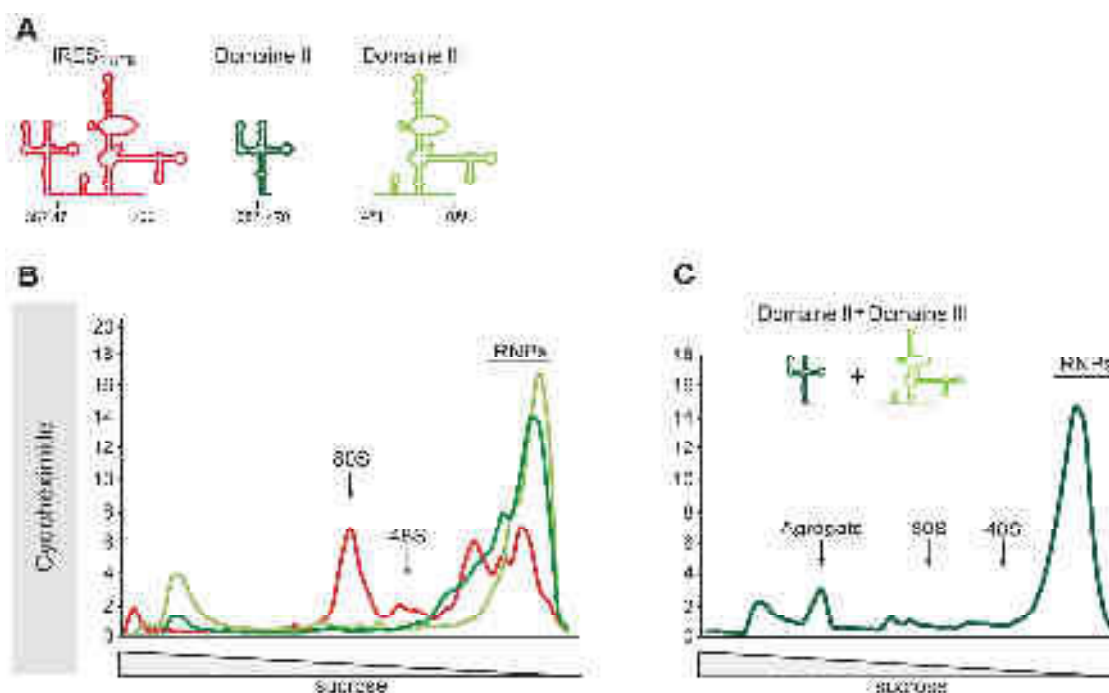


Figure 33 : Les domaines II et III de l'IRES_{5'UTR} du CrPV ne sont pas fonctionnels lorsqu'ils sont séparés l'un de l'autre. (A) Constructions ARN utilisées dans l'expérience. (B) Assemblage de complexes d'initiation bloqués à la cycloheximide sur les domaines II et III indépendants, (C) puis sur les domaines II et III mélangés en quantités équivalentes. Les complexes d'initiation de la traduction assemblés sur l'IRES_{5'UTR} sauvage puis séparés sur gradient de sucrose sont représentés en rouge, ceux obtenus après assemblage sur le domaine II sont en vert, sur le domaine III en vert clair, et sur les domaines ajoutés en *trans* en vert foncé. Dans cette dernière expérience, c'est le domaine III (en vert clair) qui est radioactif.

II.2.2. Identification des partenaires protéiques spécifiques des domaines II et III

Afin d'identifier les partenaires protéiques spécifiques des domaines II et III de l'IRES_{5'UTR} lors de l'initiation de la traduction, les complexes ribonucléo-protéiques assemblés ont été purifiés et leur composition en protéines a été déterminée par spectrométrie de masse. Je me suis focalisée sur l'étude de complexes d'initiation bloqués à l'édéine et au GMP-PNP.

Tout d'abord, dans les complexes bloqués à l'édéine (**Figure 34A et C**) les sous unités a, c et e du facteur d'initiation eIF3 sont enrichies significativement sur le domaine III. Ceci suggère que le facteur eIF3 interagit préférentiellement avec le domaine III de l'IRES_{5'UTR}. De plus, l'ITAF La est également enrichie sur le domaine III. Cette protéine est responsable du recrutement du ribosome 40S sur les IRES du poliovirus et du HCV (Costa-mattioli et al., 2004; Meerovitch et al., 1993; Meerovitch et al., 1989). Nous pouvons donc conclure de ces résultats que le domaine III recrute spécifiquement le facteur d'initiation de la traduction eIF3 et la protéine La. À partir de là, nous pouvons supposer que le domaine III est responsable du recrutement du complexe 43S, grâce à une interaction directe avec ces protéines. Curieusement, en présence de GMP-PNP le domaine III ne fixe de manière significative aucune protéine ayant un rôle dans l'initiation de la traduction (**Figure 34B et C**). Rappelons que ce fragment de l'IRES_{5'UTR} contient le codon initiateur AUG. Si le complexe 43S recruté grâce à eIF3 et La n'est pas déposé efficacement sur l'AUG du domaine III, alors cela suggère que cette étape nécessite la présence du domaine II pour y positionner correctement la particule 43S. Ensuite (**Figure 34B et C**), la sous-unité eIF3i ainsi que la protéine La ont été détectées sur le domaine II dans les complexes ribonucléo-protéiques (RNP) bloqués au GMP-PNP uniquement. Ces résultats semblent indiquer que le domaine II possède lui aussi une affinité, bien que plus faible, pour les facteurs eIF3 et La. Ceci semble corroborer le fait que les deux domaines de l'IRES_{5'UTR} du CrPV soient nécessaires à l'assemblage efficace d'une particule 48S.

En résumé, les facteurs protéiques La et eIF3 seraient donc impliqués dans l'interaction directe du 43S avec le domaine III. Cependant, sans la présence du domaine II, qui fixe également La et eIF3i, le 48S ne se forme pas sur l'AUG dans le domaine III. Ainsi, nous pouvons conclure que les domaines II et III agissent en synergie pour recruter et fixer le 48S grâce à la protéine La et au facteur eIF3.

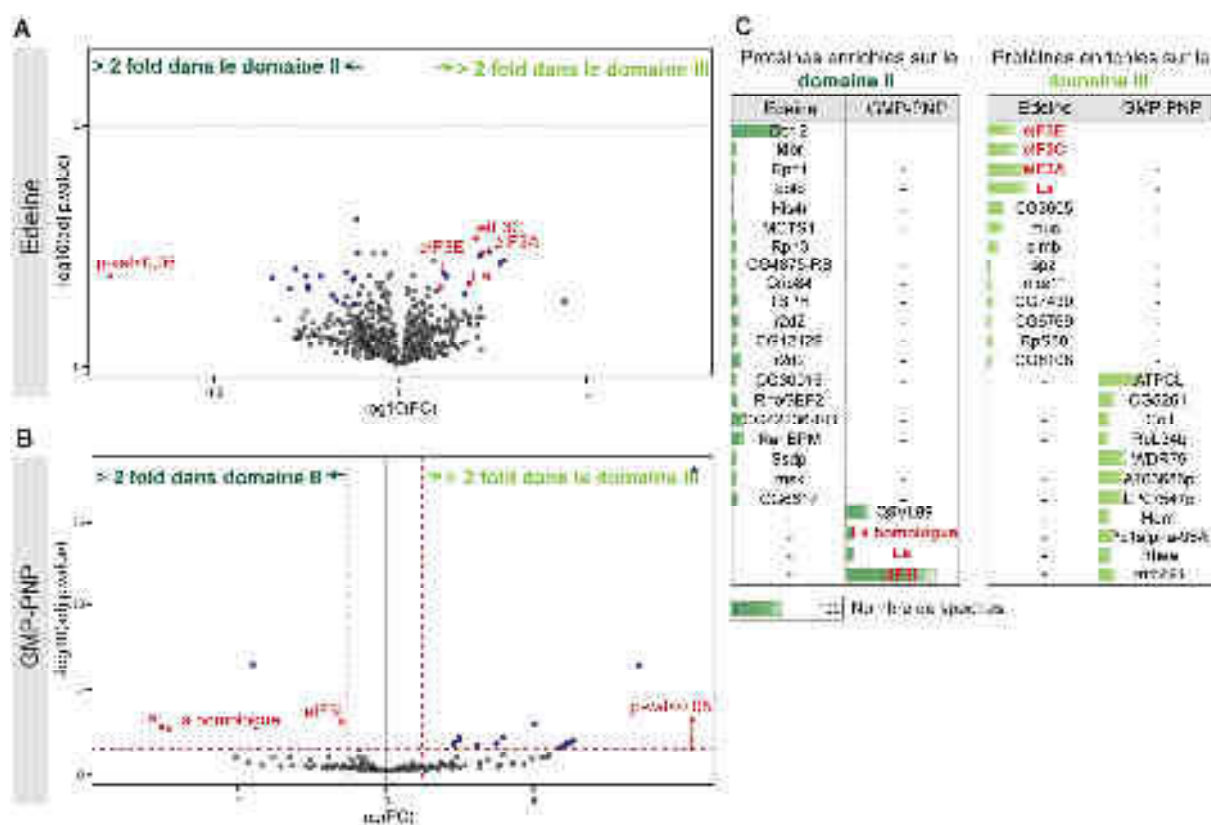


Figure 34 : Le facteur d'initiation eIF3 se fixe au domaine III de l'IRES_{5'UTR} du CrPV. (A-B) Les graphiques de type 'volcano plots' représentent les données de spectrométrie de masse obtenues sur les complexes de pré-initiation bloqués à l'édéine ou au GMP-PNP sur le domaine II (vert foncé) et le domaine III (vert clair). Les protéines enrichies plus de 2 fois dans un domaine (> 2 fold) et significativement présentes dans l'analyse (valeur statistique p ou 'p-val' < 0,05) sont en bleu foncé. En rouge sont indiqués les protéines impliquées dans les complexes de pré-initiation de la traduction 43S ou 48S. Les analyses de spectrométrie de masse sont réalisées en triplicatas pour chaque échantillon, les résultats présentés sont les moyennes des valeurs des triplicatas. **(C)** Les protéines enrichies plus de 2 fois sur les domaines II et III pour chaque condition expérimentale sont listées dans le tableau. Les histogrammes verts indiquent le nombre moyen de spectres des 3 réplicats pour chaque protéine enrichie.

II.2.3. Étude du rôle de eIF2 dans la traduction guidée par l'IRES_{5'UTR}

Le fait que le facteur d'initiation eIF2 n'apparaisse pas sur les 'volcano plots' ci-dessus est intrigant. Nous savons que le facteur d'initiation eIF2-GTP est présent dans le complexe de pré-initiation 43S et qu'il joue un rôle crucial dans la reconnaissance codon-anticodon et l'assemblage du 48S grâce à l'hydrolyse de son GTP. De plus, les auteurs Garrey *et al.*, ont montré que ce facteur serait important pour la traduction guidée par l'IRES_{5'UTR}.

Dans les résultats de spectrométrie de masse, eIF2 est bien présent dans les complexes 43S bloqués en édéine et dans les complexes 48S bloqués au GMP-PNP. Nous l'observons même avec un nombre de spectres important en présence de GMP-PNP (**Figure 35**).

Cependant, dans les deux expériences, le nombre de spectres de eIF2 identifiés sur les domaines II et III est quasiment identique. Le facteur eIF2 est donc présent de manière identique sur les deux domaines de l'IRES_{5'UTR}, ceci explique pourquoi les 'volcano plots' ne permettent pas de le visualiser. Ce résultat suggère que le facteur eIF2 est capable d'interagir à la fois le domaine II et avec le domaine III.

Nous avons vu précédemment que le domaine III permet l'assemblage des complexes 43/48S, alors que le domaine II ne le peut pas. Le facteur eIF2, de poids moléculaire élevé, est composé de trois sous-unités α , β et γ . Par conséquent il est possible qu'une ou plusieurs de ses sous-unités interagissent directement avec le domaine II, et aussi avec le domaine III de l'IRES_{5'UTR}. Les résultats de ce paragraphe confirment que le facteur eIF2 est nécessaire pour l'initiation de la traduction guidée par l'IRES_{5'UTR}.

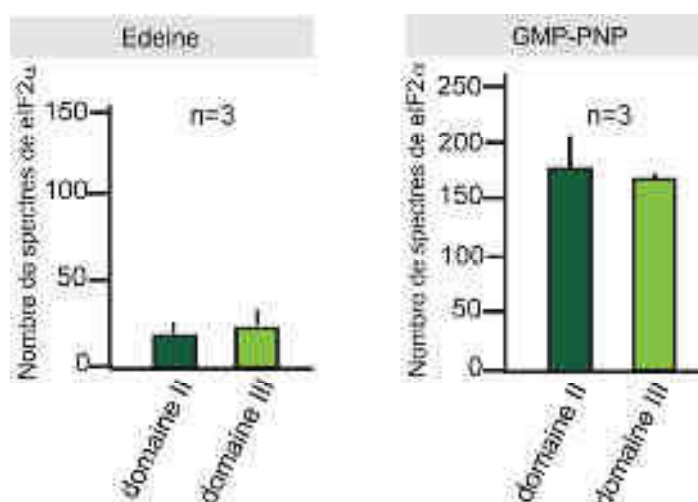


Figure 35 : Le facteur d'initiation eIF2 est identifié sur les deux domaines de l'IRES_{5'UTR} du CrPV. Nombre de spectres de eIF2- α retrouvés sur les domaines II (vert foncé) et III (vert clair) de l'IRES_{5'UTR} pour chaque condition expérimentale (édéine et GMP-PNP). Les nombres de spectres des sous-unités eIF2- β et - γ suivent la même tendance que ceux de la sous-unité eIF2- α . Les analyses de spectrométrie de masse sont réalisées en triplicatas pour chaque échantillon, les résultats présentés sont les moyennes des valeurs des triplicatas. Les barres d'erreur représentent les écarts types de ces trois valeurs.

II.2.4. Étude du rôle du pseudo-nœud dans le recrutement de complexes RNP

Nous avons montré précédemment que le pseudo-nœud (PK) présent dans le domaine II joue un rôle crucial pour permettre la formation d'une particule 48S. De plus, le recrutement de eIF3 sur l'IRES se fait même lorsque le PK est muté et ouvert (*publication 1*). Suite aux résultats de la section précédente nous avons voulu tester si l'ouverture du PK pouvait influencer le recrutement de facteurs sur le domaine II isolé. Pour ce faire, les mêmes mutations du pseudo-nœud m1, m2, m3 et m4 que celles décrites précédemment dans le

chapitre I, ont été introduites dans le domaine II isolé (**Figure 36A**). Rappelons que les mutations m1 et m2 ouvrent le pseudo-nœud et les mutations compensatoires m3 et m4 permettent de le reformer. Rappelons également que le domaine II ne contient pas le codon d'initiation AUG, il ne peut donc pas assembler de complexe 48S. La formation de complexes RNP sur les mutants a été analysée par gradients de sucrose.

Ces expériences montrent que lorsque l'ouverture du pseudo-nœud est forcée à l'aide des mutations m1, m2, il y a une augmentation de l'assemblage de complexes RNP dont la taille est comparable à une particule 43S (**Figure 36B**). En revanche, lorsque le pseudo-nœud est fermé comme dans le domaine II sauvage ou avec les mutations m3 et m4, aucune particule de cette taille ne s'assemble. En outre nous observons dans toutes les conditions la présence d'un complexe de masse inconnue mais de faible poids moléculaire (*), dont la composition n'a pas été identifiée à ce jour. Il pourrait éventuellement s'agir d'un facteur impliqué dans l'initiation de la traduction qui se fixe spécifiquement au domaine II. Ainsi, ces expériences montrent que l'ouverture forcée du pseudo-nœud permet la fixation de deux complexes RNP sur le domaine II de l'IRES_{5'UTR} dont la nature reste à déterminer. Afin de mieux comprendre ce phénomène, il conviendra de déterminer quels sont ces facteurs fixés sur les domaines II mutants m1 et m2 en purifiant ces complexes en vue de leur analyse par spectrométrie de masse. Malheureusement ces analyses n'ont pu être réalisées dans le cadre de mon doctorat par manque de temps.

Afin d'aller un peu plus loin dans la compréhension de ce mécanisme, nous avons analysé la composition des complexes assemblés sur l'IRES_{5'UTR} complète sauvage contenant les mutations m1-4 par 'western blot'. En présence d'édéine, des anticorps spécifiques des sous-unités eIF3b et eIF3j ont permis de montrer que les quatre mutants permettent de recruter efficacement le facteur eIF3, en revanche seule l'IRES mutante m2 permet de visualiser la sous-unité labile eIF3j (**Figure 36C**). Ce résultat préliminaire suggère que l'ouverture forcée du pseudo-nœud permettrait une interaction entre des régions de celui-ci et eIF3j. La sous-unité eIF3j est une sous-unité dont le rôle est encore mal compris, cependant il a été montré qu'elle occupe le même emplacement que l'ARN messager dans le ribosome. Par conséquent, la formation d'un complexe 48S requière le décrochage de 3j pour y permettre le positionnement de l'ARN messager. Nos expériences préliminaires suggèreraient que le pseudo-nœud semble impliqué dans cette étape. L'effet observé induit par des mutations introduites dans le pseudo-nœud suggère l'existence d'une structure dynamique à ce niveau qui pourrait être importante dans le mécanisme moléculaire de l'initiation. Cependant, à ce stade la validation de cette hypothèse préliminaire nécessite d'autres expériences complémentaires.

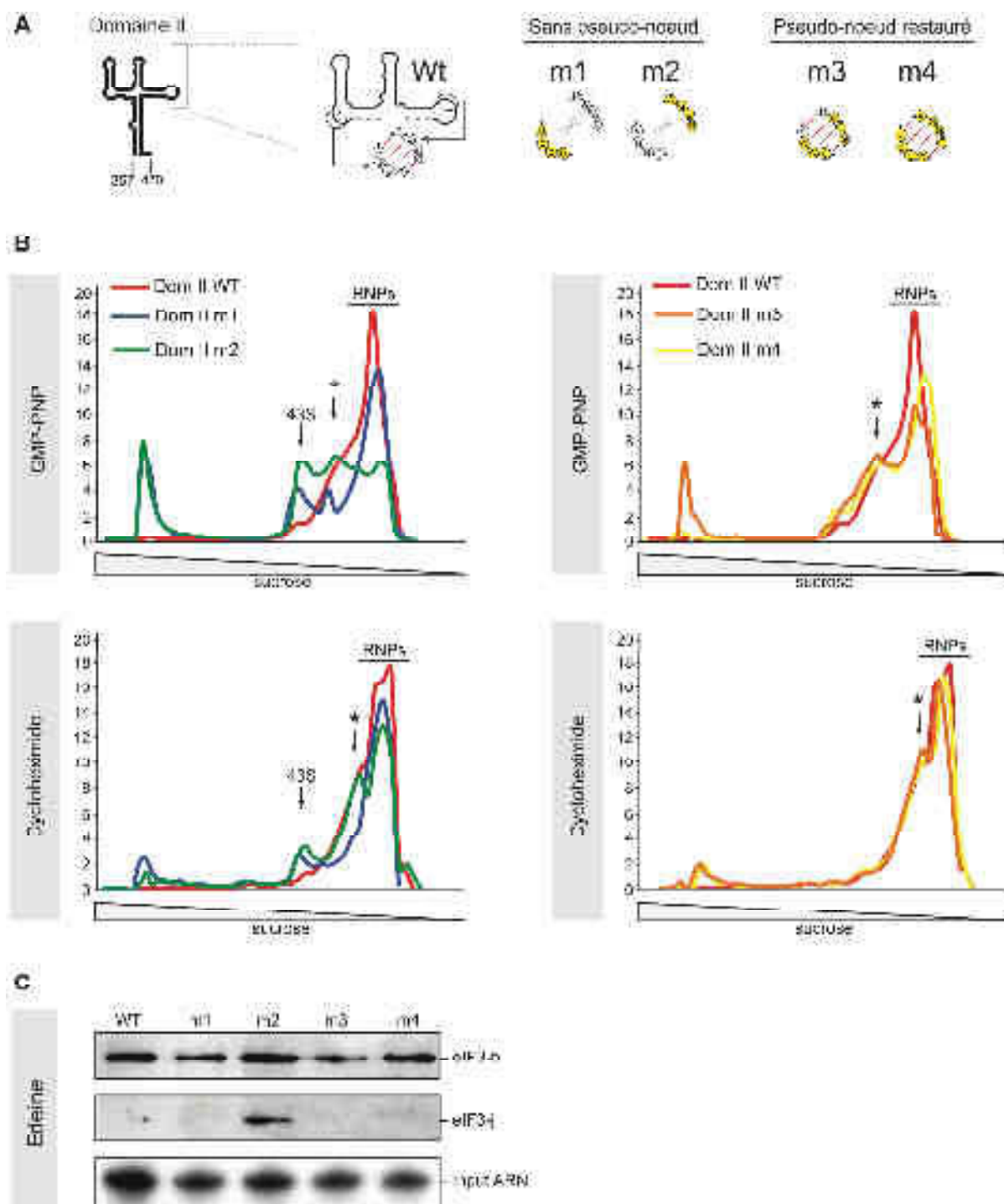


Figure 36 : Le pseudo-nœud présent dans le domaine II de l'IRES_{5'UTR} de CrPV influence la nature des complexes RNP formés. (A) Constructions utilisées pour (B) l'analyse par gradient de sucrose des complexes d'initiation associés au domaine II. Les complexes ont été bloqués au GMP-PNP et à la cycloheximide. Les complexes d'initiation 48S et 80S ainsi qu'un complexe de masse inconnue (*) sont indiqués par des flèches. (C) Analyse par western-blot des sous unités eIF3j et eIF3b assemblées sur l'IRES_{5'UTR} complète au stade d'initiation 43S. Les complexes sont formés de la même manière que ceux analysés en gradients. Résultat préliminaire non reproduit à ce jour.

II.3. Conclusion du chapitre II

En conclusion, les résultats de ce deuxième chapitre permettent de mieux comprendre le rôle des domaines II et III de l'IRES_{5'UTR} pendant l'initiation de la traduction des protéines non structurales du CrPV. Le domaine III recrute spécifiquement le facteur d'initiation eIF3. Nous avons également montré que le domaine II contient un pseudo-nœud qui semble être impliqué dans la formation du complexe 48S. Ces deux domaines semblent donc avoir des fonctions distinctes, ils doivent cependant être liés physiquement pour garantir l'assemblage d'un complexe d'initiation complet. Nos études ont montré que ces domaines agissent en synergie pour permettre de guider l'initiation de la traduction de l'ORF1 du virus.

DISCUSSION ET PERSPECTIVES

DISCUSSION GÉNÉRALE ET PERSPECTIVES

I. Mode d'action de l'IRES_{5'UTR} de CrPV et fonction des domaines structuraux qui le constituent

Les résultats obtenus lors de mon doctorat ont permis de caractériser l'IRES_{5'UTR} du CrPV et de comprendre les bases moléculaires de son mécanisme d'action. La région 5'UTR du CrPV est constituée de 3 domaines distincts qui possèdent chacun leur fonction. Par analogie avec d'autres virus, il est possible que le premier domaine soit probablement impliqué dans la réplication virale. En revanche, nous avons démontré que les domaines II et III constituent l'IRES_{5'UTR}. Cet IRES_{5'UTR} appartient à la classe III des IRES qui recrute spécifiquement et directement le facteur d'initiation eIF3. Son mode d'action s'affranchit de l'étape de 'scanning' pour reconnaître le codon initiateur (voir *publication 1*). Le domaine II de l'IRES_{5'UTR} contient une structure tertiaire de type pseudo-nœud. Ce dernier est essentiel pour la formation du complexe 48S sur l'IRES_{5'UTR}. Lorsque ce pseudo-nœud est détruit (ouvert) par des mutations, la formation du complexe 48S est abolie. Cependant le rôle exact de cette structure en pseudo-nœud reste à définir (voir *résultats chapitre I*). Le domaine III de l'IRES_{5'UTR} adopte un repliement secondaire en longue tige-boucle. L'extrémité 3' de ce domaine contient le codon d'initiation AUG de l'ORF1 du CrPV. Le domaine III est responsable du recrutement spécifique du facteur d'initiation eIF3. Ce dernier est nécessaire en début d'initiation pour recruter le ribosome puis il quitte le complexe d'initiation lorsque le ribosome complet 80S est prêt à entrer en phase d'élongation. Le domaine III fixe également l'ITAF La, une protéine cellulaire décrite comme facteur de recrutement du ribosome 40S sur les IRES. Ainsi, les deux domaines de l'IRES_{5'UTR} semblent avoir des fonctions distinctes, cependant ils doivent être liés physiquement pour garantir l'assemblage efficace d'un complexe d'initiation fonctionnel. Ces deux domaines agissent en synergie pour permettre de guider l'initiation de la traduction de l'ORF1 du virus (voir *résultats chapitre II*). Nous discuterons ci-dessous des résultats présentés dans les *chapitres I et II* de ce manuscrit.

I.1. Le facteur d'initiation eIF3 est recruté sur le domaine III de l'IRES_{5'UTR}

Nous avons constaté que le facteur d'initiation eIF3 est recruté sur l'IRES_{5'UTR} du CrPV en début d'initiation. D'après les données de spectrométrie de masse, il se décroche ensuite du complexe d'initiation une fois le ribosome 80S assemblé (*chapitre I*). Plus précisément, c'est le domaine III de l'IRES_{5'UTR} du CrPV qui est responsable du recrutement spécifique du

facteur eIF3 (*chapitre II*). Nous avons également détecté les sous-unités a, c et e de eIF3 sur ce domaine III à l'étape précoce d'initiation 43S (**Figure 34A**). Des études ont démontré que le facteur eIF3 interagit avec une grande affinité avec le domaine III de l'IRES de HCV (Sun et al., 2011). Plus précisément, la partie apicale du domaine III de l'IRES HCV, qui se replie en jonction à quatre hélices (module IIIabc), établit des contacts avec les sous-unités eIF3a et eIF3c (Erzberger et al., 2014; Hashem et al., 2013; Sun et al., 2013). Au contraire, les sous-unités eIF3b, d, g, i et j ne sont pas en contact direct avec l'IRES HCV, bien qu'elles soient essentielles pour l'intégrité du facteur eIF3 et donc pour sa fonction dans l'initiation de la traduction. Nos résultats apportent une preuve supplémentaire de la ressemblance fonctionnelle entre les IRES du CrPV et du HCV. Le site précis permettant la liaison de eIF3 à l'IRES_{5'UTR} du CrPV reste encore à déterminer. Nous constatons néanmoins que la structure secondaire du domaine III de l'IRES_{5'UTR} du CrPV adopte un repliement en longue tige boucle dont la jonction en triple-hélices apicale ressemble au module IIIabc de l'IRES HCV (**Figures 45**). Nous envisageons d'autres expériences afin de préciser la région du domaine III requise pour la fixation de eIF3 sur l'IRES_{5'UTR} du CrPV. Par exemple, l'effet de délétions de la partie basale du domaine III, ou de mutations de certains nucléotides du domaine III, sur le recrutement de eIF3 sera étudié à l'aide des extraits acellulaires de cellules S2 de drosophile. L'interaction entre l'IRES_{5'UTR} et eIF3 sera étudiée à l'aide d'anticorps spécifiques des différentes sous-unités de eIF3. Nous déterminerons également les zones de l'IRES qui sont en contact avec eIF3 en faisant des « empreintes » à l'aide de sondes RNases spécifiques sur le complexe eIF3-IRES_{5'UTR}. Pour ces expériences, nous devons entreprendre la purification du facteur eIF3 de drosophile à partir de cellules S2.

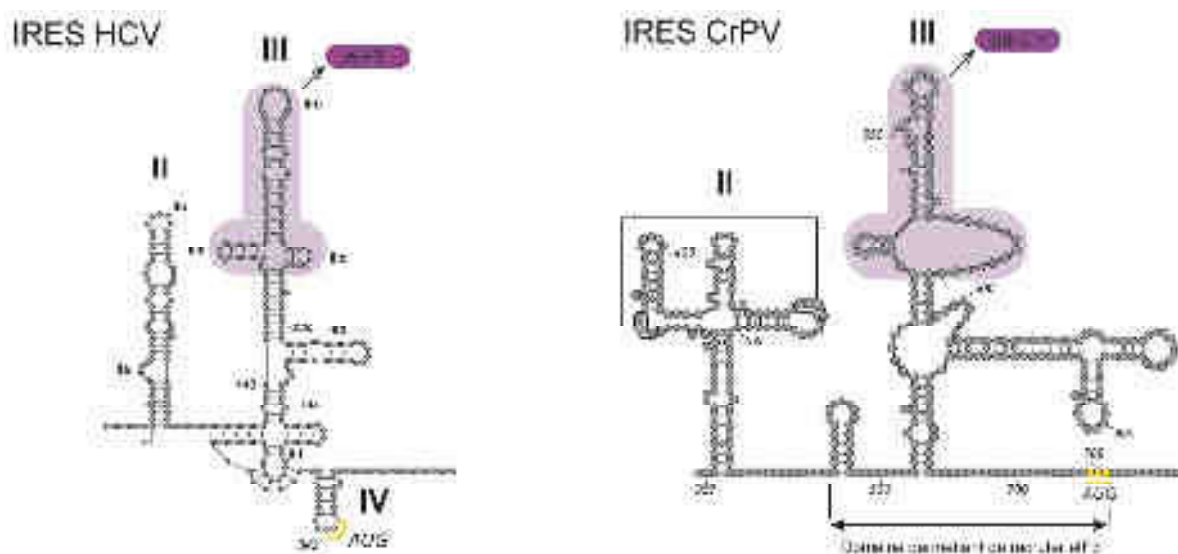


Figure 45 : Le facteur d'initiation eIF3 se fixe au domaine III des IRES du HCV et du CrPV. (À gauche) Le facteur d'initiation eIF3 reconnaît spécifiquement et précisément le module IIIabc de l'IRES du HCV qui est surligné en violet dans la figure. **(À droite)** Des expériences de spectrométrie de masse ont permis d'identifier le

facteur d'initiation eIF3 sur une construction se limitant au domaine III de l'IRES_{5'UTR} du CrPV (indiquée par la double flèche). Le site précis d'interaction entre eIF3 et l'IRES_{5'UTR} n'est pas encore défini. Par analogie avec l'IRES HCV, eIF3 pourrait se fixer sur le module apical du domaine III de l'IRES CrPV (surligné en violet), car ces deux modules sont structurellement similaires.

1.2. Les domaines II et III de l'IRES_{5'UTR} agissent en synergie pour assembler le ribosome 80S sur l'AUG.

Nous avons démontré que le domaine III de l'IRES_{5'UTR} recrute directement le facteur eIF3. Rappelons que ce fragment de l'IRES_{5'UTR} contient le codon initiateur AUG. Nous pourrions imaginer que le facteur eIF3, accroché sur le domaine III, garantisse à lui seul l'ancrage du complexe 43S grâce à cette interaction, puis le dépose sur le codon AUG. Or il n'en est rien. Lorsque le domaine II n'est pas présent, l'assemblage de la particule 48S sur le codon AUG ne peut s'effectuer (*chapitre II, figures 33 et 34*). Ainsi, le recrutement précoce de eIF3 par le domaine III n'est pas suffisant pour permettre la formation du complexe 48S et à *fortiori* le complexe 80S pour compléter l'initiation de la traduction. En effet, l'assemblage du complexe 48S nécessite la présence du domaine II pour positionner correctement la particule 43S sur l'AUG par un mécanisme que nous ignorons. Nous avons montré que le recrutement du complexe 43S ne requiert pas le domaine II (*publication 1*). Cependant, le domaine II possède lui aussi une affinité, bien que plus faible, pour le facteur eIF3 au sein du complexe 48S (*figure 34B*). Ces deux résultats nous permettent d'émettre l'hypothèse que ce domaine II intervient lors du positionnement du complexe 48S sur le codon AUG. Les deux domaines de l'IRES_{5'UTR} du CrPV sont donc nécessaires à l'assemblage efficace d'une particule 48S. Dans le cas de l'initiation guidée par l'IRES de HCV, il a été démontré que ce mécanisme utilise également de manière synergique les domaines II et III de l'IRES (Yamamoto et al., 2015 ; Locker & Lukavsky, 2007; Otto & Puglisi, 2004). Plus précisément, le domaine II de l'IRES HCV est requis pour l'hydrolyse du GTP présent sur facteur eIF2 au sein du complexe ternaire. Cette étape est indispensable à la formation efficace du complexe 48S sur le domaine III de l'IRES de HCV. Ce même domaine II serait également requis lors de la dissociation des eIF du complexe 48S afin de permettre l'association efficace de la sous-unité 60S du ribosome, une étape qui est contrôlée par le facteur d'initiation eIF5B (Locker & Lukavsky, 2007; Otto & Puglisi, 2004). L'IRES_{5'UTR} du CrPV faisant également partie de la classe III des IRES comme l'IRES de HCV, il est possible que le domaine II ait une fonction similaire. Des arguments semblent étayer cette hypothèse. Ainsi, nous avons notamment constaté que le facteur eIF2 est présent en grande quantité sur le domaine II isolé, bien que celui-ci ne soit pas capable de recruter le 43S qui contient eIF2 (*Figure 35*). Ceci suggère que le domaine II interagit peut-être directement avec eIF2 présent au sein de la particule

43S fixée sur le domaine III. Par conséquent, ces résultats préliminaires suggèrent que le domaine II de l'IRES_{5'UTR} du CrPV, à l'instar de celui du HCV, aiderait à l'hydrolyse de eIF2-GTP et donc à l'assemblage du 48S. Concernant la seconde fonction du domaine II de HCV qui contrôle la dissociation des eIF, aucune des expériences que nous avons menées ne nous permet d'affirmer aujourd'hui que le domaine II de CrPV agit de la même manière. Nous avons simplement constaté la présence en quantité équivalente du facteur eIF5B sur les domaines II et III de l'IRES_{5'UTR} au stade 48S d'initiation. Ainsi, il se pourrait que le facteur eIF5B soit recruté sur le domaine II de l'IRES_{5'UTR} pour favoriser le passage du 48S au 80S mais cette hypothèse n'est pas encore validée. Le domaine II de HCV est également impliqué dans le bon positionnement de l'ARNt^{Met}_i sur le codon AUG. Plus précisément, un important changement conformationnel du domaine II dans le site E du ribosome permet de guider la reconnaissance du codon AUG par l'ARNt^{Met}_i (Yamamoto et al., 2015). Dans le cas de CrPV, la présence du domaine II est absolument indispensable à la formation d'un complexe 48S. Là encore, par analogie avec l'IRES de HCV, il est possible que le domaine II de l'IRES_{5'UTR} de CrPV ait une fonction similaire pour positionner l'ARNt^{Met}_i. Lorsque ce domaine II n'est pas présent, l'interaction avec le codon AUG ne pourrait pas se faire, rendant ainsi la formation d'un complexe 48S impossible comme nous l'avons observé lorsque l'IRES_{5'UTR} est tronqué du domaine II.

I.3. Le pseudo-nœud du domaine II est important pour la formation du 48S et du 80S

Le domaine II de l'IRES_{5'UTR} du CrPV contient une structure de type pseudo-nœud essentielle à l'activité traductionnelle de l'IRES et plus précisément à l'assemblage du complexe d'initiation 48S (*publication 1*). Nous avons constaté que la formation du pseudo-nœud est importante pour la formation du 48S. En plus des appariements au sein du pseudo-nœud, la séquence des nucléotides appariés à distance semble également importante pour la formation cette fois du complexe 80S (*chapitre I, figure 31*). Ce constat nous permet de formuler l'hypothèse selon laquelle le pseudo-nœud du domaine II serait important pour la formation du complexe 48S. Comme nous l'avons évoqué dans le paragraphe précédent, la formation du complexe d'initiation 80S sur l'IRES HCV s'accompagne de réarrangements structuraux dans le domaine II, probablement importants pour le positionnement précis du Met-ARNt^{Met}_i dans le site P du ribosome (**Figure 18**). Ces réarrangements structuraux garantissent ainsi la transition du complexe HCV-48S vers le complexe HCV-80S (Yamamoto et al., 2015). Par analogie, nous proposons que, dans le cas de CrPV, des changements conformationnels du pseudo-nœud du domaine II de l'IRES_{5'UTR} seraient requis pour la transition du complexes d'initiation 48S vers le 80S. La structure adoptée par la région du pseudo-nœud aurait ainsi un rôle important dans le positionnement

précis du ribosome sur l'IRES_{5'UTR}, assurant de ce fait la transition du complexe CrPV-43S vers le complexe CrPV-48S puis CrPV-80S. Nous disposons de résultats préliminaires indiquant que l'ouverture artificielle du pseudo-nœud par l'introduction de mutations influence le recrutement de complexes RNP (*Chapitre II, figure 36*). Cependant, nous ignorons actuellement quels sont ces facteurs et s'ils prennent part à l'initiation de la traduction. Les perspectives qui découlent de ces expériences pour valider ces hypothèses sont les suivantes : Nous allons vérifier si le domaine II de l'IRES_{5'UTR} change de conformation et si le pseudo-nœud s'ouvre en cours d'initiation de la traduction par des analyses de sondage en solution sur les différents complexes d'initiation purifiés. Nous nous focaliserons prioritairement sur la région du pseudo-nœud afin de vérifier la présence des appariements à distance ou son ouverture pendant l'initiation. Et en parallèle, nous déterminerons, par spectrométrie de masse, la composition des complexes protéiques associés au domaine II lorsque l'ouverture du pseudo-nœud est forcée par l'introduction de mutations.

I.4. La sous-unité eIF3j n'est pas retrouvée dans les complexes d'initiation assemblés sur l'IRES_{5'UTR}

Le facteur d'initiation eIF3 interagit spécifiquement avec le domaine III de l'IRES_{5'UTR} de CrPV. Les analyses par spectrométrie de masse ont révélé la présence de toutes les sous-unités protéiques de eIF3 hormis la sous-unité eIF3j, qui est absente de tous les complexes d'initiation 43S, 48S et 80S assemblés sur l'IRES_{5'UTR} (*chapitre I, figure 29-30*). Nous savons par ailleurs que cette sous-unité eIF3j est strictement nécessaire lors de la traduction guidée par l'IRES_{5'UTR} de CrPV et l'IRES HCV (Majzoub et al., 2014). De plus, eIF3j est labile et se dissocie facilement de eIF3 pendant la purification des complexes. Par conséquent, il est possible que lors de la purification de nos complexes d'initiation, eIF3j se soit décrochée avant leur analyse par spectrométrie de masse. Nous pouvons cependant émettre une seconde hypothèse. Plusieurs études ont montré que eIF3j est essentielle à la formation de la particule HCV-48S. En outre, la sous-unité eIF3j occupe sur eIF3 le canal dans lequel va se placer l'ARN messager lors de la formation de la particule 48S. Par conséquent, 3j doit être éjectée du complexe 43S pour libérer le canal à l'ARN et permettre l'assemblage du 48S (Fraser et al., 2007). Nous pouvons penser qu'un mécanisme similaire existe lors de l'initiation guidée par l'IRES_{5'UTR} du CrPV. Nos résultats préliminaires semblent indiquer que le pseudo-nœud du domaine II de l'IRES_{5'UTR} interagirait avec la sous-unité eIF3j. Le pseudo-nœud pourrait être nécessaire au déplacement de eIF3j et permettre ainsi la formation du complexe 48S-IRES_{5'UTR}. La validation de ces hypothèses préliminaires nécessite évidemment d'autres expériences complémentaires. Par exemple, l'étude de la structure tridimensionnelle des différents complexes d'initiation de la traduction par cryo-

Microscopie Électronique nous permettrait d'observer avec précision le comportement de l'IRES_{5'UTR} et des facteurs d'initiation eIF3 et eIF2 durant l'initiation guidée par cet IRES. Nous pourrions identifier les facteurs protéiques associés à l'IRES ainsi que leurs sites d'interaction. Il sera peut-être également possible de visualiser des réarrangements structuraux de l'IRES pendant la transition du complexe 48S vers le complexe 80S. Par ailleurs, toutes les structures de complexes d'initiation assemblés sur l'IGR de CrPV publiées à ce jour ont été réalisées avec des ribosomes provenant d'organismes qui ne sont pas des hôtes de ce virus (ribosomes de mammifère, de levure et ribosome bactérien). Or, la structure du ribosome de drosophile présente des différences importantes avec ces ribosomes, notamment au niveau des segments d'expansion (Anger et al., 2013). Étant donné que les IRES se sont adaptés aux ribosomes des cellules hôtes qu'ils ciblent pendant l'infection, nous pouvons penser qu'il serait encore plus intéressant de réaliser des études structurales avec des ribosomes de drosophile, qui est un organisme qui peut être infecté efficacement par le CrPV. Par conséquent, nous proposons d'entreprendre pour la première fois, des études structurales par cryo-Microscopie Électronique de complexes avec l'IRES_{5'UTR} et l'IGR assemblés sur le ribosome de drosophile. Ces structures permettraient également de mieux comprendre les bases structurales de l'expression différentielle des deux IRES de CrPV.

1.5. Quelle serait la fonction des deux domaines de l'IRES_{5'UTR} ?

Les deux domaines de l'IRES_{5'UTR} du CrPV doivent être physiquement liés pour permettre un assemblage efficace du ribosome sur l'IRES, ces deux domaines agissent en synergie lors de l'initiation de la traduction (*chapitre II*). Nos résultats suggèrent que l'initiation de la traduction guidée par l'IRES_{5'UTR} se ferait selon le modèle suivant :

Lors d'une étape précoce de l'initiation, le facteur d'initiation de la traduction eIF3 est recruté sur le domaine III de l'IRES_{5'UTR}. L'interaction entre eIF3 et l'IRES se fait directement grâce aux sous-unités eIF3a, c et e. Puis, le recrutement du complexe de pré-initiation 43S sur le domaine III se ferait grâce à eIF3. Cette étape serait stimulée par la présence de l'ITAF La, une autre protéine reconnue spécifiquement par le domaine III de l'IRES_{5'UTR} et impliquée dans le recrutement du ribosome 40S sur les IRES. Ensuite, le domaine II permettrait d'assurer la transition entre le complexe CrPV-43S et les étapes suivantes de l'initiation. Le pseudo-nœud du domaine II participerait à l'ancrage du ribosome sur l'IRES_{5'UTR}, directement sur le codon initiateur AUG étant donné que cet IRES s'affranchit du processus de 'scanning'. Cette étape aboutirait à la formation du complexe 48S. Le relargage des eIF serait ensuite induit par l'hydrolyse du GTP de eIF2 présent entre le ribosome sur le domaine III et le domaine II. Enfin, l'association de la sous-unité 60S du ribosome pourrait aussi

dépendre du domaine II. Précisons cependant qu'il s'agit-là d'un modèle de travail qui requiert évidemment une validation expérimentale solide sur la base de plusieurs expériences évoquées plus haut.

ANNEXES

ANNEXES

ANNEXE I : MATÉRIEL ET MÉTHODES

ANNEXE II : PUBLICATION 2

Tracking the m7G-cap during translation initiation by crosslinking methods.

Gross Lauriane, Schaeffer Laure, Alghoul Fatima, Hayek Hassan, Allmang Christine, Eriani Gilbert, Martin Franck.

Methods, 2018, Volume 15 ; issue 137 : 3-10. doi: 10.1016/j.ymeth.2017.12.019.

Résumé de la publication 2 :

Dans le cadre de ma thèse je me suis également intéressée à un autre ARN messager dont l'initiation de la traduction se fait de manière non conventionnelle. Il s'agit de la traduction de l'ARNm de l'histone H4 (ARNm H4). Le mécanisme d'initiation de la traduction de l'ARNm H4 se fait par un nouveau mécanisme appelé 'ribosome tethering' (voir *paragraphe 1.3.1.1* de l'introduction). Il combine des caractéristiques de l'initiation canonique, tel que l'importance de la coiffe, avec des caractéristiques retrouvés chez les IRES viraux, comme la présence d'éléments structuraux dans l'ARN qui sont impliqués dans le recrutement de facteurs eIF lors du recrutement de la machinerie traductionnelle. Dans cet article publié dans *Methods* en 2018 nous avons suivi le destin de la coiffe au cours des différentes étapes de l'initiation de la traduction à l'aide de méthodes de pontage photo-activable ou chimique de la coiffe m⁷G. Dans un premier temps, nous avons étudié la position de la coiffe sur l'ARNm H4 libre. Pour ce faire, l'ARNm H4 a été coiffé à l'aide d'analogues de la coiffe contenant un groupement thiol photo-activable. L'irradiation aux UV induit un pontage intra-moléculaire, démontrant ainsi que la coiffe n'est pas accessible mais plutôt séquestrée dans une poche localisée dans la phase codante de l'ARN que nous avons baptisée 'Cap Binding Pochet' (CBP). Nous avons ainsi confirmé les résultats antérieurs du laboratoire qui indiquent que la coiffe n'est pas accessible avant l'initiation de la traduction. Ensuite, nous avons suivi la position de la coiffe pendant les différentes étapes de l'initiation coiffé-dépendante. La stratégie a consisté à utiliser la coiffe, marquée radioactivement, comme traceur pour suivre sa position au cours du processus d'initiation. Les protéines situées à proximité de la coiffe sont visualisées suite à un pontage chimique qui permet de créer une liaison covalente avec la coiffe radio-marquée. Ceci nous a permis de montrer que la coiffe est reconnue par la protéine eIF4E au stade 48S. En revanche, au stade 80S, la protéine eIF4E s'est décrochée et la coiffe est située à proximité des protéines ribosomiques eS26 et eS28. Ces deux protéines sont positionnées dans le passage de l'ARNm sur la sous-unité 40S du ribosome. Rappelons que l'ARNm de H4 présente une extrémité 5'UTR très réduite de seulement 9 nucléotides. Par conséquent, notre étude a permis de démontrer que la protéine eIF4E doit se décrocher de la coiffe après le stade 48S pour permettre l'assemblage de la grande sous-unité 60S. Il serait maintenant intéressant d'appliquer le protocole expérimental mis au point pour suivre le destin de la coiffe pendant l'initiation de la traduction d'autres ARNm contenant des régions 5' UTR de longueurs variables.



Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/yvmet

Tracking the m⁷G-cap during translation initiation by crosslinking methods



Lauriane Gross, Laure Schaeffer, Fatima Alghoul, Hassan Hayek, Christine Allmann, Gilbert Briani^{a,*}, Franck Martin^a

^a Université de Strasbourg, UMR, Architecture et Mécanisme de l'ARN, UMR 8001, 67082 Strasbourg, France

ARTICLE INFO

Article history:

Received 15 October 2017

Received in revised form 26 December 2017

Accepted 22 December 2017

Available online 1 January 2018

Keywords:

Cap-dependent translation

UV crosslinking

Chemical crosslinking

Ribosome

Histone H4 mRNA

© 2017

ABSTRACT

In eukaryotes, cap-dependent translation initiation is a sophisticated process that requires numerous trans-acting factors, the eukaryotic Initiation Factors (eIFs). Their main function is to assist the ribosome for accurate AUG start codon recognition. The whole process requires a 5'–3' scanning step and is therefore highly dynamic. Therefore translation requires a complex interplay between eIFs through assembly/release cycles. Here, we describe an original approach to assess the dynamic features of translational initiation. The principle is to use the m⁷Gcap located at the 5' extremity of mRNAs as a tracker to monitor RNA and protein components that are in its vicinity. Cap binding molecules are trapped by chemical and UV crosslinking. The combination of cap crosslinking methods in cell-free translation systems with the use of specific translation inhibitors for different steps such as edeine, GMP PNP or cycloheximide allowed assessing the cap fate during eukaryotic translation. Here, we followed the position of the cap in the Histone H4 mRNA and the cap binding proteins during H4 mRNA translation.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Translation initiation leads to the assembly of an elongation competent 80S ribosome on the genuine AUG start codon. In eukaryotes, this is achieved through assembly of a 43S pre-initiation complex containing the small ribosomal subunit 40S on the 5' m⁷Gcap of mRNA. Next, the so-called 43S slides on the 5'UTR by a 5'–3' nucleotide-by-nucleotide inspection that is an ATP-consuming mechanism termed scanning, until the AUG start codon is found [1]. Then, the large ribosomal subunit 60S joins in order to assemble a complete 80S ribosome on the start codon. This sophisticated process is assisted by numerous trans-acting factors called eukaryotic initiation factors (eIF) [2]. The whole process is highly dynamic with many transient complexes that are difficult to trap. The interplay between the eIF is rather complex and methods to assess the numerous assembly/release cycles of individual players are rare.

Among these, a number of techniques for studying the structure and interaction of proteins with nucleic acids depend on methods

of crosslinking. Crosslinking is the process of chemically joining two or more molecules by a covalent bond. It often implies chemical modification of one partner, or incorporation of a reactive reagent such as photo-reactive nucleotides that alter the reactivity of the original molecule. A great advantage of crosslinking approaches is to enable trapping of transient complexes by linking covalently the molecules that interact temporarily during dynamic processes such as translation initiation. Crosslinking methods provide a rapid mean of obtaining evidence for the proximity of functional groups in structurally complex RNAs and ribonucleoproteins. Accurate identification of the crosslinks is also a way to probe the conformation of the RNA of interest.

The position of the m⁷G cap during the whole translation initiation process is of special interest. Indeed, numerous studies have demonstrated that the cap first initiates the scanning process by recruiting the whole machinery at the 5' extremity of mRNA. However, the positioning of the cap during the scanning process, after 80S formation on the AUG and later on during elongation is so far unexplored. Here, we describe two crosslinking methods using the cap as a bait to trap partner molecules in order to assess the cap positioning and its fate during the whole translation process. In this study, we used the mouse histone H4 mRNA as a model to follow the cap fate during the whole initiation process. This particular mRNA is part of the cell cycle-dependent, Histone mRNAs which are

*Corresponding author.

E-mail addresses: gerard@univ-strasbourg.fr (G. Briani).

E-mail addresses: gerard@univ-strasbourg.fr (G. Briani), franck@univ-strasbourg.fr (F. Martin).

not polyadenylated at their 3' end [3]. They are massively and exclusively expressed during the S-phase of the cell cycle. They have short UTRs, H4 mRNA being the one with the shorter 5' UTR of only 9 nucleotides [4]. The translation mechanism of histone H4 mRNA is non canonical. The 5' m⁷Gcap is not accessible but instead is sequestered by a cap binding pocket located in the coding region [5,6]. This enables the internal recruitment of eIF4F on a specific element called 4E-Sensitive Element that is also located in the coding region. Then, the ribosome is dropped on the start codon by a so-called 'ribosome tethering' mechanism. Since there is no scanning, the AUG start codon is positioned accurately in the P site of the ribosome by a direct interaction between an AGG triplet in the coding region of H4 mRNA and the loop of helix h16 of the 18S ribosomal RNA [7]. In order to better characterize the role of the cap in this sophisticated process, we studied its position during the distinct steps of translation initiation and elongation.

First, we monitored the position of the cap in the mRNA alone before translation initiation starts by UV-crosslinking with thio-modified caps. As previously mentioned, we confirmed that the cap is not accessible but rather located on an internal structure formed by H4 mRNA making a cap-binding pocket. Second, we investigated the fate of the cap during H4 mRNA translation initiation and elongation by chemical crosslinking. Thereby, we detected proteins that are positioned near the 5'-end of mRNA in initiation complexes formed during histone H4 mRNA translation.

2. Material and methods

2.1. *In vitro* transcription

Synthesis of RNAs containing thio-modified nucleosides can be performed with the T7 RNA polymerase system. The gene encoding mouse histone H4 (HIST1H4C) and the human β -globin mRNA have been previously cloned into pUC19 and YpGlo respectively. Transcription templates are synthesized by PCR amplification from the plasmids. The 5' primers contain the T7 promoter sequence and the 3' primers are designed to promote *in vitro* run-off transcription at the desired position. The PCR approach is also used to synthesize transcripts of H4 mRNA with truncated 5' sequences. For that, new 5' primers containing the T7 promoter sequence connected with the desired H4 sequences are designed to synthesize the PCR fragments with variable sizes. *In vitro* transcription is performed as follows. PCR DNA template (about 20 μ g) is mixed with transcription buffer (40 mM Tris-HCl, pH 8.1, 20 mM MgCl₂, 5 mM DTT), 1 mM each ATP, CTP, GTP, UTP, 40 U of RNase inhibitor (RNasin-Promega), 0.2 mg/mL of T7 RNA polymerase, and milliQ water to 100 μ L. The reaction mixture is incubated at 37 °C for 1 h. Then, 2 μ L of pyrophosphatase (10108987001-Sigma) are added and the mix is further incubated for 30 min. Then, the DNA template is degraded by addition of 2 μ L of DNase I RNase-free (04716728001-Sigma) and incubation for 1 h at 37 °C. Usually, to check RNA integrity, an aliquot is mixed with Formamide Dye and loaded on urea-denaturing 4% PAGE and visualized under UV light after ethidium bromide staining. To eliminate unincorporated nucleotides, the remaining RNA sample is loaded on a G25 column. Transcripts are further purified by separation on denaturing 4% PAGE and electro-elution from gel slices using a Biotrap apparatus (Schleicher and Schuell). Purified RNA samples are then phenol extracted to eliminate proteins from the synthesized transcripts. The concentration of purified RNA sample is determined by absorbance measurement at 260 nm. Before use, transcripts are folded in water by incubating at 80 °C for 2 min followed by slow cooling to 35 °C and kept on ice.

2.2. Capping of RNA transcripts

RNA transcripts can be m⁷G-capped co-transcriptionally by T7 RNA polymerase in the presence of m⁷GpppG cap analogue (New England Biolabs). More efficient capping is obtained with the vaccine-capping enzyme (VCE) using the ScriptCap™ m⁷G Capping System (Epicentre Biotechnologies) which caps and methylates the 5' end of mRNAs to nearly completion (Fig. 1A). The ScriptCap™ m⁷G Capping System requires purified mRNA (purified by electro-elution, see 2.1.). When the GTP from the kit is substituted by s⁶GTP (Jena Biosciences, 1 mM), a m⁷s⁶G-capped mRNA suitable for photo-crosslinking studies is synthesized by VCE after a 1-h incubation at 37 °C (Fig. 1B). After capping, enzymes used for the reaction are eliminated by phenol-chloroform extraction and the m⁷s⁶G-capped mRNAs are precipitated with ethanol in the presence of 250 mM NaCl. The capped mRNAs are recovered by centrifugation; the pellets are dried and then resuspended in autoclaved milliQ water.

2.3. Short-range crosslinking using thio-containing 5' cap

Two pmoles of *in vitro* synthesized and capped m⁷s⁶G-histone H4 mRNA are dissolved in water, refolded at 95 °C for 2 min and then incubated at 0 °C, then are irradiated with 312 nm for 30 min at a distance of 8 cm from the bulb using a Bio-Link BLX 312 on ice as previously described [6]. Importantly, previous experiments with histone H4 mRNA demonstrated that it has to be refolded in water to favour the optimal conformation to visualize the UV-crosslink of the cap in its cap binding pocket. However, for other mRNAs, it might be critical to first set-up an optimal refolding protocol that may require Mg²⁺ to reach the optimal conformation. Then, crosslinks are detected by primer extension using fluorescently labelled H4-specific primers by reverse transcriptase. Reverse transcription is performed in final volume of 20 μ L. First, 2 pmol of cross-linked mRNA are mixed with 2 pmol of fluorescently (VIC or NED, from Integrated DNA Technologies) labelled primer (H4-rev160: 5'-TGAGGCCGGAGATGCGCTTC-3') and denatured first at 95 °C for 2 min followed by primer annealing at 65 °C for 5 min and incubation on ice for 2 min. Then, 0.2 U AMV reverse transcriptase (Life Sciences Advanced Technologies Inc.) in its buffer as recommended by the manufacturer is added and supplemented with 1.5 mM of each dNTP. Reverse Transcription extension parameters are: 42 °C for 20 min, 50 °C for 30 min, 60 °C for 10 min. Sequencing reactions are performed in parallel in similar conditions, but containing 0.5 mM of each dNTP and 16.6 μ M of ddCTP. Reactions are stopped by phenol-chloroform extraction and ethanol precipitation. Samples are resuspended in 10 μ L deionized formamide and loaded on a 96-well plate for sequencing on an Applied Biosystems 3130xl genetic analyzer. The resulting electropherograms are analysed using QuSHAPE [8], which aligns signal within and across capillaries, as well as to the dideoxy references, and corrects for signal decay. Normalized reactivities range from 0 to ~2, with 1.0–1.2 being the average reactivity for highly reactive positions. The average reactivity (*.shape file) is calculated for each H4 mRNA transcript from three independent experiments with Pearson correlation coefficients ≥ 0.9 .

2.4. 5'-end radioactive labelling of RNA transcripts

For chemical crosslinking studies, mRNA transcripts (80 pmol-10 μ g) are first 5'-end-labelled. For that, 50 μ Ci of [α -³²P]-GTP (6000 Ci/mmol, 10 mCi/mL, Hartmann Analytic) is used instead of unlabelled GTP during the capping reaction by the ScriptCap™ m⁷G Capping System (Fig. 1C). After radioactive capping, capped mRNA transcripts are separated from unincorporated GTP by denaturing PAGE (4%) and eluted overnight from gel slice at 4 °C in

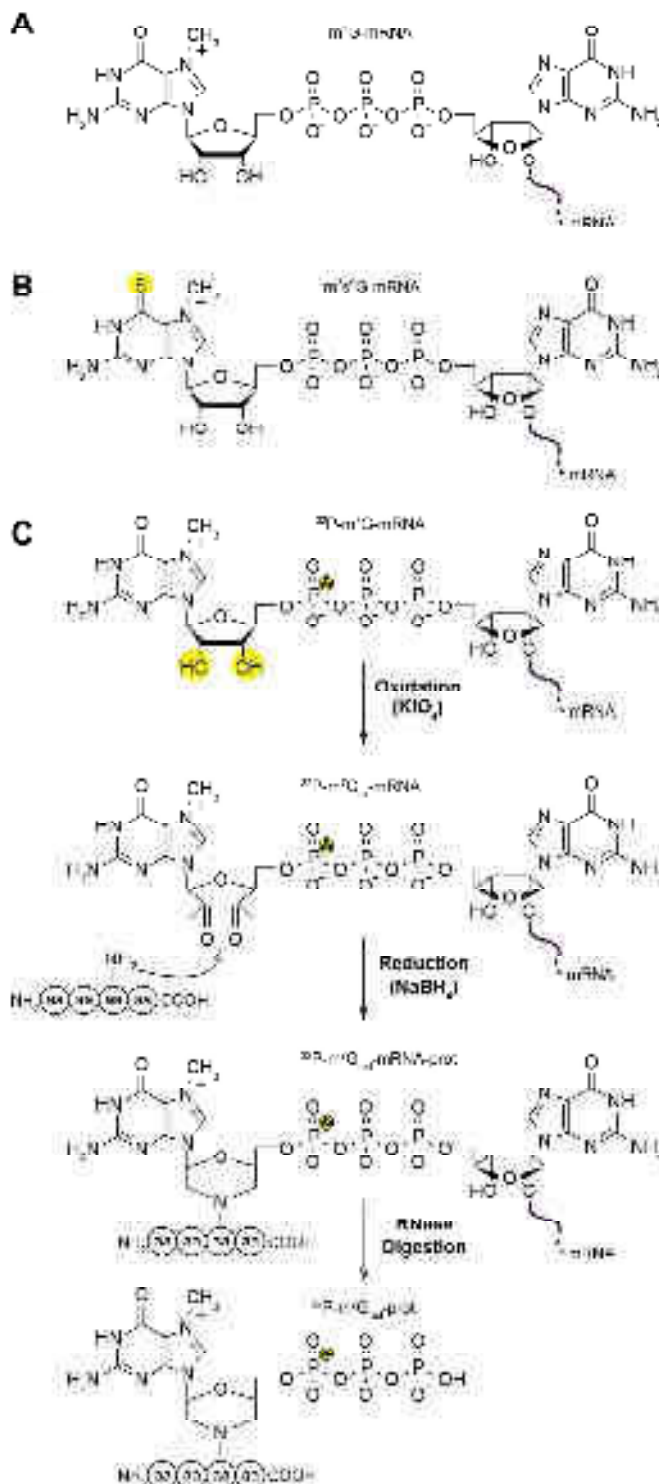


Fig. 1. Capped mRNAs used in this study. **A.** Structure of m⁷G-capped mRNA. **B.** Structure of m⁷G-capped mRNA that is site selectively labelled with ϵ -amino groups (ϵ -NH₂) over within the cap structure. **C.** Structure of radiolabelled m⁷G-capped mRNA that has been used for chemical crosslinking. The 2' and 3'-OH groups are first oxidized by KIO₄ and then reduced by NaBH₄. This creates a covalent link between a NH₂ group from the lateral chain of a protein that is in close proximity. The RNA moiety is then degraded by RNase treatment (using a ³²P-labelled protein that has been analysed by SDS-PAGE). For interpretations on the references to colour in this figure legend, the reader is referred to the web version of this article.

buffer (0.3 M NaCl, 0.05 M EDTA, 10 mM Tris-HCl pH 7.0). After passive elution, the radioactive m⁷G-capped mRNA are precipitated with ethanol in the presence of 2.0 mM NaCl. The radioac-

tive capped mRNAs are recovered by centrifugation, the pellets are dried and then resuspended in autoclaved milliQ water. Global radioactivity of the mRNA samples are measured by Cerenkov counting and adjusted to 50,000 cpm per μ L.

2.2. Chemical crosslinking of the 5' cap to ribosome initiation complexes

Chemical crosslinking of the 5' cap of mRNAs is a powerful approach to study RNA binding proteins such as initiation factors. The method was originally used to identify the cap-binding protein eIF4E [9]. The target mRNA containing a radioactively labelled cap, m⁷[³²P](3')pppG, is oxidised with sodium periodate to convert the 2',3'-diol of the 5'-terminal m⁷G to a reactive aldehyde (it is worth noting here that the diol at the 3' end of the mRNA is similarly oxidised in a dialdehyde) (Fig. 1C). Then, the oxidised mRNA is incubated in cell-free protein synthesis systems from rabbit reticulocyte lysate (RRL) or wheat germ extract (WGE) or alternatively with purified components. The resulting mRNA-ribosome initiation complexes are treated with NaIO₄ to induce and stabilize the 5eIF4E bases between mRNA 5' terminus and amino groups of neighbouring proteins, yielding to covalently linked protein-RNA conjugates. The mRNA is then further treated by RNase A (Roche Diagnostics) thereby leaving the radiolabelled cap covalently linked to its protein partners. Cell-free extracts or protein components are then separated by SDS-PAGE and radioactive crosslinked proteins are revealed by autoradiography. Usually, only polypeptides positioned near the 5' end are crosslinked to the mRNA, making this approach very useful for studying the spatial relationships between molecules in ribosomal complexes or other ribonucleoprotein (RNP) complexes that are close to the cap. In the example of histone H4 mRNA shown here, the periodate-oxidised mRNA is assembled with initiation complexes in RRL in the presence of various translation inhibitors [10]. Ecdysin is binding in the mRNA tunnel of the small ribosomal subunit and interferes with the codon-anticodon interaction between initiator Met-tRNA^{Met} and the AUG start codon in the P site. This means that ecdysin treatment impairs formation of the 48S complex. In other words, ecdysin treatment enable the accumulation of translation initiation complex prior AUG-codon recognition. GMP-PNP is a non-hydrolyzable GTP analog that prevents ribosomal subunit joining by inhibiting the GTPase activity of eIF5B subunit-joining factor [11]. The use of GMP-PNP in cell-free extracts allows the accumulation of 48S particles (small ribosome subunit) with a Met-tRNA^{Met} assembled to the AUG initiation codon. Finally, cycloheximide induces accumulation of 80S particles by blocking translocation at the translocation step. In fact, cycloheximide allows one complete round of translocation to proceed before initiating further elongation. The inhibitory effect results from binding of both cycloheximide and decylated tRNA in the E site, which leads to the arrest of the ribosome on the second codon [12]. Cycloheximide catalysis are validated by competing with an excess of ³⁵S-GDP. Experimentally, 5 μ g of purified and radiolabelled (250,000 cpm) capped H4 mRNA are incubated for 2–3 h at 0°C in 250 μ L of 100 mM sodium acetate, pH 5.3, 10 mM EDTA, 0.2 mM sodium periodate. Then, glyceral is added at 25 (v/v) final concentration for neutralization. After 10 min incubation at room temperature, the mixture is phenol extracted twice and ethanol precipitated. The RNA pellet is dissolved in 10 μ L of water. One μ L of oxidised RNA is incubated with 2.5 μ L of RRL in 10 mM HEPES-KOH, pH 7.6, 1 mM ATP, 75 mM KCl, 1 mM DTT, 1 mM Mg²⁺Ac₂ and 1 mg/mL cycloheximide, or 2 mM GMP-PNP or 10 μ M ecdysin or 1 mM m⁷GDP in a final volume of 10 μ L. After 10 min incubation at 30°C, 1 μ L of 0.2 M NaIO₄ is added and incubation is pursued for 2–3 h at 0°C. Then, RNA is digested by 1 μ L of RNase A (10 mg/mL, ThermoFisher Scientific-EN0531) for 30 min at 37°C and samples are fractionated by SDS-PAGE.

Radioactive bands correspond to proteins crosslinked to the 5' m⁷G-cap of the mRNA are revealed by autoradiography. The specificity of the crosslink is further confirmed by adding the cap analogue m⁷GDP that competes with the m⁷G-cap and clears the cross-link. Protein identification can be further performed by immuno-precipitation and/or mass spectrometry analysis.

2.6. Sucrose gradient analysis of crosslinked initiation complexes

The cap crosslinking capacity in RRI was further analysed after sucrose gradient fractionation. The protocol used to prepare the complexes with 5' labeled capped mRNA was slightly modified compared to the one described in Section 2.5. From the formation of 80S/14 initiation complexes, 30 μ l of untreated RRI were incubated for 5 min at 30 °C and 20 min in ice in a buffer containing 1 μ l of RNase inhibitor (40U/RNasin, Promega N2511), 75 mM KCl and 0.5 mM MgCl₂. In order to lock the ribosome at translation initiation, RRI were pre incubated for 1 min at 30 °C and 20 min at 0 °C in the presence of 10 μ M ecdine or 2 mM GMP-PNP or 1 mg/ml cycloheximide or without inhibitor. Finally, formation of initiation complexes was obtained by adding the oxidised histone H4 mRNA to a final concentration of 500 nM and incubating for 5 min at 30 °C. Then NaBH₄ was added to a final concentration of 20 mM and reduction was performed at 0 °C for 1 h. The resulting RRI programmed with H4 mRNA were separated on 7–17% linear sucrose gradients in buffer (25 mM Tris-HCl, pH 7.4, 75 mM KCl,

0.5 mM MgCl₂, 1 mM DTT, 1 mM cycloheximide). The reactions were loaded on the gradients and centrifuged (37,000 rpm for 2.5 h at 4 °C) in a SW41Ti rotor. Gradients are fractionated on a Bioorp gradient fractionator and fractions were analysed by Cerenkov counting in order to identify the fractions containing translation initiation complexes. Indeed, we demonstrated that the stability of translation initiation complexes is not affected by the 1 h incubation at 0 °C after reduction as attested by the radioactive profiles. Then the fractions (300 μ l) corresponding to RNP₀, 80S and polysomes were RNase A treated as described in Section 2.5 for 30 min at 37 °C before being precipitated by TCA (12.5% final concentration) for 30 min at 0 °C. Precipitated proteins were recovered by 15 min centrifugation at 15000 \times g, washed by 20 μ l acetone, dried and resuspended in 10 μ l of loading blue (400 mM Tris-HCl, pH 6.8, 4% SDS, 4% β ME, 30% glycerol, 0.05% bromophenol blue). Proteins were fractionated by SDS-PAGE and radioactive bands were detected by phosphorimager plates (70) Imaging plates). For reproducibility, we noticed that the use of freshly oxidised mRNA is critical.

3. Results and discussion

3.1. UV crosslinking of the cap to its cap-binding pocket from H4 mRNA

Although a certain number of chemical and photoreagents are available for short-range RNA-RNA crosslinking studies, the

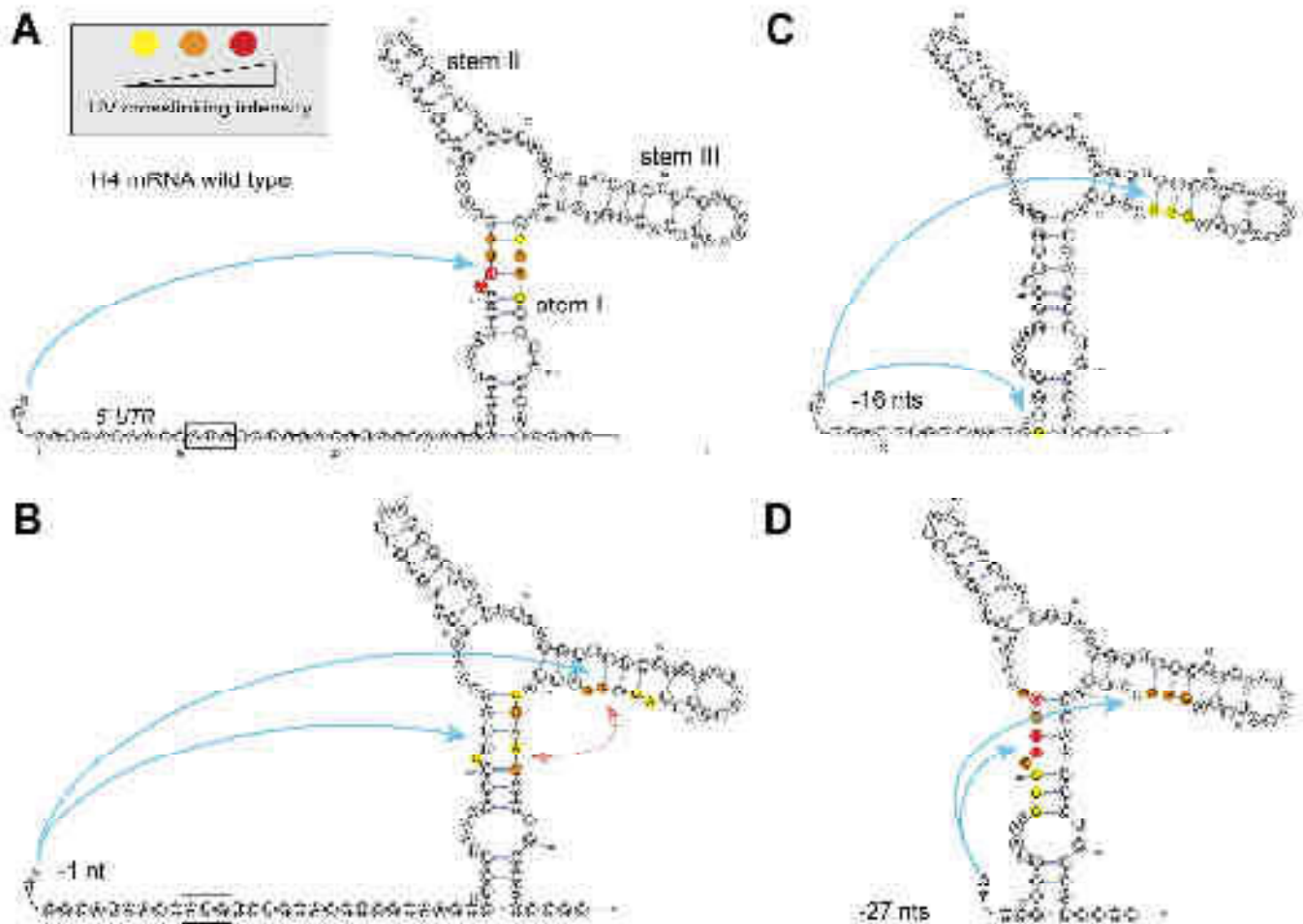


Fig. 2. UV crosslinking of m⁷G H4 mRNA. **A.** Secondary structure of the 11C E2b part of histone H4 mRNA and interaction with the m⁷G cap structure. The structure contains three helical stems (I-III) connected by a three-way helical junction (W). The initiation codon is boxed. Crosslink experiments on m⁷G-capped wild-type H4 mRNA. Crosslinks were performed at 260 nm and revealed by reverse transcription. Reverse transcription stops are indicated by colors according to the intensity scale shown in the inset. **B-D.** Similar experiments performed on mutants deleted of 1, 16 and 27 first nucleotides.

thionucleotides ^35S G and ^35S U are preferred due to their simple molecular structure, relative stability, and high reactivity. In addition, ^35S G and ^35S U only differ from the corresponding G and C nucleotides by a single atomic substitution, the replacement of a nucleobase oxygen by sulphur, thus reducing the risk for perturbation of the RNA structure. Exposure of the thionucleotides to UV light produces a sulphur radical that can react very efficiently with functional groups that are in proximity (in proteins or nucleic acids). To increase crosslinking specificity, low energy (360 nm) irradiation is preferred, yielding “zero-length” crosslinks [13] that represent direct contacts (in a 2-Å sphere) of the nucleotide of interest and protein or nucleic acid. Thionucleotides can be incorporated by either random incorporation or site-specific substitution. *In vitro* transcription is a convenient way to incorporate statistically and randomly ^35S G and ^35S U. This way of incorporation allows a rapid screening of the potential interactions over the entire RNA transcript. Site specific incorporation by transcription of RNA ligase II eliminates the ambiguity of the origin of crosslinking and generally improves the experimental signal since all the RNA molecules are containing the photo agent at a single or very few positions. It allows accurate probing of any position, and generally increases the experimental signal since all the RNA molecules contain a photo agent at a single position rather than spread throughout the molecule, however it is time- and labour-consuming since it requires the synthesis of site specifically modified RNA. A simplified variant of this approach consists in the synthesis of short RNA fragments containing a single G or U residue that is substituted by ^35S G or ^35S U during *in vitro* transcription. For instance a set of mRNAs that contained single ^35S U at unique positions between +26 and +11, relative to the A of the initiation codon, were synthesized to explore mRNA contacts in the

ribosome mRNA channel during translation initiation [14]. Similarly, TISU mRNA fragments containing single unique ^35S U between nts +4 and +8 were synthesized to identify ribosomal proteins interacting with TISU sequence [15].

Here we describe another method of site-specific incorporation of photo agent in the m⁷G cap or the 3' end of any mRNA. The thionucleotide ^35S G has been used for mapping the cap-binding pocket in the mouse histone H4 mRNA [5]. Indeed, the m⁷G-cap of eukaryotic mRNAs is a modified guanine nucleotide connected to the mRNA via an unusual 5' to 5' triphosphate linkage. *In vivo*, it is added co-transcriptionally to the mRNA. *In vitro*, the modified cap is added post-transcriptionally by the Vaccinia Capping Enzyme (VCE) using the ScriptCapTM m⁷G Capping System (Epicentre Biotechnologies) (see Section 2.2). During this process, ^35S GTP is substituted to the usual GTP, it is subsequently modified to form a m⁷ ^35S G-cap structure. In that form, the m⁷ ^35S G-cap structure is an effective structural probe that combines high crosslinking efficiency, site-specific incorporation with high levels of incorporation. Then, the modified mRNA is subjected to UV at 312 nm exposure and detection of the crosslinks is subsequently performed by primer extension with [³²P] labelled dG nucleotides and polyacrylamide gel separation as described previously [6,5] or by an alternative protocol that we are here describing in Section 2.3.

On the secondary structure representation of the three way helix junction (T₃WJ) domain of H4 mRNA we observed that capped wild-type H4 mRNA exhibits strong crosslinks in the stem I of the T₃WJ involving, in order of decreasing intensity, nts 41, 42, 43, 44 in addition to nts 103, 104, 105 and 106 on the other strand of the stem (Fig. 2). This strongly supports the hypothesis that these nts form the core of the cap-binding pocket (CBP) as previously described [5]. Next, we performed 8 progressive deletions

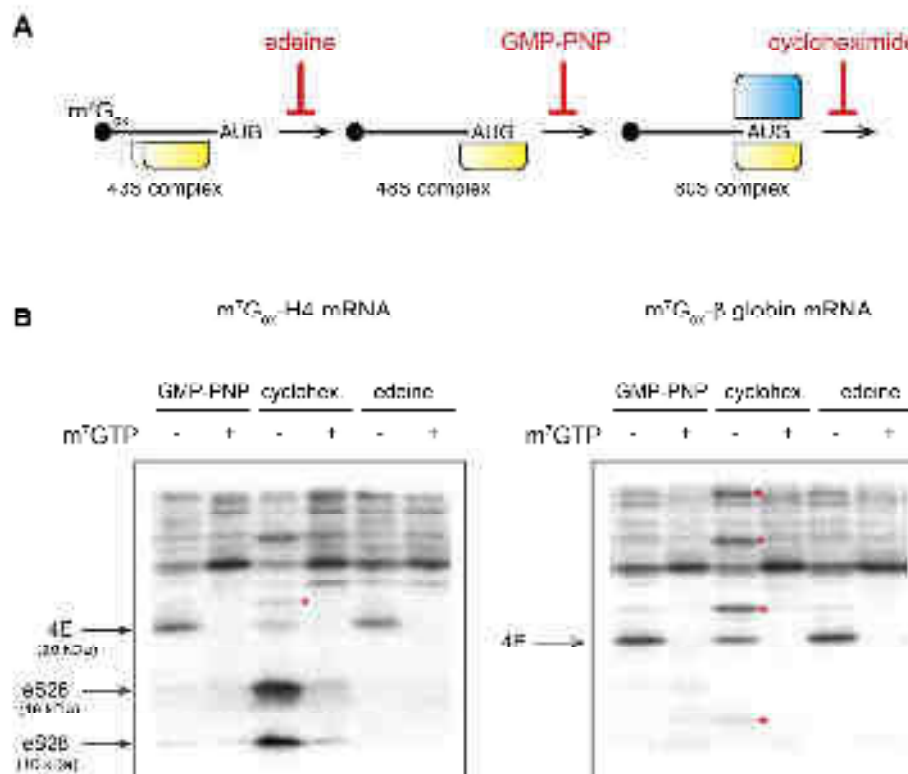


Fig. 3. Chemical crosslinking of modified histone H4 and β-globin mRNA. **A.** Cap dependent translation initiation is inhibited by specific inhibitors (shown in red). Elease prone to cycloheximide, inhibition of 43S-m⁷G_{cap} mRNA and thereby blocks 48S assembly. GMP-PNP is a non-hydrolyzable GTP analog, without being a GTPase activator, it leads to 48S accumulation on the AUG start codon. Cycloheximide inhibits translocation after 68S joining and triggers 80S formation on the AUG start codon. **B.** Chemical crosslinking of ³²P-labelled modified histone H4 mRNA and β-globin mRNA in BRL in the presence of GMP-PNP, cycloheximide (shown each and without m⁷GTP). Chemical cross-specific crosslinks are indicated by asterisk. Uncharacterized cross-specific crosslinks are indicated by red asterisks. Over-exposition of the gel allows to observe in this figure legend, the reader is referred to the web version of this article.

A

1. doi:10.1002/ajmg.b.31726.102; 1-10

of the 5' UTR and checked whether this changed the cap-crosslinking capacity. The truncated mRNA fragments were synthesized starting from shortened truncated H1 genes and capped as the wild-type mRNA with a m⁷G-cap structure using the ScriptCapTM m⁷G Capping System. One nt deletion induced a severe effect with a near disappearance of the crosslinks of nts 41–44 and

new crosslinks at nts 93–94, 96–97 whereas crosslinks at nts 103–105 and 105–106 are maintained. This suggests that nts 93–97 are spatially close of the stem 1 in the 19W and can both react with mutant +1 nt of H1 mRNA. Deletions –2 to –3 nts show that the strong crosslink at positions 41–42 is definitively lost and confirm the vicinity of nts 93–97 and nts 103–106 (Supplementary Fig. 1).

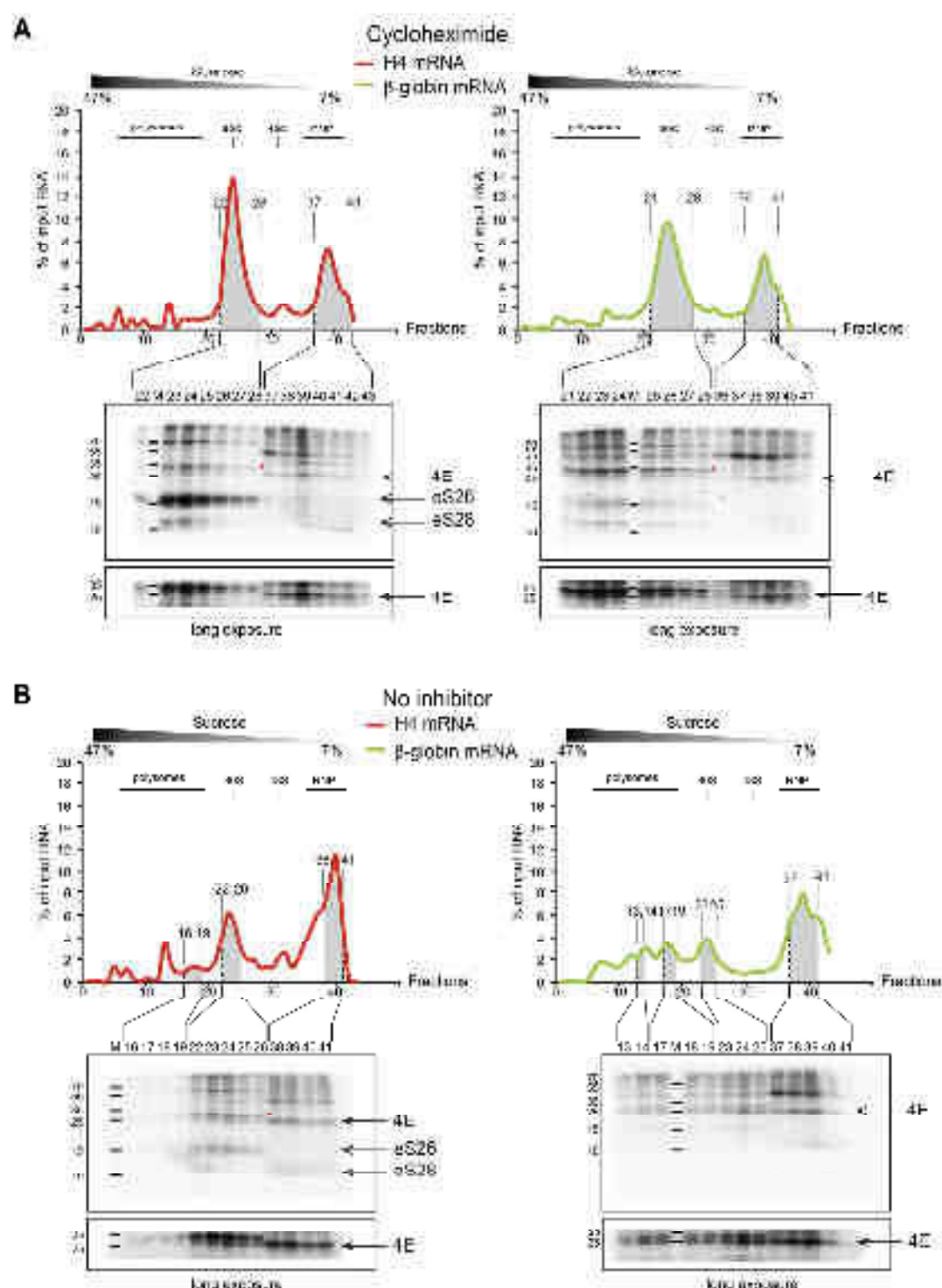


Fig. 4. Chemical crosslinking of isolated human H1 and β -globin mRNA and sucrose fractionation. **A:** Sucrose gradient fractionation of cycloheximide blocked translation in 293T. The percentage of input mRNA is measured by densitometry. The sucrose gradient fractionation profiles (shown in grey) are analysed by SDS-PAGE. Characterized cap-specific crosslinks are indicated by arrows; red arrows indicate anticardiac, cap-specific crosslinks. **B:** Sucrose gradient fractionation of translation in 293T. The percentage of input mRNA is measured by densitometry. The sucrose gradient fractions that are shown in grey are analysed and analysed by SDS-PAGE. Characterized cap-specific crosslinks are indicated by black arrows; uncharacterized cap-specific crosslinks are indicated by red arrows. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

Most of these mutants exhibit strong crosslink reactivity around nts 104–105. After a 3 nt-deletion, the mutants exhibit different reactivities spread over nts 13 and 30 that indicate that short UTRs can react with closer regions of the mRNA structure (Supplementary Fig. 1). Following these progressive deletions of the UTR, we performed six additional deletions that brought the reactive cap at positions 16, 22, 27, 29, 31 and 37 of H4 mRNA, except deletion 27, all the mutants lose their ability to react with the original cap-binding site in the pocket 41–44 and 104–106. With these mutants, new positions are crosslinked, providing new data on their position versus the modified cap. Surprisingly, mutant –27 which has the cap at the bottom of stem I, recovered the strong crosslinking activity of nts 41–44 but, no reactivity on the other strand of the stem. This indicates that the cap of mutant –27 does not bind the CBP in the native manner since it does not react as the wild-type mRNA does with nts 103–106. The low energy crosslinking experiments might be useful to detect close interactions such as the formation of the m⁷G cap in the eukaryotic context of H4 mRNA. In the H4 case, the specificity of the binding pocket was demonstrated by performing a competition assay with m⁷GpppG cap analogue [5]. Deletion experiments of nts of the 5' UTR revealed the spatial proximity of the CBP of stem I and stem III where recurrent crosslinks were found. One major drawback of the deletion approach may concern either the alteration of the folding of the mRNA structure or the severe reduction of reactivity observed with several mutants. For instance, we repeatedly observed low reactivity of several independent preparations of mutants –5, –16 and –22 nts. Whether this is due to an unusual high flexibility of the RNA sequence harbouring the 5' cap resulting in a higher exposure to the solvent side remains to be explored.

2.2. Chemical crosslinking of the cap in *ANI*

The next issue we addressed was the position of the cap in H4 mRNA during the translation initiation process. For that purpose, we performed chemical cross-linking experiments using mRNA that contain a radiolabelled m⁷Gcap that has been oxidised. This histone H4 mRNA transcript was then incubated in ERL in the presence of edeine, GMP PNP or cycloheximide and then further reduced to induce covalent bonds with proteins that are in contact with the cap. As a control, we performed the same experiment with radiolabelled oxidised capped β -globin mRNA. Crosslinking experiments in the presence of an excess of m⁷GDP as a competitor allowed us to discriminate cap-specific crosslinks that are attenuated versus unspecific bonds that are still present with the competitor (Fig. 3). In the case of H4 mRNA, we detected specific

crosslinks of eIF4E, eS26 and eS28 that were identified with specific antibodies as previously described [7]. The crosslinks of eIF4E are detected with edeine and GMP/PPNP, but are drastically decreased with cycloheximide. This suggests that eIF4E is bound on the cap on the 43S and 48S complexes but is released from the 80S complex. On the contrary, eIF4E is bound on the cap on β -globin mRNA and remained bound after 80S assembly. Concerning eS26 and eS28, no specific crosslinks are detected on β -globin mRNA. Altogether, these results indicate that in the case of H4 mRNA translation, eIF4E is bound on the cap prior and after 48S complex formation, then eIF4E is released from the cap and is replaced by eS26 and eS28 in the 80S complex. In contrast, eIF4E stayed bound on the cap of β -globin mRNA during the whole translation initiation process. However, the major drawback of this experiment is that the whole input oxidised mRNA is most probably not fully assembled in translation initiation complexes suggesting that a part of the oxidised mRNA is in fact crosslinked with soluble proteins that are also visualised here as testified by the radioactive count profiles (Fig. 4). In other words, the detected crosslinks could also be due to RNPs components in contact with the cap rather than ribosomal complex proteins.

2.3. Translation initiation step involved chemical crosslinking of the cap

In order to unambiguously assign the crosslinked proteins to translation initiation complexes, we coupled the chemical crosslinking approach with sucrose gradient sedimentation. The reduced proteins contained in each sucrose gradient fraction are precipitated and analysed by SDS-PAGE (Fig. 5). When translation was blocked with cycloheximide, the eS26 and eS28 crosslinks with H4 mRNA co-sedimented with 80S fractions confirming that the cap is indeed crosslinked to ribosomal proteins that are embedded in the fully assembled ribosome (Fig. 4A). In contrast, eIF4E crosslink was found in the soluble RNP fractions at the top of the gradient. Interestingly, with β -globin mRNA, eIF4E crosslinks were found both in the soluble RNP fractions and in the 80S fractions. We also performed the same experiments without any inhibitor. As expected, we obtained polyosomes indicating ongoing translation (Fig. 4B). eIF4E-crosslinks on β -globin mRNA were still found in soluble RNP fractions and in 80S fractions. However, the results are different for histone H4 mRNA, when translation is active, eIF4E crosslinks are also detected in 80S fractions. All together, these results allowed us to propose a dynamic model for the cap and eIF4E fates during translation initiation and elongation (Fig. 5). In the case of β -globin mRNA, eIF4E is bound to the cap

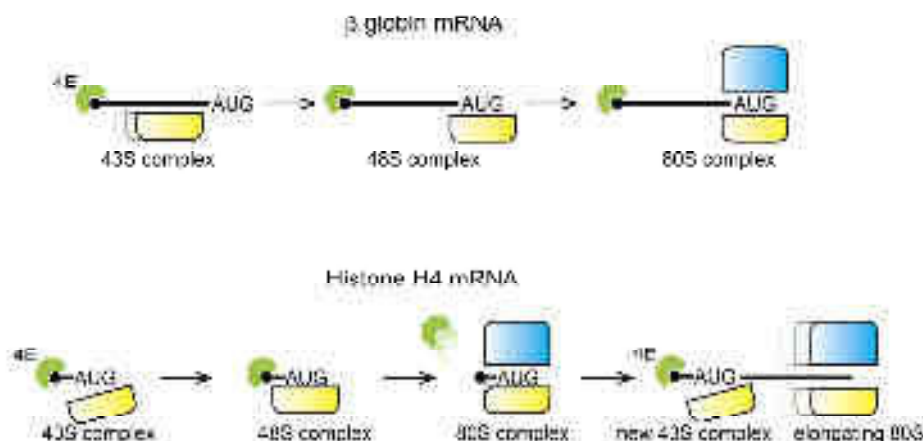


Fig. 5. Model of the cap fate during translation of histone H4 mRNA and β -globin mRNA. During β -globin mRNA translation, eIF4E remains bound on the cap during the whole process. On the contrary, H4 mRNA translation requires eIF4E binding/release cycles to allow efficient ribosome assembly.

for efficient 43S recruitment to promote 5'-3' scanning. After 48S assembly and 80S formation on the AUG start codon, eIF4E stays on the cap in order to promote further recruitment of the next 43S particle. Histone H4 mRNA is translated by a rather different mechanism. First, eIF4E is bound to the 5' cap for 43S tethering on the AUG start codon without any scanning [7,5]. At the 48S stage, eIF4E is still bound on the cap, but 80S formation requires the release of eIF4E to allow specific contacts with eS28 and eS26 on the small ribosomal subunit. This is correlated with the fact that the 5' UTR of H4 mRNA is only 9-nt long. Indeed, such a short 5' UTR is seemingly incompatible with eIF4E bound to the cap and an assembled 80S because the large subunit would clash with eIF4E. Indeed, many functional studies by toe- and footprinting experiments have demonstrated that the 80S covers ~14 to 19 nucleotides upstream of the AUG start codon. This is probably why eIF4E has to be released from the cap in order to allow 60S subunit joining for histone H4 translation initiation. However, during translation elongation, eIF4E-crosslinks are also detected in the 80S fractions. This suggests that once the 80S ribosome has proceeded to elongation, eIF4E binds to the cap that is now accessible again to enable translation initiation for the next ribosome by tethering. In conclusion our approach allowed us to monitor the 'breathing' of eIF4E which undergo successive cycles of binding/release during the translation process.

4. Conclusion

By combining chemical and UV crosslinking approaches of the cap, we have been able to track the proteins bound to cap that is present on an mRNA of interest during the translation process. By identifying the proteins that are in proximity to the cap, these methods give interesting insights on the dynamic interplay between eIF, the ribosome and mRNA. It will be of interest to further use these approaches on other mRNAs in order for example to assess the influence of 5'UTR sequence. The accurate position of the cap and cap-binding factors such as eIF4F during scanning is still poorly understood. Crosslinking of the cap is a method of choice to address such fundamental aspects of translation initiation.

Acknowledgments

We are grateful to Liz Grayhack and Jeff Collier for their kind invitation to write this article. We are also grateful to Bruno Sargueil for YpGlo plasmid. This work was supported by CNRS, Université de Strasbourg and Agence Nationale pour la Recherche (ANR-11-SVSE802501; ANR-17-CE11-0024-01; ANR-17-CE12-0025-01).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.ymeth.2017.12.019>.

References

- [1] A.G. Hinnebusch, The scanning mechanism of eukaryotic translation initiation, *Annu. Rev. Biochem.* 83 (2014) 779–812, <https://doi.org/10.1146/annurev-biochem-060713-035802>.
- [2] A.G. Hinnebusch, Structural insights into the mechanism of scanning and start codon recognition in eukaryotic translation initiation, *Trends Biochem. Sci.* 42 (2017) 589–611, <https://doi.org/10.1016/j.tibs.2017.03.004>.
- [3] S. Jaeger, S. Barends, R. Giegé, G. Eriani, F. Martin, Expression of metazoan replication-dependent histone genes, *Biochimie* 87 (2005) 827–834, <https://doi.org/10.1016/j.biochi.2005.03.012>.
- [4] V.S. Meier, R. Bohni, D. Schumperli, Nucleotide sequence of two mouse histone H4 genes, *Nucl. Acids Res.* 17 (1989) 795.
- [5] F. Martin, S. Barends, S. Jaeger, L. Schaeffer, L. Prongidi-Fix, G. Eriani, Cap-assisted internal initiation of translation of histone H4, *Mol. Cell.* 41 (2011) 197–209.
- [6] M. Nowakowska, J. Kowalska, F. Martin, A. D'Orchymont, J. Zuberek, M. Lukaszewicz, E. Darzynkiewicz, J. Jemielity, Cap analogs containing 6-thioguanosine – reagents for the synthesis of mRNAs selectively photocrosslinkable with cap-binding biomolecules, *Org. Biomol. Chem.* 12 (2014) 4841–4847, <https://doi.org/10.1039/c4ob00059e>.
- [7] F. Martin, J.-F. Ménétret, A. Simonetti, A.G. Myasnikov, Q. Vicens, L. Prongidi-Fix, S.K. Natchiar, B.P. Klaholz, G. Eriani, Ribosomal 18S rRNA base pairs with mRNA during eukaryotic translation initiation, *Nat. Commun.* 7 (2016) 12622, <https://doi.org/10.1038/ncomms12622>.
- [8] F. Karabiber, J.L. McGinnis, O.V. Favorov, K.M. Weeks, QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis, *RNA* 19 (2013) 63–73, <https://doi.org/10.1261/rna.036327.112>.
- [9] N. Sonenberg, A.J. Shatkin, Reovirus mRNA can be covalently crosslinked via the 5' cap to proteins in initiation complexes, *Proc. Natl. Acad. Sci. U.S.A.* 74 (1977) 4288–4292.
- [10] N. Garreau de Loubresse, I. Prokhorova, W. Holtkamp, M.V. Rodnina, G. Yusupova, M. Yusupov, Structural basis for the inhibition of the eukaryotic ribosome, *Nature* 513 (2014) 517–522, <https://doi.org/10.1038/nature13737>.
- [11] N.K. Gray, M.W. Hentze, Regulation of protein synthesis by mRNA structure, *Mol. Biol. Rep.* 19 (1994) 195–200.
- [12] T. Schneider-Poetsch, J. Ju, D.E. Elyer, Y. Dang, S. Bhat, W.C. Merrick, R. Green, B. Shen, J.O. Liu, Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin, *Nat. Chem. Biol.* 6 (2010) 209–217, <https://doi.org/10.1038/nchembio.304>.
- [13] A. Favre, G. Moreno, M.O. Blondel, J. Kliber, F. Vinzens, C. Salet, 4-thiouridine photosensitized RNA-protein crosslinking in mammalian cells, *Biochem. Biophys. Res. Commun.* 141 (1986) 847–854, [https://doi.org/10.1016/S0006-291X\(86\)80250-9](https://doi.org/10.1016/S0006-291X(86)80250-9).
- [14] A.V. Pisarev, V.G. Kolupaeva, M.M. Yusupov, C.U.T. Hellen, T.V. Pestova, Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes, *EMBO J.* 27 (2008) 1609–1621, <https://doi.org/10.1038/emboj.2008.90>.
- [15] O. Haimov, H. Sinvani, F. Martin, I. Ulitsky, R. Emmanuel, A. Tamarkin-Ben-Harush, A. Vardy, R. Dikstein, Efficient and accurate translation initiation directed by TISU involves RPS3 and RPS10e binding and differential eukaryotic initiation factor 1A regulation, *Mol. Cell. Biol.* 37 (2017) e00150–e217, <https://doi.org/10.1128/MCB.00150-17>.

ANNEXE III : FIGURE SUPPLÉMENTAIRE

5'-prox ↗	nb de spectres (moyenne)			
	Protein	Edeine	GMP-PNP	Cyclo
Aac11	3	-	-	-
Aats-tyr	8	-	-	-
Akap200	23,3	-	-	-
Aldh-III	3,7	-	-	-
alpha-Spec	41,3	-	-	-
AnxB9	6,7	-	-	-
Atpalpha	19,7	-	-	-
ATPsyn-beta	5	-	-	-
bel	54,3	-	-	-
beta13F	5	-	-	-
beta'Cop	3,3	-	-	-
blw	6,3	-	-	-
CaBP1	5,7	-	-	-
Ca-P60A	10,7	-	-	-
Capr	30	-	-	-
Cat	3,7	-	-	-
Cctgamma	5	-	-	-
CD98hc	7	-	-	-
CG10236-RA	15	-	-	-
CG10373-RA	3,7	-	-	-
CG10527	7,7	-	-	-
CG10576	27,3	-	-	-
CG11961	4	-	-	-
CG13364-RA	4,7	-	-	-
CG13630	4	-	-	-
CG13887	3,3	-	-	-
CG1416	3,7	-	-	-
CG17514	5,3	-	-	-
CG17544	-	6,67	-	-
CG2852	5,7	-	-	-
CG4829	4	-	-	-
CG4933	3	-	-	-
CG5525	4,7	-	-	-
CG7033	6	-	-	-
CG8963	3	-	-	-
CG9135	3,7	-	-	-
Chc	4,3	16,33	-	-
Cpr	4,7	-	-	-
Crc	13	-	-	-
CS2beta	3,7	-	-	-
Dhc64C	8,7	-	-	-
Dp1	80,7	-	-	-
eIF-4a	65,3	-	-	-
eIF-5A	7,7	-	-	-
ergic53	10,3	-	-	-
ERp60	51,3	-	-	-
Gale	5	-	-	-
Ge-1	3,7	-	-	-
GM12270p	5	-	-	-
Gp210	3	-	-	-
Hel25E	16,7	-	-	-
Hmu	3	-	-	-
Hsc70Cb	11	-	-	-
Karybeta3	11,3	-	-	-
Khc	4,3	-	-	-
Klp10A	12	-	-	-
Klp61F	-	3,00	-	-
kra	25	-	-	-
kst	43,3	-	-	-
l(1)G0320	5,3	-	-	-

5'-prox ↗	nb de spectres (moyenne)			
	Protein	Edeine	GMP-PNP	Cyclo
LanB2	11,3	-	-	-
Mec2	10	-	-	-
MSBP	3,7	-	-	-
mmps	8,3	-	-	-
mys	3,7	-	-	-
nAChRbeta3	4,3	-	-	-
NAT1	28,3	-	-	-
Not1	3,3	-	-	-
Ost48	7	-	-	-
OstDelta	3,7	-	-	-
p93	12,7	-	-	-
Past1	9	-	-	-
Pdi	52,3	-	-	-
Pen	5,3	-	-	-
poly	3,3	-	-	-
porin	3,3	-	-	-
Prosalph7	4,3	-	-	-
Psi	13,3	-	-	-
Pvr	18	-	-	-
Psi	12	-	-	-
PyK	17,7	-	-	-
Rab1	20,7	-	-	-
Rab11-RB	4,3	-	-	-
Rab6	3,7	-	-	-
Ran	6,3	-	-	-
Rbp2	18	-	-	-
RE55992p	3,7	-	-	-
RE70805p	76,7	-	-	-
Rho	8	-	-	-
RhoL	-	2,67	-	-
RnrL	3	-	-	-
Rpl36A-RA	5,7	-	-	-
Rtn1	7	-	-	-
san	3	-	-	-
SD08036p	3	-	-	-
Sec16	-	23,67	-	-
SeiD	3,7	-	-	-
Sep1	7	-	-	-
Sep2	6,7	-	-	-
smid	4	-	-	-
snf	4	-	-	-
Sra-1	4	-	-	-
Srp54	-	-	-	3,00
Stim-RA	3,7	-	-	-
Syb-RA	4,3	-	-	-
Tcp-1eta	6	-	-	-
Tctp-RA	31,7	-	-	-
TER94	23	-	-	-
Tppll	7,3	-	-	-
trsn	8,3	-	-	-
Tudor-SN	44,3	-	-	-
Ugt58Fa	3	-	-	-
und	12	-	-	-
unk	-	58,00	-	-
Usp10	3	-	-	-
Vha68-1	3	-	-	-
Vha68-2	6	-	-	-
vnc	3,3	-	-	-
yata	6	-	-	-
yin	6,7	-	-	-

Supplément Figure 29 : Analyse des données de spectrométrie de masse du chapitre I. Détail des protéines enrichies sur l'extrémité 5'-proximale (5'-prox) de la région 5'UTR du CrPV pour chaque inhibiteur testé (édéine, GMP-PNP, cycloheximide (cyclo)). Les histogrammes en vert et les chiffres indiquent le nombre de spectres moyen de chaque protéine enrichie sur l'extrémité 5'proximale. La moyenne est calculée sur les valeurs des spectres obtenus lors de trois expériences de spectrométrie de masse indépendantes.

BIBLIOGRAPHIE

BIBLIOGRAPHIE

- A**li, N., & Siddiqui, a. (1997). The La antigen binds 5' noncoding region of the hepatitis C virus RNA in the context of the initiator AUG codon and stimulates internal ribosome entry site-mediated translation. *PNAS*, *94*(6), 2249–2254.
- Altmann, M., Edery, I., Sonenberg, N., & Trachsel, H. (1985). Purification and characterization of protein synthesis initiation factor eIF-4E from the yeast *Saccharomyces cerevisiae*. *Biochemistry*, *24*(22), 6085–9.
- Andino, R., Rieckhof, G. E., Achacoso, P. L., & Baltimore, D. (1993). Poliovirus RNA synthesis utilizes an RNP complex formed around the 5' -end of viral RNA. *EMBO Journal*, *12*(9), 3587–3598.
- Andino, R., Rieckhof, G. E., & Baltimore, D. (1990). A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell*, *63*(2), 369–380.
- Andreev, D. E., Fernandez-Miragall, O., Ramajo, J., Dmitriev, S. E., Terenin, I. M., Martinez-Salas, E., & Shatsky, I. N. (2007). Differential factor requirement to assemble translation initiation complexes at the alternative start codons of foot-and-mouth disease virus RNA. *RNA*, *13*(8), 1366–1374.
- Anger, A. M., Armache, J.-P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D. N., & Beckmann, R. (2013). Structures of the human and *Drosophila* 80S ribosome. *Nature*, *497*(7447), 80–85.
- Angulo, J., Ulryck, N., Deforges, J., Chamond, N., Lopez-Lastra, M., Masquida, B., & Sargueil, B. (2016). Loop III_d of the HCV IRES is essential for the structural rearrangement of the 40S-HCV IRES complex. *Nucleic Acids Research*, *44*(3), 1309–1325.
- Asnani, M., Pestova, T. V., & Hellen, C. U. T. (2016). Initiation on the divergent Type I cadicivirus IRES: factor requirements and interactions with the translation apparatus. *Nucleic Acids Research*, *44*(7), 3390–3407.
- Au, H. H., Cornilescu, G., Mouzakis, K. D., Ren, Q., Burke, J. E., Lee, S., ... Jan, E. (2015). Global shape mimicry of tRNA within a viral internal ribosome entry site mediates translational reading frame selection. *PNAS*, *112*(47), 6446-55
- Au, H. H. T., & Jan, E. (2012). Insights into Factorless Translational Initiation by the tRNA-Like Pseudoknot Domain of a Viral IRES. *PLoS One*, *7*(12), 51477.
- B**aird, S. D., Lewis, S. M., Turcotte, M., & Holcik, M. (2007). A search for structurally similar cellular internal ribosome entry sites. *Nucleic Acids Research*, *35*(14), 4664–4677.

- 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*(2), 221–32.
- Benne, R., & Hershey, J. W. (1978). The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. *The Journal of Biological Chemistry*, *253*(9), 3078–87.
- Berlanga, J. J., Ventoso, I., Harding, H. P., Deng, J., Ron, D., Sonenberg, N., ... de Haro, C. (2006). Antiviral effect of the mammalian translation initiation factor 2 α kinase GCN2 against RNA viruses. *The EMBO Journal*, *25*(8), 1730–1740.
- Berry, K. E., Waghray, S., & Doudna, J. A. (2010). The HCV IRES pseudoknot positions the initiation codon on the 40S ribosomal subunit. *RNA*, *16*(8), 1559–69.
- Besnard-Guérin, C., Jacquier, C., Pidoux, J., Deddouche, S., & Antoniewsk, C. (2015). The Cricket Paralysis Virus Suppressor Inhibits microRNA Silencing Mediated by the *Drosophila* Argonaute-2 Protein. *PLoS One*, *10*(3), e0120205.
- Beznošková, P., Cuchalová, L., Wagner, S., Shoemaker, C. J., Gunišová, S., von der Haar, T., & Valášek, L. S. (2013). Translation Initiation Factors eIF3 and HCR1 Control Translation Termination and Stop Codon Read-Through in Yeast Cells. *PLoS Genetics*, *9*(11), e1003962.
- Bhat, M., Robichaud, N., Hulea, L., Sonenberg, N., Pelletier, J., & Topisirovic, I. (2015). Targeting the translation machinery in cancer. *Nature Reviews Drug Discovery*, *14*(4), 261–278.
- Blanchet, S., Rowe, M., Von der Haar, T., Fabret, C., Demais, S., Howard, M. J., & Namy, O. (2015). New insights into stop codon recognition by eRF1. *Nucleic Acids Research*, *43*(6), 3298–308.
- Blau, L., Knirsh, R., Ben-Dror, I., Oren, S., Kuphal, S., Hau, P., ... Vardimon, L. (2012). Aberrant expression of c-Jun in glioblastoma by internal ribosome entry site (IRES)-mediated translational activation. *PNAS*, *109*(42), 2875–2884.
- Block, K. L., Vornlocher, H. P., & Hershey, J. W. (1998). Characterization of cDNAs encoding the p44 and p35 subunits of human translation initiation factor eIF3. *The Journal of Biological Chemistry*, *273*(48), 31901–8.
- Bonning, B. C., & Miller, W. A. (2010). Dicistroviruses. *Annual Review of Entomology*, *55*(1), 129–150.
- C**hekulaeva, M., & Rajewsky, N. (2018). Roles of Long Noncoding RNAs and Circular RNAs in Translation. *Cold Spring Harbor Perspectives in Biology*, on line, a032680.
- Chen, C., & Sarnow, P. (1998). Internal Ribosome Entry Sites Tests with Circular mRNAs. In *Protein Synthesis*, *77*, 355-364

- Chen, T.-M., Shih, Y.-H., Tseng, J. T., Lai, M.-C., Wu, C.-H., Li, Y.-H., ... Sun, H. S. (2014). Overexpression of FGF9 in colon cancer cells is mediated by hypoxia-induced translational activation. *Nucleic Acids Research*, *42*(5), 2932–2944.
- Chicher, J., Simonetti, A., Kuhn, L., Schaeffer, L., Hammann, P., Eriani, G., & Martin, F. (2015). Purification of mRNA-programmed translation initiation complexes suitable for mass spectrometry analysis. *Proteomics*, *15*(14), 2417–2425.
- Clouse, K. N., Ferguson, S. B., & Schüpbach, T. (2008). Squid, Cup, and PABP55B function together to regulate gurken translation in *Drosophila*. *Developmental Biology*, *313*(2), 713–724.
- Cobbold, L. C., Spriggs, K. A., Haines, S. J., Dobbyn, H. C., Hayes, C., de Moor, C. H., ... Willis, A. E. (2008). Identification of Internal Ribosome Entry Segment (IRES)-trans-Acting Factors for the Myc Family of IRESs. *Molecular and Cellular Biology*, *28*(1), 40–49.
- Cohen-Chalamish, S., Hasson, A., Weinberg, D., Namer, L. S., Banai, Y., Osman, F., & Kaempfer, R. (2009). Dynamic refolding of IFN- γ mRNA enables it to function as PKR activator and translation template. *Nature Chemical Biology*, *5*(12), 896–903.
- Connor, J. H., & Lyles, D. S. (2005). Inhibition of Host and Viral Translation during Vesicular Stomatitis Virus Infection: eIF2 is responsible for the inhibition of viral but not host translation. *Journal of Biological Chemistry*, *280*(14), 13512–13519.
- Costa-mattioli, Mauro;Svitkin, Yuri;Sonenberg, N. (2004). La Autoantigen Is Necessary for Optimal Function of the Poliovirus and Hepatitis C Virus Internal Ribosome Entry Site In Vivo and In Vitro La Autoantigen Is Necessary for Optimal Function of the Poliovirus and Hepatitis C Virus Internal Ribosome Entry Sit. *Molecular and Cellular Biology*, *24*(15), 6861–6870.
- Costantino, D., & Kieft, J. S. (2005). A preformed compact ribosome-binding domain in the cricket paralysis-like virus IRES RNAs. *RNA*, *11*, 332–343.
- Culjkovic, B., Topisirovic, I., & Borden, K. L. B. (2007). Controlling Gene Expression through RNA Regulons: The Role of the Eukaryotic Translation Initiation Factor eIF4E. *Cell Cycle*, *6*(1), 65–69.
- Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M., & Borden, K. L. B. (2006). eIF4E is a central node of an RNA regulon that governs cellular proliferation. *The Journal of Cell Biology*, *175*(3), 415–426.
- D**es Georges, A., Dhote, V., Kuhn, L., Hellen, C. U. T., Pestova, T. V., Frank, J., & Hashem, Y. (2015). Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature*, *525*(7570), 491–5.

- Dever, T. E., & Green, R. (2012). The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harbor Perspectives in Biology*, 4(7), a013706
- Dmitriev, S. E., Terenin, I. M., Andreev, D. E., Ivanov, P. A., Dunaevsky, J. E., Merrick, W. C., & Shatsky, I. N. (2010). GTP-independent tRNA Delivery to the Ribosomal P-site by a Novel Eukaryotic Translation Factor. *Journal of Biological Chemistry*, 285(35), 26779–26787.
- Dinos, G., Wilson, D. N., Teraoka, Y., Szaflarski, W., Fucini, P., Kalpaxis, D., & Nierhaus, K. H. (2004). Dissecting the Ribosomal Inhibition Mechanisms of Edeine and Pactamycin: The Universally Conserved Residues G693 and C795 Regulate P-Site RNA Binding. *Molecular Cell*, 13(1), 113–124.
- Dupont, J. A., & Snoussi, K. (2009). Mg²⁺ modulation of EMCV IRES key activity fragment equilibria and r(G·C) base-pair kinetics. *Journal of Biological Physics*, 35(3), 231–243.
- E**Antak, L., Wagner, S., Herrmannová, A., Karásková, M., Rutkai, E., Lukavsky, P. J., & Valášek, L. (2010). The Indispensable N-Terminal Half of eIF3j/HCR1 Cooperates with its Structurally Conserved Binding Partner eIF3b/PRT1-RRM and with eIF1A in Stringent AUG Selection. *Journal of Molecular Biology*, 396(4), 1097–1116.
- Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H. S., ... Ban, N. (2014). Molecular Architecture of the 40S-eIF1-eIF3 Translation Initiation Complex. *Cell*, 159(5), 1227–1228.
- Etchison, D., Milburn, S. C., Ederly, I., Sonenberg, N., & Hershey, J. W. (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. *Journal of Biological Chemistry*, 257(24), 14806–14810.
- F**ernandez-Chamorro, J., Piñeiro, D., Gordon, J. M. B., Ramajo, J., Francisco-Velilla, R., Macias, M. J., & Martínez-Salas, E. (2014). Identification of novel non-canonical RNA-binding sites in Gemin5 involved in internal initiation of translation. *Nucleic Acids Research*, 42(9), 5742–54.
- Fernández-Miragall, O., & Martínez-Salas, E. (2003). Structural organization of a viral IRES depends on the integrity of the GNRA motif. *RNA*, 9(11), 1333–44.
- Fernandez, J., Yaman, I., Huang, C., Liu, H., Lopez, A. B., Komar, A. A., ... Hatzoglou, M. (2005). Ribosome Stalling Regulates IRES-Mediated Translation in Eukaryotes, a Parallel to Prokaryotic Attenuation. *Molecular Cell*, 17(3), 405–416.
- Fernandez, J., Yaman, I., Merrick, W. C., Koromilas, A., Wek, R. C., Sood, R., ... Hatzoglou,

- M. (2002). Regulation of Internal Ribosome Entry Site-mediated Translation by Eukaryotic Initiation Factor-2 α Phosphorylation and Translation of a Small Upstream Open Reading Frame. *Journal of Biological Chemistry*, 277(3), 2050–2058.
- Fiore JL, Nesbitt DJ. (2013). An RNA folding motif: GNRA tetraloop–receptor interactions. *Q Rev Biophys* 46, 223–264.
- Flather, D., & Semler, B. L. (2015). Picornaviruses and nuclear functions: targeting a cellular compartment distinct from the replication site of a positive-strand RNA virus. *Frontiers in Microbiology*, 6, 594.
- Fraser, C. S., Berry, K. E., Hershey, J. W. B., & Doudna, J. A. (2007). eIF3j Is Located in the Decoding Center of the Human 40S Ribosomal Subunit. *Molecular Cell*, 26(6), 811–819.
- Fuchs, G., Petrov, A. N., Marceau, C. D., Popov, L. M., Chen, J., O’Leary, S. E., ... Puglisi, J. D. (2015). Kinetic pathway of 40S ribosomal subunit recruitment to hepatitis C virus internal ribosome entry site. *PNAS*, 112(2), 319–325.
- G**araigorta, U., & Chisari, F. V. (2009). Hepatitis C Virus Blocks Interferon Effector Function by Inducing Protein Kinase R Phosphorylation. *Cell Host & Microbe*, 6(6), 513–522.
- Garreau de Loubresse, N., Prokhorova, I., Holtkamp, W., Rodnina, M. V., Yusupova, G., & Yusupov, M. (2014). Structural basis for the inhibition of the eukaryotic ribosome. *Nature*, 513(7519), 517–522.
- Garrey, J. L., Lee, Y.-Y., Au, H. H. T., Bushell, M., & Jan, E. (2010). Host and viral translational mechanisms during cricket paralysis virus infection. *Journal of Virology*, 84(2), 1124–1138.
- Ghetti, A., Piñol-Roma, S., Michael, W. M., Morandi, C., & Dreyfuss, G. (1992). hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Research*, 20(14), 3671–8.
- Gingras, A. C., Svitkin, Y., Belsham, G. J., Pause, A., & Sonenberg, N. (1996). Activation of the translational suppressor 4E-BP1 following infection with encephalomyocarditis virus and poliovirus. *PNAS*, 93(11), 5578–83.
- Golob-Schwarzl, N., Schweiger, C., Koller, C., Krassnig, S., Gogg-Kamerer, M., Gantenbein, N., ... Haybaeck, J. (2017). Separation of low and high grade colon and rectum carcinoma by eukaryotic translation initiation factors 1, 5 and 6. *Oncotarget*, 8(60), 101224-101243
- Gradi A, Svitkin YV, Imataka H, 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.

- Gross, L., Vicens, Q., Einhorn, E., Noireterre, A., Schaeffer, L., Kuhn, L., ... Martin, F. (2017). The IRES 5'UTR of the dicistrovirus cricket paralysis virus is a type III IRES containing an essential pseudoknot structure. *Nucleic Acids Research*, *45*(15), 8993–9004.
- Guo, L., Zaysteva, O., Nie, Z., Mitchell, N. C., Amanda Lee, J. E., Ware, T., ... Quinn, L. M. (2016). Defining the essential function of FBP/KSRP proteins: Drosophila Psi interacts with the mediator complex to modulate MYC transcription and tissue growth. *Nucleic Acids Research*, *44*(16), 7646–7658.
- H**alder, K., Wieland, M., & Hartig, J. S. (2009). Predictable suppression of gene expression by 5'-UTR-based RNA quadruplexes. *Nucleic Acids Research*, *37*(20), 6811–6817.
- Harding, H. P., Zhang, Y., & Ron, D. (1999). Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature*, *397*(6716), 271–274.
- Harcourt, E. M., Kietrys, A. M., & Kool, E. T. (2017). Chemical and structural effects of base modifications in messenger RNA. *Nature*, *541*(7637), 339–346.
- Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H. Y., Grassucci, R. A., ... Frank, J. (2013). Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell*, *153*(5), 1108–19.
- 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*, (February), 1–27.
- Hatin, I., Fabret, C., Rousset, J.-P., & Namy, O. (2009). Molecular dissection of translation termination mechanism identifies two new critical regions in eRF1. *Nucleic Acids Research*, *37*(6), 1789–98.
- Hellen, C. U. T. (2009). IRES-induced conformational changes in the ribosome and the mechanism of translation initiation by internal ribosomal entry. *Biochim Biophys Acta.*, *1789*(9–10), 558–570.
- Henke, J. I., Goergen, D., Zheng, J., Song, Y., Schüttler, C. G., Fehr, C., ... Niepmann, M. (2008). microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO Journal*, *27*(24), 3300–3310.
- Hentze, M. W. (1997). eIF4G: a multipurpose ribosome adapter? *Science*, *275*(5299), 500–1.
- Hershey, J. W. B. (2010). Regulation of protein synthesis and the role of eIF3 in cancer. *Brazilian Journal of Medical and Biological Research*, *43*(10), 920–30.
- Hinnebusch, A. G. (2006). eIF3: a versatile scaffold for translation initiation complexes. *Trends in Biochemical Sciences*, *31*(10), 553–562.
- Hinnebusch, A. G. (2014). The scanning mechanism of eukaryotic translation initiation.

Annual Review of Biochemistry, 83, 779–812.

- Hinnebusch, A. G. (2017). Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends in Biochemical Sciences*, 42(8), 589-611.
- Hinnebusch, A. G., & Lorsch, J. R. (2012a). The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges. *Cold Spring Harbor Perspectives in Biology*, 4(10), a011544.
- Hinnebusch, A. G., & Lorsch, J. R. (2012b). The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges. *Cold Spring Harbor Perspectives in Biology*, 4(10), a01154.
- Hoffman, M. A., & Palmenberg, A. C. (1995). Mutational analysis of the J-K stem-loop region of the encephalomyocarditis virus IRES. *Journal of Virology*, 69(7), 4399–406.
- Hunnicutt, L. E., Hunter, W. B., Cave, R. D., Powell, C. A., & Mozoruk, J. J. (2006). Genome sequence and molecular characterization of Homalodisca coagulata virus-1, a novel virus discovered in the glassy-winged sharpshooter (Hemiptera: Cicadellidae). *Virology*, 350(1), 67–78.
- Hussain, T., Llácer, J. L., Fernández, I. S., Muñoz, A., Martín-Marcos, P., Savva, C. G., ... Ramakrishnan, V. (2014). Structural Changes Enable Start Codon Recognition by the Eukaryotic Translation Initiation Complex. *Cell*, 159(3), 597–607.
- Imai, S., Kumar, P., Hellen, C. U. T., D'Souza, V. M., & Wagner, G. (2016). An accurately preorganized IRES RNA structure enables eIF4G capture for initiation of viral translation. *Nature Structural & Molecular Biology*, 23(9), 859–864.
- Jackson, R. J. (2005). Alternative mechanisms of initiating translation of mammalian mRNAs. *Biochemical Society Transactions*, 33(6), 1231.
- Jackson, R. J., Hellen, C. U. T., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews. Molecular Cell Biology*, 11(2), 113–127.
- Jaeger, S., Barends, S., Giegé, R., Eriani, G., & Martin, F. (2005). Expression of metazoan replication-dependent histone genes. *Biochimie*, 87(9–10), 827–834.
- Jaeger, S., Martin, F., Rudinger-Thirion, J., Giegé, R., & Eriani, G. (2006). Binding of human SLBP on the 3'-UTR of histone precursor H4-12 mRNA induces structural rearrangements that enable U7 snRNA anchoring. *Nucleic Acids Research*, 34(17), 4987–95.

- Jan, E., Kinzy, T. G., & Sarnow, P. (2003). Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis. *PNAS*, *100*(26), 15410–15415.
- Jan, E., & Sarnow, P. (2002). Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *Journal of Molecular Biology*, *324*(5), 889–902.
- Jang, S. K., Kräusslich, H. G., Nicklin, M. J., Duke, G. M., Palmenberg, A. C., & Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *Journal of Virology*, *62*(8), 2636–43.
- Jang, S. K., & Wimmer, E. (1990). Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. *Genes & Development*, *4*(9), 1560–72.
- Jopling, C. L., Yi, M., Lancaster, A. M., Lemon, S. M., & Sarnow, P. (2005). Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science*, *309*(5740), 1577–81.
- Jordan, R., Wang, L., Graczyk, T. M., Block, T. M., & Romano, P. R. (2002). Replication of a Cytopathic Strain of Bovine Viral Diarrhea Virus Activates PERK and Induces Endoplasmic Reticulum Stress-Mediated Apoptosis of MDBK Cells. *Journal of Virology*, *76*(19), 9588–9599.
- Jung, H., Gkogkas, C. G., Sonenberg, N., & Holt, C. E. (2014). Remote Control of Gene Function by Local Translation. *Cell*, *157*(1), 26–40.
- K**afasla, P., Morgner, N., Pöyry, T. A. A., Curry, S., Robinson, C. V., & Jackson, R. J. (2009). Polypyrimidine tract binding protein stabilizes the encephalomyocarditis virus IRES structure via binding multiple sites in a unique orientation. *Molecular Cell*, *34*(5), 556–68.
- Kafasla P, Morgner N, Robinson CV, Jackson RJ. (2010). Polypyrimidine tract-binding protein stimulates the poliovirus IRES by modulating eIF4G binding. *EMBO Journal*, *29*, 3710–3722.
- Kanamori, Y., & Nakashima, N. (2001). A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. *RNA*, *7*(2), 266–74.
- Kerr, C. H., & Jan, E. (2016). Commandeering the Ribosome: Lessons Learned from Dicistroviruses about Translation. *Journal of Virology*, *90*(12), 5538–5540.
- Kerr, C. H., Wang, Q. S., Keatings, K., Khong, A., Allan, D., Yip, C. K., ... Jan, E. (2015). The 5' Untranslated Region of a Novel Infectious Molecular Clone of the Dicistrovirus Cricket Paralysis Virus Modulates Infection. *Journal of Virology*, *89*(11), 5919–5934.
- Khong, A., Bonderoff, J. M., Spriggs, R. V., Tammpere, E., Kerr, C. H., Jackson, T. J., ...

- Jan, E. (2016). Temporal regulation of distinct internal ribosome entry sites of the dicistroviridae cricket paralysis virus. *Viruses*, 8, 25.
- Kieft, J. S. (2008). Viral IRES RNA structures and ribosome interactions. *Trends in Biochemical Sciences*, 33(6), 274–283.
- Kieft, J. S., Zhou, K., Grech, A., Jubin, R., & Doudna, J. A. (2002). Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. *Nature Structural Biology*. 9(5), 370-4.
- Kieft, J. S., Zhou, K., Jubin, R., Murray, M. G., Lau, J. Y., & Doudna, J. A. (1999). The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *Journal of Molecular Biology*, 292(3), 513–29.
- Koh, C. S., Brilot, A. F., Grigorieff, N., & Korostelev, A. A. (2014). Taura syndrome virus IRES initiates translation by binding its tRNA-mRNA-like structural element in the ribosomal decoding center. *PNAS*, 111(25), 9139–9144.
- Kolupaeva, V. G., Pestova, T. V, Hellen, C. U., & Shatsky, I. N. (1998). Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. *The Journal of Biological Chemistry*, 273(29), 18599–604.
- Komar, A. A., & Hatzoglou, M. (2005). Internal Ribosome Entry Sites in Cellular mRNAs: Mystery of Their Existence. *Journal of Biological Chemistry*, 280(25), 23425–23428.
- Komar, A. A., & Hatzoglou, M. (2011). Cellular IRES-mediated translation : the war of ITAFs in pathophysiological states. *Cell Cycle*, 10(2), 229–240.
- Kondrashov, N., Pusic, A., Stumpf, C. R., Shimizu, K., Hsieh, A. C., Ishijima, J., ... Barna, M. (2011). Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell*, 145(3), 383–397.
- Kozak, M. (1978). How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell*, 15(4), 1109–1123.
- Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 44(2), 283–92.
- Kumari, S., Bugaut, A., & Balasubramanian, S. (2008). Position and Stability Are Determining Factors for Translation Repression by an RNA G-Quadruplex-Forming Sequence within the 5' UTR of the NRAS Proto-oncogene. *Biochemistry*, 47(48), 12664–12669.
- Lacerda, R., Menezes, J., & Romão, L. (2017). More than just scanning: the importance of cap-independent mRNA translation initiation for cellular stress response and cancer. *Cellular and Molecular Life Sciences*, 74(9), 1659–1680.
- Laing, C., & Schlick, T. (2009). Analysis of four-way junctions in RNA structures. *Journal of*

- Molecular Biology*, 390(3), 547–59.
- Lamphear, B. J., Kirchweger, R., Skern, T., & Rhoads, R. E. (1995). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *The Journal of Biological Chemistry*, 270(37), 21975–83.
- Lamphear, B. J., Yan, R., Yang, F., Waters, D., Liebig, H. D., Klump, H., ... Rhoads, R. E. (1993). Mapping the cleavage site in protein synthesis initiation factor eIF-4 gamma of the 2A proteases from human Coxsackievirus and rhinovirus. *The Journal of Biological Chemistry*, 268(26), 19200–3.
- Lancaster, A. M., Jan, E. & Sarnow, P. (2006). Initiation factor-independent translation mediated by the hepatitis C virus internal ribosome entry site. *RNA*, 12(5), 894–902.
- Le Gall, O., Christian, P., Fauquet, C. M., King, A. M. Q., Knowles, N. J., Nakashima, N., ... Gorbalenya, A. E. (2008). Picornavirales, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. *Archives of Virology*, 153(4), 715–727.
- Lee, A. S. Y., Kranzusch, P. J., & Cate, J. H. D. (2015). eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature*, 522(7554), 111-4.
- Lee, A. S. Y., Kranzusch, P. J., Doudna, J. A., & Cate, J. H. D. (2016). eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature*, 536(7614), 96–99.
- Lee, J. H., Pestova, T. V., Shin, B.-S., Cao, C., Choi, S. K., & Dever, T. E. (2002). Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. *PNAS*, 99(26), 16689–16694.
- Lee, K. M., Chen, C. J., & Shih, S. R. (2017). Regulation Mechanisms of Viral IRES-Driven Translation. *Trends in Microbiology*. 25(7), 546-561.
- Leppek, K., Das, R., & Barna, M. (2017). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nature Reviews Molecular Cell Biology*, 19(3), 158–174.
- Lewis, S. M., & Holcik, M. (2008). For IRES trans-acting factors, it is all about location. *Oncogene*, 27(8), 1033–1035.
- Locker, N., Easton, L. E., & Lukavsky, P. J. (2007). HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO Journal*, 26(3), 795–805.
- López de Quinto, S., Lafuente, E., & Martínez-Salas, E. (2001). IRES interaction with translation initiation factors: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII. *RNA*, 7(9), 1213–26.
- López de Quinto, S., & Martínez-Salas, E. (1997). Conserved structural motifs located in distal loops of aphthovirus internal ribosome entry site domain 3 are required for internal

- initiation of translation. *Journal of Virology*, 71(5), 4171–5.
- López de Quinto, S., & Martínez-Salas, E. (1999). Involvement of the aphthovirus RNA region located between the two functional AUGs in start codon selection. *Virology*, 255(2), 324–36.
- López de Quinto, S., & Martínez-Salas, E. (2000). Interaction of the eIF4G initiation factor with the aphthovirus IRES is essential for internal translation initiation in vivo. *RNA*, 6(10), 1380–92.
- Lorsch, J. R. (1999). Kinetic dissection of fundamental processes of eukaryotic translation initiation in vitro. *EMBO Journal*, 18(23), 6705–6717.
- Lozano, G., & Martínez-Salas, E. (2015). Structural insights into viral IRES-dependent translation mechanisms. *Current Opinion in Virology*, 12, 113–120.
- Lukavsky, P. J. (2009). Structure and function of HCV IRES domains. *Virus Research*, 139(2), 166–171.
- Lukavsky, P. J., Kim, I., Otto, G. A., & Puglisi, J. D. (2003). Structure of HCV IRES domain II determined by NMR. *Nature Structural Biology*, 10(12), 1033–8.
- Luz, N., & Beck, E. (1991). Interaction of a cellular 57-kilodalton protein with the internal translation initiation site of foot-and-mouth disease virus. *Journal of Virology*, 65(12), 6486–94.
- M**ailliot, J., & Martin, F. (2018). Viral internal ribosomal entry sites: four classes for one goal. *Wiley Interdisciplinary Reviews: RNA*, 9(2). <http://doi.org/10.1002/wrna.1458>
- Majumdar, R., Bandyopadhyay, A., & Maitra, U. (2003). Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. *The Journal of Biological Chemistry*, 278(8), 6580–7.
- Majzoub, K., Hafirassou, M. L., Meignin, C., Goto, A., Marzi, S., Fedorova, A., ... Imler, J.-L. (2014). RACK1 Controls IRES-Mediated Translation of Viruses. *Cell*, 159(5), 1086–1095.
- Malnou CE, Poyry TAA, Jackson RJ, Kean KM. (2002). Poliovirus internal ribosome entry segment structure alterations that specifically affect function in neuronal cells: molecular genetic analysis. *J Virol*, 76, 10617–10626.
- Malygin, A. A., Kossinova, O. A., Shatsky, I. N., & Karpova, G. G. (2013). HCV IRES interacts with the 18S rRNA to activate the 40S ribosome for subsequent steps of translation initiation. *Nucleic Acids Research*, 41(18), 8706–8714.
- Martin, F., Ménétret, J.-F., Simonetti, A., Myasnikov, A. G., Vicens, Q., Prongidi-Fix, L., ... Eriani, G. (2016). Ribosomal 18S rRNA base pairs with mRNA during eukaryotic translation initiation. *Nature Communications*, 7, 12622.

- Martin, F., Barends, S., Jaeger, S., Schaeffer, L., Prongidi-Fix, L., & Eriani, G. (2011). Cap-Assisted Internal Initiation of Translation of Histone H4. *Molecular Cell*, *41*(2), 197–209.
- Martin, F., Schaller, A., Eglite, S., Schümperli, D., & Müller, B. (1997). The gene for histone RNA hairpin binding protein is located on human chromosome 4 and encodes a novel type of RNA binding protein. *EMBO Journal*, *16*(4), 769–778.
- Martínez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J., Lozano, G., & Diaz-Toledano, R. (2015). Picornavirus IRES elements: RNA structure and host protein interactions. *Virus Research*, *206*, 62–73.
- Masaki, T., Arend, K. C., Li, Y., Yamane, D., McGivern, D. R., Kato, T., ... Lemon, S. M. (2015). miR-122 Stimulates Hepatitis C Virus RNA Synthesis by Altering the Balance of Viral RNAs Engaged in Replication versus Translation. *Cell Host & Microbe*, *17*(2), 217–228.
- Matsuda, D., & Mauro, V. P. (2014). Base pairing between hepatitis C virus RNA and 18S rRNA is required for IRES-dependent translation initiation in vivo. *PNAS*, *111*(43), 15385–15389.
- McCormack, J. C., Yuan, X., Yingling, Y. G., Kasprzak, W., Zamora, R. E., Shapiro, B. A., & Simon, A. E. (2008). Structural Domains within the 3' Untranslated Region of Turnip Crinkle Virus. *Journal of Virology*, *82*(17), 8706–8720.
- McInerney, G. M., Kedersha, N. L., Kaufman, R. J., Anderson, P., & Liljeström, P. (2005). Importance of eIF2 α Phosphorylation and Stress Granule Assembly in Alphavirus Translation Regulation. *Molecular Biology of the Cell*, *16*(8), 3753–3763.
- Meerovitch, K., Pelletier, J., & Sonenberg, N. (1989). A cellular protein that binds to the 5'-noncoding region of poliovirus: implications for internal initiation. *Genes & Development*, *3*, 1026–1034.
- Meerovitch, K., Svitkin, Y. V., Lee, H. S., Lejbkowitz, F., Kenan, D. J., Chan, E. K., ... Sonenberg, N. (1993). La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *Journal of Virology*, *67*(7), 3798–3807.
- Meier, V. S., Böhni, R., & Schümperli, D. (1989). Nucleotide sequence of two mouse histone H4 genes. *Nucleic Acids Research*, *17*(2), 795.
- Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., ... Jaffrey, S. R. (2015). 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell*, *163*(4), 999–1010.
- Miller, A. and Simon A. (2011). 3' Cap-Independent Translation Enhancers of Plant Viruses. *Current Opinion in Virology*, *1*(5), 373–380.
- Miras, M., Miller, A., Truniger, V., & Aranda, M. A. (2017). Non-canonical Translation in Plant RNA Viruses. *Frontiers in Plant Science*, *8*, article 494.
- Mokrejš, M., Mašek, T., Vopálenský, V., Hlubuček, P., Delbos, P., & Pospíšek, M. (2010).

- IRESite—a tool for the examination of viral and cellular internal ribosome entry sites. *Nucleic Acids Research*, 38, 131–136.
- Mortimer, S. A., & Weeks, K. M. (2007). A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *Journal of the American Chemical Society*, 129(14), 4144–4145.
- Mortimer, S. a, & Weeks, K. M. (2009). Time-resolved RNA SHAPE chemistry: quantitative RNA structure analysis in one-second snapshots and at single-nucleotide resolution. *Nature Protocols*, 4(10), 1413–1421.
- Mosenkis, J., Daniels-McQueen, S., Janovec, S., Duncan, R., Hershey, J. W., Grifo, J. A., ... Thach, R. E. (1985). Shutoff of host translation by encephalomyocarditis virus infection does not involve cleavage of the eucaryotic initiation factor 4F polypeptide that accompanies poliovirus infection. *Journal of Virology*, 54(2), 643–5.
- Mullapudi, E., Přidal, A., Pálková, L., de Miranda, J. R., & Plevka, P. (2016). Virion Structure of Israeli Acute Bee Paralysis Virus. *Journal of Virology*, 90(18), 8150–8159.
- Mulvey, M., Arias, C., & Mohr, I. (2007). Maintenance of Endoplasmic Reticulum (ER) Homeostasis in Herpes Simplex Virus Type 1-Infected Cells through the Association of a Viral Glycoprotein with PERK, a Cellular ER Stress Sensor. *Journal of Virology*, 81(7), 3377–3390.
- 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. *Journal of Virology*, 77(20), 10917–28.
- N**icholson, B. L., & White, K. A. (2011). 3' Cap-independent translation enhancers of positive-strand RNA plant viruses. *Current Opinion in Virology*, 1(5), 373–80.
- Nicholson, B. L., Wu, B., Chevtchenko, I., & White, K. A. (2010). Tombusvirus recruitment of host translational machinery via the 3' UTR. *RNA*, 16(7), 1402–1419.
- Niepmann, M. (2013). Hepatitis C Virus RNA Translation. In: Bartenschlager R. (eds) *Hepatitis C Virus: From Molecular Virology to Antiviral Therapy*. Current Topics in Microbiology and Immunology, vol 369. Springer, Berlin, Heidelberg
- Nishiyama, T., Yamamoto, H., Shibuya, N., Hatakeyama, Y., Hachimori, A., Uchiumi, T., & Nakashima, N. (2003). Structural elements in the internal ribosome entry site of Plautia stali intestine virus responsible for binding with ribosomes. *Nucleic Acids Research*, 31(9), 2434–42.
- Norvell, A., Kelley, R. L., Wehr, K., & Schüpbach, T. (1999). Specific isoforms of Squid, a Drosophila hnRNP, perform distinct roles in Gurken localization during oogenesis. *Genes and Development*, 13(7), 864–876.

- O**h, Y. L., Hahm, B., Kim, Y. K., Lee, H. K., Lee, J. W., Song, O., ... Jang, S. K. (1998). Determination of functional domains in polypyrimidine-tract-binding protein. *The Biochemical Journal*, 331 (1), 169–75.
- Otto, G. A., & Puglisi, J. D. (2004). The Pathway of HCV IRES-Mediated Translation Initiation. *Cell*, 119(3), 369–380.
- P**aul, A. V., van Boom, J. H., Filippov, D., & Wimmer, E. (1998). Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature*, 393(6682), 280–284.
- Pelletier, J., & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, 334(6180), 320–325.
- Pérard, J., Leyrat, C., Baudin, F., Drouet, E., & Jamin, M. (2013). Structure of the full-length HCV IRES in solution. *Nature Communications*, 4, 1612.
- Perera R, Daijogo S, Walter BL, Nguyen JHC, Semler BL. (2007). Cellular protein modification by poliovirus: the two faces of poly(rC)-binding protein. *J Virol*, 81, 8919–8932.
- Pestova, T. V., & Hellen, C. U. T. (2003). Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes & Development*, 17(2), 181–186.
- Pestova, T. V., & Kolupaeva, V. G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes & Development*, 16(22), 2906–22.
- Pestova, T. V., Lomakin, I. B., & Hellen, C. U. T. (2004). Position of the CrPV IRES on the 40S subunit and factor dependence of IRES/80S ribosome assembly. *EMBO Reports*, 5(9), 906–913.
- Pfingsten, J. S., Costantino, D. A., & Kieft, J. S. (2006). Structural Basis for Ribosome Recruitment and Manipulation by a Viral IRES RNA. *Science*, 314(5804), 1450–1454.
- Pfingsten, J. S., Costantino, D. A., & Kieft, J. S. (2007). Conservation and Diversity Among the Three-dimensional Folds of the Dicrostoviridae Intergenic Region IRESes. *Journal of Molecular Biology*, 370(5), 856–869.
- Phelan, M. (2004). NMR studies of the structure and Mg²⁺ binding properties of a conserved RNA motif of EMCV picornavirus IRES element. *Nucleic Acids Research*, 32(16), 4715–4724.
- Pilipenko EV, Gmyl AP, Maslova SV, Svitkin YV, Sinyakov AN, Agol VI. (1992). Prokaryotic-like cis elements in the cap-independent internal initiation of translation on picornavirus

RNA. *Cell*, 68(88),119–131.

Prongidi-Fix, L., Schaeffer, L., Simonetti, A., Barends, S., Ménétret, J.-F., Klaholz, B. P., ... Martin, F. (2013). Rapid purification of ribosomal particles assembled on histone H4 mRNA: a new method based on mRNA–DNA chimaeras. *Biochemical Journal*, 449(3), 719–728.

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. *Nature Communications*, 6, 7646.

Roberts, A. P. E., Lewis, A. P., & Jopling, C. L. (2011). miR-122 activates hepatitis C virus translation by a specialized mechanism requiring particular RNA components. *Nucleic Acids Research*, 39(17), 7716–7729.

Robertson, M. E., Seamons, R. A., & Belsham, G. J. (1999). A selection system for functional internal ribosome entry site (IRES) elements: analysis of the requirement for a conserved GNRA tetraloop in the encephalomyocarditis virus IRES. *RNA*, 5(9), 1167–79.

Robichaud, N., & Sonenberg, N. (2017). Translational control and the cancer cell response to stress. *Current Opinion in Cell Biology*, 45, 102–109.

Rodnina, M. V., & Wintermeyer, W. (2009). Recent mechanistic insights into eukaryotic ribosomes. *Current Opinion in Cell Biology*, 21(3), 435–43.

Roux, P. P., & Topisirovic, I. (2018). Signaling Pathways Involved in the Regulation of mRNA Translation. *Molecular and Cellular Biology*, 38(12)

Rozen, F., Edery, I., Meerovitch, K., Dever, T. E., Merrick, W. C., & Sonenberg, N. (1990). Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Molecular and Cellular Biology*, 10(3), 1134–44.

Ruan, Y., Sun, L., Hao, Y., Wang, L., Xu, J., Zhang, W., ... Gu, J. (2012). Ribosomal RACK1 promotes chemoresistance and growth in human hepatocellular carcinoma. *Journal of Clinical Investigation*, 122(7), 2554–2566.

Ruehle, M. D., Zhang, H., Sheridan, R. M., Mitra, S., Chen, Y., Gonzalez, R. L., ... Kieft, J. S. (2015). A dynamic RNA loop in an IRES affects multiple steps of elongation factor-mediated translation initiation. *ELife*, 4. eLife.08146

Ryman, K. D., Meier, K. C., Nangle, E. M., Ragsdale, S. L., Korneeva, N. L., Rhoads, R. E., ... Klimstra, W. B. (2005). Sindbis Virus Translation Is Inhibited by a PKR/RNase L-Independent Effector Induced by Alpha/Beta Interferon Priming of Dendritic Cells.

Journal of Virology, 79(3), 1487–1499.

- S**agan, S. M., Chahal, J., & Sarnow, P. (2015). cis-Acting RNA elements in the hepatitis C virus RNA genome. *Virus Research*, 206, 90–8.
- Salzman, J. (2016). Circular RNA Expression: Its Potential Regulation and Function. *Trends in Genetics*, 32(5), 309–316.
- Sánchez-Eugenia, R., Goikolea, J., Gil-Cartón, D., Sánchez-Magraner, L., & Guérin, D. M. A. (2015). Triatoma Virus Recombinant VP4 Protein Induces Membrane Permeability through Dynamic Pores. *Journal of Virology*, 89(8), 4645–4654.
- Sawicka, K., Bushell, M., Spriggs, K. A., & Willis, A. E. (2008). Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein. *Biochemical Society Transactions*, 36(4), 641–647.
- Schneider-Poetsch, T., Ju, J., Eyler, D. E., Dang, Y., Bhat, S., Merrick, W. C., ... Liu, J. O. (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nature Chemical Biology*, 6(3), 209–217.
- Schüler, M., Connell, S. R., Lescoute, A., Giesebrecht, J., Dabrowski, M., Schroeer, B., ... Spahn, C. M. T. (2006). Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nature Structural & Molecular Biology*, 13(12), 1092–1096.
- Scotti, P. D. (1975). Cricket Paralysis Virus Replicates in Cultured *Drosophila* Cells. *Intervirology*, 6(6), 333–342.
- Scotti, P. D., Hoefakker, P., & Dearing, S. (1996). The Production of Cricket Paralysis Virus in Suspension Cultures of Insect Cell Lines. *Journal of Invertebrate Pathology*, 68(2), 109–112.
- Serrano, P., Gomez, J., & Martinez-Salas, E. (2007). Characterization of a cyanobacterial RNase P ribozyme recognition motif in the IRES of foot-and-mouth disease virus reveals a unique structural element. *RNA*, 13(6), 849–859.
- Shatsky, I. N., Dmitriev, S. E., Terenin, I. M., & Andreev, D. E. (2010). Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs. *Molecules and Cells*, 30(4), 285–293.
- Simon, A. E. (2015). 3'UTRs of carmoviruses. *Virus Research*, 206, 27–36.
- Simpson, R. J. (2010). Disruption of cultured cells by nitrogen cavitation. *Cold Spring Harbor Protocols*, 5(11), 1219–1222.
- Sonenberg, N., & Hinnebusch, A. G. (2009, February 20). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell*. 136(4), 731-45
- Sonenberg, N., Rupprecht, K. M., Hecht, S. M., & Shatkin, A. J. (1979). Eukaryotic mRNA cap binding protein: purification by affinity chromatography on sepharose-coupled

- m7GDP. *PNAS*, *76*(9), 4345–9.
- Soulard, A., & Hall, M. N. (2007). SnapShot: mTOR Signaling. *Cell*, *129*(2), 434.e1-434.e2. <http://doi.org/10.1016/j.cell.2007.04.010>
- Spriggs, K. A., Stoneley, M., Bushell, M., & Willis, A. E. (2008). Re-programming of translation following cell stress allows IRES-mediated translation to predominate. *Biology of the Cell*, *100*(1), 27–38.
- Stupina, V. A., Meskauskas, A., McCormack, J. C., Yingling, Y. G., Shapiro, B. A., Dinman, J. D., & Simon, A. E. (2008). The 3' proximal translational enhancer of Turnip crinkle virus binds to 60S ribosomal subunits. *RNA*, *14*(11), 2379–93.
- Sun, C., Querol-Audí, J., Mortimer, S. A., Arias-Palomo, E., Doudna, J. A., Nogales, E., & Cate, J. H. D. (2013). Two RNA-binding motifs in eIF3 direct HCV IRES-dependent translation. *Nucleic Acids Research*, *41*(15), 7512–21.
- Sun, C., Todorovic, A., Querol-Audí, J., Bai, Y., Villa, N., Snyder, M., ... Cate, J. H. D. (2011). Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). *PNAS*, *108*(51), 20473–8.
- Svitkin, Y. V., Evdokimova, V. M., Brasey, A., Pestova, T. V., Fantus, D., Yanagiya, A., ... Sonenberg, N. (2009). General RNA-binding proteins have a function in poly(A)-binding protein-dependent translation. *EMBO Journal*, *28*(1), 58–68.
- Sweeney TR, Abaeva IS, Pestova TV, Hellen CUT. (2014). The mechanism of translation initiation on type 1 picornavirus IRESs. *EMBO Journal*, *33*, 76–92.
- T**opisirovic, I., & Sonenberg, N. (2011). Translational control by the eukaryotic ribosome. *Cell*, *145*(3), 333–4.
- Toyoda H, Franco D, Fujita K, Paul AV, Wimmer E. (2007). Replication of poliovirus requires binding of the Poly(rC) binding protein to the cloverleaf as well as to the adjacent C-rich spacer sequence between the cloverleaf and the internal ribosomal entry site. *Journal of Virology*, *81*, 10017–10028.
- Trachsel, H., Erni, B., Schreier, M. H., & Staehelin, T. (1977). Initiation of mammalian protein synthesis. II. The assembly of the initiation complex with purified initiation factors. *Journal of Molecular Biology*, *116*(4), 755–67.
- Trachsel, H., & Staehelin, T. (1978). Binding and release of eukaryotic initiation factor eIF-2 and GTP during protein synthesis initiation. *PNAS*, *75*(1), 204–8.
- Trachsel, H., & Staehelin, T. (1979). Initiation of mammalian protein synthesis. The multiple functions of the initiation factor eIF-3. *Biochimica et Biophysica Acta*, *565*(2), 305–14.
- Trono, D., Pelletier, J., Sonenberg, N., & Baltimore, D. (1988). Translation in mammalian cells of a gene linked to the poliovirus 5' noncoding region. *Science*, *241*(4864), 445–8.

- V**alášek, L., Hasek, J., Trachsel, H., Imre, E. M., & Ruis, H. (1999). The *Saccharomyces cerevisiae* HCR1 gene encoding a homologue of the p35 subunit of human translation initiation factor 3 (eIF3) is a high copy suppressor of a temperature-sensitive mutation in the Rpg1p subunit of yeast eIF3. *The Journal of Biological Chemistry*, *274*(39), 27567–72.
- Valášek, L. S. (2012). 'Ribozoomin'--translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIF). *Current Protein & Peptide Science*, *13*(4), 305–30.
- Valášek, L. S., Zeman, J., Wagner, S., Beznosková, P., Pavlíková, Z., Mohammad, M. P., ... Gunišová, S. (2017). Survey and summary: Embraced by eIF3: Structural and functional insights into the roles of eIF3 across the translation cycle. *Nucleic Acids Research*, *45*(19), 10948–10968.
- Valles, S. M., Chen, Y., Firth, A. E., Guérin, D. M. A., Hashimoto, Y., Herrero, S., ... Ryabov, E. (2017). ICTV Virus Taxonomy Profile: Dicistroviridae. *Journal of General Virology*, *98*(3), 355–356.
- Van Rij, R. P., Saleh, M.-C., Berry, B., Foo, C., Houk, A., Antoniewski, C., & Andino, R. (2006). The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in *Drosophila melanogaster*. *Genes & Development*, *20*(21), 2985–2995.
- Ventoso, I., Chisari, F. V. (2006). Translational resistance of late alphavirus mRNA to eIF2 phosphorylation: a strategy to overcome the antiviral effect of protein kinase PKR. *Genes & Development*, *20*(1), 87–100.
- W**alsh, D., & Mohr, I. (2011). Viral subversion of the host protein synthesis machinery. *Nature Reviews Microbiology*, *9*(12), 860-75.
- Wang, X.-H., Aliyari, R., Li, W.-X., Li, H.-W., Kim, K., Carthew, R., ... Ding, S.-W. (2006). RNA Interference Directs Innate Immunity Against Viruses in Adult *Drosophila*. *Science*, *312*(5772), 452–454.
- Wang, Y., & Wang, Z. (2015). Efficient backsplicing produces translatable circular mRNAs. *RNA*, *21*(2), 172–179.
- Wang, Z., Treder, K., & Miller, W. A. (2009). Structure of a viral cap-independent translation element that functions via high affinity binding to the eIF4E subunit of eIF4F. *The Journal of Biological Chemistry*, *284*(21), 14189–202.
- Wilson, J. E., Pestova, T. V., Hellen, C. U., & Sarnow, P. (2000). Initiation of protein synthesis from the A site of the ribosome. *Cell*, *102*(4), 511–520.

- Wilson, J. E., Powell, M. J., & Hoover, S. E. (2000). Naturally Occurring Dicistronic Cricket Paralysis Virus RNA Is Regulated by Two Internal Ribosome Entry Sites. *Molecular and Cellular Biology*, *20*(14), 4990–4999.
- X**ue, S., Tian, S., Fujii, K., Kladwang, W., Das, R., & Barna, M. (2014). RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation. *Nature*, *517*(7532), 33-8.
- Y**amamoto, H., Collier, M., Loerke, J., Ismer, J., Schmidt, A., Hilal, T., ... Spahn, C. M. (2015). Molecular architecture of the ribosome-bound Hepatitis C Virus internal ribosomal entry site RNA. *EMBO Journal*, *34*(24), 3042–3058.
- Yamamoto, H., Nakashima, N., Ikeda, Y., & Uchiumi, T. (2007). Binding Mode of the First Aminoacyl-tRNA in Translation Initiation Mediated by *Plautia stali* Intestine Virus Internal Ribosome Entry Site. *Journal of Biological Chemistry*, *282*(11), 7770–7776.
- Yaman, I., Fernandez, J., Liu, H., Caprara, M., Komar, A. A., Koromilas, A. E., ... Hatzoglou, M. (2003). The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. *Cell*, *113*(4), 519–31.
- Z**hao, B. S., Roundtree, I. A., & He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nature Reviews Molecular Cell Biology*, *18*(1), 31–42.
- Zhou, J., Rode, K. A., & Qian, S.-B. (2016). m(6)A: A novel hallmark of translation. *Cell Cycle*, *15*(3), 309–10.
- Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S. R., & Qian, S.-B. (2015). Dynamic m6A mRNA methylation directs translational control of heat shock response. *Nature*, *526*(7574), 591–594.

Le piratage de l'appareil de traduction de la cellule hôte par le virus de la paralysie du criquet (CrPV)

Résumé en français :

Le Virus de la Paralysie du Criquet (CrPV) est un virus à ARN qui utilise des « sites d'entrée interne du ribosome » (IRES) pour détourner la machinerie traductionnelle de son hôte et synthétiser ses protéines virales. La traduction de ses deux phases codantes (ORF) est contrôlée par deux IRES différents : l'IRES de la région intergénique (IGR), qui régit la traduction de l'ORF2 et qui a été largement étudié ; et l'IRES de la région 5'UTR du génome viral (IRES_{5'UTR}), qui guide la traduction de l'ORF1 et qui était très mal connu. Nous avons démontré que cet IRES_{5'UTR} appartient à la classe III des IRES. Il recrute spécifiquement et directement le facteur d'initiation eIF3. Son mode d'action s'affranchit de l'étape de 'scanning' pour reconnaître le codon AUG. L'IRES_{5'UTR} comprend deux domaines structuraux qui ont des fonctions distinctes mais qui agissent en synergie pour recruter eIF3 et le ribosome et permettre de guider la progression de l'initiation de la traduction de l'ORF1. En outre, les deux IRES du CrPV utilisent des modes d'action différents, ils sont actifs à des moments différents du processus infectieux. La reconstitution *in vitro* des complexes d'initiation de la traduction formés sur les deux IRES en cours d'infection a permis d'identifier les protéines spécifiques de chaque IRES lors du cycle viral. À présent nous allons valider l'implication de ces protéines dans la traduction guidée par ces IRES. Cette étude aidera à la compréhension des mécanismes moléculaires d'action des traductions « IRES-dépendantes » des ARN viraux et cellulaires.

Mot clefs : CrPV, eIF, initiation de la traduction, IRES viraux, ribosome

Résumé en anglais:

The Cricket Paralysis Virus (CrPV) is an RNA virus that uses 'internal ribosome entry sites' (IRES) to hijack the host translation machinery and synthesize its own viral proteins. The translation of its two open reading frames (ORF) is controlled by two different IRES : First, the IRES located in the Intergenic Region (IGR), which governs translation of ORF2 and has been extensively studied. Second, the IRES located the 5'UTR region from viral genome (IRES_{5'UTR}), which guides the translation of ORF1 and which was very poorly known. During my PhD, we demonstrated that this IRES_{5'UTR} belongs to the class III of the IRES. It specifically and directly recruits the eIF3 initiation factor. Its mode of action is free from the 'scanning' step to recognize the AUG codon. The IRES_{5'UTR} contains two structural domains that have distinct functions, but act in synergy to recruit eIF3 and the ribosome in order to guide ORF1 translation initiation. In addition, both CrPV IRES use different modes of action, they are active at different times in the infectious process. The *in vitro* reconstitution of translation initiation complexes formed on both IRES during infection allowed us to identify proteins that are specific of each IRES during the viral cycle. Now we will validate the implication of these proteins in the translation guided by both IRES. Our study will help to understand the molecular basis of "IRES-dependent" translations in viral and cellular RNA.

Key words : CrPV, eIF, translation initiation, viral IRES, ribosome