



**HAL**  
open science

# Les mécanismes d'initiation de la traduction de la polyprotéine Gag du Virus de l'Immunodéficience Humaine (VIH-1)

Melissa Ameur

► **To cite this version:**

Melissa Ameur. Les mécanismes d'initiation de la traduction de la polyprotéine Gag du Virus de l'Immunodéficience Humaine (VIH-1). Biologie moléculaire. Université Sorbonne Paris Cité, 2016. Français. NNT : 2016USPCB125 . tel-01844836

**HAL Id: tel-01844836**

**<https://theses.hal.science/tel-01844836>**

Submitted on 19 Jul 2018

**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.

Laboratoire de cristallographie et RMN biologiques (UMR8015)

# THESE DE DOCTORAT DE L'UNIVERSITE PARIS DESCARTES

Spécialité : Biochimie et Biologie moléculaire

Présentée par : Melissa AMEUR

Pour obtenir le titre de :

DOCTEUR DE L'UNIVERSITE PARIS DESCARTES

Sujet de thèse :

Les mécanismes d'initiation de la traduction de la  
polyprotéine Gag du Virus de l'Immunodéficience Humaine  
(VIH-1)

Directeurs de thèse :  
Bruno SARGUEIL/Nathalie CHAMOND

Soutenue le 4 novembre 2016

Devant le jury composé de :

Nicolas	LOCKER	: Rapporteur
Marc	BOUVILLAIN	: Rapporteur
Josette	BANROQUES	: Examineur
Théophile	OHLMANN	: Examineur
Bruno	SARGUEIL	: Directeur de thèse
Nathalie	CHAMOND	: Co-directrice de thèse

*À la mémoire des personnes qui me sont chères et qui sont parties*

*trop tôt, je pense particulièrement à toi mon oncle chéri, à tonton Mahrez et à mes amis qui ont perdu la vie dans les événements tragiques d'il y a un an,*

*Puissiez-vous tous reposer en paix.*

# REMERCIEMENTS

J'adresse mes remerciements les plus sincères à, monsieur Marc BOUDVILLAIN et monsieur Nicolas LOCKER qui ont accepté d'évaluer ce travail en tant que rapporteurs. Je remercie également madame Josette BANROQUES et monsieur Théophile OHLMANN d'avoir accepté de faire partie de mon jury de thèse en tant qu'examineurs. Monsieur OHLMANN, merci également de m'avoir permis de passer une semaine très enrichissante au sein de votre équipe à Lyon.

J'adresse ensuite mes remerciements à monsieur Nicolas LEULLIOT pour m'avoir accueillie au sein du laboratoire de Cristallographie et RMN biologiques.

Je tiens à dire un énorme merci à mes deux directeurs de thèse Bruno SARGUEIL et Nathalie CHAMOND.

Bruno, merci de m'avoir donné ma chance et de m'avoir accueillie en stage de master 2 puis de m'avoir soutenue et encouragée depuis le concours de thèse (je me rappelle encore de la chanson d'Alister) jusqu'à aujourd'hui. Auprès de toi j'ai énormément appris, notamment sur le monde des ARN et de la traduction. Mais pas seulement, j'ai aussi appris pas mal d'expressions et je promets de ne pas «boutonner lundi avec mardi » ! Je me souviendrai également de ne pas manger trop de bonbons avant un effort physique (quoique ce ne soit pas gagné) et surtout je me rappellerai de ne jamais faire confiance à la météo !

Nathalie, je ne sais comment te dire merci, pour ton soutien depuis le début, tes encouragements lorsque (souvent) je n'avais pas le moral, ta présence dans les épreuves difficiles que j'ai traversées et pour ton aide précieuse dans la correction de ce manuscrit compliquée par mon état de santé. Merci également pour tout ce que j'ai appris à la paillasse avec toi, tous ces petits détails, certes un peu maniaques mais toujours utiles et efficaces ! Merci pour nos supers manips à 4 mains sur les tests ATPase ! On aura finalement réussi avec DDX3 !

Merci aux anciens et actuels membres de l'équipe « mécanismes moléculaires d'initiation de la traduction », Jules, Jenniffer, Felipe, Reza, Jessica, Noémie, Delphine et Grégoire pour leur bonne humeur et les moments passés ensemble dans ce bureau fournaise comme dans « la zone » seuls au monde ! Un spécial grand Merci à Nathalie Ulryck, tu as été ma seconde maman (je pense que tu l'es pour nous tous). Ta présence douce et réconfortante, ta gentillesse, ta disponibilité et ton soutien ont été précieux durant ces 4 années passées avec vous.



Un énorme merci à tous les membres du LCRB pour les bons moments que l'on a passés tous ensemble, je pense notamment aux discussions enrichissantes dans la cuisine. A mes amis des nocturnes et du travail isolé (Gilles, Assia, Karim, Clément.D). Merci à toi Magali pour tes précieux conseils notamment sur les ppt ! Merci à Martin pour ton aide et tes explications sur le SAFAS, tu m'as porté bonheur pour mes premières expériences de fluorescence! Merci à toi Carine pour m'avoir épaulée, merci à Michèle pour ton écoute et tes conseils. Merci à Franck pour ton humour et pour toujours me laisser une place dans votre bureau que j'avoue « squatter » souvent ! Merci surtout à toi Marjorie, bien plus qu'une collègue tu es devenue au cours de ces années une amie précieuse, merci d'avoir toujours trouvé un moment pour m'écouter, me réconforter et me conseiller.

Merci à toi Vivi ainsi qu'à Catherine pour votre bonne humeur !

Merci à Cathy, que ferions-nous sans toi ? Tu es une gestionnaire hors-pair !

Merci à Sylvie Gillet, pour cette expérience très enrichissante d'enseignements, pour nos discussions qu'elles soient scientifiques ou dans tout autre domaine.

Maman, Papa, comment pourrais-je vous remercier, vous êtes les meilleurs parents du monde, vous avez fait tant de sacrifices pour que je sois là où j'en suis aujourd'hui ! Merci de m'avoir inculquée vos valeurs, merci pour votre amour inconditionnel, merci de m'avoir toujours soutenue dans mes choix et de m'avoir encouragée et aidée à aller au bout de mes projets et à réaliser mes rêves. Je ne serai jamais assez reconnaissante pour tout ce que vous avez fait pour moi. Malgré la distance qui nous sépare vous avez su être toujours présents, merci *viber* de nous permettre de parler aussi souvent merci à toi maman de toujours écouter.

Je vous aime fort.

Merci à mes frères de m'avoir toujours chouchoutée et protégée et d'avoir cru en moi ! Surtout toi youyou (aamou ninou) tu es pour moi un tout ! Je vous aime fort.

Un grand Merci à ma grand-mère, c'est si difficile d'être loin de toi, de tes blagues de ta joie de vivre communicative ! Merci pour tes da3wa je pense qu'elles ont fini par aboutir, je t'aime tellement fort et j'espère t'avoir près de moi pour longtemps encore.

Merci à toi mon chouchou, d'être mon roc, d'être l'homme de ma vie et de mes rêves ! Merci d'être celui que tu es, de me soutenir et de me supporter notamment lors de ces 4 dernières années ! Merci d'avoir été patient et présent ! Nous avons traversé tant d'épreuves ensemble et nous en sommes devenus que plus fort ! Merci pour nos moments magiques, pour nos fous-rires et notre complicité ! Nhabek hayati.

Merci à ma belle-famille de toujours m'épauler, merci surtout à toi Nadia, en rejoignant votre famille j'ai gagné une sœur ! Merci d'avoir été présente, de m'avoir



# SOMMAIRE

REMERCIEMENTS .....	3
SOMMAIRE .....	6
TABLE DES MATIERES .....	8
TABLE DES ILLUSTRATIONS.....	12
LISTE DES ABBREVIATIONS .....	14
RESUME .....	16
ABSTRACT .....	18
INTRODUCTION .....	19
I.    Le Virus de l'Immunodéficience Humaine de type 1 (VIH-1) .....	20
II.   L'initiation de la traduction chez les eucaryotes.....	29
III.   Traduction de la polyprotéine Gag du VIH-1 .....	70
IV.   Le rôle de DDX3 dans la traduction.....	88
Objectifs de thèse.....	98
Résultats .....	100
I.    Étude de l'interaction entre l'IRESgag du VIH-1 et la sous-unité ribosomale 40S .....	101
II.   Etude du rôle de DDX3 dans l'initiation de la traduction du VIH-1 .....	167
Discussion et perspectives .....	184
I.    Un nouveau rôle pour l'IRES Gag ?.....	185
II.   Caractérisation biochimique de DDX3 .....	186
III.   Rôle de DDX3 dans l'initiation de la traduction dépendante de la coiffe de l'ARNg du VIH-1 .....	187
Matériels et méthodes .....	190
I. Plasmides .....	191
II.   Les mutants de DDX3 .....	193
III.   Séquence de la phase codante de DDX3-WT.....	193
IV.   Souches bactériennes .....	194
V.    Expression protéique.....	194
VI.   Purification.....	194
VII.  Clivage et purification de la protéine clivée .....	195
VIII.  Test d'activité ATPase .....	195
IX.   Test d'activité hélicase.....	195

X.	PCR .....	196
XI.	Amorces PCR.....	197
XII.	Transcription <i>in vitro</i> .....	198
XIII.	Traduction <i>in vitro</i> .....	199
XIV.	La rétention sur filtre .....	199
XV.	Gradient de densité .....	200
	Annexe.....	201
	Bibliographie.....	204

# TABLE DES MATIERES

REMERCIEMENTS .....	3
SOMMAIRE .....	6
TABLE DES MATIERES .....	8
TABLE DES ILLUSTRATIONS.....	12
LISTE DES ABBREVIATIONS .....	14
RESUME .....	16
ABSTRACT .....	18
INTRODUCTION .....	19
I.    Le Virus de l'Immunodéficience Humaine de type 1 (VIH-1) .....	20
A.    Généralités .....	20
B.    Le VIH : origine et caractéristiques .....	21
Origine .....	21
Structure du virion .....	22
Génome et protéines .....	23
C.    Cycle réplcatif du VIH-1.....	26
II.    L'initiation de la traduction chez les eucaryotes.....	29
A.    Généralités sur la traduction chez les eucaryotes .....	29
Le ribosome .....	29
La structure des ARN messagers eucaryotes nécessaire à la traduction .....	31
Les étapes de la traduction .....	35
B.    Initiation de la traduction dépendant de la coiffe .....	37
Formation du complexe ternaire .....	37
Formation du complexe de préinitiation 43S.....	42
Formation du complexe 48S.....	44
Le balayage et la reconnaissance du codon initiateur .....	48
La régulation de l'initiation de la traduction globale .....	48
C.    Initiation de la traduction par entrée interne des ribosomes .....	52
Découverte des IRES.....	52
Mise en évidence des IRES .....	53

Les IRES virales et leur classification .....	54
Les IRES cellulaires.....	64
Liaison du ribosome à des séquences complémentaires de l'ARNr 18S.....	65
Des ITAFs régulent l'activité des IRES.....	67
Le mécanisme IRES un mécanisme ancestral ? .....	68
III. Traduction de la polyprotéine Gag du VIH-1 .....	70
A. La synthèse des protéines virales.....	70
B. Initiation de la traduction de la polyprotéine Gag du VIH-1 par le mécanisme canonique dépendant de la coiffe.....	71
L'état conformationnel de la 5'UTR impliqué dans la régulation de la traduction virale .....	71
La structure TAR de la 5'UTR du VIH-1 inhibe la traduction coiffe dépendante de la polyprotéine Gag.....	74
Des hélicases pour lever l'inhibition de la traduction coiffe dépendante ?.....	75
C. Initiation de la traduction de la polyprotéine Gag par le mécanisme d'entrée interne des ribosomes .....	76
Un IRES dans la 5'UTR régulé en fonction du cycle cellulaire .....	76
Le recours au cap-binding complexe pour assurer la traduction virale en phase G2/M .....	78
Un IRES dans la région codante du gène Gag.....	78
D. Les mécanismes d'initiation de la traduction de la protéine Gag du VIH-.....	84
E. Régulation de la traduction de la protéine Gag.....	85
Par des protéines virales .....	85
Par des protéines cellulaires .....	86
IV. Le rôle de DDX3 dans la traduction.....	88
A. Généralités sur les hélicases .....	88
B. La famille des hélicases à boîte DEAD.....	88
C. DDX3, protéine aux multiples fonctions .....	90
D. Les propriétés biochimiques de DDX3 .....	91
E. Les fonctions biologiques.....	93
Rôle de DDX3 dans la traduction.....	93
DDX3 et la réplication virale.....	94
Objectifs de thèse.....	98
Résultats .....	100

I.	Étude de l'interaction entre l'IRESgag du VIH-1 et la sous-unité ribosomale 40S .....	101
II.	Etude du rôle de DDX3 dans l'initiation de la traduction du VIH-1 .....	167
A.	Obtention d'une protéine recombinante soluble et active. ....	168
	Production de rDDX3.....	168
	Purification de MBP-DDX3 .....	170
	Clivage de l'étiquette MBP .....	171
	Evaluation de l'activité ATPase des protéines recombinantes .....	174
B.	Caractérisation de l'activité ATPase.....	176
	Détermination des paramètres cinétiques de la protéine MBP-DDX3 .....	177
C.	Caractérisation de l'activité hélicase .....	180
D.	Etude fonctionnelle de DDX3 sur l'initiation de la traduction virale .....	183
	Discussion et perspectives .....	184
I.	Un nouveau rôle pour l'IRES Gag ?.....	185
II.	Caractérisation biochimique de DDX3 .....	186
III.	Rôle de DDX3 dans l'initiation de la traduction dépendante de la coiffe de l'ARNg du VIH-1 .....	187
A.	Caractérisation du mécanisme d'initiation de la traduction virale impliquant DDX3 .....	187
B.	L'interaction DDX3 – ARN du VIH-1 .....	187
C.	Interactions DDX3 avec les facteurs d'initiation de la traduction .....	188
D.	Validation par reconstruction .....	189
E.	Recherche d'ARNm cibles de DDX3 .....	189
	Matériels et méthodes.....	190
I.	Plasmides .....	191
II.	Les mutants de DDX3 .....	193
III.	Séquence de la phase codante de DDX3-WT.....	193
IV.	Souches bactériennes .....	194
V.	Expression protéique.....	194
VI.	Purification.....	194
VII.	Clivage et purification de la protéine clivée .....	195
VIII.	Test d'activité ATPase .....	195
IX.	Test d'activité hélicase.....	195
X.	PCR .....	196
XI.	Amorces PCR.....	197

XII.	Transcription <i>in vitro</i> .....	198
XIII.	Traduction <i>in vitro</i> .....	199
XIV.	La rétention sur filtre .....	199
XV.	Gradient de densité .....	200
Annexe.....		201
Bibliographie.....		204



# TABLE DES ILLUSTRATIONS

Figure 1 : Arbre phylogénétique des lentivirus de primate (d'après Ho & Bieniasz, 2008) .....	22
Figure 2 : Structure d'un virion mature du VIH-1 .....	23
Figure 3: La 5'UTR du VIH-1 et son implication dans le cycle viral.....	24
Figure 4 : Organisation du génome du VIH-1 et expression des gènes.....	26
Figure 5 : Schéma récapitulatif du cycle de réplication du VIH-1 (d'après Engelman & Cherepanov, 2012) .....	28
Figure 6: Structure des sous-unités du ribosome humain (d'après Khatter <i>et al.</i> , 2015) .....	30
Figure 7 : Structure de la coiffe des ARNs messenger eucaryotes (d'après Decroly <i>et al.</i> , 2012).....	32
Figure 8 : Représentation schématique des différentes étapes d'initiation de la traduction coiffe-dépendante .....	37
Figure 9 : Formation du complexe ternaire (adapté de Kimball, 1999). .....	38
Figure 10 : Les sous-unités d'eIF2, l'homologue d'eIF2 (d'après Schmitt <i>et al.</i> , 2010) .....	39
Figure 11 : Structure secondaire et tertiaire d'un ARNt (d'après Goldman, 2008).....	40
Figure 12 : Structure secondaire de l'ARNt <sub>i</sub> <sup>Met</sup> vs ARNt <sub>e</sub> <sup>Met</sup> chez la levure (d'après Kapp <i>et al.</i> , 2006).....	41
Figure 13 : Modèle de reconstitution du complexe de préinitiation dans les conformations ouverte et fermée (d'après Llácer <i>et al.</i> , 2015).....	42
Figure 14 : Modèle de reconstitution d'eIF3 basé sur la cryoEM (d'après des Georges <i>et al.</i> , 2015) .....	43
Figure 15 : Modèle du complexe multi facteur chez la levure (d'après Hinnebusch, 2006).....	44
Figure 16 : Structure d'eIF4E lié à la coiffe (d'après Borden <i>et al.</i> , 2011).....	45
Figure 17: Structure d'eIF4G et les différents facteurs d'initiation liés (d'après Jackson <i>et al.</i> , 2010).....	46
Figure 18 : Recyclage d'eIF2-GDP en eIF2-GTP par eIF2B et sa régulation par les kinases (d'après Klann & Dever, 2004). .....	50
Figure 19 : Régulation du facteur eIF4E par les 4E-BP (d'après Buchkovich <i>et al.</i> , 2008).....	51
Figure 20 : Schéma de l'initiation par entrée interne des ribosomes .....	53
Figure 21 : Mise en évidence expérimentale d'une activité IRES .....	54
Figure 22 : Les différentes classes d'IRES .....	55
Figure 23 : Mécanisme d'initiation de la traduction de l'ARNm cellulaire vs ARN viral lors de l'infection par le poliovirus .....	56
Figure 24 : Modèle du mécanisme d'initiation de la traduction par l'IRES du virus EMCV (d'après Chamond <i>et al.</i> , 2014) .....	57
Figure 25 : : Modèle de la structure secondaire de la 5'UTR du VHC (d'après Filbin <i>et al.</i> , 2013) .....	58
Figure 26 : Structure de l'IRES du VHC en interaction avec la sous-unité ribosomale 40S.....	59
Figure 27 : L'interaction d'eIF3 avec la sous-unité ribosomale 40S seule ou liée à l'IRES du CSFV (d'après Hashem <i>et al.</i> , 2013).....	60
Figure 28 : Structure de l'IRES de la région intergénique de dicistrovirus seule ou en interaction avec la sous-unité ribosomale 40S (d'après Fernández <i>et al.</i> , 2014 & Spahn <i>et al.</i> , 2004 ).....	62
Figure 29 : Schéma du mécanisme d'initiation de la traduction de l'IRES de l'IGR de <i>Discitroviridae</i> (adapté de Koh <i>et al.</i> 2014) .....	63
Figure 30 : Appariement de séquences entre l'IRES IGF1R et l'ARNr 18S (d'après Meng <i>et al.</i> , 2010) .....	67

Figure 31 : Formation du « kissing-complex » (d'après Paillart <i>et al.</i> , 2004) .....	72
Figure 32 : Transition d'un kissing-complex à un extended complex (d'après Paillart <i>et al.</i> , 2004) .....	73
Figure 33 : Les modèles d'appariement longue distance U5-DIS et U5-AUG (d'après Lu <i>et al.</i> , 2011).....	74
Figure 34 : P40 issue de la traduction IRES-dépendante (données non publiées;S.de Breyne and Ohlmann) .....	79
Figure 35 : L'IRES <sub>gag</sub> est conservée parmi les lentivirus (d'après Weill <i>et al.</i> , 2010) .....	80
Figure 36 : Les éléments de structure conservés pour l'IRES <sub>gag</sub> chez les lentivirus de primates.....	81
Figure 37 : Formation d'un complexe ternaire stable IRES <sub>gag</sub> -40S-eIF3 (d'après Locker <i>et al.</i> , 2011) .....	82
Figure 38 : Position des séquences INS sur le génome viral du VIH-1 .....	83
Figure 39 : Les mécanismes d'initiation de la traduction de la protéine Gag du VIH-1 .....	84
Figure 40: Régulation de la traduction de l'ARNg du VIH-1 par des protéines virales et cellulaires(adapté de Rojas-Araya <i>et al.</i> , 2015).....	87
Figure 41: Organisation structurale des protéines à boîte DEAD (D'après Linder & Jankowsky, 2011))....	89
Figure 42: Les régions conservées spécifiques à la sous-famille des hélicases à boîte DEAD, Ded1/DDX3 (d'après Floor <i>et al.</i> , 2016) .....	90
Figure 43: La structure cristallographie du core domaine de DDX3 (d'après Floor <i>et al.</i> , 2016) .....	91
Figure 44: Mécanisme hélicase « local strand separation » (adapté de Linder & Jankowsky, 2011) .....	92
Figure 45: DDX3 et la traduction du VIH-1 .....	96
Figure 46 : Expression et solubilité de DDX3.....	169
Figure 47 : Purification de MBP-DDX3.....	171
Figure 48 : Clivage de la protéine de fusion MBP-DDX3 WT par la TEV.....	172
Figure 49 : Clivage et stabilité de la protéine DDX3 .....	173
Figure 50 : Mesure de l'activité ATPase .....	174
Figure 51 : Test d'activité ATPase de la protéine clivée vs la protéine non clivée.....	175
Figure 52 : activité ATPase de MBP-DDX3 en fonction de la nature du substrat.....	176
Figure53 : spécificité de l'activité ATPase de MBP-DDX3.....	177
Figure 54 : Détermination des paramètres cinétiques de DDX3.....	178
Figure 55 : Evaluation de la capacité d'oligomérisation deDDX3.....	179
Figure 56 : Activité hélicase de DDX3 sur un duplex ARN/ARN .....	181
Figure 57 : Test d'activité hélicase de DDX3 sur un hétéroduplexe.....	182
Figure 58 : La traduction <i>in vitro</i> de la polyprotéine Gag du VIH-1 en présence de DDX3 .....	183
Figure 59: Plasmides utilisés pour l'expression protéiques .....	192

# LISTE DES ABBREVIATIONS

1M7 : 1-Méthyle 7-nitro-isatoïque anhydre.

3' ou 5' UTR : 3' ou 5' UnTranslated Region.

ADNc : Acide DésoxyRibonucléique Complémentaire.

ARE : Adenosine-Rich Elements.

ARNg : ARN génomique.

ARNm : ARNmessenger.

ARNr : ARN ribosomique.

ARNt<sup>met</sup> : ARN de transfert initiateur méthionine

CBC : Cap Binding Complex

CrPV : Cricket paralysis Virus.

CSFV : Classical Swine Fever Virus

CTE : C-Terminal Extention.

DIS: Dimerization Initiation Site.

DMSO : DiMéthylSulfOxyde.

eEF : eukaryotic Elongation Factor .

eIF : eukaryotic intiation Factor

EMCV : EncephaloMyoCarditis Virus.

eRF : eukaryotic Releasing Factor .

FMDV : Foot-and-Mouth Disease Virus.

Gag : Group antigen

HAV : Hepatitis A Virus.

IGF1R : Insulin-like Growth Factor I Receptor.

IGR : Intergenic Region.

INS : INhibitory Sequences.

IRES : Internal Ribosome Entry Site.

ISR : Integrated Stress Response.

ITAFs : IRES *Trans*-Acting Factors.

JEV : Japanese Encephalitis Virus),

LTR: long terminal repeat.

MCF : Multi Facteur Complex.

NES : Nuclear Export Sequence.

NTE :N-Terminal Extention .

ODR : Ornithine DecaRboxylase

ORF: Open Reading Frame

PABP : PolyA Binding Protein

PAIP : polyA-Binding Protein Interacting Protein.

PBS : Primer Binding Site,

PIC : Preintegration complexe.

PKR : Protein Kinase double-stranded RNA-dependent.

Psi: encapsidation Site.

PV : Poliomyelitis Virus.

RNP : RibonucléoProtein

RRE : Rev Responsive Element

RRL : Rabbit Reticulocyte Lysate

RRM : RNA Recognition Motif.

RT : Reverse Transcriptase.

SD: Splicing Donor.

SHAPE : Selective 2'-hydroxyl acylation analyzed by primer extension.

SIDA : Syndrome d'Immuno Déficience Acquise.

TAR: TransActivation Response.

TAT : Transcription Transactivator.

TEV : Tobacco Etch Virus protease.

TRBP : TAR RNA-Binding Protein.

uORF : upstream Open Reading Frame.

VHC : Virus de l'Hépatite C humaine.

VIH : Virus Immunodéficience Humaine.

VIS : Virus de l'Immunodéficience simienne.

# RESUME

L'ARN génomique du Virus de l'Immunodéficience Humaine-1 (VIH-1) est multifonctionnel. Il constitue le génome encapsidé dans les virions et sert d'ARN messenger pour la traduction des protéines virales Gag et Gag-Pol. La traduction de ces protéines dépend exclusivement de la machinerie traductionnelle cellulaire et est initiée par deux mécanismes différents : l'initiation canonique dépendante de la coiffe et l'initiation par entrée interne des ribosomes (IRES). Le VIH-1 présente deux IRES, l'un dans la région 5' non traduite (5'-UTR) qui est stimulé en phase G2/M du cycle cellulaire et l'autre dans la région codante de Gag. Ce dernier permet l'initiation de la traduction sur deux AUG en phase et conduit à la production de la protéine Gag pleine longueur mais également à la production d'une isoforme alternative de Gag, tronquée en région N-terminale. Le rôle de cette isoforme reste mal connu. Toutefois la mutation du second AUG chez VIH-1 et donc la suppression de la seconde isoforme de Gag provoque une diminution importante du taux de la réplication virale. La conservation structurelle et fonctionnelle de l'IRES Gag parmi les lentivirus suggère un rôle important de cette isoforme et de l'IRES<sub>gag</sub> dans le cycle viral.

Nos travaux visent à comprendre à un niveau moléculaire les relations hôtes-pathogènes lors de la traduction des messagers viraux. Je me suis particulièrement intéressée aux rôles de la sous-unité ribosomale 40S et de l'hélicase cellulaire DDX3 dans l'initiation de la traduction de la polyprotéine Gag du VIH-1.

La première partie de ma thèse est consacrée à l'étude de l'interaction entre la sous-unité ribosomale 40S et l'IRES<sub>gag</sub> du VIH-1. Par l'utilisation d'approches complémentaires, nous avons pu démontrer la présence de deux sites distincts de liaison au ribosome qui sont présents à proximité des deux codons d'initiation. Nous avons ensuite évalué à la fois *in vitro* et *in cellulo* (en collaboration avec l'équipe de T. Ohlmann, CIRI-ENS-Lyon) l'effet de la délétion de chacun des sites de liaison au 40S sur l'efficacité de traduction de la polyprotéine Gag. Nos résultats valident l'importance fonctionnelle des sites de liaison au ribosome pour une production optimale des deux isoformes de la polyprotéine Gag.

La seconde partie de mon travail a consisté à définir le rôle de DDX3 dans l'initiation « coiffe-dépendante » de la traduction de la polyprotéine Gag. DDX3 est une hélicase à ARN à boîte DEAD impliquée dans de nombreux processus cellulaires tels que la régulation du cycle cellulaire et la réponse immunitaire innée mais également dans tous les aspects du métabolisme de l'ARN comme la transcription, l'épissage, l'export nucléaire ou encore la traduction. Plus récemment, il a été montré que DDX3 est nécessaire à la traduction de l'ARN génomique du VIH-1, cependant son rôle exact n'a pas encore été défini. Nous avons purifié une forme recombinante de la protéine en fusion avec la MBP (Maltose Binding Protein) et effectué des cinétiques enzymatiques afin de caractériser ses propriétés biochimiques. Contrairement à ce qui a été

précédemment décrit, nos résultats montrent que DDX3 possède une activité ATPase strictement ARN-dépendante avec des constantes cinétiques similaires à celles de son homologue chez la levure, Ded1p. Nous avons également évalué l'activité hélicase de la protéine en présence de substrats de longueur et de nature variables (duplex ARN/ARN ou des hétéroduplex ADN/ARN). D'un point de vue fonctionnel, nous avons réalisé une première série d'expériences qui confirme la stimulation exercée par DDX3 sur la traduction de Gag *in vitro*. Ces résultats permettent d'envisager la caractérisation biochimique fine des interactions DDX3-ARN viral ainsi que de disséquer le rôle de DDX3 dans l'expression du génome viral.

# ABSTRACT

The Human Immunodeficiency Virus (HIV) genomic RNA is multifunctional. It acts both as a genome that is packaged within virions and as messenger RNA translated to yield the Gag and Gag-Pol polyproteins. The translation of these proteins relies exclusively on the cellular translation machinery and is initiated through two mechanisms: the canonical cap-dependent initiation pathway and the use of internal ribosome entry sites (IRESes). HIV-1 has two IRESes, one located within the 5' UTR (5' UnTranslated Region) that is stimulated during the G2/M phase of the cell cycle, and the other embedded within the Gag polyprotein coding region. The latter drives translation initiation from two AUG in frame and results in the production of the full-length Gag protein but also of an additional N-terminally truncated Gag isoform. Few things are known about this isoform, but the mutation of the second AUG causes a significant decrease in the rate of viral replication. The structural and functional conservation of Gag IRES among lentiviruses suggests an important role of this isoform and thus of the IRES in the viral cycle.

Our work aims to understand at a molecular level the host-pathogen relationships in the translation of the viral messenger RNA. My work focused on the roles of the 40S ribosomal subunit and of the cellular helicase DDX3 in the translation initiation of Gag.

During the first part of my Ph.D, I studied the interaction between the 40S ribosomal subunit and HIV-1Gag IRES. Following complementary approaches, we evidenced two distinct ribosome binding sites present close to the two the initiation sites of Gag. Then, we evaluated the effect of each 40S binding site deletion on Gag translation efficiency, both *in vitro* and *in cellulo* (in collaboration with the team of T. Ohlmann, CIRI-ENS-Lyon). Taken together, our results confirm the functional relevance of the two ribosomal binding sites to ensure optimal production of the two Gag isoforms.

The second part of my Ph.D project aims at defining the role of DDX3 in the translation initiation of Gag. DDX3 is a RNA DEAD-box helicase involved in many cellular processes such as cell cycle regulation and the innate immune response but also in all aspects of RNA metabolism such as transcription, splicing, mRNA nuclear export and translation. Recently DDX3 has been shown to favor HIV-1 Gag translation. To define its role, we first purified a recombinant form of the protein and performed kinetic experiments to analyze its biochemical properties. Contrary to what has been previously described, MBP-DDX3 displays a strictly RNA-dependent ATPase activity with kinetic constants similar to those displayed by its yeast counterpart Ded1p. We next evaluated MBP-DDX3 helicase activity towards RNA duplexes or RNA/DNA hybrids, with different length and single strand overhangs. Our preliminary results indicate that DDX3 alone is sufficient to enhance Gag translation in our *in vitro* system which paves the way to fine biochemistry experiments such as the reconstruction of functional initiation complexes assembled onto Gag RNA and evaluation of its role on Gag RNA structure.

# INTRODUCTION



## I. Le Virus de l'Immunodéficience Humaine de type 1 (VIH-1)

### A. Généralités

Le virus de l'immunodéficience humaine (VIH), est un pathogène intracellulaire obligatoire, il ne peut donc se répliquer par ses propres moyens, pour se faire il utilise la machinerie transcriptionnelle et traductionnelle de la cellule hôte dans le but d'établir une infection productive. Il se transmet suite au contact avec des fluides infectés principalement par voie sexuelle mais peut également se transmettre lors de l'utilisation de seringues contaminées, de transfusion de sang contaminé ou encore durant la grossesse, l'accouchement et l'allaitement (contamination de la mère à l'enfant). Une fois dans l'organisme, le VIH cible les cellules du système immunitaire, qui constituent les principales cellules hôtes du virus, entraînant leur destruction massive et l'apparition d'un déficit immunitaire sévère. À ce stade, on parle de syndrome de l'immunodéficience acquise, plus connu sous son acronyme SIDA. Cet état est marqué par l'apparition de maladies dites « opportunistes », souvent fatales, comme la pneumocystose pulmonaire, la tuberculose ou encore le sarcome de Kaposi (Bruchfeld et al., 2015; Grossman et al., 2006; Lucas & Nelson, 2015).

Avec plus de 34 millions de morts en 30 ans le SIDA reste l'un des problèmes majeurs de santé publique dans le monde. Selon les estimations de l'OMS et de l'ONUSIDA fin 2015 ('Le sida en chiffres 2015 | ONUSIDA), la population mondiale vivant avec le VIH se chiffrait à 37 millions de personnes. Cette même année, 1,2 million de personnes en sont mortes et près de 2 millions ont été nouvellement infectées. L'Afrique subsaharienne reste la région la plus durement touchée avec près d'un adulte sur 20 infecté, soit 26 millions de personnes infectées. En France, environ 150 000 personnes sont séropositives et 30 000 vivent avec le SIDA.

Il n'existe pas encore de traitement curatif à l'infection par le VIH, néanmoins il existe diverses approches préventives et plus de 30 traitements antiviraux efficaces qui inhibent la réplication virale et préviennent la progression au stade SIDA (Deeks, 2006; Morlat, 2014.). Cependant ces traitements restent lourds et surtout onéreux donc peu accessibles aux pays en voie de développement, les plus touchés par la pandémie. De plus, les pressions de sélection exercées par les anticorps ainsi que par les différents traitements antiviraux conduisent à l'émergence de variants d'échappement rendant ces traitements inefficaces (Bossi et al., 2001; Larder & Kemp, 1989; Richman, 1993). Pourtant, certaines personnes infectées par le VIH parviennent en l'absence de tout traitement à contrôler la réplication virale par une réponse immunitaire efficace médiée par les lymphocytes T CD8. Ainsi, chez ces patients appelés « HIV controllers » la virémie est en dessous du seuil de détection à savoir, inférieur à 50 copies d'ARN viral/ml de sang, ils ne développent donc pas la maladie (Hua et al., 2014; Lambotte et al., 2005; Sáez-Cirión et al., 2007). Le fait que ces patients « HIV controllers » arrivent

naturellement à contrôler l'infection contribue à maintenir l'espoir qu'une stratégie vaccinale puisse se concrétiser.

## B. Le VIH : origine et caractéristiques

### Origine

Le VIH, a été isolé d'un patient présentant de multiples lymphadénopathies et identifié en 1983 conjointement par une équipe française et une équipe américaine (Barré-Sinoussi et al., 1983; Gallo et al., 1983). C'est un virus de la famille des *Rétroviridae*, se caractérisant par une transcription inverse de leur génome d'acide ribonucléique (ARN) en acide désoxyribonucléique (ADN) par l'action de l'enzyme virale la transcriptase inverse (RT), qui est spécifique aux rétrovirus. Il appartient au genre *Lentivirus* caractérisé par une longue période d'incubation ainsi qu'une évolution lente de la maladie.

On distingue deux types de VIH : le VIH-1, le plus répandu, est présent dans le monde entier, et il est responsable de la pandémie du SIDA ; le VIH-2, découvert en 1986 (Clavel et al., 1986) est, quant à lui, moins virulent et sévit principalement en Afrique de l'Ouest. Les études phylogénétiques ont montré que ces deux virus partagent 42% d'homologie au niveau de leur génome, et auraient évolué à partir de souches différentes de virus de l'immunodéficience simienne (VIS). En effet, le VIH-1 est plus proche du VIScpz (Virus de l'Immunodéficience Simienne de Chimpanzé) (Gao et al., 1999), alors que le VIH-2 proviendrait du VISsm (Virus de l'Immunodéficience Simienne de Mangabeys) (Hirsch et al., 1989). Chacun de ces deux types est classé en de nombreux sous-groupes et sous-types (Figure 1). Ainsi, on dénombre 4 sous-groupes chez le VIH-1, le sous-groupe M (Major), décrit en 1983 (Barré-Sinoussi et al., 1983), a une répartition mondiale et est subdivisé en neuf sous-types A-D, F-H, J, K. Le sous-groupe O (Outlier) identifié en 1990 (De Leys et al., 1990), le sous-groupe N (New ou Non-M, Non-O) découvert en 1998 (Simon et al., 1998) et le sous-groupe P découvert en 2009 (Plantier et al., 2009), sont eux cantonnés à l'Afrique de l'Ouest.

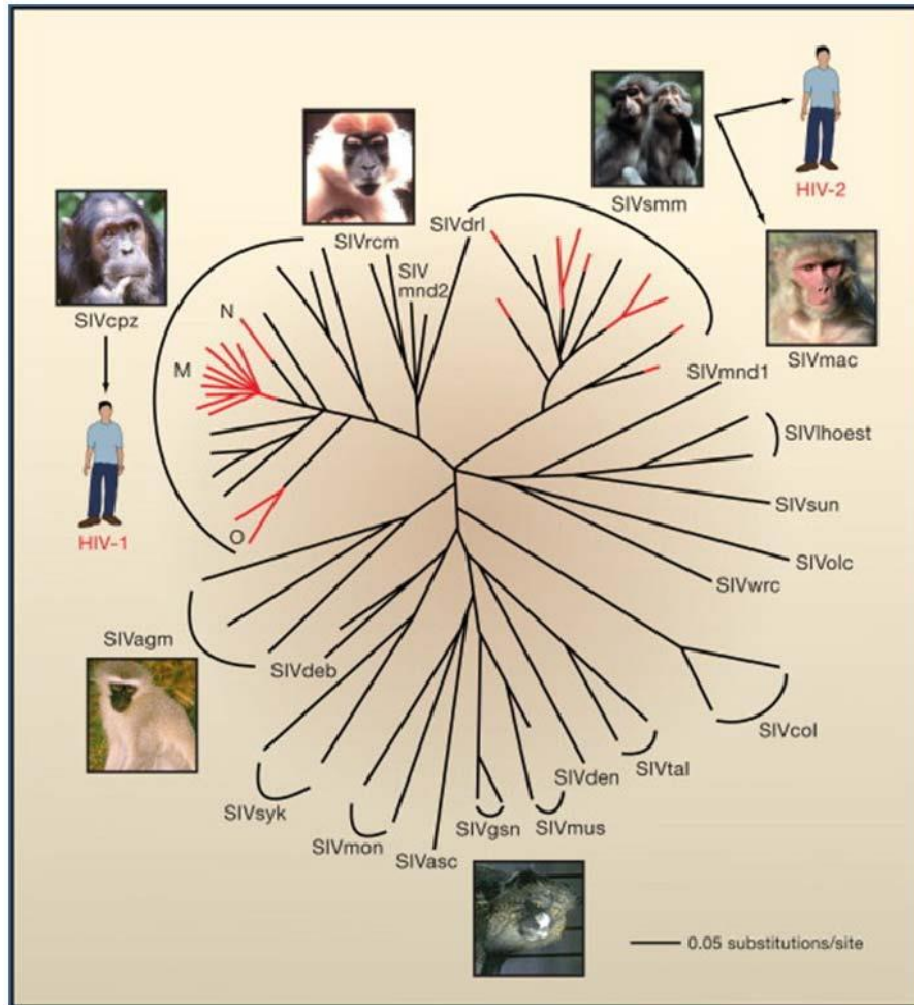


Figure 1 : Arbre phylogénétique des lentivirus de primate (d'après Ho & Bieniasz, 2008)

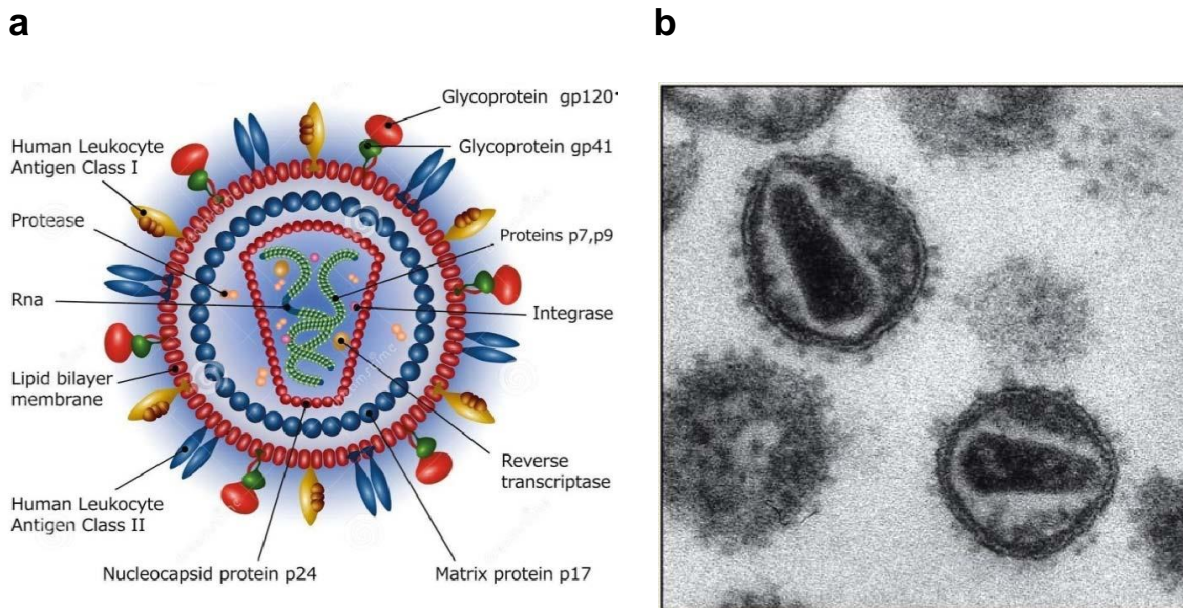
Les relations phylogénétiques entre les rétrovirus de l'immunodéficience de primates (traits noir) y compris les rétrovirus humains VIH-1 et VIH-2 (traits rouge). Les sous-groupes du VIH-1 sont indiqués par les lettres M, N et O (le sous-groupe P n'a pas été représenté).

### Structure du virion

La structure des virions du VIH-1 a été observée pour la première fois par microscopie électronique en 1987 (Gelderblom *et al.*, 1987) (Figure 2). Il apparaît de forme sphérique d'un diamètre moyen de 110 nanomètres. C'est un virus enveloppé d'une membrane plasmique provenant de la cellule hôte infectée, dans laquelle sont enchâssées les glycoprotéines d'enveloppe (Env) formées de deux sous-unités : la sous-unité de surface gp120 et la sous-unité transmembranaire gp41.

L'enveloppe comprend également des protéines cellulaires telles que l'actine, l'ubiquitine et des antigènes du complexe majeur d'histocompatibilité (HLA1 et 2). La

face interne de l'enveloppe est tapissée d'une matrice formée par l'assemblage et l'oligomérisation de la protéine de matrice p17. Cette matrice entoure une capsid conique qui forme une structure fullerène par l'assemblage de multimères de protéines p24. La capsid renferme un complexe ribonucléoprotéique constitué de deux copies d'ARN génomique viral associées à la nucléocapside p7 et aux enzymes virales : l'intégrase, la transcriptase inverse et la protéase. Les particules virales contiennent également des protéines virales auxiliaires (Nef, Vif et Vpr), des protéines cellulaires comme l'actine et la cyclophiline A, et l'ARN de transfert ARN<sup>t</sup>Lys3 (Briggs & Kräusslich, 2011; Campbell & Hope, 2015; Ganser-Pornillos et al., 2012; Gelderblom et al., 1987).



**Figure 2 : Structure d'un virion mature du VIH-1**

(a) Représentation schématique d'une particule virale. Les différentes protéines composant la particule sont indiquées. (b) particule virale observée au microscope électronique.

### Génome et protéines

Au cœur de la capsid se trouve le génome viral composé de deux copies d'ARN simple brin à polarité positive d'environ 9200 bases, liées au niveau de leurs extrémités 5' par des interactions non covalentes, sous forme d'un dimère stable. Issus de la transcription par l'ARN polymérase II cellulaire, ces ARNs viraux sont de ce fait coiffés et polyadénylés. Chaque extrémité du génome comporte une région répétée non traduite : la 5'UTR et la 3'UTR, la partie centrale correspond à la séquence codant pour les protéines virales (les protéines de structure, les enzymes et les protéines auxiliaires).

Les régions 5' et 3' UTR sont subdivisées en 3 domaines à savoir le domaine Unique 3' (U3), le domaine redondant (R) et le domaine Unique 5' (U5). Ces domaines permettent la régulation de l'expression des gènes viraux. La région U3 possède les signaux nécessaires à l'intégration de l'ADN proviral dans le génome de la cellule hôte. La région R abrite des motifs et des structures d'ARN bien caractérisés impliqués dans diverses étapes du cycle de réplication virale. Le premier élément de structure est une tige-boucle d'environ 60 nucléotides : TAR (TransActivation Response element) essentielle pour la transcription virale, suivie de la tige-boucle polyA qui doit son nom au signal de polyadénylation (AAUAAA), présente aux deux extrémités du génome mais qui n'est fonctionnelle qu'en 3'. Vient ensuite, la région U5 avec la tige-boucle PBS (Primer Binding Site) qui sert de site d'hybridation à l'ARNtLys3 afin d'initier le processus de transcription inverse. Enfin, en aval du PBS, se trouvent trois tiges-boucles (SL1 à 3) : DIS (Dimer Initiation Site) aussi appelé SL1, associée à la dimérisation de l'ARN génomique, le site donneur majeur d'épissage SD (Splicing Donor) ou SL2 et Psi (SL3) qui contient des signaux permettant l'encapsidation de l'ARN génomique (Berkhout & van Wamel, 2000; Frankel & Young, 1998) (Figure 3).

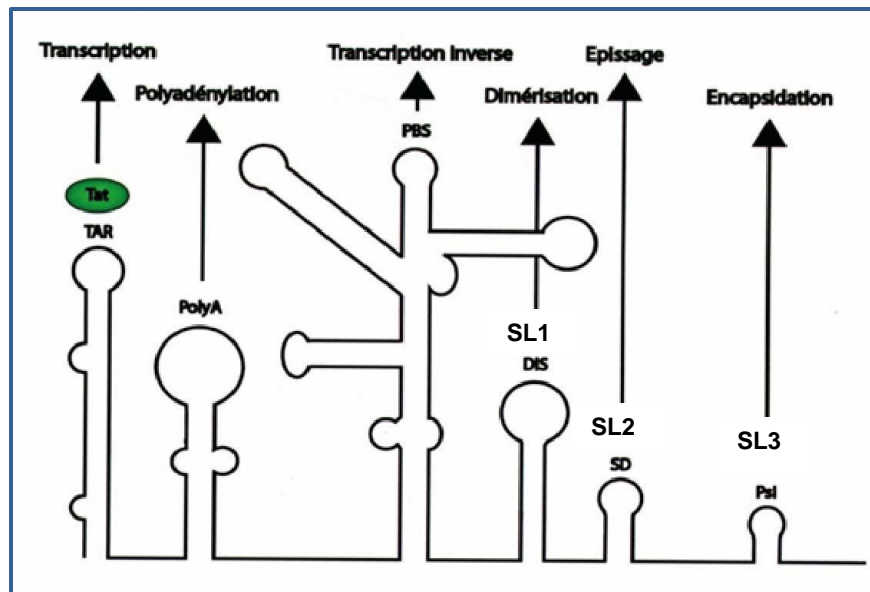


Figure 3: La 5'UTR du VIH-1 et son implication dans le cycle viral

Les différents éléments de structures secondaires de la 5'UTR du VIH-1 (335 nt) sont schématisés sur la figure et leur implication dans le cycle viral est indiquée. TAR: TransActivation Response element, PolyA : PolyA signal, PBS : Primer Binding Site, DS: dimerization initiation site, SD: splicing donor, Psi: encapsidation site.

La partie codante du génome comporte neuf gènes, trois principaux : *gag*, *pol*, *env*, qui codent pour des précurseurs polyprotéiques qui sont maturés par une protéase



virale et une protéase cellulaire afin de produire les protéines structurales et les différentes enzymes du virus (Frankel & Young, 1998; Freed, 2001) (Figure 4).

**Le gène gag** (*group specific antigen*) code pour le précurseur polyprotéique Gag (Gag p55) qui après maturation par la protéase virale donne lieu aux différentes protéines de structure : la matrice (p17), la capside (p24) et la nucléocapside (P7, P6 et deux peptides « spacer », SP1 et SP2).

**Le gène pol** (*polymérase*) code pour les enzymes virales : la transcriptase inverse (p66/p51), l'intégrase (p32) et la protéase (p12), qui sont obtenues après clivage du précurseur protéique Gag-Pol résultant d'un décalage de lecture de -1 nucléotide à la fin de la séquence codante du gène *gag*. La protéase est initialement libérée par un mécanisme d'autocatalyse afin ensuite de pouvoir assurer le clivage de la polyprotéine.

**Le gène env** (*enveloppe*) synthétise également un précurseur protéique : gp160, qui par l'action d'une protéase cellulaire, la furine, génère une glycoprotéine de surface gp120 et une glycoprotéine transmembranaire gp40.

Les six autres gènes que sont: *tat* (*Transcription transactivator*), *rev* (*Regulator of viral expression*), *nef* (*Negative factor*), *vif* (*Viral infectivity factor*), *vpr* (*Viral protein R*) et *vpu* (*Viral protein U*), sont localisés entre le gène *pol* et le gène *env* et également en 3' du gène *env*. Ils présentent des cadres de lecture chevauchants et codent pour des protéines régulatrices du même nom. Il s'agit de protéines de petit poids moléculaire qui jouent des rôles importants dans le cycle viral et dans les interactions avec la cellule hôte.

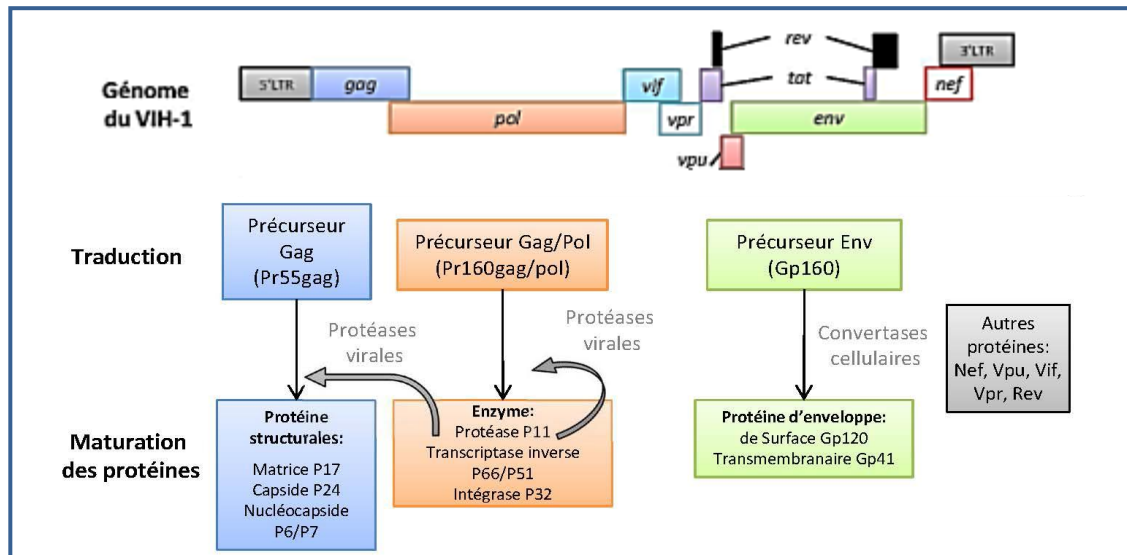


Figure 4 : Organisation du génome du VIH-1 et expression des gènes

Organisation du génome du VIH, De haut en bas : on retrouve d'abord, le génome entier du VIH-1. Puis, les précurseurs protéiques issus de la traduction de différents transcrits, et enfin, la maturation des précurseurs protéiques

### C. Cycle répliatif du VIH-1

Le cycle de réplication du VIH-1 est divisé en deux phases. La phase précoce, se réfère aux étapes allant de l'infection à l'intégration de l'ADNc viral dans le génome de la cellule, tandis que la phase tardive commence par l'expression des gènes viraux et se poursuit par la libération et la maturation du virion (Sierra *et al.*, 2005).

Le VIH-1 infecte majoritairement les cellules du système immunitaire exprimant à leur surface le récepteur CD4 tels que les lymphocytes T CD4+, les macrophages ou encore les cellules dendritiques. Ainsi, la première étape du cycle viral correspond à l'attachement de la particule virale à la membrane de la cellule cible *via* l'interaction entre la protéine virale gp120 et le récepteur CD4 entraînant des modifications de conformation de la gp120 afin de faciliter la liaison des corécepteurs (CXCR4 et CCR5). Suite à cette fixation, la membrane virale fusionne avec celle de la cellule hôte grâce à l'action de la protéine virale gp41, le matériel génétique du VIH peut alors pénétrer dans la cellule (Nisole & Saïb, 2004; Wilen *et al.*, 2012). Une fois dans le cytoplasme, la capsid se désagrège libérant les protéines viraux et l'ARN, ce dernier est rétro-transcrit en ADNc double brin sous l'action de la reverse transcriptase virale. En association avec les protéines viraux et cellulaires cet ADNc forme ce que l'on appelle le complexe de préintégration (PIC) qui est importé vers le noyau cellulaire (Craigie & Bushman, 2012). Là, l'intégrase virale catalyse l'intégration de l'ADNc au génome de la

cellule infectée formant un provirus qui fait alors partie intégrante du génome cellulaire et est en tant que tel exprimé selon les mécanismes cellulaires classiques.

Dans les phases précoces du cycle viral, le transcrit primaire du VIH est soumis à un mécanisme d'épissage complexe (Tazi *et al.*, 2010). La traduction de ces ARNs épissés aboutit à production des protéines virales vif, vpr, vpu, env, nef, tat et rev (Nisole & Saïb, 2004). La protéine Tat va favoriser la transcription d'ARN pré-messagers pleine longueur dont une partie échappe au processus d'épissage et seront exportés vers le cytoplasme durant les phases tardives du cycle viral par la protéine Rev *via* son interaction avec son élément de réponse RRE (Rev Responsive Element) situé dans l'ORF env. Une fois dans le cytoplasme, l'ARN génomique (ARNg) pleine longueur (non épissé) supporte deux fonctions : celui d'un ARNm qui sera traduit en polyprotéines gag et gag-pol, ou celui d'un ARN génomique qui sera dimérisé, encapsidé et assemblé avec les différentes protéines virales dans de nouveaux virions qui bourgeonnent de la cellule infectée en s'entourant d'une partie de sa membrane. Ceci permet la libération de virions, qui après maturation deviennent infectieux et un nouveau cycle peut recommencer (Figure 5)



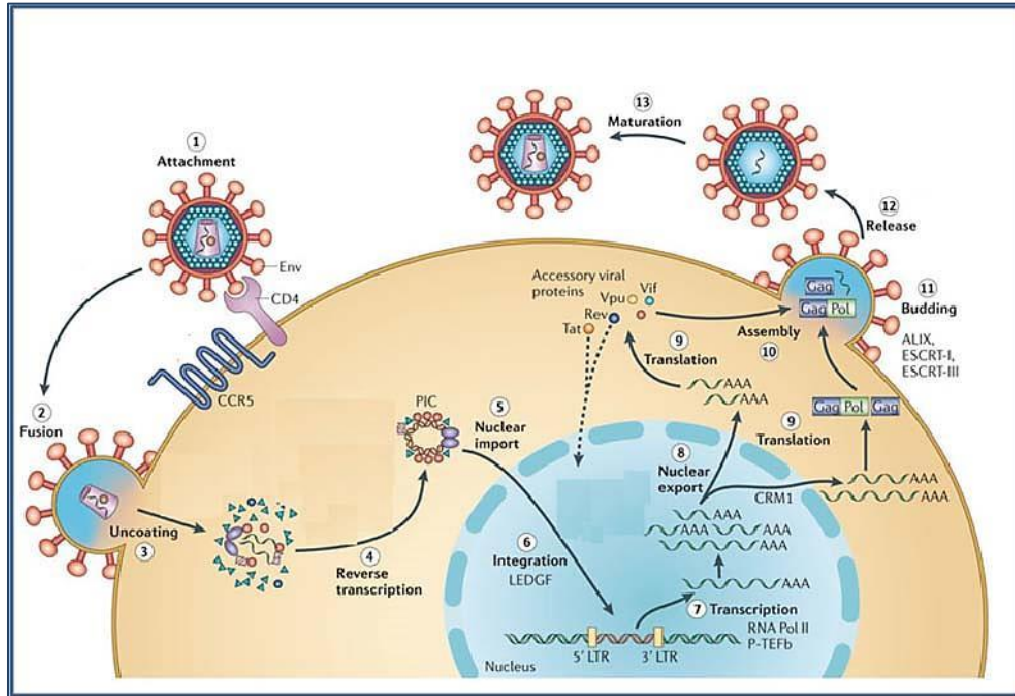


Figure 5 : Schéma récapitulatif du cycle de réplication du VIH-1 (d'après Engelman & Cherepanov, 2012)

Les différentes étapes du cycle réplcatif sont schématisées : l'attachement (1) de la particule virale au récepteur de la cellule cible puis, fusion (2) entre les membranes virale et cellulaire, pénétration suivi de la décapsidation(3) et la rétro-transcription (4) de l'ARN en ADN dans le cytoplasme de la cellule, puis (5) export nucléaire de l'ADN pour être intégré (6) au génome cellulaire par l'action de l'intégrase virale puis transcrit (7) dans le noyau en ARNm qui sera exporté (8) vers le cytoplasme pour y être traduit (9) en différentes protéines virales, lesquelles s'assemblent (10) avec deux copies d'ARN viral pour former une particule virale qui bourgeonne (11) au niveau de la membrane de la cellule infectée et qui est ensuite libérée (12) et maturée (13).

## II. L'initiation de la traduction chez les eucaryotes

### A. Généralités sur la traduction chez les eucaryotes

La biosynthèse des protéines appelée aussi traduction se déroule dans le cytoplasme et conduit à la transformation de l'information portée par un ARN messager (ARNm) en une séquence polypeptidique. Cela est rendu possible par l'action du ribosome qui positionne l'acide aminé correspondant grâce à un aminoacyl-ARN<sub>t</sub> possédant la séquence complémentaire du codon. Elle correspond également à une étape d'amplification de l'expression d'un gène, en effet, un ARNm peut être traduit de nombreuses fois successivement ou simultanément formant des regroupements de ribosomes sur un même ARNm appelés polysomes.

La traduction est un processus essentiel et conservé pour l'ensemble des règnes du vivant, nécessitant l'action contrôlée de nombreuses macromolécules. C'est donc un processus complexe et très coûteux en énergie qui est finement régulé. Cette régulation fait intervenir de multiples facteurs régulateurs ainsi qu'une certaine architecture de l'ARNm.

### Le ribosome

Le ribosome est une particule cellulaire ribonucléoprotéique constituée de l'association de molécules d'ARN appelées ARNr (ARN ribosomiques) avec des protéines ribosomales. Son rôle est d'assurer la synthèse protéique. Chez les eucaryotes le ribosome 80S mature est organisé en deux sous-unités : la grande sous-unité ribosomale 60S et la petite sous-unité ribosomale 40S (Figure 6) (S pour unité de Svedberg, qui correspond à la vitesse de sédimentation des particules).

- La sous-unité ribosomale 40S se compose de 33 protéines et de l'ARNr 18S. Elle est impliquée dans la fidélité de la traduction en permettant via l'interaction du codon de l'ARNm avec l'anticodon de l'ARN<sub>t</sub> de décoder l'information génétique portée par les ARNm.
- La grande sous-unité 60S est, elle, constituée de 49 protéines et des ARNr 5S, 28S et 5,8S. Elle contient le centre de peptidyl-transférase qui catalyse la formation de liaisons peptidiques entre les acides aminés incorporés lors de l'étape de l'élongation.

Les deux sous-unités contiennent à leur interface trois sites de liaisons aux ARN<sub>t</sub> : le site A (Aminoacyl), c'est au niveau de ce site qu'est positionné l'ARN<sub>t</sub> aminoacylé. Le site P (Peptidyl) correspond au site dans lequel se trouve l'ARN<sub>t</sub> qui porte la chaîne polypeptidique en cours de synthèse et enfin le site E (Exit), contient l'ARN<sub>t</sub> déacylé juste avant qu'il ne se dissocie de l'ARNm et du ribosome

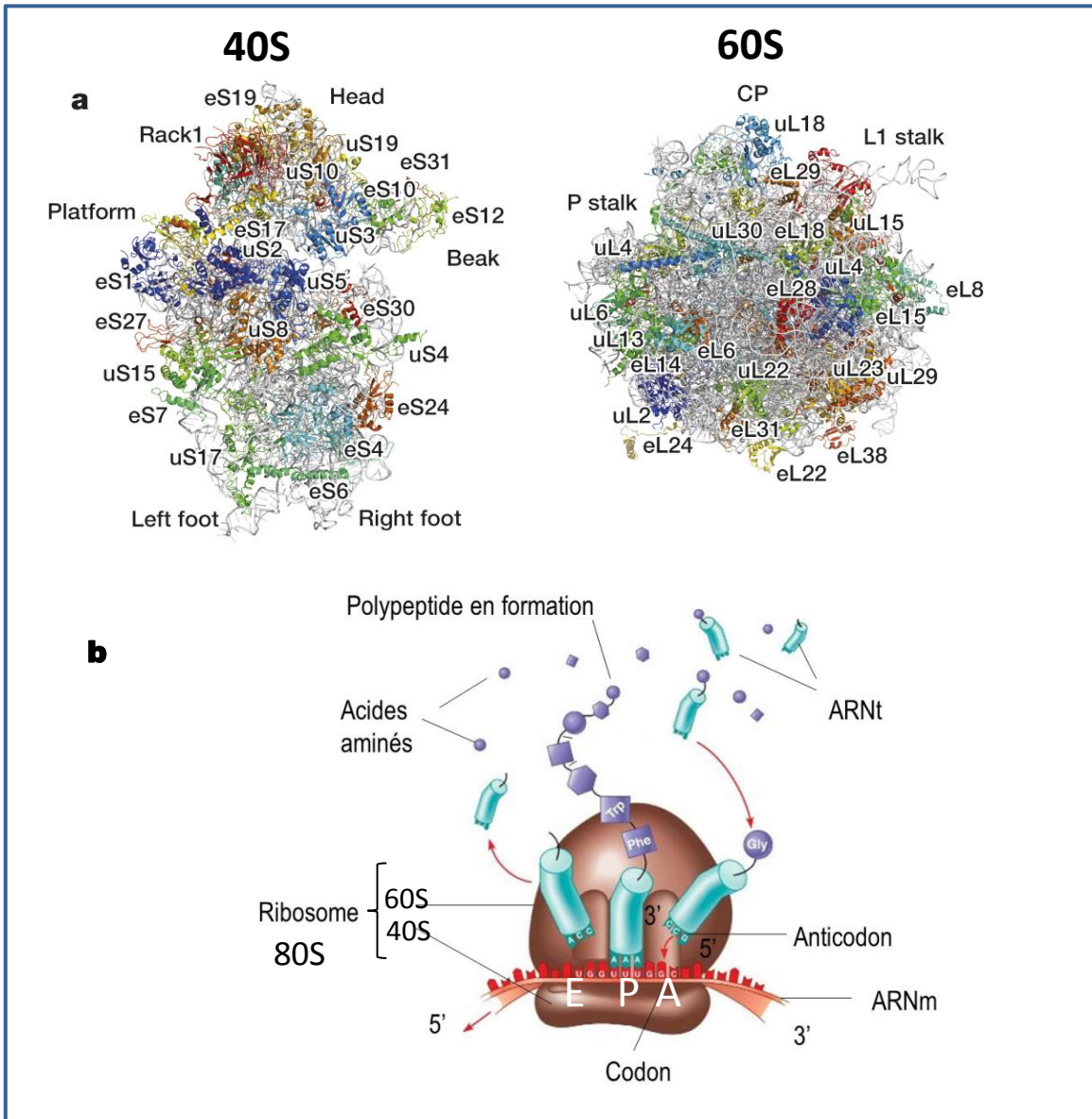


Figure 6: Structure des sous-unités du ribosome humain (d'après Khatter *et al.*, 2015)

Structure des sous-unités du ribosome humain, la sous-unité ribosomale 40S (à gauche) et 60S (à droite). Les différentes protéines ribosomales sont indiquées et colorées. Les ARN ribosomiaux sont en gris. (b) représentation schématique d'un ribosome 80S en cours de traduction. Les deux sous-unités sont indiquées ainsi que les sites A P et E.

## La structure des ARN messagers eucaryotes nécessaire à la traduction

L'ARNm est issu d'un pré-ARN messager transcrit par l'ARN polymérase II. Il est maturé de façon co-transcriptionnelle. L'ARNm est organisé en une région codante qui débute au codon d'initiation et qui se termine au codon stop, cette région est flanquée de part et d'autre, par des régions non codantes, il s'agit de la 5'UTR et de la 3'UTR. De plus, une coiffe est ajoutée en 5' de l'ARNm et une queue polyadénylée en 3' et les introns sont épissés. L'ajout de la coiffe se fait de façon co-transcriptionnelle à l'extrémité 5' d'un ARN naissant lorsqu'il atteint environ 22 à 25 nucléotides de long grâce à l'interaction d'enzymes de coiffe avec le domaine C-terminal phosphorylé de l'ARN polymérase II.

La coiffe est une guanosine méthylée sur l'azote en position 7 liée au premier nucléotide transcrit par une liaison inhabituelle **5'-5'** triphosphate (5'-m<sup>7</sup>GpppN-5'), où m<sup>7</sup>G signifie 7-méthylguanosine, p représente un groupement phosphate et N étant le premier nucléotide de la chaîne naissante d'ARN (Both *et al.*, 1975). Chez les eucaryotes supérieurs, la coiffe peut également présenter des méthylations au niveau du groupement 2'-hydroxyle du ribose des nucléotides adjacents (Figure 7). La synthèse de la coiffe se déroule dans le noyau et a lieu en trois étapes successives. La première réaction fait intervenir une ARN-triphosphatase qui retire un groupement phosphate de l'extrémité 5' du premier nucléotide transcrit, devenant ainsi un 5' diphosphate, suivie d'une ARN-guanyltransférase qui utilise le GTP comme donneur afin de catalyser le transfert de GMP sur l'ARN diphosphate. Chez l'Homme, c'est une même et unique enzyme qui assure ces deux fonctions de triphosphatase et de guanyltransférase : la protéine Hcmlp. Enfin, des méthyltransférases vont transférer des groupements méthyles, en utilisant comme donneur la S-adénosylméthionine (SAM), vers la position N7 de la guanosine et sur les nucléotides N<sub>1</sub> et/ou N<sub>2</sub>. La coiffe est une structure importante, l'inactivation d'une seule des enzymes impliquées dans sa synthèse est létale pour la cellule (Shuman, 2000; Topisirovic *et al.*, 2011).

La coiffe est impliquée dans l'initiation de la traduction, en recrutant la petite sous-unité ribosomale (40S) à l'extrémité 5' de l'ARNm par l'intermédiaire du facteur d'initiation eIF4E qui reconnaît et se lie spécifiquement à la coiffe. Elle joue également un rôle crucial dans le métabolisme des ARNs, elle permet de prolonger la demi-vie des ARNm en assurant leur protection contre la dégradation par les exonucléases cellulaires. De plus, elle participe à l'épissage et à l'export des ARNm vers le cytoplasme *via* sa liaison au complexe de fixation de la coiffe (CBC ou cap-binding complex) (Topisirovic *et al.*, 2011).

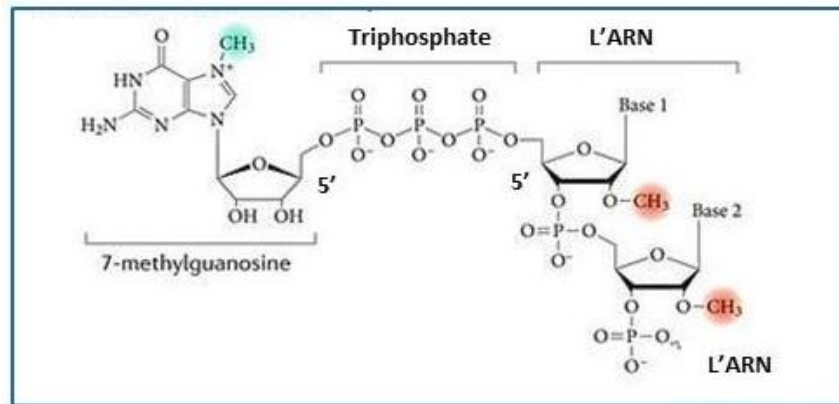


Figure 7 : Structure de la coiffe des ARNs messager eucaryotes (d'après Decroly et al., 2012)

La coiffe de l'ARNm est constituée d'un 7-méthylguanosine lié à l'extrémité 5' du premier nucléotide de la chaîne d'ARNm par une liaison 5'-5' triphosphate. Le groupe méthyle à la position N7 de la guanosine est coloré en bleu. Les groupes 2'-O-méthyle des premier et second résidus nucléotidiques de l'ARN, retrouvés chez les eucaryotes supérieurs sont colorés en rouge.

Les virus à ARN, à l'exception des rétrovirus, ne sont pas transcrits par l'ARN polymérase II et ne possèdent donc pas de coiffe. Ainsi, pour éviter d'être dégradés par la cellule, ces virus ont développé au cours de l'évolution différentes stratégies afin de contourner ce problème, c'est le cas par exemple de certains virus, appelés «cap snatcher» qui sont capables de «subtiliser» les coiffes des ARNm cellulaires déjà formées. Cette stratégie est utilisée notamment par les *Orthomyxoviridae* comme les virus de l'influenza, les *Arenaviridae* ou encore les *Bunyaviridae*, ceci implique une endonucléase qui clive les ARNm cellulaires à des sites situés 10 à 13 nucléotides en aval de la coiffe. Ces petits fragments d'ARN sont alors utilisés comme amorces pour la transcription des ARNm viraux (Dias *et al.*, 2009; Rao *et al.*, 2003). D'autres virus, notamment les virus à ADN dont le cycle réplcatif est cytoplasmique ont développé une stratégie qui consiste à coder et synthétiser la machinerie enzymatique nécessaire à la formation de la coiffe. L'exemple le plus caractérisé est celui du virus de la vaccine (Bouvet *et al.*, 2012; Decroly *et al.*, 2012; Wei & Moss, 1975).

La région comprise entre la coiffe et le codon initiateur est appelée région 5' non traduite (5'UTR pour 5' UnTranslated Region). La longueur et la structure de la 5'UTR influence l'efficacité de traduction d'un ARNm et interviennent dans le processus de contrôle traductionnel de cet ARNm. En effet, plus la 5'UTR est petite plus l'efficacité de traduction est grande, avec une longueur minimale de 12 nucléotides pour permettre une reconnaissance efficace du premier codon AUG (Kozak, 1991). Au-delà, l'efficacité



de l'initiation de la traduction augmente ce qui pourrait s'expliquer par la capacité d'une 5'UTR longue et peu structurée d'accumuler de nombreuses sous-unités ribosomales 40S sur un même ARNm lui conférant un avantage sélectif pour la traduction. Puis, lorsque la 5'UTR est trop longue et très structurée, l'efficacité de traduction diminue (Kozak, 1980; Pelletier & Sonenberg, 1985). En effet, la présence de tiges-boucles stables dans la 5'UTR d'un ARNm, diminue l'efficacité de traduction et cela est d'autant plus vrai si ces structures sont situées près de la coiffe car l'assemblage du complexe de préinitiation 43S semble alors gêné. De plus, les hélicases impliquées dans la traduction semblent peu processives et de ce fait peu efficaces pour déstabiliser de telles structures (Gray & Hentze, 1994; Pestova & Kolupaeva, 2002). Le contexte nucléotidique ainsi que la position du codon d'initiation AUG influence aussi grandement la traduction. Kozak, en comparant la séquence de près de 700 gènes a défini chez les mammifères un contexte nucléotidique favorable qui révèle l'importance d'une purine en position -3 et d'une guanine en position +4 dans la séquence consensus définie 5'-(**AG**)CCAUGG-3', dans laquelle la base A du codon AUG représente le +1 de traduction (Kozak, 1986, 1987a, b, 1997). Il a été montré par des expériences de cross-linking que l'ARNm et plus précisément les nucléotides entourant l'AUG tels que définis dans la séquence consensus, sont capables de former des interactions spécifiques avec des protéines ribosomales et avec des facteurs d'initiation. En effet, le G en position +4 est capable d'interagir avec la protéine ribosomale RPS15. La purine en position -3 permet une forte interaction avec la sous-unité  $\alpha$  du facteur eIF2, ce qui conduit, en présence du facteur eIF1, au positionnement correct du ribosome (Pisarev *et al.*, 2006).

D'autres études ont montré que le contexte optimal varie suivant les espèces. Ainsi, il correspond à la séquence GCGGC(A/C)(A/G)(A/C)CAUGGGCG chez les plantes monocotylédones et à la séquence AAAAAAAAA(A/C)AAUGGGCU pour les dicotylédones (Joshi *et al.*, 1997). Pour la drosophile la séquence consensus est ACAACCAAAAUGGC alors que pour les autres invertébrés elle correspond à AAU(A/C)AACAUG(A/G)C et à AAAAAAAAAUGUC concernant *Saccharomyces cerevisiae* (Cavener & Ray, 1991). Ces différences de contexte favorable suggèrent que la reconnaissance du codon initiateur diffère entre les espèces et que les mécanismes moléculaires et les facteurs impliqués dans la sélection du codon AUG sont également différents.

Bien que nommée région 5' non traduite, certaines 5'UTR renferment des uORF (upstream Open Reading Frame) dues à la présence de codons AUG en amont du codon initiateur. Ces uORF sont présentes dans environ 20-50% d'ARNm et sont impliquées dans la régulation de la traduction, notamment de la traduction de la phase codante en aval. En effet, lorsque l'AUG de l'uORF est présent dans un contexte favorable, il constitue un frein au balayage des ribosomes et limite les complexes

d'initiation disponibles pour la traduction de l'ORF en aval (Hinnebusch et al., 2016; Morris & Geballe, 2000).

Il existe après la phase codante d'un ARNm une autre région non traduite appelée 3'UTR (3' UnTranslated Region) qui est impliquée dans les mécanismes de régulation de la traduction et dans la localisation et la stabilité de l'ARNm. Ainsi, à l'exception de la PABP (Poly(A) Binding Protein) qui stimule la traduction (*voir ci-dessous*, « *la queue poly-adenylée* »), l'interaction de protéines régulatrices avec la 3'UTR réprime généralement la traduction des transcrits et/ou stimule leur dégradation. C'est le cas par exemple, des proto-oncogènes qui se lient à des séquences ARE (Adenosine-Uridine Rich Elements) et induisent la dégradation rapide des ARNm (Schoenberg & Maquat, 2012; Shen & Malter, 2015). La 3'UTR est également la cible de régulation par les microARN (miARN) qui régulent l'expression de presque 50% des gènes chez l'Homme. Ces miARN sont des ARNs non codant d'environ une vingtaine de nucléotides et sont complémentaires d'une séquence cible présente dans la 3'UTR. L'interaction d'un miARN à sa séquence cible va induire la dégradation de l'ARN et l'inhibition de la traduction (Chatterjee & Pal, 2009; Krol et al., 2010).

Les ARNs synthétisés par l'ARN polymérase II, sont tous polyadénylés à l'exception de ceux des histones. La polyadénylation est l'addition de façon co-transcriptionnelle d'environ 200 à 250 résidus adénine à l'extrémité 3' de l'ARNm pour former ce que l'on appelle communément une queue Poly(A). Dans un premier temps, les ARN pré-messagers sont clivés 10 à 30 nucléotides en aval du signal de polyadénylation formé par une séquence hexanucléotidique AAUAAA et à environ 30 nucléotides en amont d'une séquence riche en U ou GU nommée DSE (Downstream SequenceElement), ces régions sont respectivement reconnues par le CPSF (Cleavage andPolyadenylation Specific Factor), et le complexe CstF (CleavageStimulation Factor), le clivage produit une extrémité 3'-OH à laquelle la PAP (polyA polymérase) ajoute une dizaine de résidus adénine. Ensuite, cette séquence est allongée par la PAP stimulée par la fixation de la PABP (polyA binding protein) (Millevoi & Vagner, 2010).

La queue Poly(A) des ARNm, assure leur stabilité et leur protection contre la dégradation par des exonucléases cellulaires. Elle joue également un rôle crucial dans la traduction, elle permet, en synergie avec la coiffe, de stimuler de façon importante la traduction. En effet, la PABP est recrutée au niveau de la queue polyA et interagit ensuite avec le complexe multi-protéique eIF4F lui-même associé à la coiffe (Michel et al., 2000).

## Les étapes de la traduction

La traduction se décompose en trois étapes successives. D'abord, **l'initiation** qui conduit au recrutement et au bon positionnement du ribosome au niveau du codon initiateur. Pour la majorité des ARNm, l'initiation de la traduction se fait de manière coiffe-dépendante et commence par la reconnaissance de la coiffe située à l'extrémité 5' de l'ARNm par le facteur eIF4E et se poursuit par l'attachement du complexe de préinitiation 43S (composé de la sous-unité ribosomale 40S et des protéines cellulaires appelées facteurs d'initiation) au niveau de l'extrémité 5' de l'ARNm. Le complexe 48S ainsi formé va balayer la région 5'UTR jusqu'à rencontrer le codon initiateur. Une fois le codon AUG positionné correctement au niveau de la sous-unité ribosomale 40S, la sous-unité ribosomale 60S est recrutée pour former un ribosome 80S compétent pour l'étape d'élongation (Jackson *et al.*, 2010). Un mécanisme alternatif permet d'initier la traduction en recrutant le complexe 43S au niveau de la 5'UTR directement sur ou à proximité du codon d'initiation sans nécessiter la présence de la coiffe. Cela est rendu possible grâce à des séquences IRES (Internal Ribosome Entry Site), on parle alors de mécanisme d'entrée interne des ribosomes (Jang *et al.*, 1988; Pelletier & Sonenberg, 1988).

Vient ensuite **l'élongation**, étape au cours de laquelle la protéine est synthétisée grâce aux ribosomes et aux ARNs de transfert qui assurent le lien entre l'ARNm et la protéine produite. Le ribosome se déplace le long de l'ARNm et permet *via* des ARNt d'incorporer les uns après les autres les acides aminés correspondant aux différents codons de l'ARNm. L'élongation est donc une étape cyclique où l'incorporation de chaque acide aminé se fait en quatre étapes qui se répètent autant de fois qu'il y a d'acides aminés à incorporer. Ainsi, il y a en premier lieu reconnaissance du codon de l'ARNm par la boucle anticodon d'un ARNt aminoacylé associé au facteur eEF1A (eukaryotic Elongation Factor 1A) et à une molécule de GTP. Suite à la reconnaissance correcte codon-anticodon au niveau du site A du ribosome, le GTP est hydrolysé et le facteur eEF1A-GDP est alors relargué et recyclé en eEF1A-GTP par l'action du facteur eEF1B. Ensuite, l'activité peptidyl-transférase du ribosome catalyse la formation d'une liaison peptidique entre l'acide aminé présent au site A et la chaîne peptidique présente au site P. Enfin, l'ARNt désacétylé (ayant livré son acide aminé) est transloqué du site P vers le site E, le nouvel ARNt peptidyl passe du site A au site P et le codon suivant de l'ARNm est lui présenté au site A. Cette étape de translocation est rendue possible par l'hydrolyse du GTP associé à la GTPase eEF2 (Groppo & Richter, 2009; Kapp & Lorsch, 2004a).

Enfin, **la terminaison**, se produit lorsque le site A du ribosome rencontre un codon stop (UAA, UAG ou UGA), qui ne correspond à aucun ARNt. Le facteur eRF1 (eukaryotic Releasing Factor 1) mimant un ARNt et qui reconnaît les codons stop, est recruté par eRF3 couplé au GTP. L'hydrolyse du GTP conduit à la libération de la



chaîne peptidique néo-synthétisée et à la dissociation des deux sous-unités ribosomales. La sous-unité ribosomale 40S se lie alors à eIF3, eIF1 et eIF1A, ce qui prévient sa réassociation avec la sous-unité 60S. Les différents constituants sont ainsi recyclés, et deviennent disponibles pour traduire de nouveaux messagers (Groppo & Richter, 2009; Kapp & Lorsch, 2004a; Pisarev *et al.*, 2007).

La traduction est un processus très coûteux en énergie pour la cellule et est finement régulée, principalement au niveau de l'étape d'initiation qui représente, dans la majeure partie des cas, l'étape limitante déterminant la quantité de protéine synthétisée. L'étude de l'initiation de la traduction représente donc une étape incontournable pour comprendre l'expression d'un ARNm. Cette partie constitue l'un des sujets principaux de ce manuscrit, et sera décrite plus en détails dans les chapitres suivants.

## B. Initiation de la traduction dépendant de la coiffe

Le mécanisme d'initiation de la traduction dépendant de la coiffe, permet d'initier la traduction en assemblant un complexe d'initiation au niveau de la coiffe qui va ensuite balayer la région 5'UTR jusqu'à rencontrer le codon initiateur, puis la sous-unité ribosomale est recrutée pour commencer l'étape d'élongation (Figure 8). Il peut être divisé en différentes étapes qui sont décrites dans la suite de l'introduction. Un tableau récapitulatif de la fonction de chaque facteur d'initiation est présenté en annexe (Tableau 1)

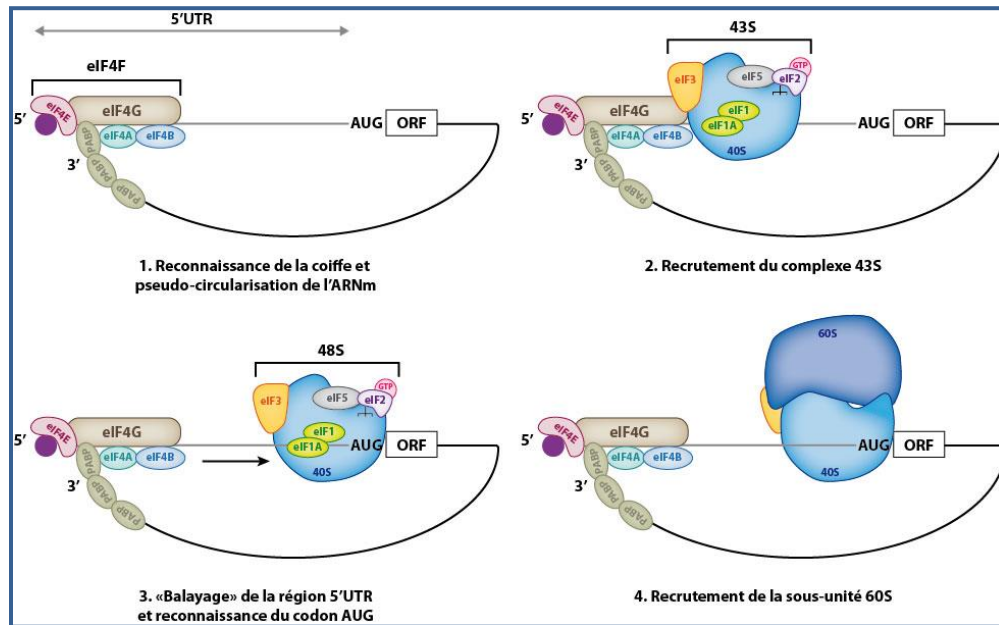


Figure 8 : Représentation schématique des différentes étapes d'initiation de la traduction coiffe-dépendante

*Schéma représentant les différentes étapes d'initiation de la traduction dépendante de la coiffe, qui commence par 1. Reconnaissance de la coiffe par eIF4E qui fait partie du complexe eIF4F (composé de eIF4G, eIF4E et eIF4A), la PABP interagit avec eIF4G provoquant le rapprochement des extrémités 5' et 3' de l'ARNm. Suivi par 2. Le recrutement du complexe d'initiation 43S puis 3. Balayage de la région 5'UTR jusqu'à rencontrer le codon initiateur AUG. Enfin 4. Recrutement de la sous-unité ribosomale 60S pour commencer l'étape d'élongation.*

### Formation du complexe ternaire

L'initiation de la traduction dépendante de la coiffe commence par la formation d'un complexe ternaire (TC, pour Ternary Complex) entre l'ARNt<sup>Met</sup> (ARN de transfert méthionine initiateur), eIF2 (eIF pour eukaryotic Initiation Factor) et une molécule de GTP qui lui est associée (Figure 9).

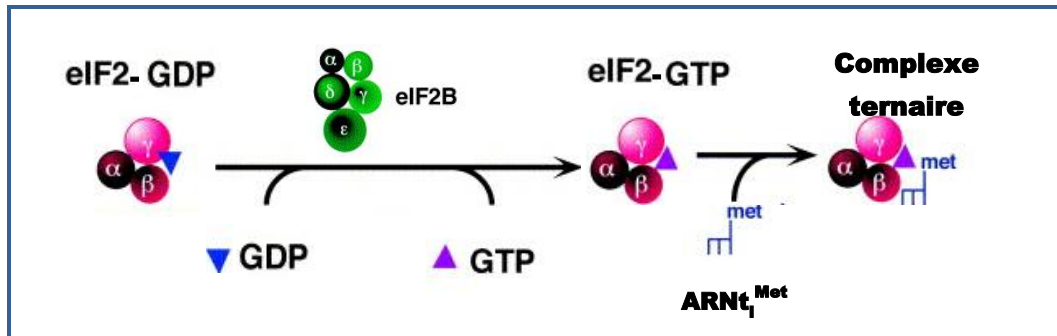


Figure 9 : Formation du complexe ternaire (adapté de Kimball, 1999).

Le GDP d'eIF2-GDP est échangé par eIF2B en GTP afin de former un eIF2-GTP actif qui s'associe à l'ARNt initiateur pour former le complexe ternaire

eIF2 est un facteur hétérotrimérique, qui se compose de deux petites sous-unités : eIF2 $\alpha$  et eIF2 $\beta$  liées à une grande sous-unité centrale eIF2 $\gamma$ . La structure d'eIF2 n'ayant pas été résolue, la structure des différentes sous-unités de son homologue d'archaebactérie (aIF2) est présentée sur la Figure 10. Les homologies entre les deux facteurs sont mises en évidence.

- **eIF2 $\gamma$**  comprend trois sites de liaison à la guanine qui sont impliqués dans la liaison directe du GTP/GDP. eIF2 $\gamma$  lie également l'ARNt<sup>Met</sup>. Sous sa forme activée, eIF2 $\gamma$ -GTP, se lie à l'ARNt<sup>Met</sup> avec une forte affinité (de l'ordre de 10 nM), cette interaction est 20 fois moins importante si le GTP est remplacé par du GDP. Cette diminution d'affinité est directement corrélée à la présence de la méthionine. En effet, eIF2-GTP et eIF2-GDP ont la même affinité pour un ARNt<sub>i</sub> désacétylé, ce qui suggère une interaction directe entre la méthionine de l'ARNt<sup>Met</sup> et eIF2-GTP, et suite à l'hydrolyse du GTP, eIF2-GDP n'interagit plus avec la méthionine et se dissocie de l'ARNt<sub>i</sub>. Ceci permet de s'assurer que l'initiation de la traduction ne sera pas inhibée par un ARNt<sub>i</sub> non chargé puisque seul l'ARNt<sub>i</sub> chargé peut être pris en charge par eIF2 $\gamma$ -GTP. Il a également été montré que la sous-unité  $\gamma$  d'eIF2 se lie au 40S au niveau de l'hélice 44 de l'ARNr 18S, ce qui contribue au recrutement du complexe ternaire sur la sous-unité ribosomale 40S (Kapp & Lorsch, 2004b; Shin *et al.*, 2011; Sonenberg & Hinnebusch, 2009).
- **eIF2 $\beta$**  est organisée en trois domaines : un domaine N-terminal contenant trois segments riches en lysines, entre 6 et 8 résidus consécutifs chacun, qu'on appelle les boîtes K, impliquées à la fois dans la liaison à l'ARNm et de l'ARNt<sup>Met</sup> mais également dans la liaison d'eIF5 (une GTPase Activating Protein) et d'eIF2B (facteur d'échange de GDP en GTP), un domaine central qui interagit avec eIF2 $\gamma$  et un domaine C-terminal en doigt de zinc impliqué dans la sélection du codon initiation AUG. En effet, des mutations au niveau ou autour de ce motif à doigt de zinc, conduit à l'initiation de la traduction sur un codon UUG (Cho & Hoffman, 2002;

Donahue *et al.*, 1988; Lorsch & Dever, 2010; Pestova, T. V., Lorsch, J. R., and Hellen, C. U., 2007).

- **eIF2 $\alpha$**  peut être considéré comme la sous-unité régulatrice du trimère. Elle contient un site important de phosphorylation : une sérine en position 52. Elle permet de recruter le facteur d'échange du GDP en GTP : eIF2B, ce qui permet de faire passer eIF2 de sa forme inactive eIF2-GDP à sa forme active eIF2-GTP qui pourra de nouveau être recruté dans un complexe ternaire et participer à une nouvelle initiation de la traduction (Pestova, T. V., Lorsch, J. R., and Hellen, C. U., 2007; Sonenberg & Dever, 2003).

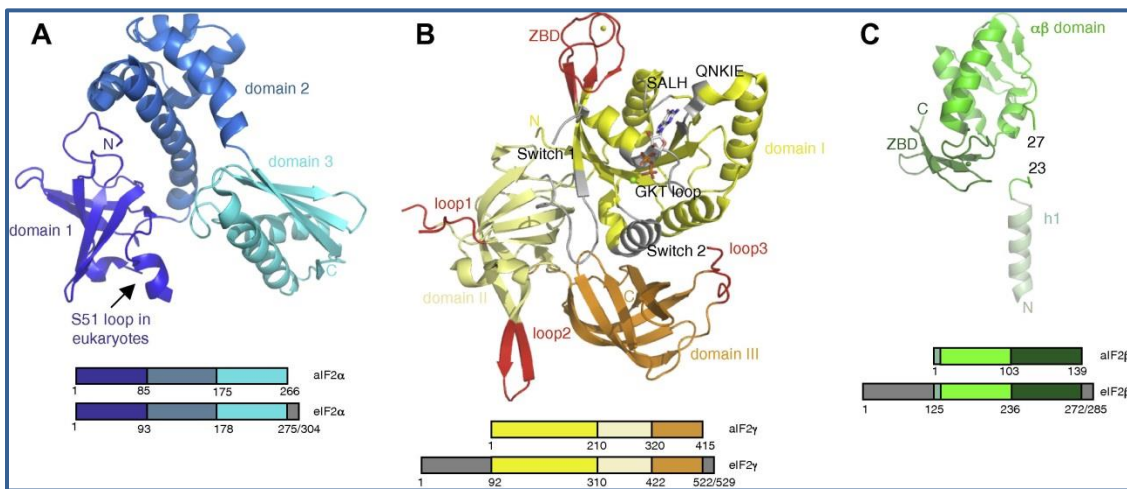


Figure 10 : Les sous-unités d'aIF2, l'homologue d'eIF2 (d'après Schmitt *et al.*, 2010)

Les Sous-unités d'aIF2. A, la sous-unité  $\alpha$  et ses trois domaines structuraux: domaine 1 en bleu foncé, domaine 2 en bleu marin, et le domaine 3 en cyan. La boucle portant le résidu Sérine51 dans eIF2 $\alpha$  est indiquée par une flèche. En B, la sous-unité  $\gamma$  et ses trois domaines structuraux: domaine I en jaune, domaine II en jaune pâle, et le domaine III en orange. Le GDP est représenté en bâtons, Mg<sup>2+</sup> en sphère verte, et Zn<sup>2+</sup> en sphère jaune. Les régions impliquées dans la liaison des nucléotides sont colorées en gris. Régions caractéristiques du facteur d'initiation sont colorées en rouge. En C, la sous-unité  $\beta$  et ses trois domaines structuraux: 1'hélice 1 en vert pâle, domaine 2 en vert, et le domaine 3 en vert foncé. Zn<sup>2+</sup> est représentée comme une sphère verte. En dessous des structures, sont comparées les sous-unités d'e/aIF2. Le code couleur des boîtes est le même que celui des domaines structuraux. Pour les sous-unités d'archaebactéries, la numérotation est celle d'aIF2 de *S. solfataricus* et pour les sous-unités eucaryotes, la numérotation est celle d'eIF2 provenant de *S. cerevisiae*. Les domaines spécifiques des sous-unités eucaryotes sont indiqués en gris.

Les ARNt sont de petits ARNs très structurés transcrits par l'ARN polymérase III. Leur structure secondaire est organisée en trois branches. La première branche appelée bras accepteur est formée suite à l'appariement des régions 3' et 5'. À l'extrémité 3' les trois dernières bases CCA ne sont pas appariées et lient un acide aminé *via* une liaison ester. La tige-boucle D doit son nom à la base modifiée dihydrouridine (D) présente dans la boucle. De même, la tige-boucle T possède une ribothymidine (T) et une pseudouridine ( $\psi$ ). Enfin, la dernière branche correspond à la boucle A qui porte l'anticodon complémentaire d'un codon sur l'ARNm. La structure tertiaire quant à elle, se replie en forme de « L » suite à l'empilement du bras accepteur avec la tige T, et de la tige D avec la tige A. La stabilité de la structure tertiaire est assurée par les bases modifiées. Ces dernières contribuent également à l'efficacité de la traduction en augmentant notamment la surface de contact avec les différentes protéines impliquées (Goldman, 2008) (Figure 11).

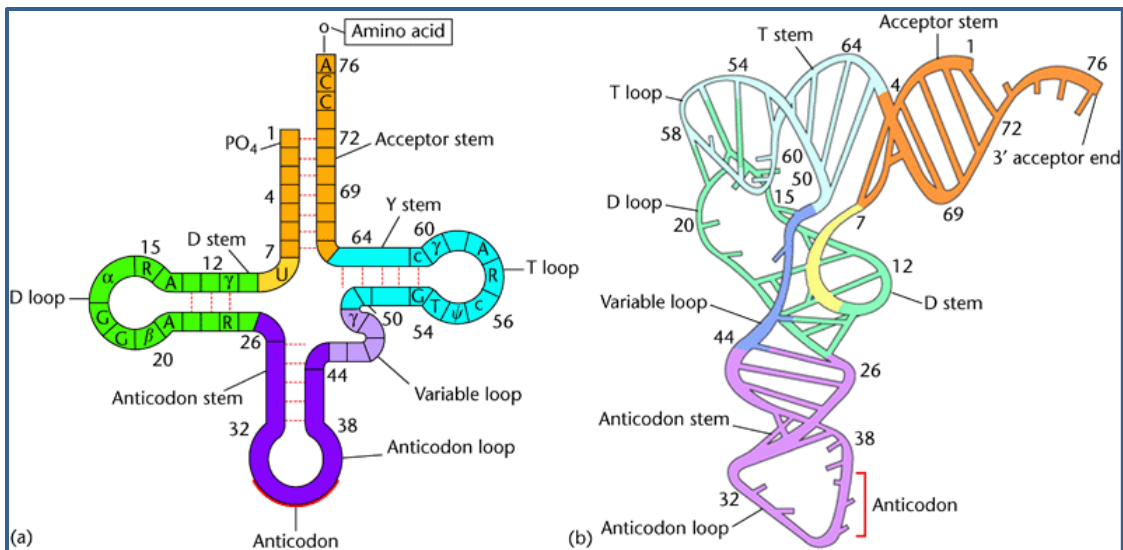


Figure 11 : Structure secondaire et tertiaire d'un ARNt (d'après Goldman, 2008)

(a) Schéma de la structure secondaire de l'ARNt<sup>Phe</sup> chez la levure. Les bases nucléotidiques qui sont indiquées se retrouvent presque universellement dans la même position dans toutes les séquences d'ARNt. Le nombre de nucléotides dans les différentes tiges et boucles est généralement constant à l'exception des deux parties de la boucle D (en vert) désignée  $\alpha$  et  $\beta$ , qui peut comporter 1 à 3 nucléotides différents suivant les ARNt) et la boucle variable (en violet). Abréviations: A : adénosine; G : guanosine; C : cytidine; U : uridine; R : adénosine ou guanosine; Y : cytidine ou uridine; T : ribothymidine; et  $\gamma$  : pseudouridine. (b) Schéma de la structure tertiaire de l'ARNt<sup>Phe</sup> chez la levure. Le code couleur de la figure (a) est conservé.

L'ARNt<sup>Met</sup> possède une organisation générale similaire aux ARNt mais contrairement à ces derniers l'ARNt<sup>Met</sup> se lie au niveau du site P et non au site A du ribosome. Il est impliqué dans la reconnaissance du codon initiateur méthionine mais ne sera pas utilisé

pour l'incorporation de résidus méthionine lors de l'élongation. En effet, l'ARNt<sub>i</sub><sup>Met</sup> et l'ARNt<sup>Met</sup> possèdent des séquences et des modifications post-transcriptionnelles différentes qui leur permettent d'être reconnus par des protéines différentes selon l'étape de la traduction concernée l'initiation ou l'élongation. Ainsi, l'ARNt<sub>i</sub><sup>Met</sup> possède une séquence et des caractéristiques structurales uniques qui lui permettent d'être spécifiquement reconnu par eIF2 telle que la paire de base A-U au niveau du bras accepteur (Farruggio et al., 1996). On retrouve également les paires de bases G-C au niveau de la boucle A qui jouent un rôle crucial dans la liaison du complexe ternaire à la sous-unité ribosomale 40S. Enfin, plusieurs bases conservées dans le boucle D et la boucle T lui permettent d'assurer ses fonctions et font de cet ARNt<sub>i</sub><sup>Met</sup> une structure unique parmi les ARNt (Pestova, T. V., Lorsch, J. R., and Hellen, C. U., 2007) (Figure 12).

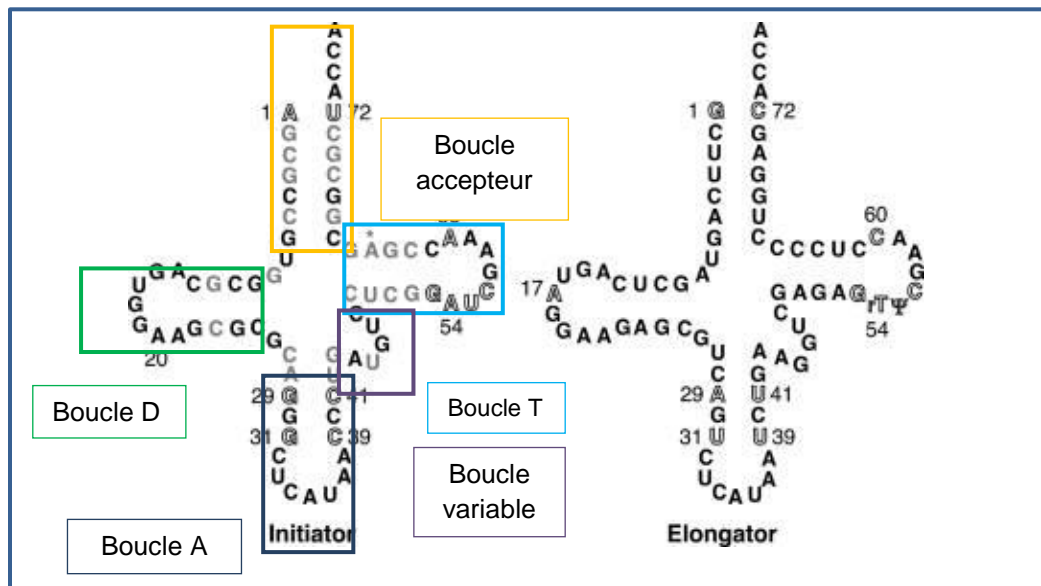


Figure 12 : Structure secondaire de l'ARNt<sub>i</sub><sup>Met</sup> vs ARNt<sub>e</sub><sup>Met</sup> chez la levure (d'après Kapp et al., 2006)

Structure secondaire de l'ARNt initiateur méthionine (à gauche) et de l'ARNt méthionine d'élongation (à droite) chez la levure. Les positions numérotées correspondent aux bases importantes pour distinguer l'ARNt initiateur de l'ARNt de l'élongation. Les nucléotides en gris indiquent ceux qui diffèrent entre les deux ARNt et qui pourraient être impliqués dans la discrimination entre les deux ARNt mais qui n'ont pas été démontrés. Les bases modifiées ne sont pas représentées, sauf celles qui sont impliquées dans la discrimination entre les deux ARNt. L'astérisque indique la position de la modification 2'-O-phosphoribosyl sur l'A64 présente chez les ARNt initiateur de plantes et de champignons. Cette modification est un déterminant de discrimination entre les ARNt initiateur et les ARNt de l'élongation.



## Formation du complexe de préinitiation 43S

Le complexe de préinitiation appelé aussi complexe 43S est formé de la sous-unité ribosomale 40S associée aux facteurs d'initiation eIF1A, eIF3 et eIF5, et du complexe ternaire (eIF2-GTP- ARNt<sup>Met</sup>).

eIF1 et eIF1A sont des facteurs d'initiation qui lient directement la sous-unité ribosomale 40S, eIF1 se lie près du site P alors qu'eIF1A se lie près du site A. Ces deux facteurs jouent un rôle important dans le recrutement du complexe ternaire au niveau du 40S et dans la reconnaissance du codon initiateur. En effet, en l'absence de ces facteurs le complexe d'initiation peut sélectionner un codon d'initiation non-AUG (Pestova & Kolupaeva, 2002). L'interaction d'eIF1 et eIF1A avec le 40S induit un changement de conformation qui est dite « ouverte », le tunnel d'entrée de l'ARNm devient alors accessible suite au basculement de la « tête » du 40S sur le « bras » (Figure 13). De façon coopérative ces deux facteurs permettent de maintenir le ribosome sous sa forme « ouverte » ce qui stimule la liaison du complexe ternaire et favorise l'étape de balayage de l'ARNm. Le facteur eIF1 accélère, aussi bien, le recrutement que le relargage du complexe ternaire alors que le facteur eIF1A accélère son recrutement mais ralentit son relargage stabilisant ainsi la liaison du complexe ternaire sur le complexe de préinitiation (Aitken & Lorsch, 2012; Majumdar *et al.*, 2003).

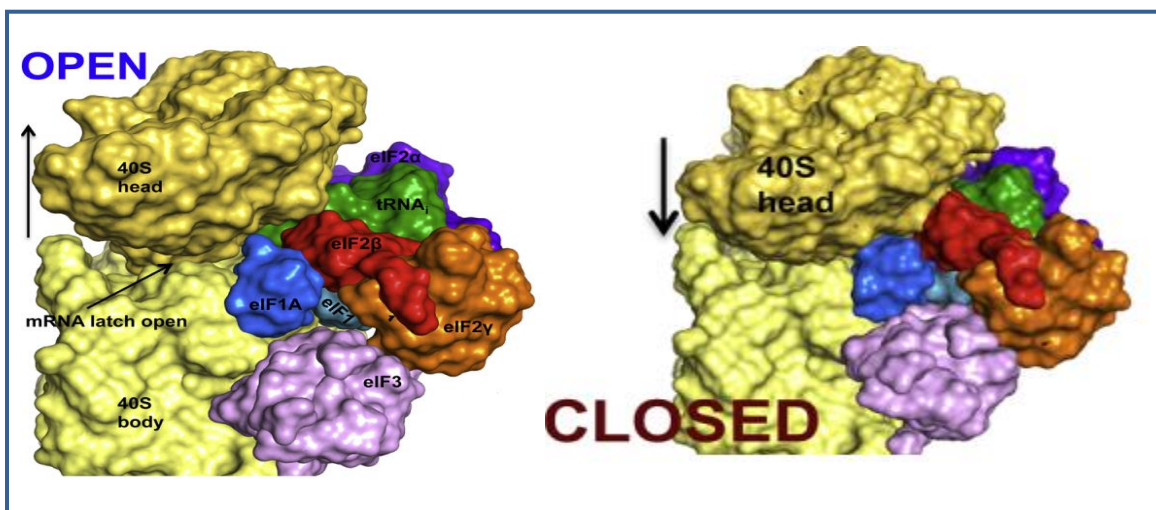


Figure 13 : Modèle de reconstitution du complexe de préinitiation dans les conformations ouverte et fermée (d'après Llácer *et al.*, 2015)

La structure du complexe d'initiation et de ces différents constituants est représentée. La sous-unité ribosomale 40S (jaune), le facteur eIF1A (bleu foncé), eIF1 (bleu clair), l'ARNt<sub>i</sub> (vert), et les trois sous-unités d'eIF2 α (violet), β (rouge) et γ (orange). Dans la conformation fermée (closed), la tête (head) de la sous-unité 40S bascule sur le corps (body).

eIF5, quant à lui ne se lie pas directement au ribosome mais est recruté au complexe 43S *via* son interaction avec eIF2 et eIF3. Sa partie N-terminale qui se lie à eIF2 agit comme une GAP (GTPase Activating Protein) et conduit à l'augmentation du taux de GTP hydrolysé au sein du complexe de préinitiation (Hinn, 2007). Il a été montré chez la levure qu'en plus de son activité GAP, eIF5 possède une activité GDI (GDP dissociation inhibitor) indépendante de sa fonction GAP. En effet, eIF5 stabilise la liaison du GDP à eIF2 et empêche sa régénération en GTP par eIF2B (Jennings & Pavitt, 2010). Par ailleurs, les travaux de Maag chez la levure, ont montré qu'eIF5 était impliqué dans la reconnaissance du codon initiateur, pour cela il interagit avec eIF1A conduisant ainsi au passage du complexe 43S de l'état « ouvert » à l'état « fermé » et donc à l'arrêt de l'étape de balayage (Maag *et al.*, 2006).

Chez les mammifères eIF3 est un facteur multi-protéique d'environ 750 kDa composé de 13 sous-unités protéiques (d'eIF3a à eIF3m) (Figure 14), contrairement à son homologue de levure *Saccharomyces cerevisiae* qui ne contient que 6 sous-unités (a, b, c, g, i, et j). Il est impliqué dans différentes étapes de l'initiation. Ainsi, eIF3 se lie directement à la sous-unité ribosomale 40S et empêche l'association de la sous-unité 60S, il stimule le recrutement du complexe ternaire, l'assemblage et la stabilisation du complexe de préinitiation 43S. Il est également impliqué dans le recrutement de l'ARNm par l'intermédiaire de domaines spécifiques de liaison à l'ARN. Il joue aussi un rôle dans le balayage de la 5'UTR et dans la reconnaissance du codon initiateur par l'action de sa sous-unité eIF3c qui interagit avec eIF5 et eIF1 (Hinnebusch, 2006). De plus, eIF3 est impliqué dans la formation du complexe 48S du fait de son interaction directe avec le facteur eIF4G, permettant ainsi de faire le lien entre le complexe eIF4F et la sous-unité ribosomale 40S (Hinnebusch, 2006; Siridechadilok *et al.*, 2005). Par ailleurs, eIF3 agit également comme un facteur de ré-initiation et de recyclage des ribosomes en se liant à la sous-unité 40S dès la fin de la terminaison (Kolupaeva *et al.*, 2005).

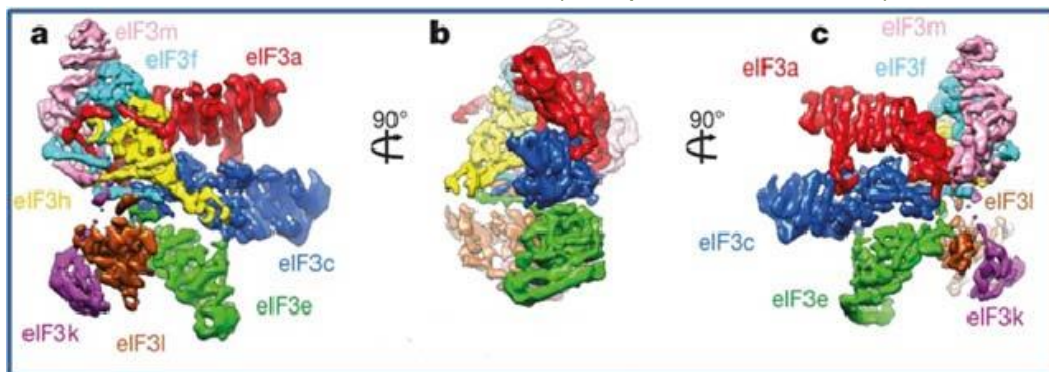


Figure 14 : Modèle de reconstitution d'eIF3 basé sur la cryoEM (d'après des Georges *et al.*, 2015)

La structure des différentes sous-unités d'eIF3 vue sous différents angles (a-c), sont représentées en différentes couleurs.



Chez *Saccharomyces cerevisiae*, il a été montré que même en l'absence de la sous-unité 40S, eIF3, eIF1, eIF5 ainsi que le complexe ternaire pouvaient s'associer pour former un complexe multi-facteur (MCF). eIF3 étant le pilier central, s'associe à eIF1 et à eIF5. eIF5 lie eIF2 servant ainsi de lien entre eIF2 et eIF3. Cependant, il est difficile de savoir si chaque facteur se lie indépendamment à la sous-unité 40S ou si le MCF se forme pour être recruté sur le 40S ou bien encore si ces deux événements coexistent (Hinnebusch *et al.*, 2007) (Figure 15).

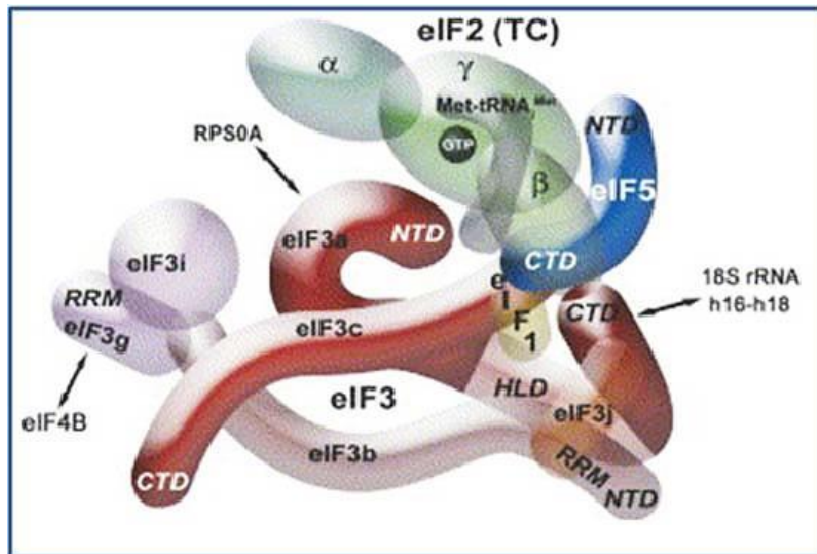


Figure 15 : Modèle du complexe multi facteur chez la levure (d'après Hinnebusch, 2006)

Modèle du MCF chez la levure basé sur l'étude des interactions *in vitro* entre les différents facteurs d'initiation

### Formation du complexe 48S

Le complexe multi-protéique eIF4F composé d'eIF4E, eIF4G et eIF4A aidé d'eIF4B et de la PABP, va permettre de recruter le complexe 43S à l'extrémité 5' de l'ARNm et former ainsi un complexe 48S. Le complexe eIF4F est composé de trois facteurs dont eIF4E qui reconnaît spécifiquement la coiffe, eIF4A une hélicase à ARN et eIF4G qui possède un rôle architectural.

### eIF4E, facteur de liaison à la coiffe

eIF4E est un facteur de 25 kDa, essentiel chez les eucaryotes. Il est exprimé de façon ubiquitaire et est localisé à la fois dans le cytoplasme et le noyau de la cellule où l'on retrouve jusqu'à 68% d'eIF4E et ce, chez une grande variété d'espèces allant de la levure à l'homme. Dans le cytoplasme, eIF4E se fixe spécifiquement à la coiffe (5'-m<sup>7</sup>GpppN-5') située à l'extrémité 5' des ARNm. Ceci stimule la liaison d'eIF4G et d'eIF4A et permet de former le complexe eIF4F prêt à recruter le complexe 43S (von der Haar *et al.*, 2004). Dans le noyau, eIF4E sert à l'export de l'ARNm d'un sous-ensemble

spécifique d'ARNm présentant un motif de 50 nucléotides au niveau de leur 3'UTR connu sous le nom de 4E-SE (eIF4E Sensitivity Element). Ce sont généralement des ARNs impliqués dans le cycle cellulaire et dans la survie et la prolifération cellulaire tels que l'ARNm codant la cycline D1 ou encore l'ARN codant pour l'ODR (ornithine decarboxylase) (Culjkovic *et al.*, 2006, 2007).

La structure d'eIF4E lié à la coiffe a été résolue par cristallographie aux rayons X pour le facteur humain et murin et par RMN en solution pour son homologue de levure. Elle se compose d'un feuillet bêta comportant huit brins antiparallèles soutenus par trois hélices alpha formant une structure incurvée en forme de « gant de baseball ». Son interaction avec la coiffe est assurée par une liaison de type  $\pi$ - $\pi$  stacking entre le résidu guanine de la coiffe et deux résidus tryptophane d'eIF4E (Trp56 et Trp102 chez l'Homme) (Borden *et al.*, 2011; Fischer, 2009) (Figure 16).

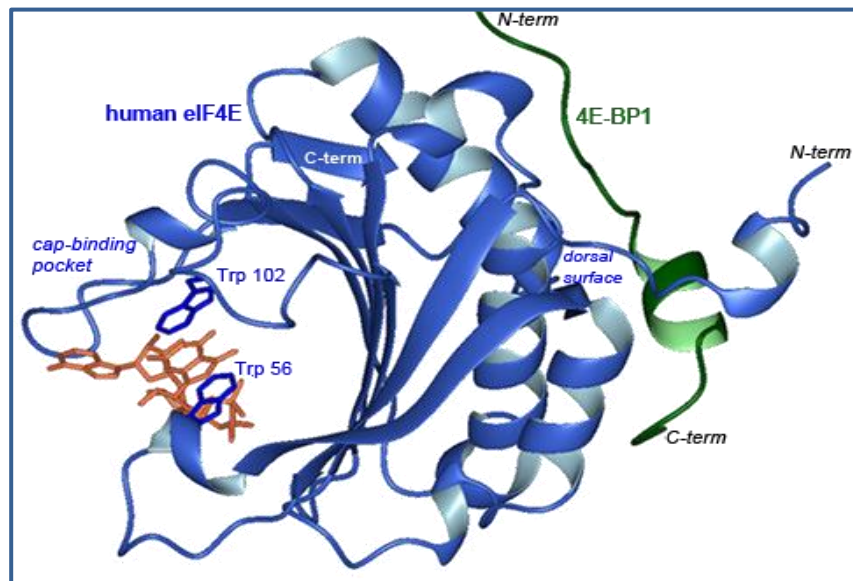


Figure 16 : Structure d'eIF4E lié à la coiffe (d'après Borden *et al.*, 2011)

*eIF4E* (en bleu) est représenté lié à la coiffe (en orange) intercalée entre les tryptophanes 102 et 56. Le domaine d'interaction d'eIF4E avec le peptide 4E-BP1 (en vert) est représenté.

Bien qu'ils soient fonctionnellement conservés, eIF4E humain et son homologue chez la levure *S. cerevisiae*, ne partagent que 30% d'identité de séquence. Cependant, l'alignement d'acides aminés d'eIF4E de mammifères avec eIF4E de plantes et de levures, révèle que le noyau d'eIF4E représenté par environ 160 acides aminés (de l'histidine 37 à l'histidine 200 chez eIF4E humain) est conservé et est suffisant pour la reconnaissance de la coiffe et pour la liaison à eIF4G et à 4E-BPs (eIF4E Binding Proteins) (Borden *et al.*, 2011). Ces dernières interviennent dans la régulation de la

traduction (ce point est détaillé dans la partie régulation globale de l'initiation de la traduction)

## eIF4G, la protéine de liaison

eIF4G est une protéine dite d'échafaudage qui fait le lien entre l'ARNm et les différents facteurs impliqués dans la formation d'un complexe 48S, (Figure 17).

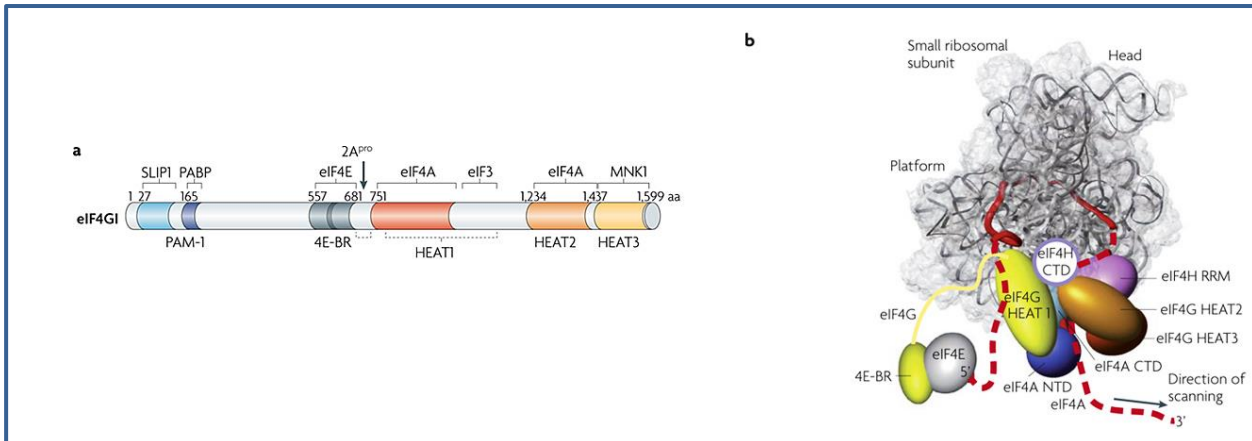


Figure 17: Structure d'eIF4G et les différents facteurs d'initiation liés (d'après Jackson et al., 2010)

a : Représentation schématique de la structure primaire d'eIF4G. Ses domaines de liaison pour les partenaires protéiques sont indiqués en couleurs et en pointillés pour l'ARN. Le site de clivage d'eIF4G par la protéase des picornavirus est indiqué par une flèche. b : Modèle de structure du complexe préinitiation 43S représentant les interactions entre les domaines d'eIF4G et ses différents partenaires protéiques.

eIF4G se lie à l'extrémité N-terminale d'eIF4E par son extrémité N-terminale et plus précisément au niveau du motif très conservé Y(X4)L $\phi$  (où  $\phi$  est un acide aminé hydrophobe, X correspond à n'importe quel résidu, Y pour la tyrosine et L pour la leucine)., cette interaction va induire un changement de conformation du complexe. eIF4E va alors adopter une structure en forme de « poing » autour de laquelle le motif d'eIF4G qui prend une structure en hélice  $\alpha$  s'enroule tel un « bracelet moléculaire » (Gross et al., 2003; von der Haar et al., 2004). eIF4G possède au niveau de son domaine médian deux sites de liaison à l'ARNm. Ces domaines seraient notamment impliqués dans une reconnaissance efficace de la coiffe par eIF4E (Yanagiya et al., 2009). eIF4G, possède également au niveau de son extrémité C-terminale un site de liaison à la sous-unité eIF3e, ce qui permet de recruter le complexe 43S au niveau de l'extrémité 5' de l'ARNm (LeFebvre et al., 2006). Le domaine C-terminal et le domaine médian d'eIF4G permettent de recruter eIF4A et de stimuler son activité hélicase ce qui facilite le recrutement du ribosome (Imataka & Sonenberg, 1997). Enfin, eIF4G lie

également la PABP grâce au site de liaison situé dans sa partie N-terminale, ce qui permet de circulariser l'ARNm et de favoriser l'initiation (Hinnebusch *et al.*, 2007).

### **eIF4A, l'hélicase à ARN**

eIF4A est une hélicase à ARN ATP- dépendante appartenant à la famille des hélicases à boîte « DEAD », qui partagent des motifs et des séquences conservés impliqués dans la liaison à l'ARN, dans la liaison et l'hydrolyse de l'ATP ce qui leur permet d'assurer leurs activités ATPase et hélicase. eIF4A est structurée en deux domaines RecA-like reliés entre eux par un « linker » (Caruthers *et al.*, 2000; Cordin *et al.*, 2006; Pestova, T. V., Lorsch, J. R., and Hellen, C. U., 2007). Son interaction avec le domaine central HEAT d'eIF4G permet de maintenir la conformation « fermée » active d'eIF4A ce qui va permettre de stimuler l'étape de balayage en déstabilisant les structures secondaires présentes dans la région 5'UTR. eIF4A seule n'est pas suffisante pour déstabiliser ces structures, cependant son activité hélicase est stimulée de façon synergique par son recrutement au sein du complexe eIF4F et par les deux cofacteurs eIF4B et eIF4H (Caruthers *et al.*, 2000; Jackson *et al.*, 2010; Pestova *et al.* 2007; Rozen *et al.*, 1990).

### **eIF4B, cofacteur d'eIF4A**

eIF4B est une protéine de 69 kDa qui possède un domaine C-terminal riche en arginine qui permet la liaison de l'ARNm, et un domaine N-terminal qui contient un motif RRM (RNA Recognition Motif) qui lie l'ARNr 18S du ribosome. Elle possède aussi un site de liaison à la sous-unité eIF3a. Ces interactions lui permettent, en parallèle du lien eIF4G-eIF3, de recruter le complexe 43S au niveau de l'ARNm. eIF4B est plus connu comme cofacteur stimulant l'activité hélicase d'eIF4A. Son mécanisme d'action reste peu clair, cependant, il semblerait qu'elle agisse en coopération avec eIF4G et stimule la liaison à la fois de l'ATP et de l'ARN à eIF4A ce qui conduit à une augmentation de l'activité hélicase d'eIF4A (Özeş *et al.*, 2011).

### **PABP, la "polyA binding protein"**

La protéine PABP, est un facteur de 70 kDa qui, par l'intermédiaire de son domaine N-terminal contenant les motifs RRM 1 et 2, reconnaît la queue polyA. Elle interagit avec le facteur eIF4G et permet ainsi la pseudo-circularisation de l'ARNm et le rapprochement de la région 3'UTR de la région 5'UTR ce qui est à la base de l'effet synergique de la coiffe et de la queue polyA sur l'efficacité de traduction (Borman *et al.*, 2000; Michel *et al.*, 2000). Le domaine C-terminal quant à lui est impliqué dans la dimérisation de la PABP. Ainsi, chez les mammifères jusqu'à 6 molécules de PABP peuvent se lier à la queue polyA (Baer & Kornberg, 1980; Mangus *et al.*, 2003). La PABP interagit aussi avec deux partenaires protéiques PAIP 1 et 2 (poly(A)-binding protein Interacting Protein) qui respectivement active et inhibe la traduction (Craig *et al.*,

1998). PAIP-1 interagit avec la PABP et eIF3g et stabilise la liaison de la PABP à eIF4G favorisant ainsi la pseudo-circularisation de l'ARNm ce qui conduit à la stimulation de la traduction (Martineau *et al.*, 2008). PAIP-2 quant à elle dissocie la PABP de la queue polyA, la liaison à eIF4G n'est plus assurée, ce qui conduit à l'inhibition de la traduction (Karim *et al.*, 2006).

PABP interagit également avec le facteur de terminaison eRF3, et facilite le recyclage des ribosomes libérés au niveau du codon stop vers l'extrémité 5' de l'ARNm pour initier la traduction (Hoshino *et al.*, 1999).

### **Le balayage et la reconnaissance du codon initiateur**

Le complexe d'initiation 43S balaye l'ARNm en allant de la région 5' vers 3' à la recherche du codon initiateur AUG. Le mécanisme de balayage se décompose en deux étapes étroitement liées, le déroulement des structures secondaires de la 5'UTR par l'hélicase à ARN eIF4A et le mouvement de la sous-unité ribosomale 40S tout au long de cette région 5'. Les facteurs eIF1 et eIF1A maintiennent le ribosome dans sa conformation « ouverte » propice à l'étape de balayage. Le tunnel alors formé par cette conformation « ouverte » permet de recruter l'ARNm dont les triplets seront décodés au niveau du site P. Lorsque le codon initiateur est présent dans un contexte de Kozak, il est reconnu et forme un appariement de type Watson-Crick avec la boucle anticodon de l'ARNt<sub>i</sub><sup>Met</sup>. Suite à cette interaction le facteur eIF1 qui est présent près du site P, crée un encombrement stérique avec l'ARNt<sub>i</sub><sup>Met</sup>. eIF1 est alors relargué, ce qui provoque le passage du ribosome de sa conformation « ouverte » à « fermée ». Ceci arrête l'étape de balayage et conduit à l'immobilisation de la sous-unité ribosomale 40S sur l'ARNm et à l'hydrolyse du GTP associé à eIF2 stimulée par eIF5. eIF2-GDP ayant moins d'affinité pour le complexe d'initiation se dissocie de ce dernier. À cette étape, seul le facteur eIF1A demeure associé au 40S et permet par l'intermédiaire de sa partie C-terminale de s'associer au facteur eIF5B-GTP qui recrute alors la grande sous-unité ribosomale 60S pour former un ribosome 80S prêt à démarrer l'étape d'élongation (Hinnebusch, 2011, 2014; Jackson *et al.*, 2010; Pestova, T. V., Lorsch, J. R., and Hellen, C. U., 2007).

### **La régulation de l'initiation de la traduction globale**

De nombreux niveaux de régulation de la traduction sont possibles jouant entre autre sur l'intégrité, la disponibilité et l'état de phosphorylation des facteurs impliqués. Ceci permet de moduler leur activité afin de répondre rapidement aux stimuli extracellulaires (hormones, facteurs de croissance) ou intracellulaires (disponibilité en acides aminés, en énergie) ou bien suite à un stress cellulaire. Quelques exemples de régulation traductionnelle dans le contexte d'une infection virale sont abordés ci-dessous.

## La phosphorylation du facteur eIF2

L'un des mécanismes principalement ciblé par la cellule pour réguler sa traduction est la phosphorylation du facteur eIF2. En effet, suite à une infection virale, PKR (Protein Kinase double-stranded RNA-dependent) une kinase inductible par les interférons, est activée par la présence d'un ARN double brin. Une fois activée, PKR va phosphoryler le facteur eIF2 au niveau de la sérine 51 de sa sous-unité régulatrice alpha. Ceci a pour conséquence d'augmenter l'affinité d'eIF2-GDP pour le facteur eIF2B et d'empêcher le recyclage du GDP en GTP par ce dernier. Ce qui conduit à la séquestration d'eIF2 par le facteur eIF2B et donc à la diminution du stock d'eIF2-GTP actif disponible dans la cellule, induisant alors l'inhibition de la traduction globale (Figure 18). (Jackson *et al.*, 2010; Sonenberg & Hinnebusch, 2009).

Chez les eucaryotes, 4 kinases permettent de phosphoryler eIF2. Ces kinases possèdent un domaine catalytique conservé qui leur permet de phosphoryler la sérine 51 d'eIF2 $\alpha$ . En revanche, leurs domaines effecteurs étant différents, l'activation de ces kinases est différente en fonction du type de stress. L'ensemble des réponses induites par l'activation de ces kinases est regroupé sous le nom de réponse intégrée au stress (ISR, pour integrated stress response) qui conduit soit à la restauration de l'homéostasie cellulaire soit à l'apoptose selon le contexte cellulaire (Donnelly *et al.*, 2013; Proud, 2005) (Figure 18).



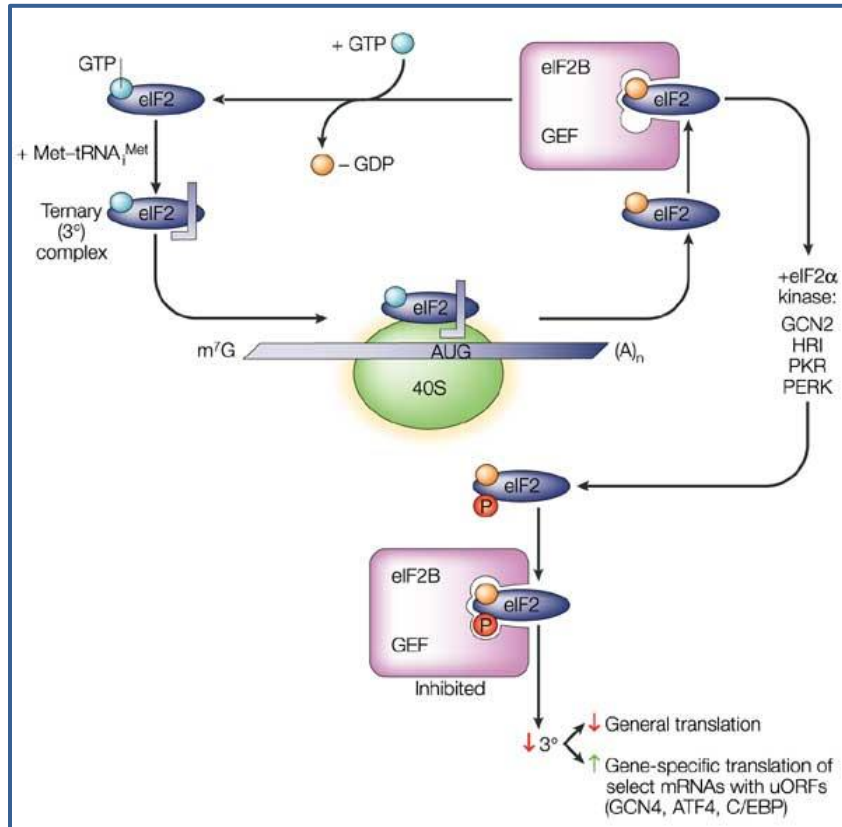


Figure 18 : Recyclage d'eIF2-GDP en eIF2-GTP par eIF2B et sa régulation par les kinases (d'après Klann & Dever, 2004).

*La phosphorylation de la sous-unité eIF2α par des kinases cellulaires empêche l'échange du GDP par le GTP assuré par le facteur eIF2B et inhibe la traduction des messagers cellulaires.*

### **Régulation d'eIF4E par eIF4E-BP**

En réponse à une infection virale, la kinase m-TOR (mammalian Target Of Rapamycin) qui phosphoryle eIF4E-BP (eIF4E Binding protein) est inhibée. eIF4E-BP n'étant pas phosphorylé, se lie à eIF4E au niveau du site de liaison d'eIF4G. En effet, eIF4E-BP partage avec eIF4G le même site de liaison à eIF4E. La liaison d'eIF4E à eIF4G n'est alors plus possible et par conséquent, la traduction d'ARNm est inhibée. En revanche, lorsqu'eIF4E-BP est phosphorylé son interaction à eIF4E est inhibée, ce dernier est alors libre d'interagir avec eIF4G permettant de lever l'inhibition de la traduction. (Borden *et al.*, 2011; Marcotrigiano *et al.*, 1999; Raught & Gingras, 1999) (Figure 19)

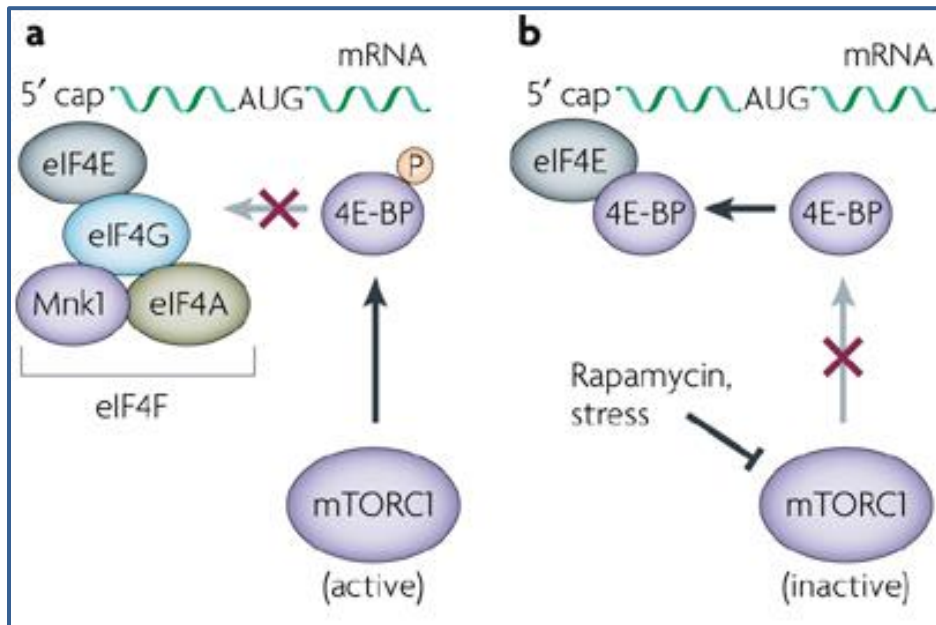


Figure 19 : Régulation du facteur eIF4E par les 4E-BP (d'après Buchkovich et al., 2008)

Le complexe *mTOR* 1 (*mTORC1*) contrôle la liaison de 4E-BP à eIF4E en contrôlant sa phosphorylation. Dans des conditions de croissance positive (a), *mTORC1* est active, elle phosphoryle 4E-BP, qui est par conséquent incapable de se lier à la protéine eIF4E. Ainsi, eIF4E est libre de se lier sur la coiffe en 5' et de lier eIF4G, afin de former avec eIF4A un complexe eIF4F complet, prêt à recruter le PIC pour initier la traduction. Dans des conditions de croissance négatifs(b), *mTORC1* est inactive; elle ne peut donc pas assurer la phosphorylation d'4E-BP, qui se lie alors efficacement à eIF4E, inhibant de ce fait de la traduction coiffe-dépendante.

### Le clivage d'eIF4G et de PABP par des protéases

Lors d'une infection virale, l'intégrité d'eIF4G est compromise par l'action de protéases virales telles que la protéase 2A de certains virus de la famille des picornavirus (poliovirus, rhinovirus, coxsachievirus) (Novoa & Carrasco, 1999; Ventoso *et al.*, 1998), les protéases L et 3C de (FMDV) (Belsham *et al.*, 2000; Devaney *et al.*, 1988; Medina *et al.*, 1993) et la protéase des rétrovirus VIH-1 et VIH-2 (Ventoso *et al.*, 2001). Ces protéases clivent le facteur eIF4G séparant ainsi son extrémité N-terminale liant eIF4E de son extrémité C-terminale liant eIF3, ceci empêche ainsi l'interaction de ce facteur avec la machinerie de traduction ce qui a pour conséquence d'inhiber la traduction coiffe-dépendante, souvent dans le but de privilégier une expression protéique virale par un mécanisme alternatif dit d'entrée interne des ribosomes (Garrey *et al.*, 2010; Ohlmann *et al.*, 2002; Ventoso *et al.*, 2001). eIF4G est également clivée par des caspases cellulaires au cours de l'apoptose (Marissen & Lloyd, 1998; Schneider & Mohr, 2003).



Comme pour eIF4G, la PABP est une cible des protéases virales qui clivent la PABP au niveau de son domaine C-terminal ce qui conduit à l'inhibition de la traduction (Alvarez *et al.*, 2006; Kuyumcu-Martinez *et al.*, 2002, 2004).

Les régulations présentées ci-dessus ne représentent que quelques exemples des mécanismes mis en place par la cellule pour maintenir l'homéostasie cellulaire et freiner la réplication virale. Cependant, comme décrit dans la suite de l'introduction, les virus ont développé différentes stratégies pour contourner ces obstacles. Parmi les stratégies adoptées par les virus, l'initiation de la traduction par entrée interne des ribosomes est largement utilisée.

### C. Initiation de la traduction par entrée interne des ribosomes

#### **Découverte des IRES.**

En 1988, de nombreuses études sur la traduction de picornavirus ont permis de mettre en évidence un mécanisme alternatif de traduction. Les ARNm de picornavirus sont incompatibles avec une traduction canonique coiffe-dépendante. En effet, ils sont dépourvus de coiffe (Frisby *et al.*, 1976; Nomoto *et al.*, 1976), et possèdent des régions 5'UTR longues, fortement structurées qui contiennent plusieurs codons AUG en amont du codon initiateur. Cela suggère que la traduction de ces ARNs se fait par un mécanisme qui ne nécessite ni la présence de la coiffe ni d'étape de balayage de la région 5'UTR. Diverses expériences ont permis de conclure que l'initiation de la traduction chez les picornavirus se fait par l'intermédiaire d'une séquence IRES (Internal Ribosome Entry Site) présente au niveau de la région 5'UTR (Jang *et al.*, 1988; Pelletier & Sonenberg, 1988) et qui est capable de recruter directement le complexe d'initiation (à l'exception d'eIF4E) à proximité du codon initiateur. Ce mécanisme alternatif est connu sous le nom d'**entrée interne des ribosomes** (Figure 20). Depuis, plusieurs autres IRES virales et cellulaires ont été découvertes.

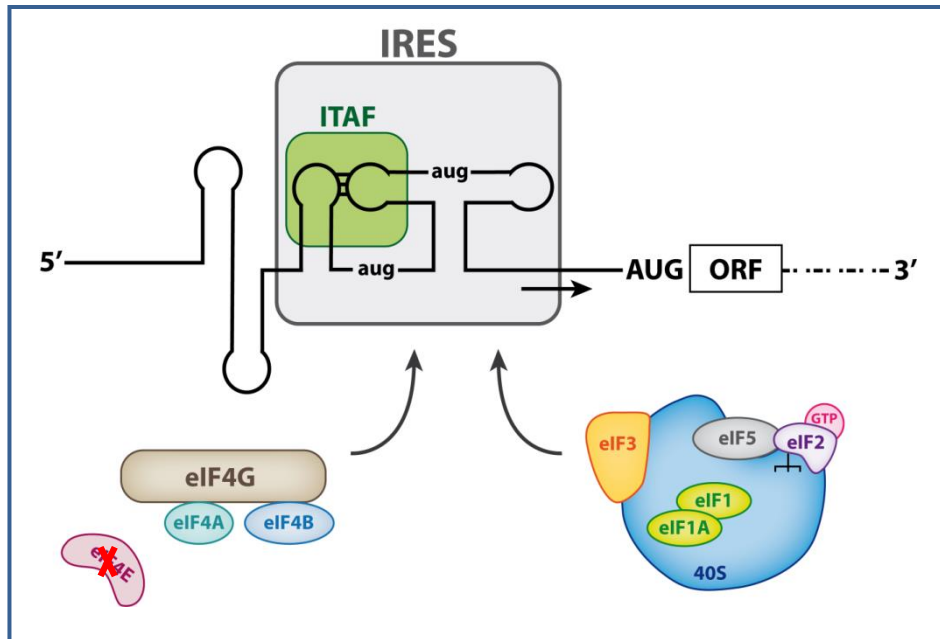


Figure 20 : Schéma de l'initiation par entrée interne des ribosomes

Les différents constituants du complexe d'initiation sont schématisés. La séquence IRES est encadrée. Les ITAFs sont représentés en vert. L'ORF (Open reading Frame) est encadrée. Les deux flèches indiquent que le complexe d'initiation, à l'exception d'eIF4E (barré d'une croix rouge) est recruté par la séquence IRES à proximité de l'AUG initiateur représenté sur le schéma en majuscule.

### Mise en évidence des IRES

Expérimentalement la mise en évidence d'une activité IRES, se fait principalement en utilisant une construction d'ARN bicistronique. Cette construction consiste en deux gènes rapporteurs séparés par la séquence d'intérêt à tester pour une activité IRES potentielle. La traduction du premier rapporteur est assurée par le mécanisme coiffe-dépendant, tandis que la traduction du second rapporteur n'a lieu que si la séquence insérée comporte bien un site d'entrée interne des ribosomes qui permettrait donc de recruter le complexe d'initiation de la traduction. Les séquences IRES candidates peuvent également être testées en utilisant des ARNs monocistroniques et en vérifiant l'indépendance de leur traduction vis-à-vis de la coiffe. Pour cela, l'efficacité de traduction de l'ARNm contenant la séquence IRES potentielle est testée en présence de compétiteur de la coiffe ou encore en présence de protéases virales. L'ARN monocistronique peut donc permettre de valider une séquence IRES mais également de comparer la contribution de la traduction coiffe-dépendante à celle IRES dépendante et cela est d'autant plus intéressant dans le cas d'ARNm passant d'un mécanisme à un autre en fonction des conditions cellulaires et environnementales ou encore chez les ARNm pour lesquels ces deux mécanismes coexistent (Balvay *et al.*, 2009) (Figure 21).

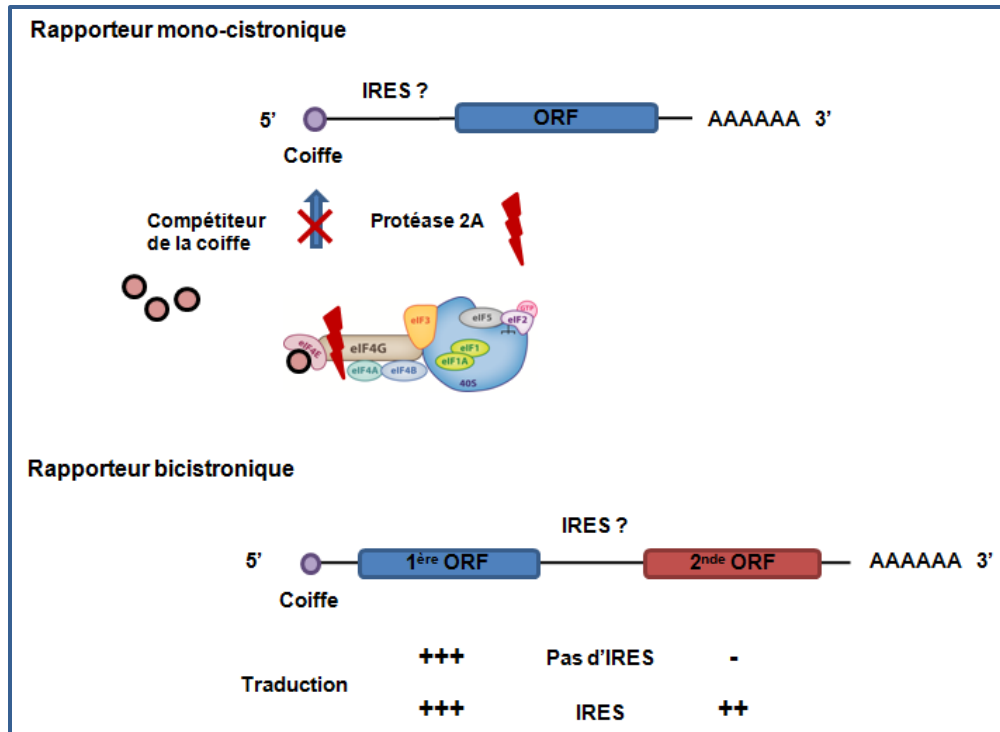


Figure 21 : Mise en évidence expérimentale d'une activité IRES

Les deux extrémités 5' et 3' de l'ARN sont indiquées ainsi qu'une (ARN monocistronique) ou deux ORF (ARN bicistronique). La coiffe est représentée par un rond violet à l'extrémité 5' et la queue poly A en 3'. Les compétiteurs de la coiffe (ronds roses) se lient à eIF4E et empêchent sa liaison à la coiffe. La protéase virale (en rouge) clive eIF4G en deux domaines : un domaine qui lie eIF4E et un domaine qui lie eIF3, ce qui inhibe la traduction dépendante de la coiffe. La traduction (indiquée par un ++ ) de la 2ème ORF de l'ARN bicistronique est dépendante de la présence d'une séquence IRES entre les deux.

### Les IRES viraux et leur classification

Les IRES viraux sont pour la plupart localisés dans la région 5'UTR et adoptent des structures tridimensionnelles stables essentielles au recrutement d'un ou de plusieurs constituants de la machinerie de traduction. Ainsi, lors de conditions défavorables au mécanisme canonique, les IRES viraux permettent d'assurer la traduction par le mécanisme alternatif d'entrée interne des ribosomes en détournant la machinerie de traduction cellulaire au profit de la traduction virale. C'est le cas notamment avec des virus dépourvus de coiffe ou ceux possédant au niveau de leur 5'UTR des structures tridimensionnelles stables incompatibles avec l'étape de balayage.

De nombreuses IRES ont été identifiées mais aucune séquence ou déterminant structural commun n'ont été retrouvés, rendant difficile leur classification suivant ces

critères. Toutefois, les IRES ont été classées en quatre types (Figure 22), qui diffèrent par leur dépendance vis-à-vis des facteurs d'initiation canoniques et des protéines cellulaires auxiliaires appelées ITAFs (IRES *Trans-Acting Factors*).

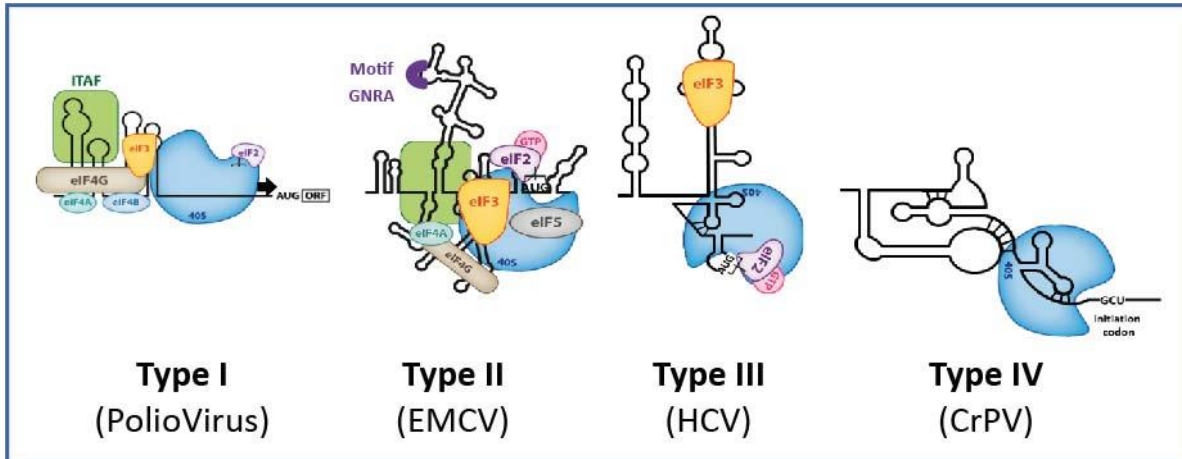


Figure 22 : Les différentes classes d'IRES

Représentation schématique des quatre classes d'IRES. La structure secondaire d'un exemple de virus par classe est représentée. Les différents facteurs nécessaires à l'activité IRES sont schématisés. Les IRES de type I et II nécessitent tous les facteurs ainsi que des ITAFs. Les IRES de type I, contrairement au type II effectuent une étape de balayage indiquée par la flèche noire. Le motif GNRA est indiqué en violet. Les IRES de type III nécessitent moins de facteurs (40S, eIF3 et le complexe ternaire) et les IRES de type IV n'ont besoin que la sous-unité ribosomale 40S.

### Les classes I et II

Les IRES de classe I et II, ont majoritairement été décrites chez les picornavirus comme par exemple le virus de la poliomyélite (PV) ou l'aphtovirus FMDV (Foot-and-Mouth Disease Virus). Ces deux classes d'IRES nécessitent la présence de tous les facteurs d'initiation à l'exception d'eIF4E et nécessitent dans la plupart des cas la présence d'ITAFs. Pour initier la traduction virale, ces IRES lient directement la partie C-terminale d'eIF4G, qui va ensuite *via* son interaction avec le facteur eIF3, permettre le recrutement du complexe d'initiation (Figure 23).

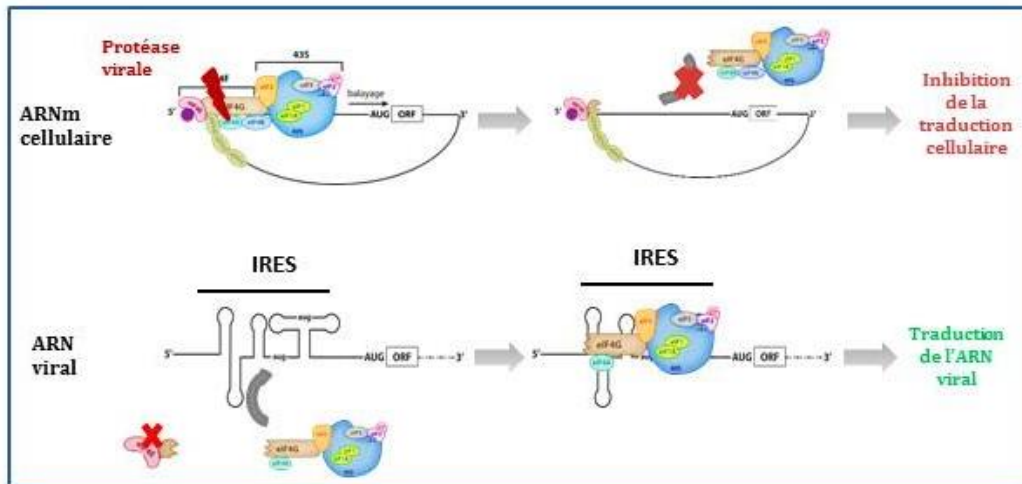


Figure 23 : Mécanisme d'initiation de la traduction de l'ARNm cellulaire vs ARN viral lors de l'infection par le poliovirus

*La traduction de l'ARNm cellulaire est inhibée suite à l'action de la protéase virale (en rouge) qui clive eIF4G (en marron) en deux, le complexe d'initiation ne peut plus être recruté. L'ARN viral est traduit indépendamment d'eIF4E et de la partie N-terminale d'eIF4G, par un mécanisme IRES.*

Pour le cardiovirus EMCV (EncephaloMyoCarditis Virus), une récente étude menée au laboratoire a permis de mettre en évidence une liaison directe entre la sous-unité ribosomale 40S et l'IRES d'EMCV plus précisément au niveau des domaines H-I de cette IRES. Ce qui a permis de proposer un nouveau modèle d'assemblage du complexe de préinitiation sur cet IRES de type II (Figure 24) et d'affiner le rôle attribué au facteur eIF4G proposé maintenant comme un facteur qui promeut et stabilise le changement de conformation nécessaire à l'activité IRES (Chamond *et al.*, 2014).

Ces deux classes d'IRES partagent certaines caractéristiques communes, comme la tétra-boucle GNRA (où R représente une purine et N n'importe quel nucléotide) essentielle à l'organisation structurale de l'IRES (représentée en violet, Figure 21). Elles possèdent également une courte région conservée riche en pyrimidines située 20 à 25 nucléotides en amont du triplet AUG (Balvay *et al.*, 2009; Doudna, 2007). Chez les IRES de classe I comme pour le virus de la poliomyélite, l'AUG en aval de la séquence riche en pyrimidines n'est pas utilisé comme codon initiateur, le complexe d'initiation est recruté à ce niveau-là mais doit ensuite procéder au balayage sur environ 100 à 150 nucléotides pour initier la traduction au prochain AUG rencontré (Jang, 2006). Alors que chez les IRES de classe II, comme c'est le cas du virus EMCV, le complexe d'initiation est recruté directement sur ou à proximité immédiate de cet AUG et ne nécessite donc pas d'étape de balayage pour initier sa traduction (Kaminski *et al.*, 1990) (Figure 21).

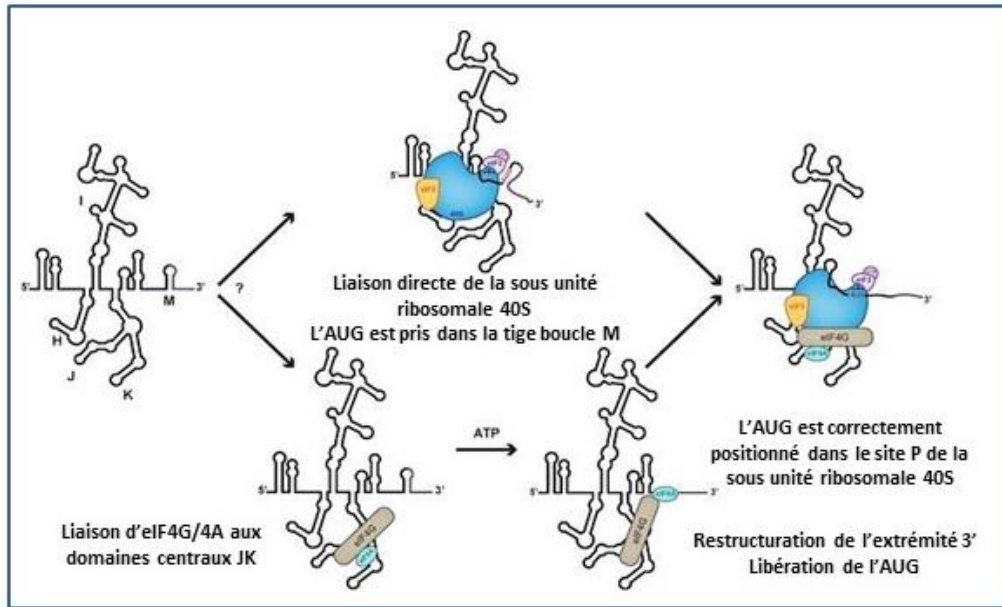


Figure 24 : Modèle du mécanisme d'initiation de la traduction par l'IRES du virus EMCV (d'après Chamond et al., 2014)

La structure secondaire de l'IRES d'EMCV est représentée, les différents domaines sont indiqués par des lettres. Le modèle du mécanisme d'initiation de la traduction par l'IRES d'EMCV propose que l'étape précoce d'initiation soit le recrutement de la sous-unité ribosomale 40S par l'IRES puis le recrutement d'eIF4G/eIF4A qu'il soit simultané ou non permet de remodeler la structure de l'IRES en 3' afin de libérer l'AUG pris dans la tige-boucle M et de l'accommoder au site P du ribosome afin de permettre l'initiation.

### La classe III

Cette troisième classe est représentée par les IRES de *Flaviviridae*, dont l'archétype est l'IRES du VHC (Virus de l'Hépatite C humaine). Cette classe d'IRES est capable de lier directement la sous-unité ribosomale 40S et le facteur eIF3, puis les autres facteurs d'initiation, à l'exception des facteurs du complexe eIF4F, sont recrutés pour initier la traduction.

Située au niveau de la 5'UTR de l'ARN viral, l'IRES du VHC est longue d'environ 330 nucléotides et adopte des structures secondaires et tertiaires qui remplissent différentes fonctions lors du processus d'initiation. Ainsi, le domaine II de l'IRES est impliqué dans l'hydrolyse du GTP d'eIF2 et dans l'assemblage de la sous-unité 60S du ribosome. Le domaine III, lui, se lie à la sous-unité ribosomale 40S et eIF3, tandis que le domaine IV abrite le codon d'initiation AUG (Kieft et al., 2001; Lozano et al., 2016) (Figure 25).



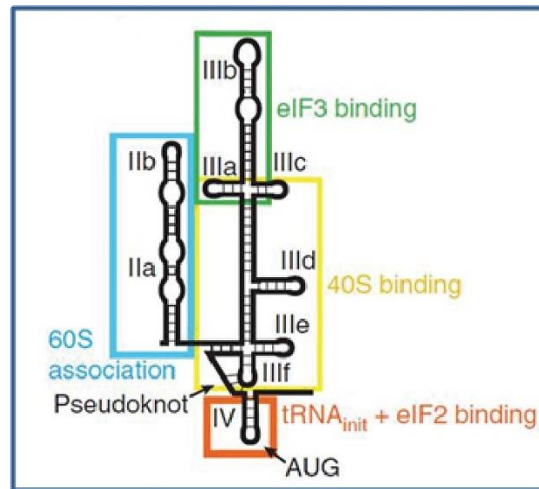


Figure 25 : Modèle de la structure secondaire de la 5'UTR du VHC (d'après Filbin et al., 2013)

La structure secondaire de l'IRES du VHC est représentée. Les différents domaines et l'AUG initiateur sont indiqués. Les sites de liaisons aux facteurs d'initiations sont encadrés : en vert le site de liaison à eIF3, en jaune le site de liaison à la sous-unité ribosomale 40S, en bleu le site de liaison à la sous-unité ribosomale 60S et en orange le site de liaison au complexe ternaire.

Des expériences de cross-linking ont permis de mettre en évidence de nombreuses interactions entre l'IRES du VHC et les protéines ribosomales telles que les protéines RPS2, RPS3, RPS5, RPS10, RPS15, RPS16, RPS18, RPS27 (Laletina *et al.*, 2006; Otto *et al.*, 2002; Pestova *et al.*, 1998). L'interaction du domaine II de l'IRES du VHC avec la sous-unité ribosomale 40S, entraîne une rotation de la tête par rapport au corps qui se traduit par l'ouverture du canal d'entrée de l'ARNm au niveau de la sous-unité ribosomale 40S et conduit au positionnement du codon initiateur au niveau du site P sans passer par une étape de balayage. Ce changement conformationnel n'a pas lieu si l'IRES est dépourvue du domaine II validant ainsi l'importance du domaine II dans le changement de conformation du ribosome (Boehringer *et al.*, 2005; Hellen, 2009; Spahn *et al.*, 2001). De plus, la boucle apicale du domaine II (boucle IIb) est impliquée dans la liaison de la protéine ribosomique RPS5, stabilisant ainsi le ribosome dans une conformation propice pour la translocation (Filbin *et al.*, 2013). Ce réarrangement a également été présenté comme étant à l'origine de l'augmentation du taux d'hydrolyse du GTP lié à eIF2 mais aussi à la faible affinité de liaison d'eIF2-GDP à la sous-unité 40S permettant le relargage d'eIF2 du complexe d'initiation (Locker *et al.*, 2007).

Le domaine III de l'IRES du VHC, qui représente le domaine principal, est subdivisé en sous-domaines nommés IIIa à IIIf et contient également un pseudo-nœud (Figure 25). Ce domaine est impliqué dans le recrutement direct et avec une forte affinité de la sous-unité ribosomale 40S ( $k_d=1,9nM$ ) et du facteur d'initiation eIF3 ( $k_d=35nM$ ) (Kieft *et al.*,



2001). En effet, la partie apicale du domaine III comprenant les sous-domaines IIIa,b,c et la jonction entre ces trois sous-domaines (jIIIabc), est impliquée dans la liaison à eIF3. Des mutations au niveau du domaine IIIa impactent d'un facteur 6 la liaison à eIF3. Les déterminants structuraux de liaison à la sous-unité ribosomale 40S ont été cartographiés *in vitro* par des expériences de «footprinting » et des expériences de mutagenèse dirigée et montrent que les sous-domaines IIIId et IIIe ainsi que le pseudo-nœud sont impliqués dans le recrutement de la sous-unité ribosomale 40S (Angulo *et al.*, 2016). De façon surprenante, les mutations affectant la formation du pseudo-nœud n'altèrent que légèrement la liaison au ribosome mais diminuent fortement l'efficacité de traduction, suggérant que le pseudo-nœud n'est pas impliqué directement dans le recrutement du ribosome mais est plutôt, du fait de sa proximité avec le domaine IV portant l'AUG, impliqué dans le positionnement correct du ribosome (Berry *et al.*, 2010).

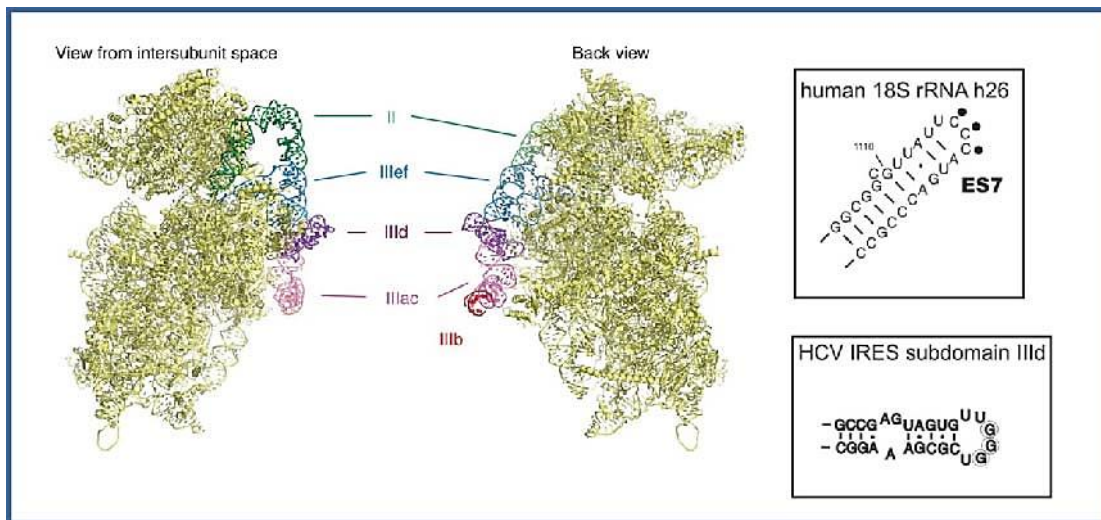


Figure 26 : Structure de l'IRES du VHC en interaction avec la sous-unité ribosomale 40S

A gauche : structure de l'IRES du VHC liée à la sous-unité ribosomale 40S (en jaune). Les différents domaines impliqués dans la liaison au 40S sont indiqués. (d'après Quade *et al.*, 2015). A droite : représentation de la structure secondaire de la tige-boucle de l'hélice h26 de l'ARNr18S et du domaine IIIId l'IRES du VHC. Les nucléotides qui pourraient s'apparier sont notés par un point (18S, en haut) ou encadrés en gris (IRES HCV, en bas) (d'après Malygin *et al.*, 2013)

Des expériences de « footprinting » et de Cryo-microscopie électronique montrent une large surface d'interaction entre la sous-unité ribosomale 40S et l'IRES du VHC. Cette dernière se lie sous une conformation étendue à la surface exposée au solvant de la sous-unité 40S. Le domaine II interagit avec la tête de la sous-unité ribosomale 40S tandis que le domaine III se lie à la plate-forme et au corps de la sous-unité ribosomale 40S. Plus précisément trois résidus guanines conservés au niveau du sous-domaine IIIId de l'IRES du VHC sont impliqués dans une interaction Watson Crick avec trois résidus

cytosines de l'hélice h26 de l'ARNr 18S du ribosome (Angulo *et al.*, 2016; Malygin *et al.*, 2013; Matsuda & Mauro, 2014; Quade *et al.*, 2015) (Figure 26).

D'un point de vue structural, une étude par Cryo-microscopie électronique de la liaison d'eIF3 à un IRES de classe III (l'IRES du CSFV pour Classical Swine Fever Virus), montre qu'eIF3 occupe une place différente de celle qu'il occupe d'habitude au sein du complexe 43S. En effet, la surface lui permettant d'interagir avec le ribosome est retrouvée en interaction avec la boucle apicale du domaine III de l'IRES virale. Ce qui suggère que l'interaction entre l'IRES et eIF3 ne soit pas impliquée dans le mécanisme d'initiation en soi mais plutôt dans la prévention de l'association du ribosome à eIF3 afin d'inhiber la traduction cellulaire canonique au profit de la traduction virale IRES dépendante. En quelque sorte, cela permet de titrer eIF3 et ainsi réduire la compétition entre l'IRES virale et eIF3 pour la liaison au site commun sur la sous-unité ribosomale 40S (Hashem *et al.*, 2013) (Figure 27).

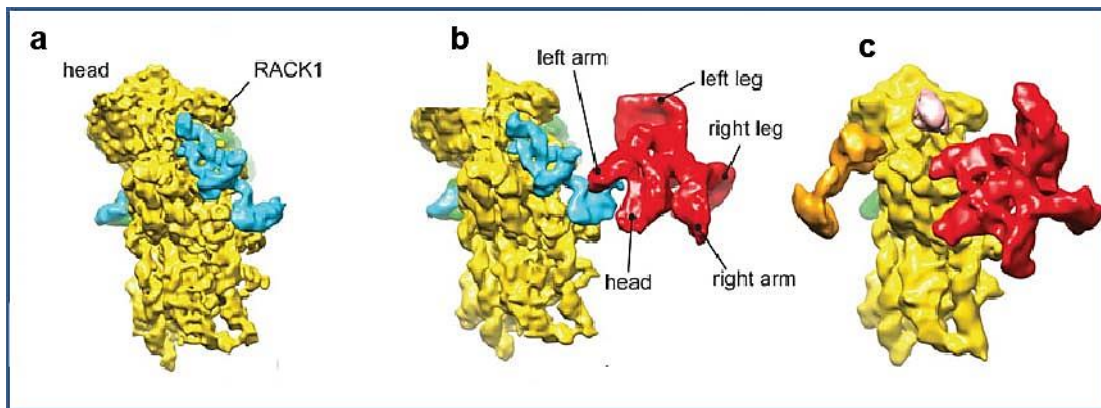


Figure 27 : L'interaction d'eIF3 avec la sous-unité ribosomale 40S seule ou liée à l'IRES du CSFV (d'après Hashem *et al.*, 2013)

*Structures en Cryo-EM, vue de dos, de (a) : liaison de la sous-unité ribosomale 40S (en jaune) à l'IRES-ΔII du CSFV (en cyan) et à DHX29 (en vert). (b) : +eIF3 (en rouge). (c) : le complexe de préinitiation (en jaune le 40S, en rouge le core domaine d'eIF3 et en rose les sous-unités périphériques, en orange le complexe ternaire) +DHX29 (en vert)*

L'initiation de la traduction du VHC commence donc par la formation d'un complexe stable entre le ribosome et l'IRES du VHC, qui se lie ensuite au complexe ternaire (eIF2-GTP-ARN<sub>i</sub><sup>Met</sup>) formant un complexe de préinitiation au niveau du codon initiateur, puis l'hydrolyse du GTP associé à eIF2 catalysée par eIF5 va permettre à la sous-unité ribosomale 60S de se lier et de former un ribosome 80S compétent pour l'étape d'élongation (Fraser & Doudna, 2007). Cependant, de façon intéressante, lorsqu'eIF2 est phosphorylée et n'est donc pas disponible pour initier la traduction, l'IRES du VHC est capable de promouvoir une initiation indépendante d'eIF2 et d'eIF5 et de délivrer

l'ARNt<sub>i</sub><sup>Met</sup> en utilisant des mécanismes alternatifs faisant intervenir eIF5B (Pestova *et al.*, 2008; Terenin *et al.*, 2008), eIF2D (Dmitriev *et al.*, 2010; Skabkin *et al.*, 2010) ou encore eIF2A (Kim *et al.*, 2011) offrant au VHC un avantage sélectif pour sa réplication virale.

### La classe IV

Les IRES de type IV se replient en une structure très compacte qui est suffisante pour assurer la liaison à la sous-unité 40S et initier la traduction sans nécessiter aucun facteur d'initiation ni même l'ARNt<sub>i</sub><sup>Met</sup>. Ce qui permet d'assurer la production de protéines même dans des conditions cellulaires défavorables à la traduction comme par exemple lors d'une infection virale.

Cette classe d'IRES est retrouvée dans la famille des *Dicistroviridae*, dont l'exemple le plus connu est celui du virus de la paralysie du cricket (CrPV). Il possède un génome ARN simple brin positif, qui contient deux grandes ORFs (bicistronique naturel) dont la traduction est assurée par un mécanisme d'entrée interne des ribosomes (Wilson *et al.*, 2000a). La première IRES située dans la région 5'UTR, assure l'initiation de la traduction de la première ORF et conduit à la production de protéines non structurales, alors que la deuxième IRES est retrouvée dans la région intergénique (IGR) qui sépare ces deux ORFs et conduit à la production de protéines structurales. Cette deuxième IRES, plus active, fait environ 200 nucléotides et adopte une structure complexe nécessaire et suffisante à la liaison directe du ribosome avec une forte affinité (Kd~30nM) (Jan & Sarnow, 2002). De plus, cette IRES possède la caractéristique remarquable d'initier la traduction de la seconde ORF sur un codon non pas méthionine mais alanine (GCU, GCA, GCG ou GCC) (Jan *et al.*, 2001; Sasaki & Nakashima, 1999), et au niveau du site A du ribosome et non au site P comme c'est le cas pour tous les autres mécanismes d'initiation (Wilson *et al.*, 2000b) (Figure 28-b).

La structure de l'IRES IGR résolue par cristallographie et par CryoEM, montre qu'elle est organisée en trois domaines qui abritent trois pseudo-nœuds (PK I, II et III) ainsi que deux tiges-boucles conservées (SL IV et V) (Figure 28-a), formant une structure tridimensionnelle compacte qui permet de recruter le ribosome et d'initier la traduction en absence de tout autre facteur (Pfingsten *et al.*, 2006; Schüler *et al.*, 2006; Spahn *et al.*, 2004). Les tiges-boucles assurent le recrutement du ribosome en interagissant directement avec les protéines ribosomales RPS5 et RPS25 (Figure 28-d). Les PK II et III se regroupent pour former un grand domaine stable qui recrute le ribosome. Ce grand domaine est rattaché par un linker au PK I, lequel forme le petit domaine 3. Le PK I précède le codon d'initiation non-AUG et mime l'interaction d'un ARNt initiateur à son codon sur l'ARN (Figure 28-c) (Costantino *et al.*, 2008). Ce mimétisme peut être observé dans la structure du complexe ribosome IRES où PKI est positionné dans le site A du ribosome avec la partie restante de l'IRES comprises entre les sites P et E (Figure 28-b) (Fernández *et al.*, 2014).

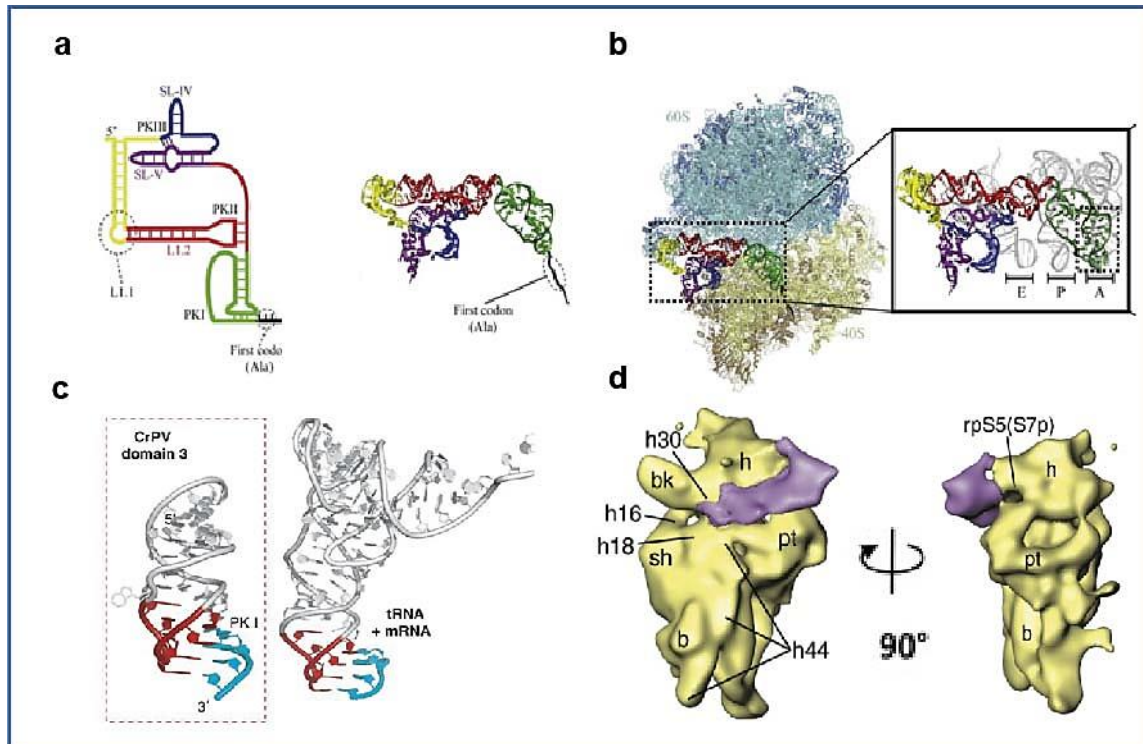


Figure 28 : Structure de l'IRES de la région intergénétique de dicistrovirus seule ou en interaction avec la sous-unité ribosomale 40S (d'après Fernández et al., 2014 & Spahn et al., 2004 )

(a) représentation de la structure secondaire et tertiaire de l'IRES de la région intergénétique du CrPV. Les différents pseudo-nœuds (PKI-III) et les différentes tiges-boucles (SL IV-V), ainsi que le codon initiateur alanine (Ala) sont indiqués. Le code couleur est respecté sur la structure tertiaire. (b) liaison de l'IRES du CrPV au ribosome 80S. Le premier encadrement montre la superposition de l'IRES avec les sites A, P et E du ribosome. Le pseudo-noeud PKI de l'IRES (en vert) se trouve dans le centre de décodage (A) normalement occupé par la boucle anticodon d'un ARNt. (c) Mimétisme structural du domaine 3 de l'IRES du CrPV à un ARNt lié à l'ARNm dans le site P du ribosome (anticodon coloré en rouge et l'ARNm en cyan). (d) structures cryo-EM de l'IRES de CrPV (en violet) en interaction avec la sous-unité ribosomale 40S (en jaune). Les annotations sur le 40S correspondent à : b, body; bk, beak; h, head; lf, left foot; rf, right foot; pt, platform; and sh, shoulder. Les positions des 5 hélices (h16, h1, h30, h34, h44) de l'ARNr 18S, ainsi que la protéine rpS5 sont indiquées.

L'initiation de la traduction de l'IRES du CrPV commence d'abord par la formation d'un complexe entre l'IRES et la sous-unité ribosomale 40S. Le PKI qui mime l'interaction intermoléculaire d'un ARNt initiateur avec le codon initiateur sur l'ARNm, est lui retrouvé au niveau du site A du ribosome (Figure 29-A). Suivi de l'association de la grande sous-unité ribosomale 60S pour former un ribosome 80S. Des interactions supplémentaires



sont formées entre l'IRES et la grande sous-unité, notamment avec L1 (*Figure 29-B*). A ce stade PKI est encore au niveau du site A, ceci suggère qu'un facteur permet de déplacer le PKI du site A vers le site P avant que l'aminocyl ARNt soit recruté, c'est le facteur eEF2 qui permet cette première translocation (*Figure 29-C*). Le site A étant libre, le premier ARNt chargé est accommodé au site A par le facteur eEF1A (*Figure 29-D*). Un second événement de translocation a lieu, et le deuxième aminocyl-ARNt est accommodé (*Figure 29-E*). L'étape d'après correspond à la formation de liaison peptidique et au début de l'étape d'élongation.

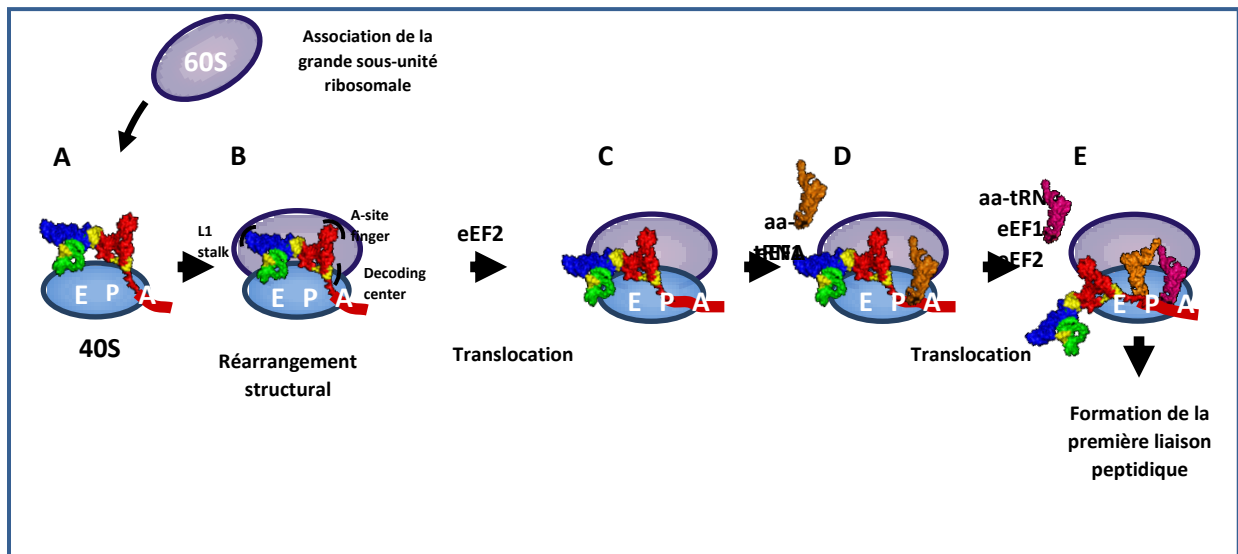


Figure 29 : Schéma du mécanisme d'initiation de la traduction de l'IRES de l'IGR de *Discitroviridae* (adapté de Koh *et al.* 2014)

Les cinq étapes représentent A: la formation du complexe 40S-IRES avec le PKI lié au centre de décodage A du ribosome. Suivi en B, de la liaison de la sous-unité 60S et la formation d'un 80S lié à l'IRES. C: Première translocation catalysée par eEF-2 et D: accommodation du premier aminocyl-ARNt dans le site A. E: Deuxième événement de translocation et accommodation du deuxième aminocyl-ARNt. L'étape d'après correspond à la formation d'une liaison peptidique et au début de la phase d'élongation. La sous-unité 40S est colorée en bleu, la sous-unité 60S en violet. La structure de la région intergénique est colorée de telle sorte à distinguer les trois domaines. Le PKI est représenté en rouge. Les séquences qui interagissent avec la tige L1, le site A et le centre de décodage sur le ribosome sont indiquées. Le premier et le deuxième ARNt sont colorés en orange et rose, respectivement.

## Autres classes

Cette classification a tout de même ses limites puisque de nombreux autres IRES ont été découverts mais n'ont pas pu être rattachés à l'une de ces quatre classes. C'est le cas par exemple de l'IRES du virus à ADN KSHV (Kaposi's sarcoma-associated herpesvirus) dont le mécanisme d'action est assez mal connu. L'une des protéines virales, V-FLIP est produite par un mécanisme IRES (Grundhoff & Ganem, 2001). La première caractérisation moléculaire de cette IRES a révélé certaines caractéristiques inhabituelles. De façon intéressante, l'IRES de KSHV se lie spécifiquement à la sous-unité 40S du ribosome mais l'initiation de la traduction à partir de cette IRES nécessite un complexe eIF4F complet, y compris eIF4E, ce qui est inhabituel puisque les IRES normalement assurent l'initiation de la traduction indépendamment de la coiffe et par conséquent indépendamment d'eIF4E (facteur de reconnaissance de la coiffe) (Othman *et al.*, 2014). Le virus de l'hépatite A (HAV) nécessite lui aussi l'ensemble du complexe eIF4F, plus précisément l'interaction entre eIF4E et eIF4G pour initier sa traduction (Ali *et al.*, 2001; Borman *et al.*, 2001).

Les lentivirus dont les ARNm contiennent une coiffe mais abritent aussi des IRES, peuvent ainsi être traduits par un mécanisme canonique coiffe-dépendant et/ou par un mécanisme d'entrée interne des ribosomes. Par exemple, le VIH-1 possède non seulement un IRES situé dans la 5'UTR de l'ARNm viral (Brasey *et al.*, 2003), mais possède également un IRES atypique situé dans la région codante du gène gag et qui est conservé parmi les lentivirus de primates (Herbreteau *et al.*, 2005; Locker *et al.*, 2011; Weill *et al.*, 2010). Il présente donc aussi bien les propriétés de la classe III (liaison directe et indépendante de la sous-unité ribosomale 40S et d'eIF3 (Locker *et al.*, 2011)) que les propriétés de la classe I et II en exigeant la présence de tous les facteurs d'initiation de la traduction à l'exception d'eIF4E.

## Les IRES cellulaires

C'est en 1991 que la première IRES cellulaire a été découverte dans l'ARNm codant pour la protéine BiP de l'hémoglobine humaine (human immunoglobulin heavy chain-binding protein). En effet, dans les cellules infectées par le PV la traduction coiffe-dépendante est inhibée, néanmoins, la traduction de l'ARNm de BiP est maintenue ceci suggère qu'il est traduit par un mécanisme indépendant de la coiffe (Macejak & Sarnow, 1991). Depuis, on en recense une centaine d'autres. Environ 3% des ARNm cellulaires possèdent une séquence IRES au niveau de leur 5'UTR, ce sont pour la plupart des gènes impliqués dans des processus finement régulés tels que les voies de transduction du signal, la différenciation cellulaire ou encore le cycle cellulaire. Cette voie alternative de traduction IRES-dépendante va permettre à la cellule d'assurer la traduction dans des conditions où le mécanisme canonique dépendant de la coiffe est compromis comme c'est le cas suite à un stress oxydatif, un choc thermique, une carence en

nutriments, une infection virale ou encore lors de la mitose ou de l'apoptose. Cela va permettre de fournir un niveau supplémentaire de régulation de l'expression des gènes cellulaires lorsque la traduction canonique coiffe-dépendante se trouve compromise (Orna et al., 2007; Vagner et al., 2001).

Contrairement aux IRES virales, le mécanisme d'action des IRES cellulaires reste encore mal compris, l'existence même d'IRES au sein de certains gènes est controversée du fait notamment de leur faible efficacité de traduction.

### **Liaison du ribosome à des séquences complémentaires de l'ARNr 18S**

Il existe des ARNs capables de lier directement le ribosome non pas par l'intermédiaire de structures qu'adopterait cet ARN mais par des appariements de séquences entre les ARNm et l'ARNr 18S du ribosome, à l'instar de l'interaction retrouvée chez les procaryotes entre l'ARNr 16S de la petite sous-unité ribosomale 30S et une séquence de l'ARNm riche en purine localisée 5 à 8 nucléotides en amont du codon d'initiation appelée séquence de Shine-Dalgarno (5'- AGGAGGUAA-3') (Shine & Dalgarno, 1974).

#### **L'IRES du gène Gtx chez la souris.**

Chez la souris, l'IRES de la 5'UTR de l'ARNm Gtx qui code pour une protéine à homéodomaine, abrite une séquence de 9 nucléotides (5'-CCGGCGGGU-3') complémentaire des nucléotides 1124 à 1132 de l'ARNr 18S du ribosome. Cette complémentarité de séquence joue un rôle déterminant dans l'activité IRES. En effet, elle est capable de stimuler l'efficacité de traduction *in cellulo* d'un rapporteur placé en aval et elle est d'autant plus forte que le nombre d'insertion de cette séquence augmente, pouvant aller jusqu'à une stimulation de 570 fois lors de l'insertion de 10 copies de cette séquence de 9 nucléotides (Chappell et al., 2000; Hu et al., 1999). Chez la levure cette séquence est incapable de stimuler la traduction, puisque la complémentarité de séquence entre les 9 nucléotides de l'IRES et l'ARNr 18S de levure n'est pas conservée. L'activité IRES de cette séquence peut être retrouvée chez la levure si des mutations ponctuelles permettant de restaurer la complémentarité de séquence sont introduites, la séquence de 9 nucléotides retrouve son effet stimulateur (Deforges et al., 2015). Ces résultats suggèrent un mode d'initiation similaire à celui retrouvé chez les procaryotes.

#### **La 3'UTR du HCRSV**

La 3'UTR du HCRSV (Hibiscus Chlorotic Ringspot Virus) abrite une séquence hexanucléotidique (5'-GGGCAG-3') conservée parmi les carmovirus, complémentaire des nucléotides 1626 à 1631 de l'extrémité 3' de l'ARNr 18S qui est localisée près du



« tunnel » de l'ARNm entre le « bec » et « l'épaule » de la sous-unité ribosomale, et à proximité immédiate du site E. Des expériences de mutations/délétions ont permis de mettre en évidence et de valider l'effet de cette séquence sur la stimulation de l'efficacité de la traduction assurée par la 3'UTR du HCRSV. Cependant, en plus de cette interaction, une longueur minimale de 180 nucléotides est nécessaire pour permettre à la 3'UTR de fonctionner comme un activateur de la traduction (CITEs) (Koh *et al.*, 2002).

### **L'IRES IGF1R**

L'IGF1R (pour Insulin-like Growth Factor I Receptor) est un récepteur hétérotétramérique à activité tyrosine kinase impliqué notamment dans la prolifération cellulaire et dans l'oncogenèse. IGF1R est codé par un ARNm contenant une 5'UTR longue, riche en G-C et très structurée qui contient un site d'entrée interne des ribosomes qui joue un rôle crucial dans le contrôle de la prolifération et la différenciation cellulaire (Giraud *et al.*, 2001). L'IRES de l'IGF1R correspond à une séquence de 90 nucléotides présente immédiatement en amont du codon initiateur et présente une complémentarité de séquence avec l'ARNr 18S humain présent prêt du site E du ribosome. Plus précisément, c'est la tige-boucle 2 de l'IRES de l'IGF1R qui présente une séquence de 14 nucléotides (5'-GGGAAUUUCAUC-3') complémentaire de la séquence 5'-958GGUGAAAUCCU969- 3' de l'ARNr 18S qui se situe au niveau de la boucle G961 de l'hélice 23b. Toute mutation impactant cette complémentarité de séquence engendre une diminution drastique de l'activité IRES suggérant que l'IRES de l'IGF1R pourrait recruter le ribosome d'une manière équivalente à celle retrouvée chez les procaryotes (séquence Shine-Dalgarno : 5' -AGGAGGUAA-3')(Meng *et al.*, 2010) (Figure 30)

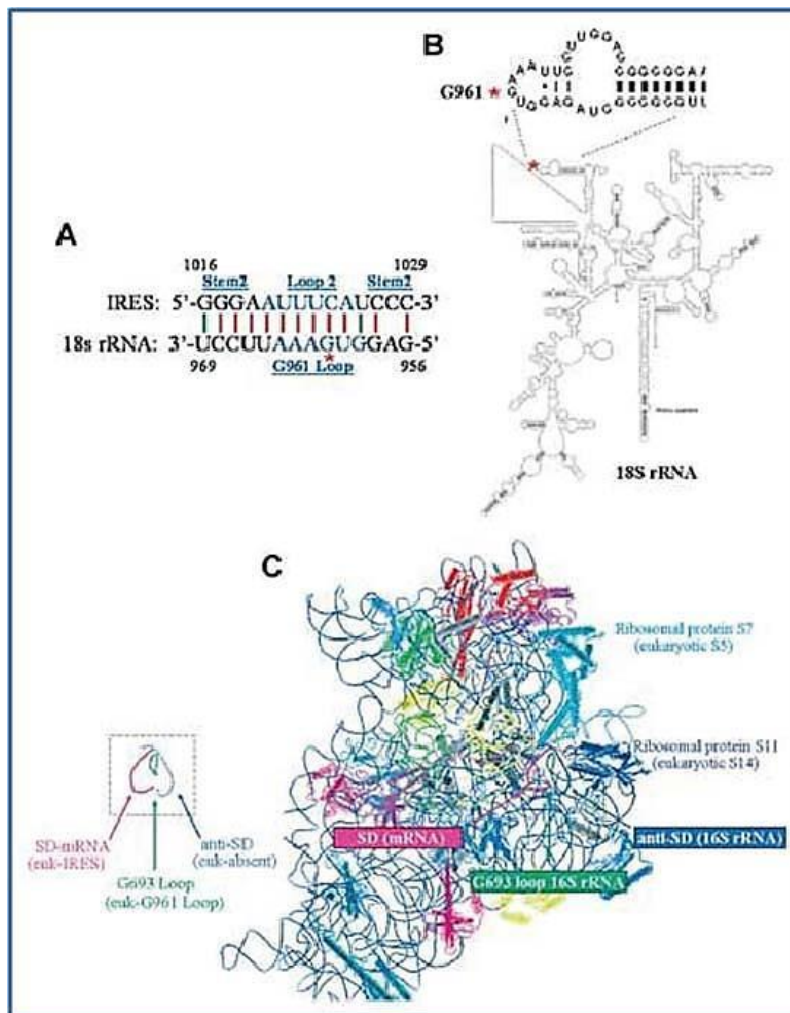


Figure 30 : Appariement de séquences entre l'IRES IGF1R et l'ARNr 18S (d'après Meng *et al.*, 2010)

A) Schéma de l'appariement possible entre la 5' UTR de l'IGF1R et l'ARNr 18S humain. B) La position de la tige-boucle G961 dans le schéma de la structure secondaire de l'ARNr 18S est indiquée par un astérisque rouge. C) Structure cristallographique de la sous-unité ribosomale 30S de *T. thermophilus* indiquant en rose, la position de la séquence de Shine-Dalgarno d'un ARNm. La boucle de G693 de l'ARNr 16S (homologue à la G961 boucle de 18S ARNr eucaryotes), est mise en évidence en vert. La séquence « anti-SD » de l'ARN 16S qui n'est pas conservée chez les eucaryotes est représentée en bleu. A gauche, l'encadré représente un schéma simplifié de ces trois éléments

### Des ITAFs régulent l'activité des IRES

En plus des facteurs d'initiation canoniques (eIFs), les IRES requièrent parfois d'autres facteurs non canoniques : les ITAFs, qui agissent en trans des IRES et facilitent le recrutement du complexe d'initiation. Diverses approches expérimentales, comme l'effet de la spécificité cellulaire sur la traduction de messagers viraux, ont permis de mettre en

évidence des ITAFs. Par exemple, lors de l'étude de la traduction du poliovirus, ce dernier apparaît faiblement traduit en cellules BHK (Baby Hamster Kidneys) et en lysat de réticulocytes de lapin contrairement aux cellules HeLa où sa traduction est nettement plus active (Roberts *et al.*, 1998). De même, l'IRES du gène cellulaire c-myc est traduit vingt fois plus efficacement en cellules HeLa qu'en cellules MCF (Michigan Cancer Foundation, lignée de cellules tumorales mammaires) (Stoneley *et al.*, 2000). Une étude protéomique a permis l'identification de nombreuses protéines interagissant avec l'IRES située en 5'UTR du VIH-1 et qui agiraient comme ITAFs stimulant la traduction virale dépendante de cet IRES (Vallejos *et al.*, 2011).

Un grand nombre d'ITAFs, régulant aussi bien l'activité d'IRES virales que cellulaires ont été identifiés. Parmi ces ITAFs, on retrouve : les protéines de la famille des hnRNPs comme par exemple hnRNP A qui favorise la traduction des IRES *fgf2* (Bonnal *et al.*, 2005), mais inhibe celle de l'IRES d'*apaf1* (Cammass *et al.*, 2007). La protéine PCBP (Poly(C)Binding Protein) qui se lie à l'IRES du poliovirus et à l'IRES cellulaire Bag-1 (Pickering *et al.*, 2004). La protéine « La » qui stimule l'activité de l'IRES du virus de l'hépatite C (Costa-Mattioli *et al.*, 2004), mais régule négativement la traduction médiée par l'IRES du virus de l'hépatite A (Cordes *et al.*, 2008). Ou encore PTB, elle est l'une des mieux connues et des plus étudiées, elle est considérée comme un ITAF « universel » car elle est impliquée dans la majorité des mécanismes de traduction IRES dépendant qu'ils soient cellulaires ou viraux (Bushell *et al.*, 2006). Il s'agit d'une protéine de liaison à l'ARN au niveau de séquences riches en pyrimidines. Elle est impliquée dans de nombreux processus cellulaires comme l'épissage ou la stabilité des ARNm, et est proposée comme une protéine induisant un réarrangement structural tridimensionnel des IRES lorsqu'elle s'y fixe, ceci permettant de stimuler le recrutement du complexe d'initiation comme c'est le cas pour l'IRES Bag-1 (Pickering *et al.*, 2004). Plusieurs ITAFs peuvent être recrutés au sein d'une même séquence IRES et réguler de manière synergique la traduction : c'est le cas de l'IRES du gène c-myc qui fixe les ITAFs YB-1 et PTB, des études *in vitro* montrent que ces protéines, prises individuellement multiplient par deux l'activité de IRES c-myc alors qu'ensemble, elles augmentent cette activité d'un facteur 6 (Cobbald *et al.*, 2010).

Les ITAFs semblent donc être impliqués dans la régulation de l'activité des IRES, vraisemblablement en assurant le bon repliement des ARNs à la manière de chaperonnes, ce qui permet d'assurer en synergie avec les facteurs canoniques, le recrutement du complexe d'initiation.

### **Le mécanisme IRES un mécanisme ancestral ?**

Depuis la mise en évidence du mécanisme IRES en 1988 chez les picornavirus, de nombreux autres IRES ont été découvertes et continuent d'être découvertes aujourd'hui. C'est notamment le cas d'une récente étude qui a mené un criblage à haut débit dans le

but d'identifier et de caractériser de nouvelles IRES et qui a permis de mettre en évidence plus de 1000 séquences IRES et ainsi d'enrichir de 50 fois le nombre de séquences IRES connues à ce jour. Ces séquences IRES sont aussi bien cellulaires que virales, et sont retrouvées dans les régions non traduites 5' et 3' UTR, mais également dans des régions codantes notamment chez les virus. Elles agissent par l'intermédiaire d'éléments fonctionnels qui peuvent être aussi bien structuraux permettant ainsi le recrutement du ribosome directement ou par l'intermédiaire de facteurs d'initiation et/ou de protéines ITAFs, mais, peuvent également être des motifs spécifiques qui recrutent le ribosome *via* des appariements de type Watson-Crick avec l'ARNr 18S (Weingarten-Gabbay *et al.*, 2016).

Malgré des différences notables dans leurs séquences et leurs structures secondaires et tertiaires les rendant difficilement comparables, les IRES peuvent tout de même être définies comme des entités fonctionnelles capables de recruter la machinerie de traduction d'une façon coiffe-indépendante permettant ainsi aux virus comme à certains gènes cellulaires d'assurer la traduction dans des conditions physiologiques défavorables à la traduction coiffe-dépendante (stress cellulaire ou infection virale). Le mécanisme d'entrée interne des ribosomes offre ainsi une échappatoire à la cellule ou encore aux virus dans des conditions où le mécanisme coiffe-dépendant est compromis, suggérant donc que l'existence du mécanisme IRES serait postérieure à celui du mécanisme coiffe-dépendant. Cependant, la simplicité relative du mécanisme IRES pourrait, au contraire, indiquer une existence ancestrale des IRES, laissant ainsi la porte ouverte au débat (Hernández, 2008).

Plusieurs ARNm très éloignés phylogénétiquement arborent un mimétisme structural et/ou fonctionnel de l'ARNt : c'est le cas par exemple des éléments TSS à l'extrémité 3' de certains virus de plantes qui adoptent une structure comparable à celle des ARNt ou encore chez l'IRES du virus de la paralysie du cricket qui mime fonctionnellement l'ARNt. Ceci a conduit à l'hypothèse de la présence d'une structure en forme d'ARNt conservée pour tous les IRES leur permettant de recruter le ribosome. Cette hypothèse a été testée par des expériences de clivage des IRES par la RNase P dont les substrats naturels sont les précurseurs d'ARNt. Les résultats montrent que cette RNase est capable de cliver des IRES différents et distants phylogénétiquement comme le VHC, CrPV, FMDV, CSFV, ce qui suggère la présence d'un motif commun « tRNA-like » reconnu par la RNase P et conservé à différents degrés au cours de l'évolution (Lyons & Robertson, 2003). Ainsi, d'un point de vue évolutif, les IRES auraient originellement été capables de lier le ribosome uniquement par l'intermédiaire de la structure tridimensionnelle adoptée et certains IRES auraient au cours de l'évolution perdu cette capacité et l'auraient compensée par l'action de facteurs d'initiation et de protéines auxiliaires (ITAFs) qui leur permettent par des réarrangements structuraux d'adopter une conformation active capable de recruter le ribosome ou encore d'assurer le lien

entre l'IRES et le ribosome. Cependant, compte tenu d'une large spécificité de substrat de la RNaseP, cette hypothèse reste à vérifier.

Une meilleure compréhension des mécanismes moléculaires et structuraux de liaison des IRES à la machinerie de traduction en général et au ribosome en particulier permettrait sans doute d'utiliser les IRES comme cibles thérapeutiques et aiderait au développement de nouvelles stratégies contre certains pathogènes.

### III. Traduction de la polyprotéine Gag du VIH-1

#### A. La synthèse des protéines virales

La traduction des polyprotéines Gag et Gag-Pol se fait grâce à un décalage du cadre de lecture de -1 en fin de traduction du précurseur gag. Ce décalage d'un nucléotide instaure un ratio d'environ 1/10 de production du précurseur Gag-Pol par rapport à Gag, et permet à environ 10% des ribosomes de contourner le codon stop de la protéine Gag et d'initier la traduction du précurseur Gag-Pol. Il est rendu possible grâce à la « slippery séquence » et à la tige-boucle FSS (pour Frameshift Stimulatory Signal) retrouvées dans la région codante du gène gag (Charbonneau et al., 2012; Dulude et al., 2002). La production des autres protéines virales est quant à elle assurée à partir d'ARNm viraux épissés. Le processus d'épissage alternatif du génome du VIH-1 fait intervenir 4 sites donneurs d'épissage et 8 sites accepteurs, et permet de produire une quarantaine d'ARNm matures différents (Martin Stoltzfus, 2009; Tazi et al., 2010). Ces différents transcrits du VIH-1 partagent les 270 premiers nucléotides de la 5'UTR allant de l'extrémité 5' de l'ARN jusqu'au premier site donneur d'épissage.

Les ARN pré-messagers viraux partiellement épissés (classe 4kb) donnent lieu aux ARNm codant Vif, Vpr, Tat ainsi que l'ARNm codant à la fois Vpu et Env. vpu et env possèdent un cadre de lecture chevauchant, la traduction de la protéine Env a lieu par un mécanisme de leaky scanning. De plus, vpu est situé dans un site d'initiation à faible efficacité (contexte de Kozak faible) favorisant ainsi un événement de leaky scanning et l'initiation sur l'AUG de env (Schwartz et al., 1990, 1992a). Les ARN pré-messagers viraux totalement épissés (classe 2kb) conduisent eux à la production des protéines Rev et Nef ainsi que Tat. La proportion des différents transcrits du VIH-1 dépend entre autre de la régulation par des protéines SR et hnRNP.

Le mécanisme d'initiation de la traduction des ARNs messagers épissés du VIH-1 est encore relativement mal connu, tandis que l'initiation de la traduction de l'ARNg a été largement étudié et a permis de mettre en évidence le fait que la traduction des polyprotéines Gag et Gag-Pol peut s'effectuer suivant différents mécanismes d'initiation aussi bien dépendant de la coiffe que de séquences IRES. Ces différents mécanismes seront décrits plus en détails dans la suite de l'introduction.

## B. Initiation de la traduction de la polyprotéine Gag du VIH-1 par le mécanisme canonique dépendant de la coiffe

L'ARNg du VIH-1 est produit par l'ARN polymérase II de la cellule hôte, il présente donc les caractéristiques d'un ARNm cellulaire à savoir une coiffe à son extrémité 5' et une queue polyA à l'extrémité 3'. De plus, la 5'UTR du VIH-1 ne présente pas de uORF qui risquerait d'inhiber l'étape de balayage. L'initiation de la traduction par le mécanisme canonique coiffe-dépendant est donc théoriquement possible. Pourtant, la présence de longues structures en tiges-boucles stables au niveau de la 5'UTR de l'ARNg, limite fortement le recrutement du complexe d'initiation au niveau de la coiffe ainsi que l'étape de balayage. Des mécanismes alternatifs et/ou des acteurs cellulaires permettent certainement de réguler la traduction des protéines virales.

### **L'état conformationnel de la 5'UTR impliqué dans la régulation de la traduction virale**

#### **Le kissing complex et l'extended duplex**

La 5'UTR du VIH-1 adopte deux conformations qui permettent de réguler la traduction. Une forme monomère favorable à la traduction de l'ARNg et une forme dimère qui inhibe la traduction. Les mécanismes conduisant au passage de l'une des conformations à l'autre ne sont pas encore complètement connus. C'est la séquence palindromique très conservée (5'-<sub>274</sub>GUGCAC<sub>279</sub>-3') située sur la boucle DIS de l'ARNg du VIH-1 qui permet de former ces dimères. En effet, la boucle conservée DIS peut s'apparier par des interactions de type Watson-Crick avec la même séquence d'une autre molécule d'ARN viral. Cette interaction boucle-boucle forme ce que l'on appelle communément un « kissing-complex » qui est considérée comme l'étape initiale du mécanisme de dimérisation. Ainsi, lorsque cette structure est formée, l'ARNg sera certainement impliqué dans le mécanisme de dimérisation et par conséquent, il ne pourra pas être traduit. (Ennifar *et al.*, 2001; Marquet *et al.*, 1994; Paillart *et al.*, 1996, 2004) (Figure 31).



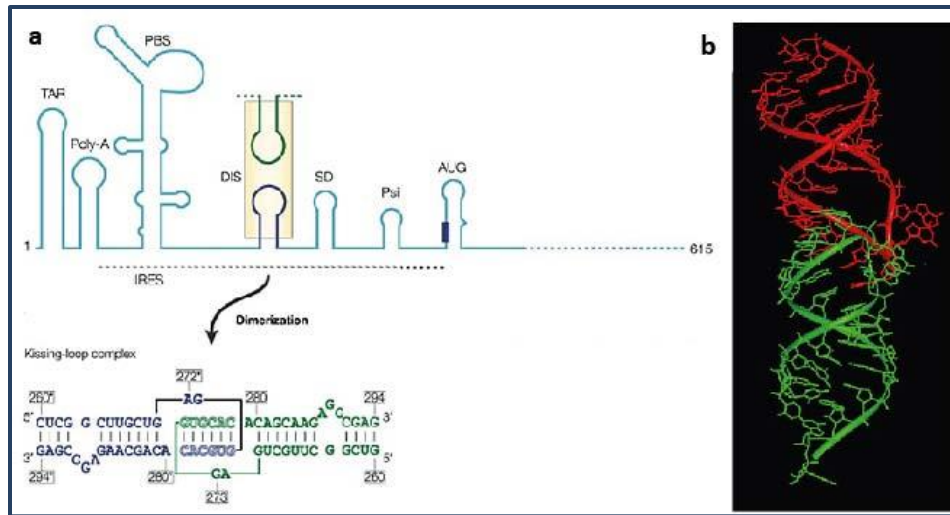


Figure 31 : Formation du « kissing-complex » (d'après Paillart *et al.*, 2004)

(a) Schéma de la structure secondaire des 500 nucléotides à l'extrémité 5' de l'ARN génomique du VIH-1. Les principaux éléments de structure secondaire sont indiqués : TAR: TransActivation Response element, PolyA : PolyA signal, PBS : Primer Binding Site, DS: dimerization initiation site, SD: splicing donor, Psi: encapsidation site. Les deux régions impliquées dans le « kissing-complex » sont mises en évidence par l'encadré jaune. (b) Structure cristallographique du « kissing-complex » formé par deux molécules d'ARN contenant la séquence de la tige-boucle DIS

Le « kissing-complex » peut être converti en duplex plus stable « l'extended duplex », correspondant à un appariement intermoléculaire de deux boucles DIS dans lequel la région 5' de la tige DIS serait appariée avec la partie 3' de l'autre molécule du dimère. C'est la présence de la nucléocapside qui permettrait de favoriser la formation de l'extended duplex, un tel duplex a été mis en évidence par RMN et cristallographie aux rayons X en utilisant une courte séquence d'ARN contenant la tige-boucle DIS uniquement (Ennifar *et al.*, 1999) (Figure 32). Cependant, cette transition en extended duplex ne semble pas avoir lieu dans le contexte d'un ARN synthétique contenant la région 5'-UTR complète (Paillart *et al.*, 1996). Il est donc difficile de savoir si ce phénomène a réellement lieu avec l'ARNg entier.



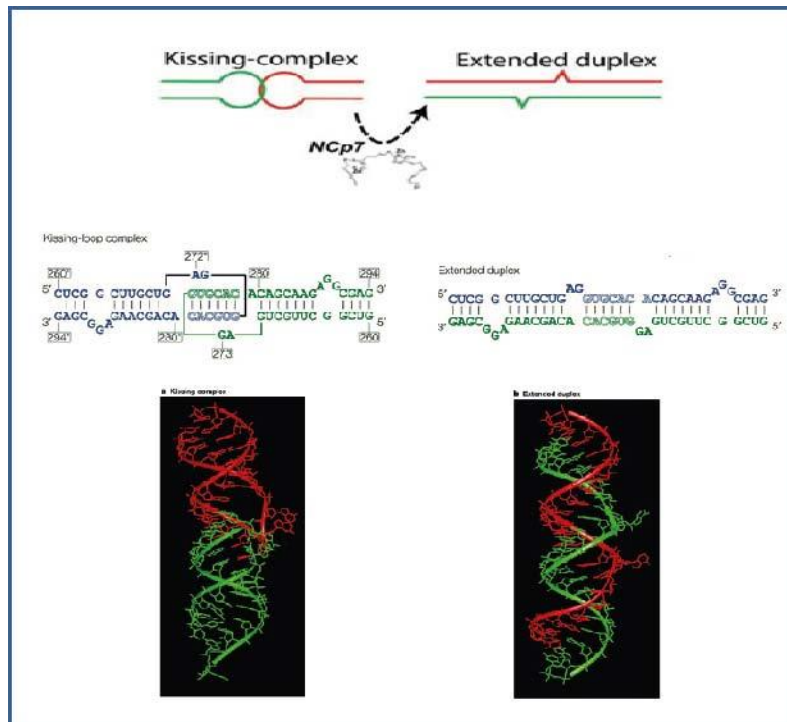


Figure 32 : Transition d'un kissing-complex à un extended complex (d'après Paillart *et al.*, 2004)

De haut en bas : Schéma représentant la conversion du « kissing-complex » vers l'« extended duplex » observée en présence de la nucléocapside virale Ncp7. Séquences nucléotidiques des deux tiges-boucles DIS impliquées dans la formation du « kissing-complex » et de l'« extended duplex ». Structure cristallographique du « kissing-complex » et de l'« extended duplex ».

### Des appariements longues distances U5-AUG et U5-DIS régulent la traduction

Des études menées aussi bien *in vitro* qu'*in vivo*, ont permis de mettre en évidence un autre modèle impliqué dans la régulation entre dimérisation et traduction. Ce modèle permet d'alterner entre deux conformations faisant intervenir des interactions longues distances soit entre la région U5 de la 5'UTR et l'AUG, conduisant ainsi à l'inhibition de la traduction de l'ARNg au profit de sa dimérisation et de son encapsidation. Le deuxième appariement longue distance se fait entre U5 et la boucle DIS. C'est le plus retrouvé dans les cellules où il permettrait, d'inhiber la dimérisation et favoriserait la traduction (Kenyon *et al.*, 2013; Lu *et al.*, 2011) (Figure 33).

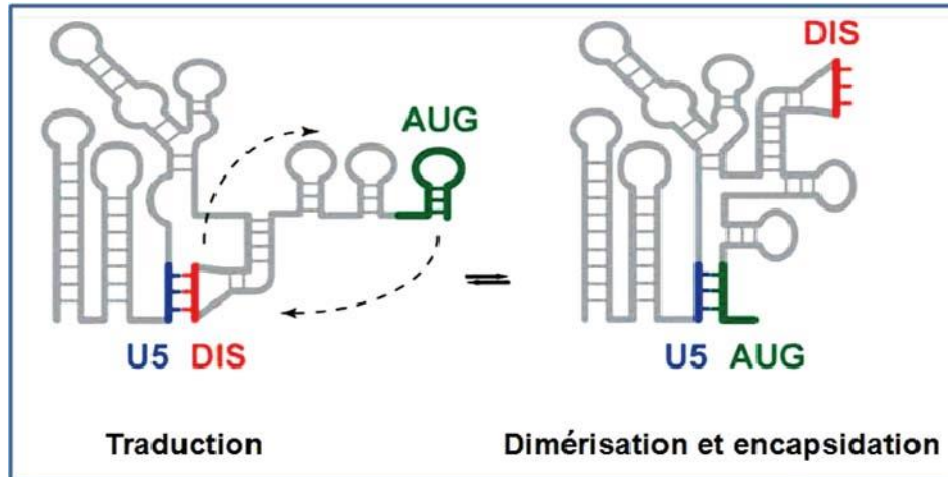


Figure 33 : Les modèles d'appariement longue distance U5-DIS et U5-AUG (d'après Lu *et al.*, 2011)

A gauche : l'appariement entre la région U5 (en bleu) et la région DIS (en rouge) permet de libérer l'AUG (en vert) et de favoriser la traduction. A droite : l'appariement entre l'AUG et U5 conduit à la dimérisation et à l'encapsidation de l'ARNg

### La structure TAR de la 5'UTR du VIH-1 inhibe la traduction coiffe dépendante de la polyprotéine Gag.

L'ensemble des tiges-boucles présentes au niveau de la 5'UTR et particulièrement la première tige-boucle très stable TAR, rend très difficile le recrutement et la progression du ribosome et compromet le mécanisme d'initiation coiffe-dépendant. En effet, la traduction *in vitro* d'un gène rapporteur cloné en aval de la 5'UTR du VIH-1 est fortement inhibée (Soto-Rifo *et al.*, 2012a). Cette inhibition est levée lorsque la tige-boucle TAR est délétée ou lorsqu'une répétition du motif non structuré CAA est insérée à la base de TAR. Ceci suggère que TAR est responsable de l'inhibition observée. Cela peut s'expliquer de deux façons : soit par le fait que sa structure secondaire très stable s'oppose au mécanisme de balayage des ribosomes, soit par le fait que l'inaccessibilité de la coiffe située à la base de TAR, qui n'est alors pas reconnue par le facteur eIF4E empêche la formation du complexe d'initiation. Pourtant, l'effet inhibiteur de la 5'UTR n'est plus observé lorsque, des constructions contenant la 5'UTR du VIH-1 en amont d'un gène rapporteur sont transfectées dans des cellules HeLa ou lorsque des extraits de cellules HeLa sont ajoutés au RRL (Soto-Rifo *et al.*, 2012a). Ces résultats suggèrent que des protéines cellulaires spécifiques sont recrutées pour contourner l'effet inhibiteur de TAR sur la traduction.

## Des hélicases pour lever l'inhibition de la traduction coiffe dépendante ?

Malgré l'inhibition exercée par la 5'UTR, des expériences de traduction menées *in vitro* et *ex-vivo* ont montré que la traduction de l'ARNg du VIH-1 était dépendante de la coiffe. Des expériences de traduction conduites dans du RRL, en présence de la protéase-L du virus FMDV, ont montré une inhibition de la traduction de l'ARNg du VIH-1 (Ricci *et al.*, 2008). Des expériences similaires menées *ex-vivo* en co-exprimant la protéase 2A du poliovirus, ont conduit aux mêmes résultats (Berkhout *et al.*, 2011). De même, l'ajout d'une quantité importante d'analogue de la coiffe qui séquestre eIF4E, a révélé une inhibition de la traduction de la polyprotéine Gag du VIH-1 (Berkhout *et al.*, 2011; Monette *et al.*, 2013; Ricci *et al.*, 2008). Ces résultats indiquent que dans ces conditions, la traduction de l'ARNg du VIH-1 repose sur un mécanisme dépendant de la coiffe.

Afin de résoudre le paradoxe entre l'inhibition de la traduction dépendante de la coiffe par les structures stables de la 5'UTR de l'ARNg du VIH-1 et sa dépendance vis-à-vis d'un mécanisme dépendant de la coiffe, plusieurs études ont proposé que des hélicases pouvaient être recrutées pour déstabiliser les structures secondaires de la 5'UTR et ainsi lever l'inhibition de la traduction. C'est le cas, par exemple de RHA (DHX9) et de DDX3, deux hélicases recrutées par le VIH-1 et qui agissent en coopération avec l'hélicase eIF4A pour améliorer la processivité du complexe 43S lors du balayage de la 5'UTR.

### L'hélicase à ARN A (RHA)

L'hélicase RHA (RNA Helicase A), également appelée DHX9, est une hélicase à ARN ATP-dépendante, impliquée dans divers processus de l'expression génique allant de la transcription à la traduction. Chez le VIH-1, RHA est impliquée dans la traduction virale (Bolinger *et al.*, 2010; de Breyne *et al.*, 2013; Hartman *et al.*, 2006). En effet, elle interagit spécifiquement avec la région U5 de la 5'UTR du VIH-1, plus précisément au niveau de la tige-boucle PBS (Boeras *et al.*, 2016). L'utilisation d'un ARN interférant (siRNA) qui éteint l'expression du gène codant pour RHA dans des cellules infectées, a révélé une diminution importante de la production de la polyprotéine Gag du VIH-1. De façon intéressante, l'ajout exogène de l'hélicase RHA permet de restaurer la traduction de Gag alors que le mutant ATPase de la RHA dans lequel le domaine d'hydrolyse de l'ATP est altéré, ne peut pas restaurer la traduction de Gag. Ceci suggère que l'activité de la RHA dans la traduction virale est dépendante de son activité ATPase. L'une des hypothèses qui pourrait expliquer la traduction virale propose que l'hélicase RHA active déstabilise les structures secondaires de la 5'UTR du VIH-1 et favoriserait ainsi l'étape de balayage.

## L'hélicase à ARN DDX3

L'hélicase à ARN DDX3 est également impliquée dans la traduction de la polyprotéine Gag du VIH-1. Plusieurs expériences de traduction menées aussi bien *in vitro* qu'*ex-vivo* dans des cellules HeLa et dans des cellules en utilisant des constructions avec un gène rapporteur placé en aval de la 5'UTR du VIH-1, ainsi que des expériences de surexpression/extinction de l'expression de DDX3 endogène, ont révélé, d'une part que la présence de DDX3 était requise à la traduction virale et d'autre part, qu'en sa présence, la traduction virale était fortement stimulée et que cette stimulation était dépendante de l'activité ATPase de DDX3. DDX3 pourrait être comme la RHA impliquée dans la déstabilisation des structures de la 5'UTR. Toutefois, la co-localisation de l'ARNg du VIH-1 dans des granules cytoplasmiques avec DDX3, eIF4G et PABP (Soto-Rifo *et al.*, 2013b) mais en absence d'eIF4E, suggère que DDX3 est recrutée par l'ARN viral pour arrimer le complexe de traduction sur l'ARN viral soit à l'extrémité 5' ou au niveau interne, cela permettrait d'assurer la traduction de Gag dans des conditions où eIF4E serait inactif ou lorsque la coiffe n'est pas accessible et n'est donc pas reconnu par eIF4E.

L'implication de DDX3 dans la traduction virale sera détaillée dans le chapitre suivant.

### C. Initiation de la traduction de la polyprotéine Gag par le mécanisme d'entrée interne des ribosomes

Lors de l'infection par le VIH-1, la protéine virale Vpr bloque le cycle cellulaire en phase G2/M. Cette phase est associée à une hypo-phosphorylation d'eIF4E-BP qui est alors complexé à eIF4E, en conséquence la traduction dépendante de la coiffe est inhibée (Sharma *et al.*, 2012). Pourtant, même dans ces conditions la traduction virale est observée. L'existence de site d'entrée interne des ribosomes sur l'ARN viral pourrait expliquer le maintien de la traduction virale en phase G2/M.

## Un IRES dans la 5'UTR régulé en fonction du cycle cellulaire

### Découverte d'un IRES dans la région 5'UTR du VIH-1

Des tests de traduction *in cellulo* en utilisant des constructions bicistroniques dans lesquelles la 5'UTR du VIH est placée entre deux gènes rapporteurs, ont révélé que la traduction du second gène est efficacement assurée par la région 5'UTR du VIH-1. De même, l'efficacité de traduction de l'ARN viral n'est pas affectée par l'infection du poliovirus exprimant la protéase virale 2A ou par l'ajout exogène de protéases virales. Ces résultats mettent en évidence la présence d'un IRES dans la région 5'UTR du VIH-1 (Brasey *et al.*, 2003; Gendron *et al.*, 2011; Vallejos *et al.*, 2012). L'activité de cet IRES est conservée chez les homologues félins (VIF) et simiens (VIS) mais aucun IRES n'a

été retrouvée au sein de la 5'UTR du VIH-2 (Camerini *et al.*, 2008; Ohlmann *et al.*, 2000).

L'IRES de la 5'UTR du VIH-1 (IRES<sub>5'UTR</sub>) a été localisée entre les nucléotides 104 et 336, ce qui correspond à la région allant de U5 jusqu'au codon AUG non compris (Brasey *et al.*, 2003; Gendron *et al.*, 2011; Vallejos *et al.*, 2012). Une récente étude a permis de mettre en évidence que le « core IRES » qui est impliqué dans l'activité IRES correspond aux tiges-boucles PBS DIS et SD (Carvajal *et al.*, 2016). Par ailleurs, l'activité IRES de la 5'UTR est retrouvée au niveau des différents transcrits du VIH-1. Par exemple, l'ARNm codant pour la protéine tat est traduit selon un mécanisme d'initiation par entrée interne, dépendant de la présence de la protéine tat (Charnay *et al.*, 2009). Cette IRES permet en outre le rétrocontrôle de la protéine tat sur la traduction de son propre ARNm.

L'activité IRES de la 5'UTR est modulée en fonction du type cellulaire mais également en fonction de la séquence codante située en aval (Plank *et al.*, 2013) ce qui permet un niveau supplémentaire de régulation de la traduction virale.

### **L'IRES de la région 5'UTR est stimulée en phase G2/M**

L'activité de l'IRES<sub>5'UTR</sub> a été démontrée dans de nombreux types cellulaires tels que les cellules HeLa, les cellules lymphocytaires Jurkat ou encore les ovocytes de xénopes, de façon intéressante, l'efficacité de traduction maximale est obtenue avec des cellules lymphocytaires qui sont les cellules cibles du VIH-1. Cependant, L'IRES<sub>5'UTR</sub> est inefficace pour la traduction *in vitro* de Gag. Il a été montré que la traduction de Gag pouvait être restaurée par l'ajout d'extraits de cellules HeLa et que cette traduction était stimulée lors de l'ajout d'extraits de cellules HeLa arrêtées en phase G2/M. Ces résultats suggèrent la présence de protéines cellulaires spécifiques de la phase G2/M qui sont recrutées par l'IRES<sub>5'UTR</sub> pour stimuler la traduction de Gag (de Breyne *et al.*, 2013; Ohlmann *et al.*, 2014; Vallejos *et al.*, 2011).

### **Des ITAFs régulent la traduction par l'IRES de la 5'UTR**

Peu de choses sont connues quant au mécanisme moléculaire régissant le fonctionnement de l'IRES de la 5'UTR. Toutefois, des expériences de pull-down réalisées avec des extraits de cellules HeLa synchronisées en phase G2/M, ont permis de révéler une série de protéines en interaction avec l'IRES de la 5'UTR du VIH-1 et qui pourrait agir comme des ITAFs lors de la traduction médiée par cet IRES. Cependant, ces résultats restent préliminaires car le rôle de ces protéines dans l'initiation de la traduction n'a pas encore été vérifié par des tests fonctionnels (Vallejos *et al.*, 2011).

D'autres expériences ont permis de mettre en évidence le rôle de certaines protéines dans la traduction IRES-dépendante. On retrouve par exemple, une protéine stimulatrice

de la traduction virale hnRNP A1 (Heterogeneous Nuclear RibonucleoProtein A1) qui peut se fixer à l'ARNg viral dans le noyau et participer à l'export nucléaire et à la traduction virale (Monette *et al.*, 2009). La protéine HuR, quant à elle, inhibe la traduction (Rivas-Aravena *et al.*, 2009).

## **Le recours au cap-binding complexe pour assurer la traduction virale en phase G2/M**

Durant la phase G2/M, le CBC (Cap Binding Complex) est préférentiellement associé à l'ARN viral non épissé. Ce complexe consiste en deux protéines, la CBP80 (Cap Binding Protein 80) et CBP20 (Cap Binding Protein 20). Il joue un rôle important dans divers aspects du métabolisme de l'ARN. Dans le noyau, le CBC est impliqué entre autre dans l'épissage et l'export des ARNs. Dans le cytoplasme, son rôle serait de vérifier la qualité des ARNm en assurant ce qui est connu sous le nom de « pioneer round of translation ». À savoir, l'ARNm subit un premier tour de traduction, suite au recrutement de la machinerie de traduction par l'intermédiaire de la liaison du CBC à la coiffe. Après ce premier tour le CBC se décroche et est remplacé par eIF4E. Pour l'ARNg du VIH-1, il a été proposé que le CBC restait accroché à la coiffe, et permettait d'assurer une traduction virale coiffe-dépendante et donc de lever l'inhibition imposée par l'infection virale lors de la phase G2/M (Maquat *et al.*, 2010; Sharma *et al.*, 2012).

## **Un IRES dans la région codante du gène Gag**

### **Découverte de l'IRES gag du VIH-1 et mise en évidence d'une isoforme de Gag (p40)**

Des travaux sur la traduction menés par l'équipe de Siliciano en utilisant des ARNs dépourvus de la 5'UTR ont permis de mettre en évidence une seconde IRES située dans la région codante du gène gag appelé IRES<sub>gag</sub>. Cette IRES permet d'initier la traduction sur deux AUG en phase. Le premier conduit à la production de la protéine Gag pleine longueur (p55) alors que le deuxième AUG va permettre la traduction d'une isoforme de Gag (p40) tronquée en N-terminal de tout le domaine matrice (Buck *et al.*, 2001). Cette isoforme p40 ne peut être produite que par l'IRES<sub>gag</sub>, sa traduction est coiffe-indépendante et est maintenue lors de l'ajout de protéases virales (L-protéase de FMDV) (Ricci *et al.*, 2008).

La traduction virale à partir de séquences du virus sauvage issues d'une cohorte de 185 patients naïfs de tout traitements, a mis en évidence que p40 était exprimée à partir de toutes les séquences étudiées. De façon intéressante, même lorsque le second AUG (codon 142) était muté, une nouvelle mutation faisait apparaître un autre codon AUG en phase non loin de la position 142, de façon à maintenir l'expression de p40, suggérant un rôle important de cette isoforme. Le mécanisme d'initiation dépendant de l'IRES<sub>gag</sub> est un mécanisme de traduction original, puisqu'en plus du fait que l'IRES<sub>gag</sub> soit située



dans l'ORF de gag, elle permet d'initier la traduction à partir de deux AUG en phase dont l'un est situé en amont de la séquence IRES, ce qui implique probablement des réarrangements structuraux afin de permettre le bon positionnement du complexe d'initiation au niveau de son site d'initiation.

### L'isoforme p40 de Gag résulte d'une néosynthèse et non d'une protéolyse

Pour confirmer que l'isoforme p40 était bien produite par un mécanisme d'entrée interne et afin de la différencier de p41 un produit intermédiaire de la maturation de Gag p55 par la protéase virale, des cellules Jurkat (cellules de lignée lymphocytaire) ont été infectées avec la souche NL4.3 du VIH-1 en présence ou non d'antiprotéases. Les protéines produites ont ensuite été analysées par western-blot, en utilisant des anticorps dirigés soit contre la capsid (p24) ou contre la matrice (p17). Les résultats avec l'anticorps dirigé contre p24 montrent qu'en absence d'antiprotéases, seule p24 est détectée alors qu'en présence d'antiprotéases trois bandes sont détectées : p55, p40 et une bande intermédiaire (marquée d'un\* sur la Figure 34) tandis que p24 seule n'est plus détectée. En utilisant l'anticorps dirigé contre p17 et en présence d'antiprotéases seule la bande correspondant à la protéine p40 n'est pas observée, confirmant qu'il s'agit bien de p40 puisque cette dernière ne possède pas la protéine p17. L'ensemble de ces résultats confirme que l'expression de p40 se fait *de novo* et non par protéolyse de p55 (Figure 34). (données non publiées;S.de Breyne and Ohlmann)

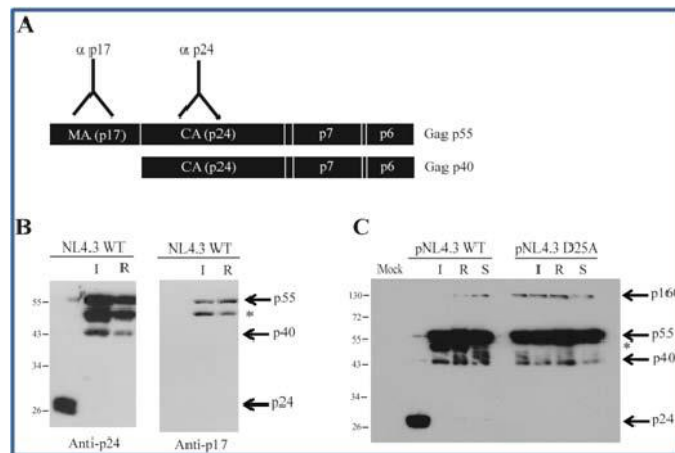


Figure 34 : P40 issue de la traduction IRES-dépendante (données non publiées;S.de Breyne and Ohlmann)

(A) Représentation schématique des épitopes reconnus par les anticorps anti-p17 et anti-p24. (B) Western Blot analysé avec un anticorps dirigé contre p24 ou p17, des protéines virales issues de cellules Jurkat infectées avec la souche NL4.3 du VIH-1 et traitées ou non avec l'indinavir (I) ou la ritonavir (R). Les migrations sur le gel des protéines p55, p40 et p24 ainsi que la protéine intermédiaire (\*) issue d'un clivage partiel sont indiqués. (C) Western Blot analysé avec un anticorps anti-p24, des protéines



virales issues de cellules Jurkat (Cellules 293T transfectées avec les constructions pNL4.3 ou pNL4.3 D25A) en présence ou non d'antiprotéase l'indinavir (I), ritonavir (R), saquinavir (S). La protéine Gag-Pol (p160) est détectée en présence d'antiprotéases.

### Un IRES conservé parmi les lentivirus de primates

Ce mécanisme atypique d'entrée interne est conservé chez plusieurs lentivirus de primates comme le VIH-2 et VIS<sub>MAC</sub> et conduit également à la synthèse d'isoformes tronquées de la région N-terminale de Gag (Figure 35). Chez le VIH-2, l'IRES de la région gag conduit à l'initiation de la traduction sur trois AUG situés dans le même cadre de lecture et jusqu'à trois complexes d'initiation peuvent être assemblés au niveau de cet IRES (Weill *et al.*, 2010). Ceci va permettre en plus de la traduction de la protéine pleine longueur, la traduction de deux isoformes tronquées en N-terminale.

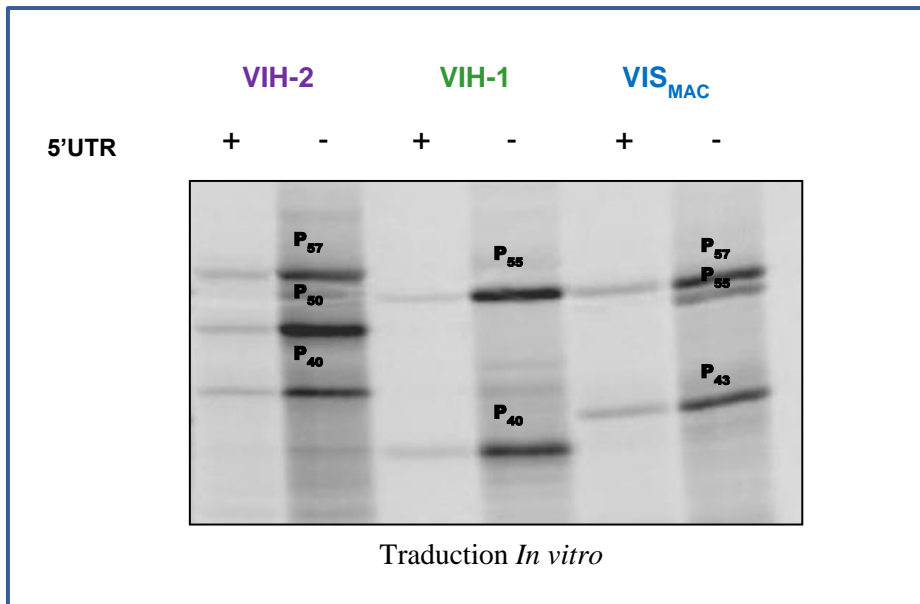


Figure 35 : L'IRES<sub>gag</sub> est conservée parmi les lentivirus (d'après Weill *et al.*, 2010)

Tests de traduction *in vitro* en RRL en présence de méthionine marquée au <sup>35</sup>S, d'ARNg du VIH-2 rod, du VIH-1 pNL4.3 et du VIS<sub>mac</sub> Mm251, contenant la séquence codante des gènes gag du VIH-1, VIH-2 respectivement et précédée (+) ou non (-) de la 5' UTR. La traduction des lentivirus est plus efficace en absence de leur 5'UTR. Le poids moléculaire des isoformes est indiqué sur le gel.

Si peu de choses sont connues sur la fonction de ces isoformes ; leur présence est importante et modifie la virulence du virus. En effet, la mutation du second AUG (codon142) chez VIH-1 et donc la suppression de la deuxième isoforme de Gag provoque une diminution importante du taux de la réplication virale dans les cellules

lymphoïdes, suggérant un rôle important de cette isoforme dans le cycle viral (Buck *et al.*, 2001).

Malgré des divergences de séquence, la modélisation de la région Gag de ces différents virus à l'aide de données biochimiques (« probing » chimique et enzymatique) et phylogénétiques montre qu'elle peut se replier suivant une structure secondaire semblable qui comporte notamment quatre éléments structuraux : les appariements P2, P3, P4 et P5 qui définissent une région centrale de l'IRES conservée pour ces trois virus (Chamond *et al.*, 2010; Weill *et al.*, 2010) (Figure 36).

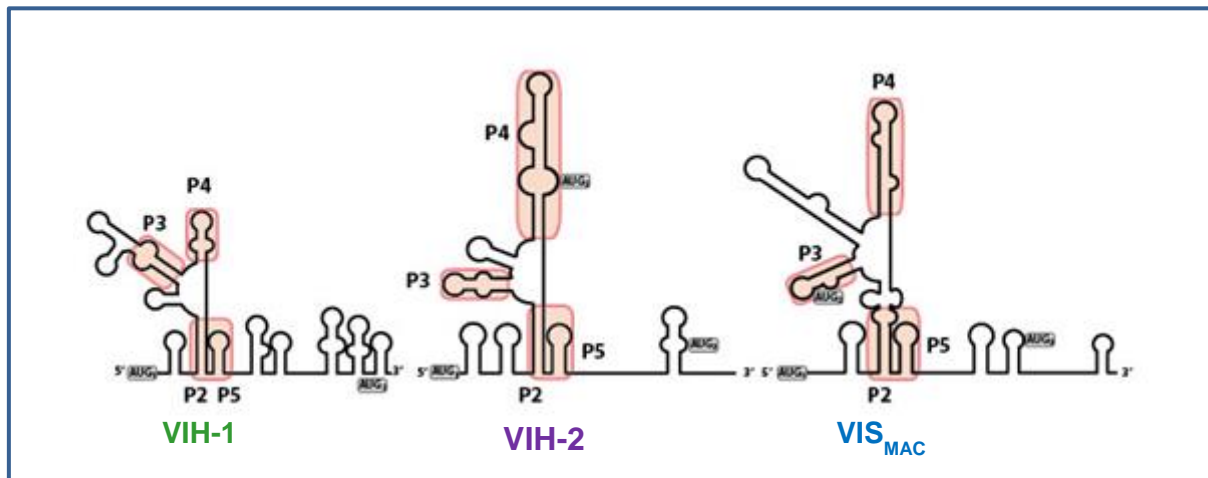


Figure 36 : Les éléments de structure conservés pour l'IRES<sub>gag</sub> chez les lentivirus de primates

Schéma représentant la structure secondaire des IRES<sub>gag</sub> des virus : VIH-1, VIH-2 et VIS<sub>MAC</sub>. Les appariements conservés chez les trois virus, sont encadrés en rose.

### L'étude du mécanisme moléculaire de l'IRES<sub>gag</sub>

Des expériences menées au laboratoire ont montré que l'IRES<sub>gag</sub> est capable de lier indépendamment et avec une forte affinité la petite sous-unité ribosomale 40S ( $k_d$  de 20 nM) et le facteur d'initiation eIF3 ( $k_d$  de 125 nM). La reconstitution des complexes d'initiation *in vitro* ont mis en évidence la formation d'un complexe ternaire stable entre l'IRES<sub>gag</sub>, la sous-unité ribosomale 40S et le facteur eIF3. Les profils obtenus avec l'ARN du VIH-1 en présence d'eIF3 ou de la sous-unité ribosomale 40S (Figure 37c) présentent des pics ayant une mobilité de 3ml et 5ml respectivement (en partant du haut du gradient) ce qui est compatible avec la formation d'un complexe d'ARN-eIF3 ou d'ARN-40S. De façon intéressante lorsque les sous-unités 40S et eIF3 sont assemblées ensemble sur l'ARN du VIH-1, on observe un complexe supplémentaire (Figure 37c-courbe rouge) avec une mobilité réduite (4 ml en partant du haut du gradient) par

rapport au complexe formé avec l'ARN et la sous-unité ribosomale 40S. La liaison d'eIF3 au complexe ribosomique (40S+ARN) est connue pour réduire la mobilité des complexes ribosomiques dans des gradients de densité de sucrose. L'explication proposée pour cette observation est que les changements conformationnels de la sous-unité 40S, ou des changements dans la forme du complexe vers une forme moins globulaire, pourrait augmenter son coefficient de frottement et réduire sa mobilité dans le gradient de sucrose (Kolupaeva *et al.*, 2005).

Le complexe ternaire formé par l'IRES<sub>gag</sub>, la sous-unité ribosomale 40S et le facteur eIF3 constitue probablement l'étape précoce de la formation des complexes d'initiation (Locker *et al.*, 2011) (Figure 37).

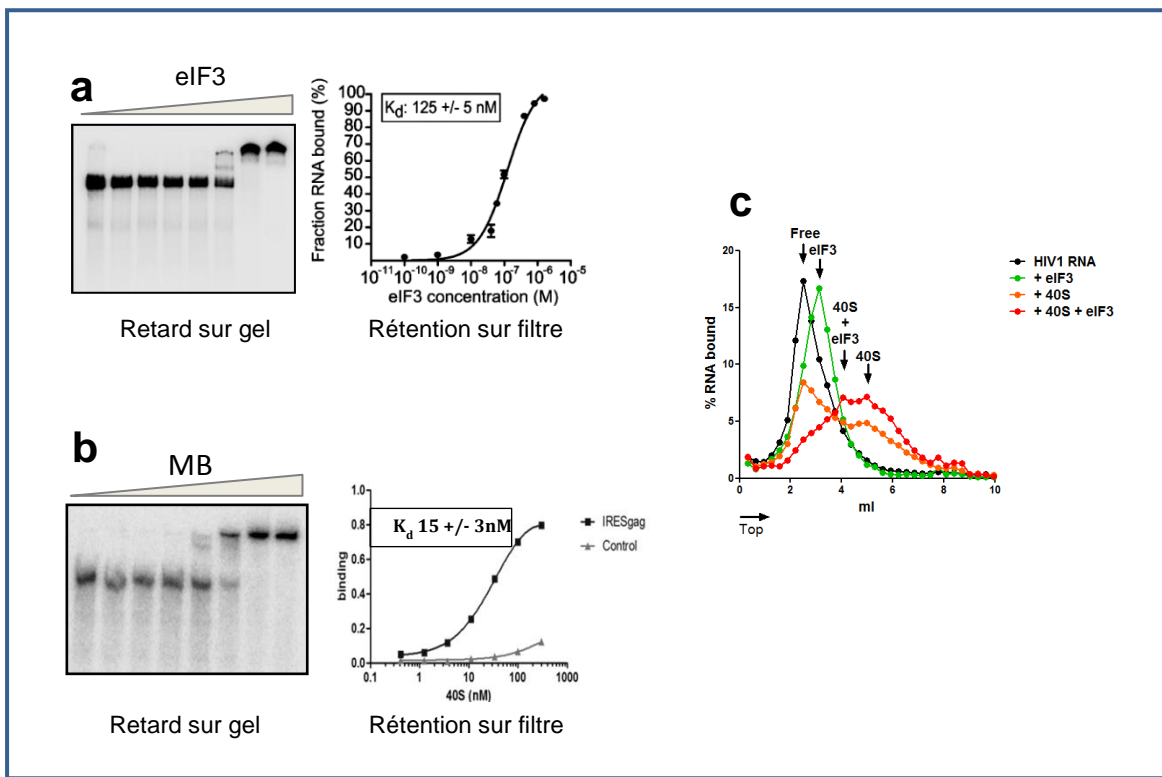


Figure 37 : Formation d'un complexe ternaire stable IRES<sub>gag</sub>-40S-eIF3 (d'après Locker *et al.*, 2011)

Expériences de retard sur gel réalisées en conditions natives avec l'IRES<sub>gag</sub> et des concentrations croissantes du facteur eIF3 (a) ou de la sous-unité ribosomale 40S (b). A côté de chaque gel est représentée la proportion d'ARN lié en fonction de la concentration en eIF3 ou en 40S gag. Le  $K_d$  calculé est indiqué en nM. (c) La mobilité des complexes ribosomiques pendant l'ultracentrifugation en gradient 10-30% de densité de sucrose de l'ARN du VIH-1 seul (courbe noire) ou en présence d'un excès de sous-unités 40S (courbe orange), ou eIF3 (courbe verte) ou 40S +eIF3 (courbe rouge).

Les flèches indiquent si l'ARN est libre (free) ou associé à la sous-unité 40S et / ou au facteur eIF3. L'incubation de l'IRES<sub>gag</sub> avec eIF3 et la sous-unité ribosomale 40S permet la formation d'un complexe ternaire stable.

Le domaine nécessaire à ces liaisons chez le VIH-2 est compris entre les éléments conservés P2 et P5, la présence du domaine P5 étant requise pour la liaison du ribosome mais pas pour celle du facteur eIF3. La détermination du site de liaison du 40S sur l'IRES<sub>gag</sub> du VIH-1 a constitué l'un des objectifs de mon travail de thèse et sera abordé dans le chapitre résultats.

### Des séquences INS au sein de Gag régulent la traduction virale ?

Des séquences inhibitrices de la traduction nommées INS pour INhibitory Sequences sont présentes au niveau de la région codante des gènes Gag et Gag-Pol (*Figure 38*). Cette inhibition est due au moins en partie à l'instabilité des ARNm produits (Cochrane *et al.*, 1991; Maldarelli *et al.*, 1991; Schwartz *et al.*, 1992b; Valiente-Echeverría *et al.*, 2013).

Il faut noter que ces séquences INS inhibitrices sont présentes au niveau de l'IRES<sub>gag</sub>, ceci paraît a priori contradictoire avec l'activité de l'IRES<sub>gag</sub> dans l'initiation de la traduction de la protéine Gag. Cependant, la présence de la protéine virale Rev permet de contrecarrer cet effet inhibiteur. En effet, l'inactivation des séquences INS par mutagenèse en introduisant des mutations qui altèrent la séquence nucléotidique mais pas la séquence protéique, conduit à une forte expression de la protéine Gag p55 indépendamment de la présence de Rev (Schneider *et al.*, 1997). Ainsi, en l'absence de la protéine Rev, le recrutement direct du ribosome par l'IRES<sub>gag</sub> pourrait expliquer cette inhibition. En effet, l'IRES<sub>gag</sub> serait inactif et titrerait les ribosomes disponibles pour la traduction. Rev permettrait de lever cette inhibition selon un mécanisme encore indéfini. Une des hypothèses possibles serait que Rev puisse se lier à ces séquences INS et libérer les ribosomes pour initier la traduction coiffe et IRES dépendante.

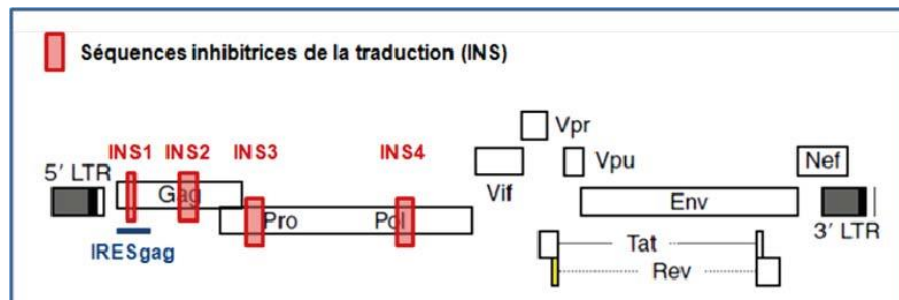


Figure 38 : Position des séquences INS sur le génome viral du VIH-1

#### D. Les mécanismes d'initiation de la traduction de la protéine Gag du VIH-

Le VIH-1 peut initier sa traduction selon deux mécanismes distincts (Figure 39) : un mécanisme canonique coiffe-dépendant et un mécanisme d'entrée interne des ribosomes via deux IRES, l'une située dans la région 5'UTR et activée en phase G2/M du cycle cellulaire (l'IRES<sub>5'UTR</sub>) et l'autre plus atypique (l'IRES<sub>gag</sub>) située au niveau de la région codante du gène gag qui est le seul mécanisme permettant l'expression de la protéine p40. La production de la protéine Gag entière (p50) peut théoriquement avoir lieu suivant le mécanisme coiffe et IRES dépendante. Cependant, la contribution relative de chacun des mécanismes dans la production de Gag n'est pas connue.

De manière intéressante, ces mécanismes ne nécessitent pas les mêmes facteurs pour un fonctionnement optimal, suggérant que le VIH-1 peut adapter et réguler sa propre traduction en fonction de différents facteurs disponibles (de Breyne *et al.*, 2012). Ces différents mécanismes confèrent au virus un avantage sélectif pour assurer un niveau important de traduction de l'ARNg du VIH-1 tout au long du cycle viral mais également dans des conditions où la traduction coiffe-dépendante est inhibée.

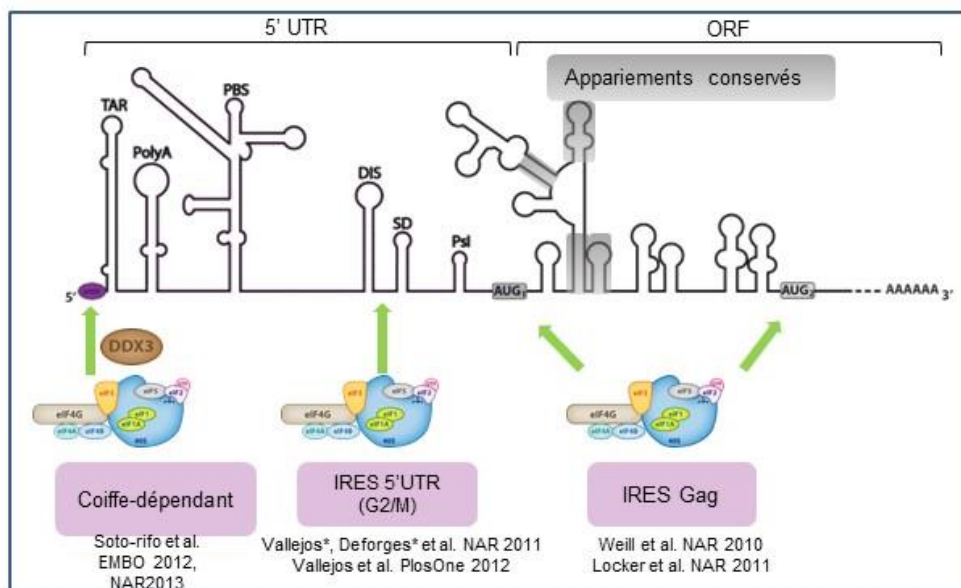


Figure 39 : Les mécanismes d'initiation de la traduction de la protéine Gag du VIH-1  
Les mécanismes d'initiation de traduction du VIH-1 et les complexes d'initiation nécessaires à chaque mécanisme sont représentés. Les structures secondaires de la

5'UTR et de l'IRES<sub>gag</sub> sont représentées. Les appariements conservés de l'IRES<sub>gag</sub> au sein des lentivirus sont encadrés en gris.

### E. Régulation de la traduction de la protéine Gag

L'infection par le VIH-1 est associée à un important changement dans l'homéostasie cellulaire et une diminution globale de la traduction. Le VIH-1 a développé de nombreux mécanismes afin de dépasser les contraintes liées à l'infection virale et à la réponse immunitaire et de continuer à assurer une traduction virale. Quelques exemples de régulation de la traduction par des protéines virales et par des protéines cellulaires sont présentés ci-dessous (Figure 40).

#### Par des protéines virales

La protéine virale Tat est impliquée dans le contrôle de la traduction en inhibant l'activation de PKR (Protein Kinase R). Lors des phases précoces de la réplication, la tige-boucle TAR présente à l'extrémité 5' de l'ARN viral interagit avec la PKR, ce qui induit sa dimérisation, son autophosphorylation et donc son activation. PKR ainsi activée va phosphoryler eIF2 $\alpha$  et induire l'inhibition de la traduction globale (Ederly *et al.*, 1989; SenGupta & Silverman, 1989). Cependant, la protéine virale Tat empêche l'activation de PKR. En effet, Tat interagit soit avec TAR, ce qui empêche la dimérisation de PKR et par conséquent son activation, soit Tat se lie directement à PKR et agit donc en compétiteur vis-à-vis de son substrat naturel eIF2. Dans les deux cas l'inactivation de PKR par Tat conduit à la stimulation de la traduction (Braddock *et al.*, 1990, 1993; Clerzius *et al.*, 2011; McMillan *et al.*, 1995; SenGupta *et al.*, 1990).

Rev est, elle aussi, impliquée dans la traduction. En effet, non seulement Rev permet l'export des ARNm viraux contenant la séquence RRE du noyau vers le cytoplasme mais en plus, Rev favorise la traduction en permettant l'ARNm pleine longueur et les ARNm partiellement épissés (vif, vpr, env et vpu) d'être associés au sein de polysomes (Arrigo & Chen, 1991; D'Agostino *et al.*, 1992; Groom *et al.*, 2009).

Enfin, Gag permet la régulation de sa propre traduction en exerçant un double effet : à faible concentration, Gag régule positivement sa traduction alors qu'à forte concentration elle inhibe sa traduction. Cette inhibition dépend de la capacité de Gag, à travers le domaine à doigt de zinc de la nucléocapside, de se lier au signal d'encapsidation sur l'ARN du VIH-1. Ceci pourrait conduire au revêtement de l'ARN viral par des molécules de Gag et à la formation d'un complexe initial d'encapsidation (Anderson & Lever, 2006).



## Par des protéines cellulaires

Plusieurs protéines cellulaires sont également impliquées dans la régulation de la traduction virale. Il a été montré que la surexpression des protéines SRp40 et SRp55 stimule la traduction de l'ARNg viral d'un facteur 20 et 100 respectivement (Swanson *et al.*, 2010). En plus des protéines SR, la protéine hnRNP A1 augmente l'efficacité de la traduction médiée par l'IRES<sub>5'UTR</sub>, d'autres hnRNP comme hnRNP U et hnRNP L ont été retrouvées dans les extraits cellulaires HeLa synchronisées en phase G2/M du cycle cellulaire (Monette *et al.*, 2009; Vallejos *et al.*, 2011). Tandis que la protéine hnRNP E1 inhibe la traduction de protéines virales (Woolaway *et al.*, 2007).

D'autres protéines cellulaires comme TRBP (TAR RNA-Binding Protein), stau1 ainsi que la protéine La (Lupus antigen) stimulent la traduction d'ARNm portant la tige-boucle TAR. En effet, TRBP se lie à TAR préférentiellement à travers ses domaines de liaison à l'ARN et active l'expression protéique à partir de la 5'UTR (Gatignol *et al.*, 1991, 1993). En outre, des niveaux réduits de TRBP altèrent l'expression et la production de protéines virales (Christensen *et al.*, 2007; Ong *et al.*, 2005). De façon intéressante, TRBP se lie à PKR et inhibe son activité (Daher *et al.*, 2001; Park *et al.*, 1994). Indépendamment de sa capacité à supprimer l'activité de PKR, TRBP est en plus capable de stimuler la traduction des ARNs contenant la tige-boucle TAR (Dorin *et al.*, 2003).

Au cours du cycle de réplication virale, Stau1 forme des complexes ribonucléoprotéiques avec Gag et l'ARN viral, et est incorporée dans des particules virales (Abrahamyan *et al.*, 2010; Chatel-Chaix *et al.*, 2008; Mouland *et al.*, 2000). Stau1 est capable de stimuler la traduction des ARNm dont la traduction est réprimée par TAR *in vitro* et dans les cellules humaines selon un mécanisme qui nécessite la liaison de la protéine à l'extrémité 5' des transcrits (Dugré-Brisson *et al.*, 2005). *In vitro*, la protéine La lève l'inhibition causée par TAR. Cependant, ces données *in vitro* montrent que la protéine La ne se lie pas à la 5'UTR du VIH-2 indiquant un rôle spécifique au VIH-1 (Craig *et al.*, 1997; Svitkin *et al.*, 1994; Waysbort *et al.*, 2001).

De même, les hélicases à ARN RHA, (décrites plus haut dans ce chapitre) et DDX3 sont impliquées dans la stimulation de la traduction virale. L'hélicase DDX3 et son implication dans la traduction seront détaillées dans le chapitre suivant.

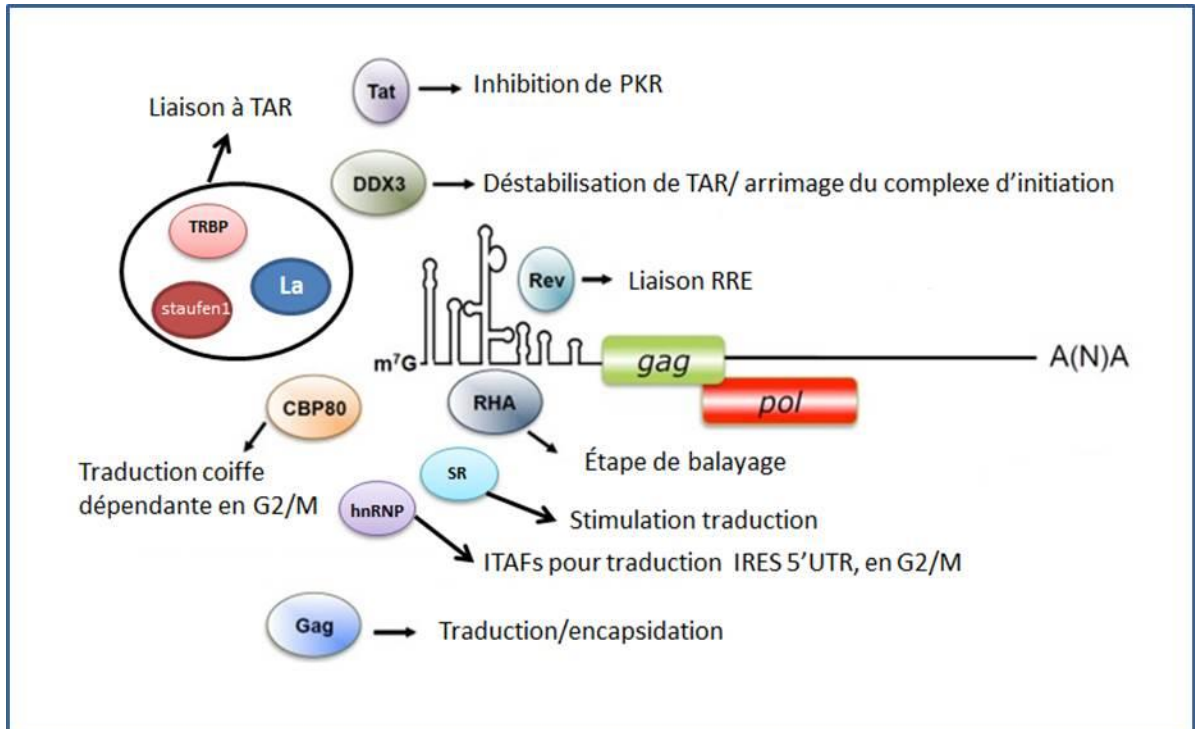


Figure 40: Régulation de la traduction de l'ARNg du VIH-1 par des protéines virales et cellulaires(adapté de Rojas-Araya et al., 2015)

Représentation schématique de protéines virales (Gag, Rev, Tat) et cellulaires (RHA, DDX3, CBP80, TRBP, La, Staufen1, Sr, hnRNP) impliquées dans la régulation de la traduction de l'ARNg du VIH-1.

## IV. Le rôle de DDX3 dans la traduction

### A. Généralités sur les hélicases

Les hélicases sont présentes dans l'ensemble des règnes du vivant et sont classées sur la base de leur séquence et de leur structure protéique en superfamilles (SF1 à SF6) (Gorbalenya & Koonin, 1993; Singleton *et al.*, 2007). Deux grandes catégories sont retrouvées parmi ces superfamilles : les hélicases hexamériques (SF 3 à SF6), présentes généralement chez les bactéries et les virus (Patel & Picha, 2000) et les hélicases non hexamériques (SF1 et SF2) (Fairman-Williams *et al.*, 2010) qui présentent une même organisation structurale avec deux domaines RecA-like qui ressemble à la structure de l'ATPase bactérienne Rec A (Cordin *et al.*, 2006; Story *et al.*, 2001; Story & Steitz, 1992). Les deux domaines représentent environ 350 à 400 résidus et constituent l'hélicase-core domaine. La SF2 comprend des hélicases à ARN parmi lesquelles on retrouve les familles d'hélicases à boîte DEAD, DEAH, DExH et DExD, communément appelées la famille d'hélicases DExD/H. Nous nous sommes intéressés à une hélicase à boîte DEAD : DDX3.

### B. La famille des hélicases à boîte DEAD

Les protéines à boîte DEAD sont des hélicases à ARN ATP-dépendante présentes chez de nombreux procaryotes et chez tous les eucaryotes (Fairman-Williams *et al.*, 2010; Rocak & Linder, 2004). Elles sont impliquées dans le métabolisme de l'ARN et dans le remodelage des structures d'ARN et des complexes protéines-ARN (Linder, 2006; Linder & Jankowsky, 2011; Rocak & Linder, 2004).

Les hélicases à boîte DEAD ont été regroupées dans une même famille pour la première fois en 1989 (Linder *et al.*, 1989) suite à la mise en évidence de plusieurs motifs conservés grâce à l'alignement de la séquence peptidique de huit protéines homologues de l'hélicase eIF4A. C'est la séquence caractéristique D.E.A.D (Asp-Glu-Ala-Asp) du motif Walker B (motif II), qui a inspiré leur nom. Depuis, de nombreuses autres protéines à boîtes DEAD ont été découvertes et représentent aujourd'hui la plus grande famille d'hélicases à ARN avec plus de 37 membres chez l'Homme (Fairman-Williams *et al.*, 2010).

Les hélicases à boîtes DEAD partagent douze motifs conservés caractéristiques. Ces motifs sont impliqués dans la liaison et l'hydrolyse de l'ATP (Q, I, II et VI), la liaison à l'ARN (Ia, Ib, Ic, IV, IVa et V), et dans le couplage entre liaison à l'ATP et liaison à l'ARN (III, Va) (Linder & Jankowsky, 2011) (Figure 41). Ces motifs sont contenus dans l'hélicase core domaine qui s'étend de la phénylalanine en amont du motif Q jusqu'à environ 35 nucléotides en aval du domaine VI (Banroques *et al.*, 2011).

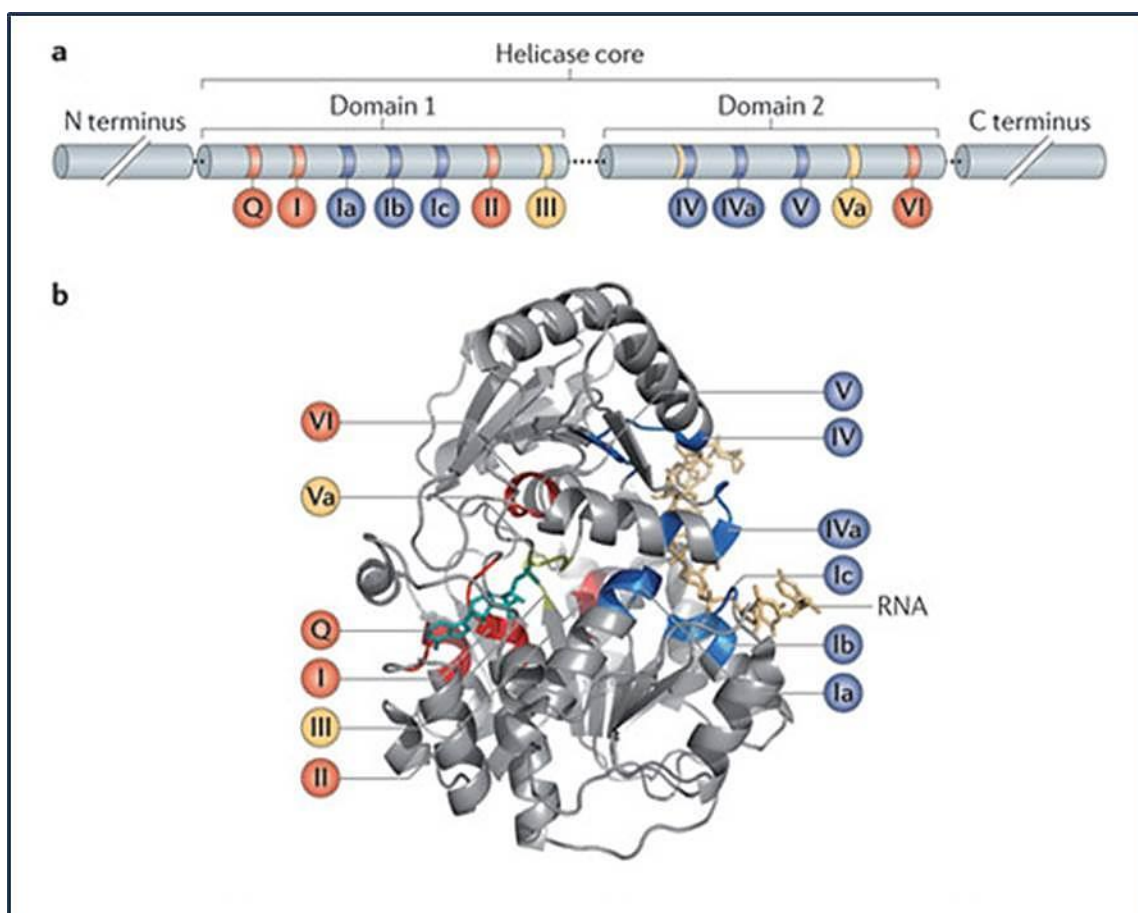
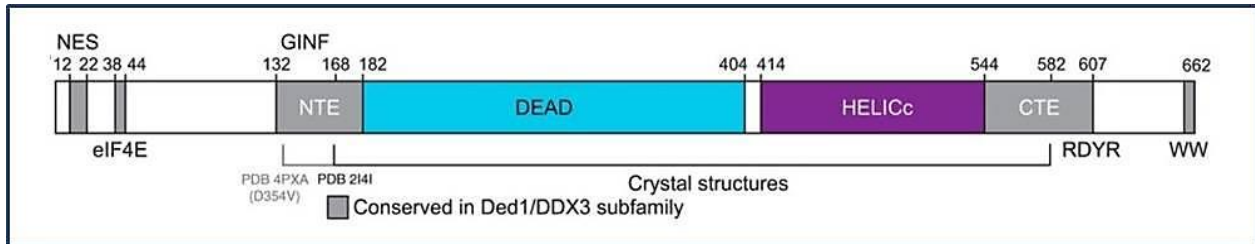


Figure 41: Organisation structurale des protéines à boîte DEAD (D'après Linder & Jankowsky, 2011))

(a) Le core hélicase est constitué de deux domaines RecA-like (Domaines 1 et 2). Les motifs conservés sont représentés par un code couleur correspondant à la fonction principale de chaque domaine ; liaison et hydrolyse de l'ATP (rouge), liaison de l'ARN (bleu), couplage entre liaison de l'ATP et liaison de l'ARN (jaune). (b) Structure des core domaines de l'hélicase à boîte DEAD Vasa de la *Drosophila melanogaster*. Le code couleur des domaines conservés est le même qu'en (a) sont. L'ARN est représenté en beige.

Mise à part ces motifs conservés, chaque sous-famille de protéines à boîte DEAD possède des séquences flanquantes aux extrémités N et C-terminales de longueur variable, qui comportent des éléments définissant les propriétés uniques de chacune de ces sous-familles. Par exemple, la sous-famille Ded1/DDX3 comprend plusieurs motifs conservés au niveau des extrémités N et C-terminales impliqués dans des interactions avec d'autres protéines et avec des ARNs cibles (Figure 42). En effet, le domaine N-terminal contient une séquence riche en leucine (NES) essentielle à l'export nucléaire

médiée par CRM-1 (Yedavalli *et al.*, 2004). Le domaine N-terminal contient aussi un motif de liaison à eIF4E (Hilliker *et al.*, 2011; Shih *et al.*, 2008, 2012). Au niveau du domaine C-terminal des motifs conservés sont également présents mais leurs fonctions restent à découvrir bien qu'ils semblent essentiels à l'oligomérisation (Putnam *et al.*, 2015; Sharma & Jankowsky, 2014). Plus récemment, deux régions adjacentes à l'hélicase core domaine nommés NTE et CTE (pour, N/C-Terminal Extension) ont été montrées comme essentielles aux activités de DDX3.



**Figure 42: Les régions conservées spécifiques à la sous-famille des hélicases à boîte DEAD, Ded1/DDX3 (d'après Floor *et al.*, 2016)**

Schéma représentant les caractéristiques de la sous-famille Ded1/DDX3, dont le core hélicase DEAD (bleu) et le domaine HELICc RecA-like (violet). La séquence d'export nucléaire CRM-1-dépendante (NES), et le site de liaison à eIF4E Les régions conservées spécifiques à la sous-famille Ded1 / DDX3 sont indiquées en gris; GINF, RDYR, et WW indiquent les motifs d'acides aminés. Les régions NTE et CTE sont indiqués.

### C. DDX3, protéine aux multiples fonctions

DDX3, appartient à la sous-famille Ded1/DDX3 des protéines à boîte DEAD. C'est une hélicase de 661 acides aminés et comme les membres de la famille des DEAD-box, elle est organisée en deux domaines RecA-like et abrite les douze motifs hautement conservés décrits précédemment (Figure 43). Le génome humain possède deux gènes DDX3 paralogues, DDX3Y et DDX3X (Kim *et al.*, 2001) qui partagent 92% d'homologie de séquence et qui codent respectivement pour deux protéines : DDX3Y (DBY), produite dans les spermatocytes où elle est essentielle pour la spermatogenèse (Ditton *et al.*, 2004; Ferlin *et al.*, 2003; Foresta *et al.*, 2000; Rauschendorf *et al.*, 2011) et DDX3X (DBX) qui est exprimée de façon ubiquitaire dans une large gamme de tissus et se retrouve chez tous les organismes allant de la levure à l'Homme. DDX3X est impliquée dans de nombreux processus cellulaires tels que la progression du cycle cellulaire, l'apoptose, le métabolisme de l'ARN allant de la transcription à la traduction en passant par l'épissage, l'export nucléaire ou encore la dégradation des ARNm (Ariumi, 2014; Sharma & Jankowsky, 2014; Soto-Rifo & Ohlmann, 2013). Elle est également impliquée dans certains cancers et dans la réplication de nombreux virus qui infectent l'Homme (Schröder, 2011; Valiente-Echeverría *et al.*, 2015). Seules ses



implications dans la traduction et dans la réplication virale seront abordées dans la suite de cette introduction.

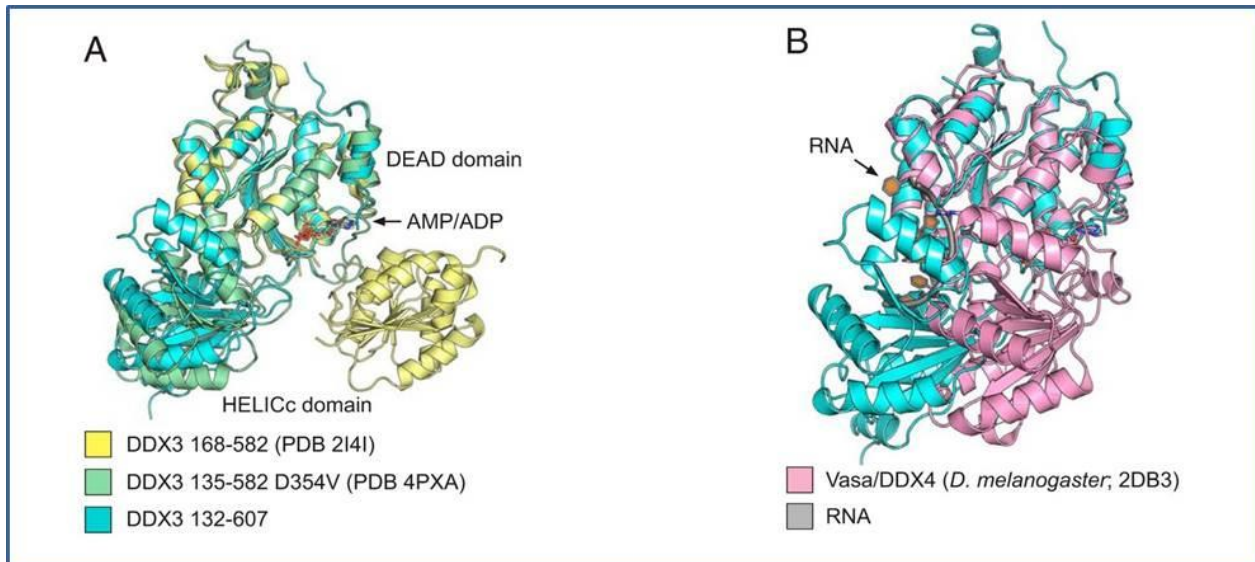


Figure 43: La structure cristallographie du core domaine de DDX3 (d'après Floor et al., 2016)

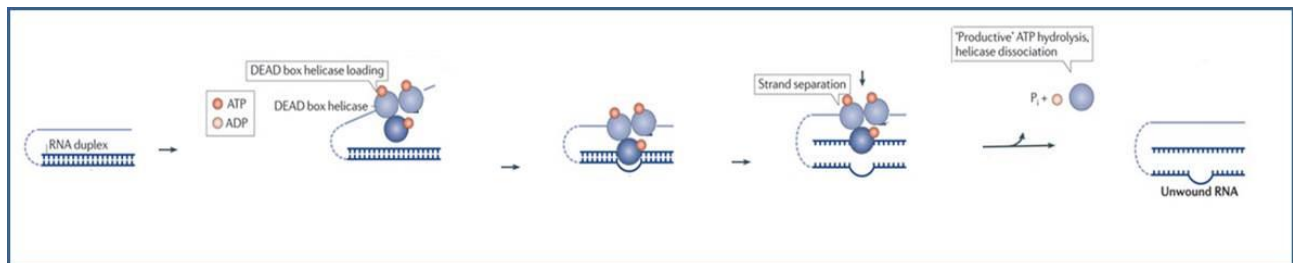
(A) La structure de DDX3 132-607 lié à l'AMP (bleu) est représenté le long de la structure de DDX3 135-582 D354V (vert; PDB 4PXA) et de DDX3 168-582 (jaune; PDB 2I4I). Les structures sont alignées par le domaine DEAD, mettant en évidence la rotation du domaine C-terminal HELICc entre les trois structures. (B) l'état partiellement fermé de DDX3 (bleu) se heurte au site de liaison à l'ARN basée sur une comparaison avec la protéine à boîte DEAD Vasa lié à l'ARN (Vasa: rose; ARN: gris).

#### D. Les propriétés biochimiques de DDX3

Contrairement à DDX3, Ded1p son homologue chez *S. cerevisiae* a été largement étudiée et ses activités biochimiques caractérisées. Ainsi, comme toutes les protéines à boîte DEAD, Ded1p hydrolyse l'ATP. Cette activité ATPase est dépendante de l'ARN (Cordin et al., 2004). Cependant, les ARNs court (10 nt) stimulent moins bien l'activité ATPase que des ARNs longs et probablement structurés (Yang & Jankowsky, 2005). Hormis la taille de l'ARN, la stimulation de l'activité ATPase par des séquences ou des structures spécifiques d'ARN n'a pas été démontrée. Ded1p déstabilise les duplexes d'ARNs d'une manière dépendante de l'ATP, aussi bien avec le rATP qu'avec dATP ce qui indique que le 2'-OH de la base n'est pas critique pour l'activité hélicase (Yang & Jankowsky, 2005). Les données cristallographiques montrent que c'est la glutamine du motif Q qui assure la spécificité de reconnaissance de l'adénosine (Cordin et al., 2004). Ded1p déstabilise les duplex d'ARN à extrémités franches et les duplex d'ARN présentant des extrémités sortantes non appariées (Yang & Jankowsky, 2006), cependant la longueur des extrémités sortantes et la stabilité du duplex affectent



fortement la processivité de l'enzyme. L'activité hélicase optimale est observée avec des duplex d'ARN présentant des extrémités sortantes d'environ 25 nucléotides (Yang *et al.*, 2007). Récemment, il a été mis en évidence que l'activité hélicase de Ded1p était coopérative et faisait intervenir trois protomères. Ainsi, les deux premiers protomères interagissent avec un segment d'ARN simple-brin et stimulent la fixation du troisième protomère sur un duplex d'ARN voisin (Putnam *et al.*, 2015). Le mécanisme par lequel les protéines à boîte DEAD déroulent les duplex ARN et ADN-ARN est principalement dérivé de l'étude du mécanisme de Ded1p (Yang *et al.*, 2007). Ce mécanisme, est appelé " local strand separation " (Figure 44), est un mécanisme en trois étapes qui ne comporte pas d'étape de translocation de l'hélicase. D'abord, la protéine à boîte DEAD lie le duplex (en interne ou en terminal) et cette liaison est facilitée par la présence de régions simple brin. Pour Ded1p le chargement de la protéine sur l'ARN implique plusieurs protomères. Pour d'autres protéines à boîte DEAD le chargement se fait par des domaines liant l'ARN dans les domaines C- ou N-terminal de l'hélicase, et ne nécessite qu'un seul protomère (Grohman *et al.*, 2007; Jarmoskaite & Russell, 2014). Une fois chargés sur le duplex, Ded1p ouvre un nombre limité de paires de bases (Yang *et al.*, 2007) et le reste des paires de bases est séparé de façon non processive. Le déroulement local nécessite la liaison mais pas l'hydrolyse de l'ATP (Liu *et al.*, 2008; Yang *et al.*, 2007). Après l'ouverture du duplex, l'ATP est hydrolysé et l'hélicase se dissocie du substrat.



**Figure 44: Mécanisme hélicase « local strand separation » (adapté de Linder & Jankowsky, 2011)**

Les brins d'ARN sont représentés sous forme de lignes ; le duplex (ligne bleu foncé) et la région simple brin (bleu clair). Les pointillés soulignent que cette région n'a pas besoin d'être physiquement connecté au duplex d'ARN pour aider au chargement de l'enzyme Ded1p (cercles bleu). L'ATP (rouge) et ADP (rose) sont représentées en petits ronds. La réaction hélicase commence par le chargement de deux protomères de l'enzyme sur l'ARN simple brin. Ce qui stimule la fixation du troisième protomère sur le duplex d'ARN. Duplex. Une fois chargé, Ded1p ouvre un nombre limité de paires de bases et le reste des paires de bases. Ce déroulement local nécessite la liaison mais pas l'hydrolyse de l'ATP. Après l'ouverture du duplex, l'ATP est hydrolysé et l'hélicase ainsi que les produits d'hydrolyse de l'ATP (ADP+Pi) se dissocient.

Les activités hélicase ATP-dépendante et ATPase ARN-dépendante de Ded1p ont été démontrées en 1999 (Iost *et al.*, 1999). Depuis, d'autres activités de la protéine ont été mise en évidence et incluent, le remodelage des complexes ribonucléoprotéique (Bowers *et al.*, 2006; Fairman *et al.*, 2004), la liaison stable de l'ARN (Liu *et al.*, 2014), la détection de l'AMP (Putnam & Jankowsky, 2013) mais également la capacité de promouvoir l'hybridation des brins d'ARN (Yang & Jankowsky, 2005).

Bien que la protéine DDX3 ait été découverte en 1998 (Park *et al.*, 1998), peu de données sur la caractérisation des activités ATPase et hélicases de DDX3 sont disponibles (Franca *et al.*, 2007; Garbelli *et al.*, 2011b; Yedavalli *et al.*, 2004). La caractérisation des activités biochimiques a mis en évidence des propriétés différentes de celles décrites pour son homologue Ded1p. Par exemple, il a été montré que DDX3 présentait une activité ATPase indépendante de l'ARN et cette activité ATPase serait stimulée aussi bien par l'ARN ou l'ADN (Franca *et al.*, 2007; Garbelli *et al.*, 2011b), ce qui est inhabituel pour les protéines à boîte DEAD. De façon étonnante, les 410 premiers acides aminés de DDX3 qui ne contiennent ni les motifs de liaison à l'ARN ni le motif VI, présentent encore une activité ATPase intrinsèque (Garbelli *et al.*, 2011b; Högbom *et al.*, 2007).

Concernant l'activité hélicase de DDX3, elle a été décrite comme capable de dérouler un duplex d'ARN long de 50 nucléotides (Franca *et al.*, 2007; Yedavalli *et al.*, 2004). Cette activité hélicase est dépendante de l'activité ATPase. En effet, des mutants engendrant une perte partielle (K230E, motif I) ou totale (S382L, motif III) de l'activité ATPase abolissent l'activité hélicase (Yedavalli *et al.*, 2004). Les autres membres de la famille des DEAD-box y compris Ded1p, ne parviennent pas à dérouler des duplex de plus de 20 paires de bases (Yang & Jankowsky, 2005). Enfin, DDX3 serait capable d'hydrolyser tous les NTP. Ceci contraste avec les données publiées pour toutes les autres hélicases à boîte DEAD, qui montrent que la présence du motif Q assure une spécificité de reconnaissance de l'adénosine (Högbom *et al.*, 2007).

#### E. Les fonctions biologiques

### Rôle de DDX3 dans la traduction

Bien que Ded1p soit un facteur général d'initiation traduction essentiel pour la traduction de tous les ARNm de levure (Chuang *et al.*, 1997; Hilliker *et al.*, 2011), le rôle de DDX3 dans la synthèse des protéines reste sujet à débat. En raison de sa capacité à sauver le phénotype létal produit par la suppression de Ded1, DDX3 a d'abord été proposée comme un facteur général d'initiation de la traduction dans les cellules de mammifères (Tarn & Chang, 2009). En effet, dans les cellules HeLa, la déplétion de DDX3 par siRNA inhibe la traduction dépendante de la coiffe (Lee *et al.*, 2008a). De même, dans les cellules Huh7 déplétée en DDX3, la traduction coiffe-dépendante mais

pas IRES dépendante, est inhibée (Geissler *et al.*, 2012). Cependant, une autre étude a démontré que l'expression de DDX3 dans les cellules Huh7 inhibe l'expression protéique générale (Shih *et al.*, 2008), ce qui suggère que DDX3 pourrait agir comme un répresseur de la traduction et ce, indépendamment de ses activités ATPase et hélicase. En effet, les mutants de l'activité ATPase DQAD (motif II) et AAA (motif III) sont toujours capables d'inhiber la traduction. Cet effet inhibiteur a été ensuite relié à la présence du motif hautement conservé (YxxxxLφ) à l'extrémité N-terminale de DDX3 qui ressemble au motif de liaison à eIF4E présent chez eIF4G et 4E-BP. Il a ainsi été suggéré que DDX3 interfère avec la traduction coiffe-dépendante en perturbant l'interaction d'eIF4E avec eIF4G de façon analogue à ce qui se passe avec 4E-BP. Récemment, DDX3 a été proposée comme participant à la traduction d'un sous-ensemble d'ARNm cellulaires et viraux qui possèdent une tige-boucle à l'extrémité 5' de leur ARN (Lai *et al.*, 2010; Soto-Rifo *et al.*, 2012b; Soto-Rifo & Ohlmann, 2013) comme c'est le cas pour l'ARN du VIH-1. Malgré des données controversées sur l'effet de DDX3 sur la traduction, il existe encore peu de doute concernant son rôle comme un facteur d'initiation de la traduction (Soto-Rifo & Ohlmann, 2013). DDX3 a été co-purifiée avec plusieurs facteurs d'initiation de traduction tels qu'eIF4A, eIF4G, eIF4E, PABP et eIF3 (Lee *et al.*, 2008a; Shih *et al.*, 2012) et son interaction directe avec eIF4G a été démontrée.

### **DDX3 et la réplication virale**

Divers virus détournent DDX3X pour l'utiliser comme cofacteur cellulaire pour leur réplication ou pour l'inactiver par des protéines spécifiques (Schröder, 2011; Valiente-Echeverría *et al.*, 2015).

#### DDX3 et la réplication du VIH-1

DDX3 est essentielle à la réplication virale du VIH-1. Elle est associée à deux étapes du cycle viral : l'export nucléaire et la traduction du génome viral. Ainsi, suite à l'infection virale, DDX3 se lie à Rev et à CRM1 (Chromosomal Maintenance 1, aussi connue sous le nom d'Exportin 1) pour l'export de l'ARNg du VIH-1 du noyau vers le cytoplasme. Néanmoins, le domaine d'interaction avec CRM-1 reste controversé (Fröhlich *et al.*, 2016; Yedavalli *et al.*, 2004).

DDX3 est une protéine navette entre le cytoplasme et le noyau, elle est impliquée dans l'assemblage ATP-dépendant de granules cytoplasmiques nécessaires à l'ARNg pour recruter le complexe d'initiation 43S au niveau de la coiffe (Soto-Rifo *et al.*, 2013b). Ainsi DDX3 est impliquée dans l'initiation de la traduction du VIH-1, cependant, son rôle exact est difficile à déterminer. Trois hypothèses pourraient être avancées (Figure 45). D'une part, DDX3 pourrait être impliquée dans la traduction du VIH-1 par le mécanisme dépendant de la coiffe par sa liaison à la coiffe de façon indépendante d'eIF4E ou par son interaction à TAR. En effet, l'insertion d'un motif non structuré de 15 nucléotides

(longueur nécessaire pour fixer le complexe 43S) entre la coiffe et la base de TAR a révélé une indépendance de la traduction virale vis-à-vis de DDX3. Ceci suggère que DDX3 interagit avec TAR et est plutôt impliquée dans les étapes précoces de l'initiation de la traduction plus que dans l'étape de balayage de l'ARNm. En effet, il a été proposé que DDX3 déstabilise la structure secondaire TAR, ce qui a pour conséquence de rendre la coiffe accessible au complexe d'initiation (Lai *et al.*, 2013; Soto-Rifo *et al.*, 2012b). Par ailleurs, DDX3 a également été retrouvée, au niveau de granules cytoplasmiques, en interaction avec les facteurs d'initiations de la traduction eIF4G et PABP mais pas avec eIF4E. De plus, *in vitro*, DDX3 est capable de lier spécifiquement une résine de <sup>7</sup>GTP (Soto-Rifo *et al.*, 2012b). Ces deux résultats soulèvent l'hypothèse que DDX3 pourrait se substituer à eIF4E et être recrutée par l'ARN viral au niveau de la coiffe puis formerait un complexe avec eIF4G et la PABP permettant ainsi au complexe 43S d'être recruté pour initier la traduction. Un tel mécanisme pourrait permettre à l'ARNg d'assurer la traduction de protéines virales même dans les conditions où eIF4E est inactif ou lorsque la coiffe n'est pas accessible et n'est donc pas reconnue par eIF4E (Rojas-Araya *et al.*, 2015; Soto-Rifo *et al.*, 2012b, 2013b). Enfin, elle pourrait également être impliquée dans un mécanisme IRES-dépendant en servant de plateforme d'arrimage du complexe d'initiation au niveau de ses sites de liaison sur la 5'UTR du VIH-1. Des expériences ont montré qu'en plus de sa liaison à TAR, DDX3 se lie également à deux autres sites au niveau de la 5'UTR du VIH-1 correspondant à des régions simples brins et qui pourraient servir de plateforme d'arrimage du complexe de préinitiation sur l'ARN viral (Soto-Rifo *et al.*, 2012b).

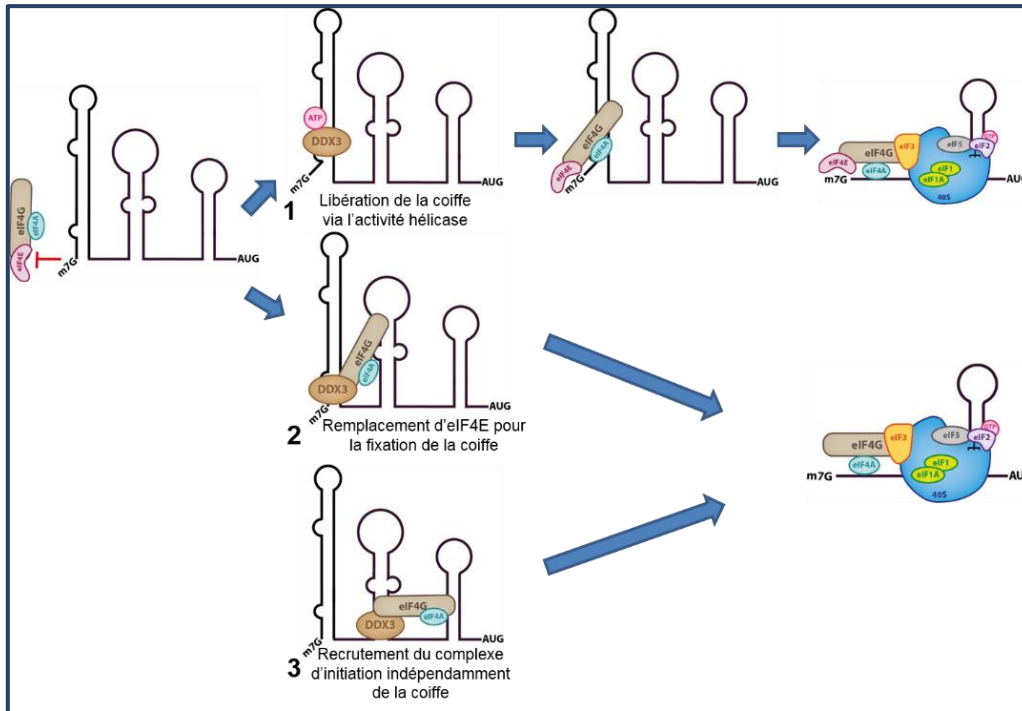


Figure 45: DDX3 et la traduction du VIH-1

Représentations schématiques des trois hypothèses avancées pour l'implication de DDX3 dans la traduction du VIH-1. Lorsque la coiffe est inaccessible, DDX3 pourrait soit rendre accessible la coiffe via son activité hélicase ATP-dépendante (1). Soit arrimer le complexe d'initiation au niveau de l'extrémité 5' ou au niveau de TAR (2) ou encore DDX3 pourrait arrimer le complexe d'initiation de façon interne directement au niveau de séquences IRES (3)

### DDX3 et la réplication du VHC

Plusieurs études ont mis en évidence le rôle de DDX3 dans la réplication du virus de l'hépatite C Humaine (VHC) (Ariumi *et al.*, 2007). Deux modes d'action de DDX3 sur la réplication virale du VHC sont proposés. Le premier implique l'interaction de DDX3 avec la 3'UTR du VHC pour activer IKK $\alpha$ , qui subit une translocation vers le noyau et induit le programme transcriptionnel bénéfique pour la réplication du VHC médié par CBP /p300 (Li *et al.*, 2013). L'autre mode implique une interaction directe entre DDX3 et la protéine de capsid du VHC. Cette interaction interférerait avec la fonction de DDX3 dans l'induction de l'IFN- $\beta$  induction, supprimant ainsi sa production et donc la réponse antivirale innée (Schröder, 2011). En effet, 59 résidus de l'extrémité N-terminale de la protéine core du VHC interagissent directement avec l'extrémité C-terminale de DDX3 et inhibent ses activités catalytiques. En outre, la protéine core du VHC induit une relocalisation de DDX3 à des sites d'assemblage du VHC autour des gouttelettes lipidiques (Ariumi *et al.*, 2007).

### DDX3 et la réplication d'autres flavivirus

DDX3 est également impliquée dans la réplication d'autres flavivirus. Au cours de la réplication du virus JEV ( Japanese Encephalitis Virus), DDX3 favorise la traduction de l'ARN viral (Li *et al.*, 2014). De même, DDX3 est impliquée dans la réplication du virus du Nil occidental (Chahar *et al.*, 2013). Ces données soulèvent la possibilité que DDX3 joue un rôle similaire dans la réplication de différents flavivirus.

### DDX3 et le virus de la vaccine

DDX3 est ciblée par une protéine du virus de la vaccine. La protéine VACV K7, qui adopte une structure  $\beta$ -cell lymphoma-2-like la, se lie à DDX3 et inhibe ainsi la réponse immunitaire innée (Oda *et al.*, 2009). K7 interagit directement avec le domaine N-terminal de DDX3 et inhibe l'activation de la voie TLR et l'induction de l'IFN- $\beta$ . Les données structurales et fonctionnelles suggèrent que K7 masque l'épitope de DDX3 nécessaire pour son interaction avec IKKe.

Les rôles des DDX3 dans la réplication virale, notamment du VIH-1 et du VHC, ont fait de DDX3 une cible thérapeutique potentielle. Plusieurs composés ont été rapportés comme inhibant l'activité ATPase de DDX3 *in vitro* et la réplication virale dans les cultures cellulaires (Brai *et al.*, 2016; Garbelli *et al.*, 2011a; Radi *et al.*, 2012).



## Objectifs de thèse

De nombreux mécanismes potentiellement responsables de la traduction de la polyprotéine Gag ont été décrits. En ce qui concerne les mécanismes dépendants de la coiffe, l'intervention d'hélicases permettrait certes de surmonter l'obstacle des structures présentes dans la 5'UTR, mais pas d'assurer la traduction dans des cellules en phase G2/M. A l'opposé, le mécanisme inspiré du « *pioneer round* », permettrait lui la traduction dans une cellule bloquée en phase G2/M mais ne donne aucun éclairage sur la levée de l'inhibition due aux structures secondaires. En ce qui concerne les mécanismes d'entrée interne des ribosomes, ils permettent de s'affranchir des deux problèmes mais l'absence de codon d'initiation dans la région 5' non traduite laisse tout de même penser que la traduction majoritaire est dépendante de la coiffe. Par ailleurs, la présence d'une IRES conservée dans la région codante permettrait de produire une isoforme tronquée de la polyprotéine Gag qui, d'après les travaux de Buck, est indispensable à la réplication virale (Buck *et al.*, 2001), mais dont l'existence est contestée par certains laboratoires. Ces mécanismes interviennent-ils à différents stades du cycle viral ? Dans différents types cellulaires ? En fonction de la progression des cellules dans le cycle viral ? Tous ces mécanismes interviennent-ils conjointement pour maximiser la production de Gag dans les phases tardives du cycle viral ?

Pour essayer de répondre à ces questions, nous avons suivi une démarche « inverse » qui consiste à étudier le mécanisme pour le re-situer ensuite dans un contexte physiologique pertinent. Ainsi, lors de ma thèse je me suis intéressée à deux aspects des mécanismes de traduction de la polyprotéine Gag : La liaison de la sous-unité 40S dans l'IRES<sub>gag</sub> d'une part, et le mécanisme permettant à l'hélicase DDX3 de stimuler la traduction coiffe-dépendante d'autre part.

A mon arrivée au laboratoire, un mécanisme original d'initiation de la traduction médié par l'IRES<sub>gag</sub> était proposé. L'une des étapes précoces de ce mécanisme est la liaison directe de deux acteurs centraux de la machinerie traductionnelle: la petite sous-unité ribosomale 40S et le facteur d'initiation eucaryote eIF3. Afin de mieux caractériser l'interaction entre la petite sous-unité ribosomale et l'IRES<sub>gag</sub>, nous avons utilisé un ensemble d'approches moléculaires qui nous ont permis de définir le(s) site(s) d'interaction de la sous-unité 40S sur l'IRES<sub>gag</sub>. Nous avons ensuite évalué les conséquences fonctionnelles de la délétion de ces domaines d'interaction sur l'initiation de la traduction de Gag aussi bien *in vitro* qu'*in cellulo*.

En ce qui concerne le rôle de DDX3X et compte-tenu de ses propriétés enzymatiques plus que particulières, nous avons entrepris la caractérisation biochimique de cette protéine tant du point de son activité ATPase que de son activité hélicase. Nos résultats nous permettent d'invalider les études précédentes et de conclure que DDX3X présente des propriétés proches de celle de son homologue Ded1p. Nous avons également pu évaluer l'effet de DDX3 sur l'initiation de la traduction de Gag et ainsi valider notre système d'étude *in vitro*.

En outre, durant ma thèse j'ai participé (en collaboration avec l'équipe de M.Lopez-Lastra, université de Santiago du Chili) à la mise en évidence d'une interaction entre PTB (Polypyrimidine Tract-binding Protein) et l'IRES du virus MMTV (Mouse Mammary Tumor Virus) (voir figure 3 de l'article présenté en annexe).

## Résultats

## I. Étude de l'interaction entre l'IRES<sub>gag</sub> du VIH-1 et la sous-unité ribosomale 40S

En collaboration avec l'équipe de T. Ohlmann, notre laboratoire a pour la première fois proposée l'idée novatrice et provocante d'un IRES dans une région codante qui permet la traduction non seulement à partir d'un codon d'initiation en aval, mais aussi à partir d'un codon AUG en amont (Herbreteau *et al.*, 2005). Cet IRES permettrait donc la production de la polyprotéine Gag pleine longueur p55 et celle d'une isoforme tronquée p40. Le rôle de cette isoforme qui n'est pas détectée par tous les laboratoires reste inconnu. Pourtant, notre laboratoire a montré que cet IRES est conservé dans plusieurs lentivirus de primates (VIH-1, VIH-2 et VISmac) et qu'il permet de produire diverses isoformes tronquées de la polyprotéine Gag (Weill *et al.*, 2010). Une deuxième étude émanant du laboratoire, a ensuite montré que la région codante de Gag est capable de recruter indépendamment la sous-unité 40S du ribosome, et le facteur d'initiation eIF3 avec, dans les deux cas, une forte affinité spécifique. Là encore il s'agit de propriétés conservées parmi les lentivirus de primates (Locker *et al.*, 2011). Il semble donc exclu que ces propriétés associées à la fonction IRES soient un hasard ou un artefact. Poursuivant notre démarche, nous avons entrepris de mieux caractériser le site de liaison au ribosome, afin de mieux comprendre l'évènement d'initiation interne et d'être capable de le re-situer dans un contexte physiologique.

Dans une première partie de ce travail, nous avons cherché à définir précisément le site d'interaction de la sous-unité 40S sur l'IRES<sub>gag</sub>. Pour ce faire, nous avons utilisé diverses approches complémentaires telles que le footprint, le toeprint, et la délétion progressive de l'IRES à partir des deux extrémités. Cette stratégie ne nous a pas permis de définir de site précis, mais à notre grande surprise, nous avons découvert que la sous-unité 40S pouvait être recrutée à deux endroits de l'IRES<sub>gag</sub>. Nous avons ensuite montré que le recrutement du ribosome sur ces deux endroits n'était pas mutuellement exclusif, autrement dit qu'il est possible de recruter deux sous-unités 40S sur le même morceau d'ARN. Enfin, nous avons étudié l'impact de la présence de ces régions sur la traduction des deux isoformes de la polyprotéine Gag, *in vitro* et *in cellulo*. Nous avons ainsi montré qu'un des deux sites est indispensable à une bonne expression de la polyprotéine p40 alors que la production de la polyprotéine p55 semble insensible à la présence des sites de recrutement du ribosome, au moins dans les conditions que nous avons testées.

L'analyse des régions responsables du recrutement du ribosome fait apparaître qu'elles ne sont probablement pas ou peu structurées, et qu'elles sont très riches en adénosine (A). Une étude récente a montré que des régions riches en A étaient capables de recruter la sous-unité 40S qui est ensuite capable de balayer l'ARN dans les deux directions à la recherche d'un codon d'initiation. Ce mécanisme permet d'expliquer l'essentiel des observations effectuées sur cet IRES à ce jour.

La compréhension de ce mécanisme nous permet de formuler une nouvelle hypothèse quant à la fonction physiologique de cette IRES. En effet, ce mécanisme permet non seulement d'initier la traduction des polyprotéines p55 et p40, mais aussi d'initier la traduction sur des triplets AUG qui ne sont pas en phase avec l'ORF Gag. La traduction à partir de ces triplets permettait la traduction de petits peptides, que nous avons trouvés conservés dans la séquence de nombreux isolats collectés sur le site du NIH de los Alamos (<http://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html> ). La production de certains de ces peptides a été postulée et recherchée. Pour au moins l'un d'entre eux, dont le site d'initiation se trouve juste en aval de l'IRES, il a été montré qu'il existe dans le sang de patients infectés, des lymphocytes T activés qui le reconnaissent. Ainsi, nous émettons l'hypothèse qu'un des rôles de l'IRES serait de permettre la production de petits peptides variés qui influenceraient la réponse du système immunitaire.

**ARTICLE: Two ribosome recruitment sites direct multiple translation initiation events within HIV-1 Gag open reading frame**



## **Two ribosome recruitment sites direct multiple translation initiation events within HIV-1 Gag open reading frame**

Melissa Ameur<sup>1,5</sup>, Jules Deforges<sup>1,2,5</sup>, Sylvain de Breyne<sup>3,5</sup>, Nathalie Ulryck<sup>1</sup>, Nathalie Chamond<sup>1</sup>, Afaf Saadi<sup>4</sup>, Yann Ponty<sup>4</sup>, Theophile Ohlmann<sup>3</sup>, Bruno Sargueil<sup>1</sup>

- (1) CNRS UMR 8015, Laboratoire de cristallographie et RMN Biologiques, Université Paris Descartes, 4 avenue de l'Observatoire, 75270 Paris Cedex 06, France.
- (2) Present adress : Université de Lausanne, Département de biologie moléculaire végétale Bâtiment Biophor, CH-1015 Lausanne, Suisse
- (3) CIRI (International Center for Infectiology Research), INSERM U1111, Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5308, 46 allée d'Italie, 69364 Lyon Cedex 07, France
- (4) LIX
- (5) These authors contributed equally

Key Words: HIV, IRES, translation, ribosome, footprint



## ABSTRACT

In the late phase of the HIV virus cycle, the full length unspliced genomic RNA is exported to the cytoplasm and serves as mRNA to translate the Gag and Gag-pol polyproteins. Three different translation initiation mechanisms responsible for Gag production have been described. However to date a rationale for the involvement of as many translation pathways for gRNA translation is still missing. The Gag-IRES has the singularity to be located within the Gag open reading frame and to directly recruit the 40S ribosomal subunit. To further characterize this interaction, we first probed the Gag-IRES RNA structure. We then developed an innovative integrative modelling approach and propose a novel secondary structure model for the Gag-IRES. The minimal 40S ribosomal subunit binding site was further mapped using different assays. To our surprise, we found that at least two regions within Gag-IRES can independently recruit the ribosome. Next, we validated that these two regions influence Gag translation both *in vitro* and *in cellulo*. These binding sites are mostly unstructured and highly A-rich, such sequences have previously been shown to be sufficient to recruit the ribosome and to support an IRES function. A combination of biochemical and functional data provide compelling evidences for the importance of the Gag IRES. Hypothesis about its physiological role reflecting its conservation amongst primate lentiviruses are proposed.

## **INTRODUCTION :**

Translation in eukaryotic cells largely relies on a strictly cap-dependent initiation mechanism. During this process the ribosome is recruited at the 5' terminus of the mRNA through a series of interactions that indirectly attach the 40S subunit of the ribosome to the cap structure. A minimal set of at least ten cellular proteins, known as eukaryotic Initiation Factors (eIF), is involved. The cap binding protein eIF4E interacts with both the 5' terminus of the mRNA and eIF4G, a central protein binding the RNA helicase eIF4A and its activator eIF4B. The preformed 43S pre-initiation complex (PIC) comprising eIF3, eIF2, the Met-tRNA<sub>i</sub>, eIF1, eIF1A and eIF5 bound onto the 40S small ribosomal subunit is then recruited on the mRNA owing to the eIF4G-eIF3 interaction. The 40S ribosomal subunit does not have any specific affinity for these mRNA, and eIF4E-eIF4G-eIF3 can be seen as a "molecular bridge" linking the small ribosomal subunit to the 5'-cap structure of the RNA. Part or all of this ribonucleoprotein complex then scans the 5'UTR until it reaches the first AUG triplet, where the large ribosomal subunit joins the complex to form a ribosome competent for translation (Jackson et al, 2009). Viruses are obligatory parasites, and evolution has selected many different mechanisms allowing the recruitment of the cellular translation machinery to translate viral mRNA. Of peculiar interest, RNA viruses that are generally replicated by viral RNA-dependent RNA polymerase are not capped. Some of these mRNAs, such as picornaviruses, are translated through a cap-independent initiation mechanism known as the "internal entry of the ribosome" (Jang et al, 1988; Pelletier & Sonenberg, 1988). In this phenomenon, the translation machinery is recruited directly on the initiation codon, or at immediate vicinity. This process has been shown to rely on the secondary and

tertiary structure of the 5'UTR which is then called the Internal Ribosome Entry Site (IRES)(Jackson et al, 2009). IRESes have been identified in many different viruses and have been classified in four major classes. Type I and type II IRESes, epitomized by the Poliovirus and the Encephalomyocarditis virus (EMCV) IRESes respectively, require all the initiation factors but eIF4E. These IRESes have a specific affinity for eIF4G/4A, which in turn recruits the 43S PIC through the interaction between eIF4G and eIF3(de Breyne et al, 2009; Kolupaeva et al, 2003; Lomakin et al, 2000; Pestova et al, 1996; Sweeney et al, 2014). Type III IRES which archetype is found in the Hepatitis C Virus (HCV) directly binds the 40S subunit of the ribosome placing the initiation codon in the P site (Angulo et al, 2016; Pestova et al, 1998; Quade et al, 2015; Yamamoto et al, 2014). In terms of initiation factors, it requires only eIF2 to deliver the Met-tRNA<sub>i</sub>, eIF5 and eIF5B to promote the 60S ribosomal subunit joining(Pestova et al, 1998). Type III IRESes also directly bind eIF3, but do not necessitate this factor to initiate translation *per se*. The suggested role of this interaction is to displace eIF3 from the 40S ribosomal subunit, and to titrate it away from the translation machinery (Hashem et al, 2013). Finally, type IV IRES which was first characterized in the Cricket Paralysis Virus (CrPV), also directly binds the 40S ribosomal subunit of the ribosome and does not require the Met- tRNA<sub>i</sub>. A region of the IRES that mimics the codon-anticodon interaction within the P site promotes the first translocation event leaving the P site available for initiation at a non-AUG triplet(Costantino et al, 2008; Fernandez et al, 2014; Koh et al, 2014; Wilson et al, 2000). Since this classification was first issued, exceptions, and additional IRESes that do not readily fall into these classes have been described(Abaeva et al, 2016; Ali et al, 2001; Borman et al, 2001; Caceres et al, 2016; Olivares et al, 2014; Othman et al, 2014; Terenin et al, 2005; Vallejos et al, 2010). Interestingly type III and IV IRESes have

the characteristic to directly interact with the 40S ribosomal subunit without any initiation factors. In addition, our lab has Recently shown that type II IRESes molecular mechanisms could be slightly amended knowing that the recruitment of the PIC on EMCV IRES has been found to be independent of the eIF3/eIF4G interaction (Lomakin et al, 2000) but rather rely on direct binding of the 40S ribosomal RNA(Chamond et al, 2014).

This ability to directly bind the 40S ribosome has also been reported to RNA sequences located in the human immunodeficiency virus (HIV) genomic mRNA (gRNA) (Locker et al, 2011). The HIV unspliced genomic RNA is exported to the cytoplasm in the late phase of the virus cycle and allows expression of both Gag and Gag-Pol polyproteins. As retroviruses, lentivirus mRNAs are produced by the cellular RNA polymerase II and are therefore capped and polyadenylated. Remarkably, their long 5' UTR (335 and 545 nucleotides for HIV-1 and HIV-2 respectively) conceals some non-cognate initiation codons (CUG, UUG, GUG) but is deprived of AUG triplets (Weill et al, 2010). The counter selection of initiation codons within the 5'UTR is a good indication for a cap-dependent initiation pathway involving ribosome scanning of this region. These two characteristics favour a canonical cap-dependent translation initiation pathway; however additional elements could impede cap-dependent initiation and could therefore be indicative of alternative pathways. First, lentiviral 5'UTR contain many stable structures that have distinct roles along the virus cycle(Baudin et al, 1993; Berkhout, 1996; Dirac et al, 2002; James & Sargueil, 2008; Kenyon et al, 2008). Particularly the 5'UTR begins immediately with the long and very stable TAR stem-loop which is very detrimental for translation (Guerrero et al, 2015; Parkin et al, 1988; Soto-Rifo et al, 2012; Svitkin et al,



1994). Second, the viral Vpr protein induces the cell cycle arrest of infected cells in phase G2/M where the hypo-phosphorylation of eIF4E is known to severely downregulate cap-dependent translation (Pyronnet & Sonenberg, 2001). In this context, several mechanisms allowing cap-dependent translation of HIV-1 gRNA to overcome these detrimental factors have been proposed. In particular, Soto-Rifo et al. demonstrated that the cellular helicase DDX3 is required for translating mRNA which 5' terminus is involved in a stable stem loop, and specifically allows translation of transcripts beginning with TAR (Soto-Rifo et al, 2012). RNA helicase A (also known as DHX9) has also been shown to stimulate cap-dependent translation of HIV-1 mRNA (Bolinger et al, 2010). Independently, it has been hypothesized that HIV-1 gRNA could be translated in G2/M phase, using the cap-binding complex, composed of CBP80/CBP20, instead of eIF4E (Sharma et al, 2012). In parallel, internal ribosome entry within HIV-1 5'UTR has been described and extensively documented (Amorim et al, 2014; Brasey et al, 2003; Buck et al, 2001; Carvajal et al, 2016; Monette et al, 2013), it is stimulated by an oxidative stress (Gendron et al, 2011), in specific cell lines (Plank et al, 2013), and in G2/M phase of the cell cycle (Brasey et al, 2003; Vallejos et al, 2011). In contrast, HIV-2 gRNA does not contain an IRES in the 5'UTR but within the 350 first nucleotides of Gag ORF. It has been shown to direct translation from the *bona fide* Gag initiation codon, but also from two in-frame AUG triplets yielding N-terminally truncated isoforms of Gag (Herbreteau et al, 2005). This occurs in infected Jurkat cells and shorter isoforms can be incorporated in virions (Herbreteau et al, 2005). This IRES known as HIV Gag-IRES has been identified in other primate lentiviruses such as HIV-1 and SIV<sub>MAC</sub> (Buck et al, 2001; Chamond et al, 2010; de Breyne et al, 2012; Locker et al, 2011; Nicholson et al, 2006; Weill et al, 2010). The HIV1 Gag-IRES can direct initiation

at <sup>336</sup>AUG yielding the full length Gag polyprotein (Gag p55) and at <sup>759</sup>AUG to translate a N-truncated isoform named Gag p40. The role of this short isoform is unknown but <sup>759</sup>AUG mutation decreases the mutant virus infectivity ((Buck et al, 2001) and C. Swanson personal communication). Interestingly, the presence of <sup>759</sup>AUG and its nucleotide context are very conserved in clinical isolates derived from HIV-1 positive patients. Furthermore the expression of a p40 truncated Gag isoform has been observed for many of these clinical isolates in both reticulocytes lysate and Jurkat T cells(Daude et al, 2016; de Breyne et al, 2012). Initiation mechanisms involved for both IRESes remain elusive, but the Gag-IRES was shown to have the conserved ability to directly and independently bind the 40S subunit of the ribosome and eIF3 (Locker et al, 2011). Analysis of the initiation complexes paused on HIV-2 Gag-IRES revealed that in addition to the 40S, they contain eIF2, eIF3, eIF4A, eIF4G and eIF5 but neither eIF1 nor eIF4E(Locker et al, 2011).

In summary, translation initiation on the unspliced gRNA can clearly be achieved through at least three different mechanisms (de Breyne et al, 2013; Guerrero et al, 2015; Rojas-Araya et al, 2015). However a role for each of them is still lacking to date. Are they alternatively used at different stage of the virus cycle? Are they cell-type specific? What is the role of the different isoforms? Are they all used indifferently to maximise Gag production? Because the three mechanisms are intricate, it is difficult to assay them separately. For this reason, IRES-dependent Gag translation remains controversial (Miele et al, 1996; Smirnova et al, 2016; Yilmaz et al, 2006). Here, following a "reverse" strategy, we further characterize the Gag-IRES structure and properties focusing on its ability to bind the 40S ribosome to gain insight on its function. This led us to propose an

original mechanism for ribosome entry on this IRES and to suggest a potential physiological role for this non-conventional initiation mode.

## MATERIAL AND METHODS:

### DNA constructs/

The HIV-1 sequence used in this study corresponds to the 5'UTR followed by the *gag* coding region from the NL-4.3 strain (Genbank: AF324493.2). Deletion of the site 1 (nt. 363 to nt. 435), the site 2 (nt. 632 to nt. 736) or both was performed by site-directed mutagenesis with corresponding primers (supplementary table) on the pHIV-1 UTR-AUG (761)-Renilla plasmid (de Breyne et al., 2012) and on the NL-4.3 strain. Fragments of the IRES Gag used for filter-binding assays were generated by PCR using the primers described in the supplementary table. The bicistronic construct pT7-Puromycin/Renilla was built by inserting into the pRenilla vector digested by PvuII and BamHI (Soto-Rifo et al., 2007), a PCR fragment, flanking with the Pvu II restriction site in 5' and a multiple restriction site (Pst I/Sal I/Bam HI) in 3', and coding for the T7 promoter followed by the globin 5'UTR and the puromycin coding region. The IRES *gag* coding region (nt. 331 to nt. 851, WT) and associated deletions (named  $\Delta$ P1,  $\Delta$ P1-P3,  $\Delta$ P1-P4,  $\Delta$ N1,  $\Delta$ P1-N7,  $\Delta$ P8-P10,  $\Delta$ N5-P10,  $\Delta$ N1-P10,  $\Delta$ P4-P10,  $\Delta$ P1-P4/P8-P10 and  $\Delta$ N1/P8-P10) were amplified by PCR with corresponding primers (Supplementary table) to be inserted into the intergenic region (IGR) of the pT7-Puromycin/Renilla plasmid digested by Sal I and Bam HI restriction enzymes. DNA template production for *in vitro* transcription was performed either by digestion of pT7-Renilla and pT7-Puromycin/Renilla plasmids with the Eco RI restriction enzyme, either by PCR using the T7 5'UTR HIV-1 and the Hiv1-3'UTR primer (Supplementary Table).

### ***In vitro* transcription**

Transcription reaction was performed in a final volume of 100  $\mu$ l, using T7 RNA polymerase, 5 mM DTT, 5 mM rNTPs, 1X transcription buffer (40 mM Tris-HCl pH 8.0, 25 mM MgCl<sub>2</sub>, 1 mM spermidine) and 20 U of RNAsin (Promega) and incubated 2 h at 37°C. Upon synthesis, RNAs were treated with 2 U of DNase RQ1 (Promega) for 1 h at 37°C. RNA was precipitated for one hour with 2.5 M LiCl, centrifuged at 16 000 *g* for 30 min at 4°C, washed with 70% ethanol, resuspended in 50  $\mu$ l of nuclease-free water and purified through Sephadex G50 columns. RNAs concentrations were determined spectrophotometrically (Biospec-shimadzu), and RNA integrity was monitored by electrophoresis on agarose gels. The cap and polyA addition were as advised by the manufacturer (New England Biolabs). RNAs were then purified by size exclusion chromatography. Radiolabeled uncapped monocistronic RNAs were transcribed as above but in presence of 1 mM of UTP and 3000 mCi/mmol of  $\alpha$ -<sup>32</sup>P-UTP.

### **SHAPE analysis of HIV1 Gag-IRES**

The secondary structure of HIV-1 Gag-IRES was probed using 1-methyl-7-nitroisatoic anhydride (1M7) or N-methylnitroisatoic anhydride (NMIA) and modifications were revealed by Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) experiments as described in (Deforges et al, 2012; Karabiber et al, 2013; Wilkinson et al, 2006). Briefly, 6 pmol of RNA were resuspended in 48  $\mu$ l of water (for a final RNA concentration of 100 nM) and denatured for 2 min at 80 °C. Then, 6  $\mu$ l of 10X buffer containing 400 mM Hepes pH 8.0, 1 M KCl and 50 mM MgCl<sub>2</sub> were added and the mixture was cooled down to 37 °C for 10 min. 6  $\mu$ l of 20 mM 1M7, or 6  $\mu$ l of DMSO

(negative control) was added and the reaction was incubated for 3 min at 37°C. The RNA was then immediately precipitated at -20°C with ethanol, 0.5 M ammonium acetate and 20 µg of glycogen. Modifications were revealed by elongating fluorescently labelled primers using RNase H minus M-MLV RT (Promega). cDNA fragments were resolved by capillary electrophoresis (Beckman Coulter CEQ 8000). Raw electrophoregrams were converted to normalized SHAPE reactivity using the quSHAPE software (Karabiber et al, 2013). All experiments were repeated at least three times, the mean and the standard deviation value for each nucleotide were reported and analyzed.

### **T1 or V1 RNase footprinting assays**

12 pmol of RNA were resuspended in 96 µl of water (for a final concentration of 100 nM), and were denatured and renatured under the same conditions in the same buffer as described previously for the SHAPE analysis. After a 10 min incubation at 37°C, 3 molar equivalent of the purified 40S ribosomal subunit (300 nM final of 40 S), or the same volume of buffer (sample - 40S) was added, and the mixture was incubated for 10 min at 37 °C. For each condition, the mixture was split in two tubes. In one, 1 µl of RNase V1 (Ambion) diluted at 1/1000 from the commercial stock tube or RNase T1 (Ambion) diluted at 1/10000 was added. In the other tube, 1 µl of DMSO was added (negative control). The mixture was incubated for 10 min at 37 °C. The reaction was blocked by the addition of SDS 0.2 %, EDTA 5 mM, AcONH<sub>4</sub> 0.5 M and 20 ng of glycogen. After phenol chloroform extraction, the RNA was precipitated with 2.5 volumes of ethanol. The cleavage sites were revealed by elongating fluorescently labelled primers, exactly as described for the SHAPE experiment. All experiments were repeated three times and the mean and the standard deviation value for each nucleotide was

reported and analysed. For V1 the thresholds of moderate and strong hit were 0.5 and 1 respectively, while for T1 they were 0.7 and 1. When comparing two profiles, statistically significant increase or decrease of cleavage was determined according to a bilateral student test with an acceptance value of 0.05. Among these, we selected nucleotides that showed at least a 2-fold modification of reactivity and for which the reactivity difference was at least 0.1.

### ***In vitro* translation**

Uncapped and capped mRNAs (200 fmol) were translated for 45 min at 30°C in a 50 % volume of nuclease-treated rabbit reticulocyte lysate (RRL, Promega) in the presence of 75 mM KCl, 0.5 mM MgCl<sub>2</sub>, 20 µM each amino acid (except for methionine), and 0.5 mCi of [<sup>35</sup>S]-methionine per ml.

The reaction was stopped with 90 µl of protein loading buffer (10% SDS, 50 mM Tris pH 6.8, 10% glycerol, 100 mM DTT and 0.1% bromophenol blue). Translation products were then separated by 12 % SDS-PAGE. The gel was fixed for 10 min in a solution of 30% of ethanol and 10% of acetic acid, and dried for 45 min in a vacuum drying system at 75°C. The labeled products were visualized and quantified using a Storm PhosphorImager (GE Healthcare). Each experiment was repeated independently three times using different RNA preparations.

### **Preparation of 40S ribosomal subunits**

Ribosomal subunits were prepared following previously established procedures (Locker et al, 2011). Cytoplasmic extracts from HeLa cells (Ipracell) or rabbit reticulocyte lysate (Green Hectares) were centrifuged in Buffer A (20 mM Hepes-KOH pH 7.5, 50 mM KCl,



4 mM MgCl<sub>2</sub>, 2 mM DTT and 0.25 M sucrose) at 14 000 × *g* for 15 min to remove mitochondria. The supernatant was layered onto a 50% sucrose cushion and centrifuged for 5 h at 44 000 rpm in a 45 Ti rotor. The pellet (P1) was further resuspended in Buffer A (OD<sub>260</sub> ± 100) and treated with 1 mM puromycin (Sigma) for 10 min on ice followed by 10 min at 37°C. KCl was then added to a final concentration of 500 mM and the solution was incubated on ice for 30 min. The suspension was centrifuged for 1 h 40 at 70 000 rpm in a 70 Ti rotor. The obtained pellet (P2) was resuspended in Buffer B (20 mM HEPES–KOH pH 7.5, 0.5 M KCl, 4 mM MgCl<sub>2</sub> and 2 mM DTT) and layered onto a 10–30% sucrose gradient prepared in Buffer B. Gradients were run for 17 h at 22 000 rpm in a SW32 Ti rotor. Fractions were collected and 10 µl of each were loaded onto an agarose gel. Fractions corresponding to 40S ribosomal subunits were pooled, dialyzed and concentrated in Buffer C (20 mM HEPES KOH pH 7.5, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 2 mM DTT and 0.25 M sucrose).

### **Filter binding assay**

Radiolabeled HIV-1 5'UTR\_Gag mRNA wild type and mutants (10 nM) was denatured by heating to 80°C for 2 min followed by cooling to room temperature in FB buffer (20 mM Tris–Cl pH 7.5, 100 mM KOAc, 200 mM KCl, 2.5 mM MgCl<sub>2</sub>, 1 mM DTT). Serial dilutions from a 2 µM solution of 40S were prepared extemporaneously, added to a 10 µl reaction and incubated at 37°C for 30 min. Filter binding was accomplished essentially as previously described using two filters(Locker et al, 2011). From top to bottom a nitrocellulose filter and a charged nylon filter. The filters were presoaked in FB buffer, assembled in the dot blot apparatus and the reactions were applied and directly vacuum

filtered. The filters were then rinsed with FB buffer, removed and radioactivity was quantified using a storm phosphorImager (GE Healthcare).

To determine the apparent  $K_d$ , the data was fit to the Langmuir isotherm described by the equation  $\theta = P/[P + K_d]$  where  $\theta$  is the fraction of RNA bound and P is 40S subunit concentration. Reported values are the average of three repetitions with standard errors. All calculations were performed with Prism 5 (GraphPad).

### **Sucrose density gradients**

Ribonucleoprotein complexes were assembled on  $^{32}\text{P}$ -labeled HIV-1 5'UTR\_Gag mRNA wild type and mutants. 2 pmol of RNA (10  $\mu\text{l}$ ) was denatured by heating the RNA to 85°C for 2 min, followed by cooling to room temperature for 10 min. Determination of complex formation with purified 40S ribosomal subunits (HeLa or RRL) was performed by incubating the RNA (20 nM) with 400 nM of 40S subunits for 15 min at 37°C in buffer (20 mM Tris pH 7.5, 100 mM KOAc, 200 mM KCl, 1 mM DTT, 2.5 mM  $\text{MgCl}_2$ ). The reactions were stopped on ice and layered over 10–30% sucrose gradients (25 mM Tris pH 7.6, 6 mM  $\text{MgCl}_2$ , 75 mM KCl) and sedimented by ultracentrifugation at 39 000 rpm for 4 h at 4°C in a SW40 Ti rotor (Beckman Coulter). Fractions of 300  $\mu\text{l}$  were collected. 100  $\mu\text{l}$  of each fraction were vacuum blotted onto a Hybond N+ membrane (Amersham), exposed and quantified using a Storm PhosphorImager (GE Healthcare). The amount of RNA in each fraction was determined and expressed as the percentage of total counts.

### **Toeprinting assays**

6 pmol of RNA was resuspended in 10  $\mu\text{l}$  of water, denatured for 2 min at 80 °C, then cooled down for 1 min at 37 °C. After addition of 3  $\mu\text{l}$  of fluorescently labelled primers (2

$\mu\text{M}$ ), the mixture was incubated for 5 min at 65 °C, 10 min at 35 °C and cooled down for 1 min on ice. 20  $\mu\text{l}$  of water and 20  $\mu\text{l}$  of 5 X M-MLV buffer was added. The sample was incubated for 10 min at 37 °C. 3 molar equivalent of the 40S ribosomal subunit (final concentration of 300 nM), or an equal volume of buffer was added. The mixture was incubated for 10 min at 37 °C. 5  $\mu\text{l}$  of dNTP mix, and 2  $\mu\text{l}$  of RNase H minus M-MLV RT were finally added. The reaction was incubated for 2 min at 35 °C, 30 min at 42 °C, 5 min at 55 °C and finally cooled down on ice for 1 min. The (+) and (-) 40S subunit reactions, as well as a RNA sequencing reaction were combined in a single tube containing 20  $\mu\text{l}$  of a stop solution (NaOAc 1.2 M, EDTA 50 mM, glycogen 20 ng ). The cDNAs were immediately precipitated with ethanol at -20 °C and resolved by capillary electrophoresis, as described previously. The raw electrophoregrams were processed and analysed using quSHAPE to determine the area under each peak. For each nucleotide, the difference between the area under the peak (+) 40S subtracted from the area under the peak (-) 40S was calculated. Positions with a difference above 30000 A.I. and a p-value < 0.025 were considered as “moderate stop”, while positions with a difference above 50000 A.I. and a p-value < 0.025 were reported as “strong stops”.

### **Stoichiometry Assay:**

The stoichiometry of 40S ribosomal subunit binding to the entire Gag-IRES or the first ribosome binding site was determined by filter binding assay, using conditions previously described (Hellman & Fried, 2007). Briefly, radiolabelled RNA was denatured in water for 1 min at 80 °C and cooled down to at 37 °C in FB buffer (20 mM Tris pH 7.5, 100 mM AcOK, 200 mM KCl, 2.5 mM  $\text{MgCl}_2$  and 1 mM DTT) and the solution was equilibrated for 15 min at 37 °C. Then, different concentrations of the purified 40 S ribosomal subunit

were added (for a final 40S concentration of 750, 375, 187.5, 93.8, 46.9, 23.4 or 0 nM) to 250 nM final of radiolabelled RNA. After a 15 min incubation at 37 °C, the mixture was filtered through a nitrocellulose and a nylon membrane. For each condition, the radioactivity was quantified on both membranes to determine the proportion of free and 40S bound RNA. Results were treated with Prism5 (Graphpad software).

#### **Cell culture - RNA transfection and western blot analysis**

Jurkat T-cell line, obtained from ATCC, was cultured in RPMI-1640 medium supplemented with 10% fetal calf serum, penicillin, streptomycin, Hepes pH 7.5, L Glutamine and sodium pyruvate.

For RNA transfection, 5.10<sup>5</sup> cells were resuspended in 10µl of buffer R in presence of capped mRNAs, and immediately transfected using the Neon transfection kit (Invitrogen). Cells were incubated in RPMI medium for 90min at 37°C and then resuspended in RIPA buffer (150 mM NaCl, 1% deoxycholate, 1% Triton X-100, 0.1% SDS, 10 mM Tris pH 7.8) supplemented with protease inhibitor cocktail (Roche). Cellular extracts were sonicated; equal amount of protein was resolved on a 12% SDS-PAGE and transferred to a PVDF membrane. Proteins of interest were detected by western blot with antibody targeting the HA tag or the HIV-1 p24 (183-H12-5C)(Chesebro et al, 1992; Toohey et al, 1995; Wehrly & Chesebro, 1997) and revealed with the Supersignal West Dura extended Duration signal kit (Pierce).

## RESULTS

### Footprinting the Gag-IRES structure with the 40S ribosome

In order to identify the determinants involved in ribosome binding, we undertook to footprint the 40S subunit on the IRES using chemical and enzymatic probes. First, the isolated HIV-1 Gag-IRES (nucleotide 336 to 851) was incubated with T1 or V1 RNAses, N-Methyl Isatoic Anhydride (NMIA) or 1 Methyl -7 Isatoic Anhydride (1M7). T1 RNase specifically cleaves RNA 3' of single stranded guanosines, whereas V1 RNase cleaves double-stranded positions essentially independently of the nucleotide identity (Lowman & Draper, 1986; Sobczak et al, 2010). NMIA and 1M7 are SHAPE reagents that react with flexible ribose, thus detecting unpaired nucleotides (Mortimer & Weeks, 2007; Wilkinson et al, 2008). The main difference between the two SHAPE reagents is their half-life, resulting in essentially compatible but slightly different modification profiles (Steen et al, 2012). Experiments with each probe were carried out in triplicate, and the results obtained were reported on the secondary structure model proposed by Weill et al. (Weill et al, 2010). The reactivity patterns obtained (Supplementary Figures S1-S4) are essentially in agreement with the model proposed that had been obtained with other probes (Weill et al, 2010). In order to further improve the model, we followed an integrative modelling strategy (See Material and Methods) in which the secondary structure was independently modelled with RNAfold using RNase T1, RNase V1, NMIA or 1M7 reactivity maps as constraints. Two thousand models issued from each modelling batch were naively sampled and the (8000) models gathered were clustered. Six families of structural models were obtained and represented by their centroid. Figure 1A shows the structural model that better fits with all our data. Our novel model features

most of the pairings originally proposed but P5, P6 and P7 for which a new arrangement that takes better the data into account was identified (Pairings N5, N6, N7). In addition, we proposed new pairings (N2 and N3) forming a three way junction between P2 and P3. This new model is compatible with most of the experimental data, meaning that nucleotides in double strand are poorly reactive to 1M7 and are in some occurrences substrate for RNase V1; while unpaired positions are reactive to 1M7 among which some of the Gs are cleaved by RNase T1. However, this observation suffers some exceptions. Most notably, the stretch of nucleotides from 450 and 470 which has been modelled as single stranded contains many positions that are susceptible to V1 while mildly to highly reactive to 1M7. This most probably reflects the presence of unstable pairings or the involvement of this region in an alternative folding. We found some local potential pairings such as  $_{460}\text{AAC}_{463}$  with  $_{471}\text{GUU}_{473}$  or  $_{447}\text{AGC}_{449}$  with  $_{455}\text{GCU}_{457}$ , which are included in the model but are likely unstable. In addition, some nucleotides at the edge of helical regions are susceptible to RNaseT1 and/or reactive to 1M7, such as  $\text{G}_{639}$  or  $\text{G}_{738}$ , this most probably reflect the “breathing” of terminal pairings. Strikingly, nucleotides  $_{377}\text{UAAAUG}_{382}$  in the N2 loop are very poorly reactive to 1M7. This may reflect the existence of a pseudoknot yet to be identified.

The same experiments were repeated but in presence of saturating purified 40S ribosomal subunit. Cleavages or reactivities obtained were compared with those observed in absence of 40S ribosomal subunit. Both experiments were carried in triplicate, positions considered to be more or less reactive toward each of the probes in presence of the ribosome were defined following a statistical approach detailed in the material and methods section. Modification of the reactivity patterns, known as

*footprints*, may have several significations. First the ribosome bound on the RNA may hinder the access of the reagent to the nucleotide, in which case a protection reveals the actual binding site. In this case, steric hindrance from large probes such as RNAses map the global binding site, while small SHAPE reagents reveal short distance interactions. Alternatively, a modification of the reactivity pattern may reveal a structural rearrangement of the RNA upon ribosome binding. This is clearly the case for reactivity enhancement. Note that positions more reactive toward the SHAPE reagent may be close to the ribosome if not involved in an atomic interaction (see examples in (Angulo et al, 2016)). Reactivity alterations in presence of the ribosome were reported on the structure and are shown in Figure 1B (and S1 to S4). Although they appear more concentrated in some regions, the footprints are scattered all over the secondary structure and do not allow to readily define one binding site. However some regions retained our attention such as nucleotides 448 to 458 which clusters six reactivity drop of the three different probes, or N1 loop in which five nucleotides are protected from 1M7 reactivity and one from T1 cleavage. These two regions could conceal important ribosome binding determinants. In contrast, regions such as 419 to 443 where we observe five nucleotides more reactive to 1M7, two less susceptible to V1 cleavage and one more susceptible to T1 are more likely to reveal a RNA conformational change upon ribosome binding. Dispersed footprints could reflect a three dimensional determinant or else, a relatively direct structural rearrangement.

### **Toeprinting the 40S ribosome on Gag-IRES**

To further characterise the interaction between the 40S and the IRES, we undertook to localise the ribosome on the IRES by toe-printing assay. In this assay, the HIV-1 Gag-



IRES RNA is reverse transcribed in the presence or in the absence of the small ribosomal subunit. The reverse transcription premature stops observed only in the presence of the 40S ribosomal subunit are known as “toeprints” and are assumed to reflect the reverse transcriptase bumping against the ribosome. When 48S or 80S complexes are paused with the initiation codon properly accommodated within the P site, toeprints are usually observed 15 to 18 nucleotides downstream the A of the AUG triplet (Dmitriev et al, 2003; Kozak, 1998). In addition, as mentioned above for the footprinting assay, ribosome binding may also modify the secondary structure of the RNA stabilizing some structures and thus inducing premature reverse transcription stops which do not directly reflect the 40S position. This experiment was repeated five times covering each nucleotide at least three times. Twelve unambiguous toeprints of moderate and strong intensity (see material and methods) were observed (Figure 1B and Supplementary Figure S5). They are very reproducible, but the intensity of each of them with the exception of the super strong toeprint at U<sub>748</sub> varies from an assay to another (Sup Fig.S5). Surprisingly, we did not detect any toeprint within the twenty first nucleotides downstream any of the two initiation codons. Consistent with what was observed for the footprinting assay, a first toeprint is observed in the N2 loop while six cluster in the P3 loop (L3). This further suggests that these regions are in close contact with the 40S subunit. Additional three premature RT stops were detected at U<sub>526</sub>, C<sub>537</sub> and C<sub>540</sub> in this region as well as one on C<sub>633</sub>, and a very strong toeprint was consistently observed few nucleotides upstream of the second AUG, at U<sub>748</sub>. All these toeprints are in close proximity of the scattered footprints.

### **Minimal ribosome binding sites as defined by progressive deletion**

Toeprinting and footprinting assays suggest that N2 and P3 loops and probably some sequences downstream are important determinants for ribosome recruitment. However, they do not permit to define the minimal region necessary for binding. We undertook to map the minimal binding site by progressive deletions from both the 5' and the 3' terminus (Figure 2A). In a previous report, we had shown that a fragment of the Gag-IRES ending at nucleotide 572 efficiently binds the ribosome (Locker et al, 2011). To extend these results, we made several truncations ending at C<sub>540</sub> (3' of P2), C<sub>477</sub> (3' end of L3), C<sub>458</sub>, C<sub>433</sub> and C<sub>409</sub>. These RNA fragments were assayed for binding of purified 40S according to filter binding assay. Transcripts terminating at C<sub>540</sub> and C<sub>477</sub> bind the 40S subunit likewise a fragment encompassing the full length IRES (336-851) (Figure 2B). Further elimination of twenty nucleotides within L3 (fragment A<sub>336</sub>-A<sub>458</sub>) increases the K<sub>d</sub> by about two fold (Table 1). For transcripts further truncated within L3, saturation is not reached at 300 nM of 40S impeding the determination of a precise K<sub>d</sub>, but it is evaluated to be superior to 100 nM. The same strategy was carried out from the 5' terminus, and transcripts beginning at A<sub>475</sub>, G<sub>501</sub>, G<sub>547</sub>, G<sub>630</sub> and U<sub>728</sub> were assayed for their ability to bind the 40S ribosome. Surprisingly, transcripts starting at the end of L3 (C<sub>475</sub>-U<sub>851</sub>) which lacks the above defined binding site (A<sub>336</sub> to C<sub>475</sub>) still recruit the 40S ribosome with only a twofold drop of the K<sub>d</sub> (Figure 2C and Table 2). Even more strikingly, further 5' deletion up to G<sub>630</sub> does not have any notable effect on ribosome recruitment. Further deletion in the 3' direction yields an RNA fragment (U<sub>728</sub> to U<sub>851</sub>) that does not bind the ribosome over the background. These two series of experiment defined two non-overlapping RNA fragments A<sub>336</sub>-C<sub>477</sub> and G<sub>501</sub>-U<sub>851</sub> that specifically bind the 40S ribosomal subunit with a significant affinity and specificity (K<sub>d</sub> = 23 ± 2 nM and 48 ± 3 nM respectively). This suggests that the Gag open reading frame actually

harbours two 40S binding sites. In order to precisely map the first binding site, we constructed a series of transcript ending at C<sub>540</sub> while starting 3' to A<sub>336</sub> which were assayed for 40S subunit recruitment. Ribosome binding is not affected for transcripts preserving N2 (A<sub>368</sub> - C<sub>540</sub>), but significantly drops upon N2 deletion (U<sub>396</sub>-C<sub>540</sub>), and is completely lost upon further deletion within L3 (Figure 2D and Table1). This circumscribes a first minimal binding region between A<sub>368</sub> and C<sub>477</sub> (Site 1). In parallel 3' truncations of transcripts starting with A<sub>475</sub> were assayed for purified 40S subunit binding. While deletion of all nucleotides up to G<sub>752</sub> does not significantly alter the K<sub>d</sub>, further truncation within P9 rapidly leads to the loss of significant binding activity (Figure 2E and Table 2). Together with the first series of deletion, these results define a second 40S binding site lying between G<sub>630</sub> to G<sub>752</sub> (Site 2). To confirm these conclusions, the two putative binding sites were synthesized as independent fragments ("site 1" and "site 2"). The fragment lying between the two sites (intersite region: A<sub>500</sub> -G<sub>633</sub>) and the globin gene were used as negative controls (figure 3A and table 3). As expected, the transcript corresponding to Site 1 (G<sub>367</sub>-U<sub>486</sub>) binds the 40S ribosome subunit with an affinity similar to the full length Gag IRES (Figure 3A), while the fragment corresponding to Site 2 (G<sub>630</sub>-A<sub>751</sub>) also independently recruits the ribosome but with a slightly reduced affinity as compared to the full length Gag-IRES. The intersite fragment (A<sub>500</sub>-G<sub>633</sub>) does not show a significant affinity for the 40S subunit, although we observed some more RNA-ribosome complex than when using the globin mRNA (Figure 3A and table 3). In parallel, we followed the complementary approach which consists in deleting the nucleotides corresponding to Site 1 and/or Site 2 from the full length IRES, and evaluate the ability of these fragments to recruit the ribosome. As can be observed in Figure 3B, Site 2 deletion ( $\Delta$  Site 2: A<sub>336</sub>to U<sub>881</sub> $\Delta$ 632-756) does not significantly influences 40S binding,

while a transcript devoid of the region corresponding to Site 1 ( $\Delta$  Site 1: A<sub>336</sub> to U<sub>881</sub> $\Delta$ 363-485) is still able to significantly bind the ribosome although with a slightly reduced affinity. Finally a transcript from which both sites have been deleted ( $\Delta$  Site 1  $\Delta$  Site 2) is no longer able to recruit the 40S subunit (Figure 3B).

### **Two regions can recruit the 40S subunit independently**

The identification of two distinct binding sites raises the question of whether two 40S subunits can be recruited on the same RNA, or else if these two sites are mutually exclusive. In order to answer this question, we determined the maximal ratio of 40S subunit over the full length RNA in a 40S/IRES complex (Ryder et al, 2008). An RNA concentration 10 fold higher than the dissociation constant for the 40S subunit ( 250 nM) was incubated in the presence of increasing amount of 40S subunit. The fraction of RNA bound to the 40S was determined by filter binding assay (Figure 3C). The experiment was carried out with an RNA containing the minimal site one and with the full length RNA. Using the isolated Site 1, 100% of binding would be reached with 1.05 molar equivalent of 40S ribosome which is consistent with the presence of only one site. In contrast, when using the full length IRES, 1.5 molar equivalent would be required to reach saturation. This result is compatible with the presence of two 40S binding sites on this RNA with one site having less affinity than the other.

The 40S subunit-RNA complexes were then analysed on sucrose gradients using the full length IRES, transcripts in which the sequences corresponding to one or the other site had been deleted ( $\Delta$  Site 1 or  $\Delta$  Site 2) and transcripts in which both sites have been removed ( $\Delta$  Site 1  $\Delta$  Site 2) (Figure 3D). In addition to the peak corresponding to the free

RNA, the full length IRES sediments essentially as one 40S subunit, but we also observed one peak in a heavier fraction that could correspond to two 40S ribosomal subunit bound on the same RNA. Deletion of Site 2 yields an RNA that sediments as a complex with a single 40S subunit although with a lower efficiency than the WT construct. In contrast,  $\Delta$  Site 1 RNA essentially runs as a free RNA and only a small amount of complex was repeatedly observed. Finally no high molecular weight complex is observed with the  $\Delta$  Site 1  $\Delta$  Site 2 RNA. Altogether these results show that there are multiple determinants of 40S subunit recruitment along the 400 first nucleotides of Gag open reading frame. Site 1 and Site 2 can recruit the ribosome independently, although site 2 presents a lower affinity and yields complexes that appear unstable on a sucrose gradient.

#### ***Impact of the ribosome binding site on Gag translation in vitro and in cellulo:***

We then assayed the impact of Site 1 and Site 2 ribosome binding sites on Gag translation. First we intended to monitor only the IRES activity *in vitro*, to this end the whole IRES, from <sup>336</sup>AUG to <sup>759</sup>AUG was inserted in the intergenic region of a bicistronic reporter. In this construct <sup>759</sup>AUG is fused to the HA-tagged Renilla luciferase ORF. When this mRNA was used to program a Rabbit Reticulocyte Lysate containing <sup>35</sup>S methionine, we observed the production of the two isoforms confirming once again that the Gag-IRES can direct initiation to the upstream AUG (figure 4A). The intergenic fragment was then progressively truncated from the 5' or the 3' end eliminating ribosome binding site 1 or 2 (figure 4B). All the constructs were designed to preserve an open reading frame encoding a shortened full length protein, but an intact isoform translated from <sup>759</sup>AUG. As observed on figure 4A, progressive deletion in site 1 stimulate

translation at <sub>336</sub>AUG. This is specifically obvious after the elimination of the 5' highly structured domain comprising N1, P2, N2, N3, P3 and P4. In contrast, site 2 deletion does not strongly influence initiation at <sub>336</sub>AUG. Translation from <sub>759</sub>AUG was gradually impaired concomitantly with the progressive deletion of both site 1 and site 2. The impact of the two ribosome binding sites was then evaluated in monocistronic constructs encompassing the 5'UTR and the Gag IRES followed by either the HA-tagged Renilla luciferase ORF (figure 4C) or the Gag ORF (figure 4D). In this context, translation from <sub>336</sub>AUG remains unaffected by the deletion of both sites independently or together (figure 4C and 4D), confirming that under such conditions the translation of p55 Gag is mainly driven by the 5'UTR. In contrast translation of the p40 isoform is slightly reduced upon site 1 deletion, strongly affected after site 2 elimination and almost abolished when both ribosome binding sites were removed. Translation efficiency of the same constructs *in cellulo* was then addressed by directly transfecting the different capped and polyadenylated RNA in Jurkat cells and analysing Renilla (figure 5A) or Gag (figure 5B) production by western blot analysis. Results were similar to those observed *in vitro*, in summary, Site 1 deletion does not affect translation, while deletion of Site 2 abrogates the shortest isoform translation without affecting Gag p55 translation. Finally, only the Full length Gag protein is produced when using an artificial construct deleted for both sites ( $\Delta$  Site 1  $\Delta$  Site 2)(Figure 5A and 5B). In conclusion, Site 2 is strictly required for translation from <sub>759</sub>AUG which is expected to be essentially IRES dependent. Regarding the full length Gag, most of the translation observed using 5'UTR-containing constructs rather reflect cap-dependent initiation preventing the observation of a IRES-dependent translation.

## **DISCUSSION:**

### **A novel secondary structure model obtained by an integrative approach:**

In this study, the structure of an RNA corresponding to the 515 first nucleotides of HIV-1 Gag open reading frame has been probed with two different chemical reagents and two specific RNAses. Although these data were originally produced to evidence and characterize 40S ribosomal subunit binding to this RNA, we took advantage of their availability to refine the Gag-IRES secondary structure model. Several software, such as mfold(Zucker, 2003), RNAstructure(Xu & Mathews, 2016) or RNAfold (Hofacker, 2009) which are based on thermodynamical data generate secondary structure models using experimental data as constraints. These software generate many different models, rank them according to their predicted free energy and yield one or several models for the experimenter to choose. Most of the time, the resulting models are mostly, but not completely in agreement with the probing data. This can be due to model inaccuracy or to experimental artefacts, but in many cases most probably reflects the inherent property of RNA to fold into several conformations. In addition, each probe has its own specificity and can reflect slightly different things. For example some “single strand” specific reagents react with nucleotides involved in non-canonical base-pairs while others do not (Steen et al, 2012). As a result, the modelling process often leaves the experimenter with different RNA structure models among which it is often difficult to choose. Here, we



have developed an integrative approach which identifies the model that fits the best with the reactivity map obtained with different structural probes. The structure probing and modelling of this portion of Gag ORF has been reported in two previous studies (Siegfried et al, 2014; Watts et al, 2009). Both rely on the SHAPE technology and have been conducted in the context of the full length genomic RNA. In the first report, nucleotides reactive to 1M7 were mapped through electrophoresis as in this study, while the second used the recently developed SHAPE-Map technology which relies on Next Generation Sequencing (Siegfried et al, 2014). The 1M7 reactivity maps presented in both reports are in good agreement although not identical. Both are also globally comparable to the 1M7 reactivity map reported in this study, but for two noticeable exceptions. The nucleotides between U<sub>559</sub> and U<sub>570</sub> and those involved in N2 stem were found to be mostly reactive in both studies while we found them poorly reactive. We have no explanation for such discrepancy although it could reflect an alternative folding of these regions in the isolated IRES as compared to the whole gRNA. Although grounded on similar reactivity maps, the three studies proposed three different models. Four pairings are identical (P3 and N5) or similar (N6 and N7) in the three models. The model proposed by Watts et al. which is more similar to ours also predicted P4 and P8. In both models proposed by the Weeks laboratory, the sequence <sub>363</sub>GGGGG<sub>368</sub> is base paired <sub>746</sub>CCUCC<sub>750</sub> instead of forming P2 with <sub>535</sub>CCCUU<sub>539</sub> as in our model. This roots the difference between their models and ours, but it is interesting to note that both possibilities are essentially compatible with the three reactivity maps. In summary, the main difference is the modelling strategy and not the experimental results. Here we propose a novel integrative strategy that reflects the results obtained with various probes, and smooths the effect of potential artefacts that may reside in one particular

experimental set up. As the model obtained does not give a rationale to all the probing data, it should be considered as a working model to be further refined.

#### **40S ribosomal subunit recruitment:**

We had previously shown that 40S recruitment by the Gag open reading frame is a conserved property among primate lentiviruses. Here we set out to better define the minimal binding site and its structure. Using chemical and enzymatic probes, we observed ribosome footprints dispatched all over the sequence although they are more abundant in some regions. In addition to reactivity protection that could reveal sites of direct contact with the ribosome, we observed reactivity enhancement to 1M7 and T1 nuclease which rather indicate the destabilisation of the structure upon ribosome binding. The dispersion of the footprints could also suggest the existence of a large contact surface between the ribosome and a compactly folded IRES as observed for type III and IV IRESes. Therefore we were surprised to observe that non-overlapping fragments of the Gag-IRES can independently recruit the 40S ribosomal subunit. These two fragments appear to have a slightly different affinity for the small ribosomal subunit, but are not cooperative as expected if these were two fragments of a same site. Finally, stoichiometry determination and sucrose gradients showed that the full length IRES is able to recruit two 40S subunits while Site 1 or Site 2 deleted fragments are not. The two independent sites thus defined are Site 1 lying from A<sub>368</sub> to C<sub>477</sub>, and Site 2 encompassing nucleotides from G<sub>634</sub> to G<sub>752</sub>. Consistent with this finding, these two regions concentrate most of the toeprints and footprints observed. Most notably, L3

gather 18 footprints and 6 toeprints, while site two comprises U<sub>748</sub> where we observed the strongest toeprint. Somehow confusing, we noticed that except if they fold in alternative conformation when isolated, the two sites are mostly unstructured. Site 1 would retain N2 and few unstable short stem-loops in L3, while Site 2 features P8 and P9. This suggests that binding determinants may rather lie in the sequence than in the RNA structure. We therefore looked for sequences that could base pair with the 18S rRNA. Among others, we found a twelve nucleotide long sequence in N2 <sub>370</sub>AAUUAGAUAAAU<sub>381</sub> complementary to <sub>1368</sub>UUGGUCUGUUUUAG<sub>1356</sub> of the 18S rRNA. Interestingly, a similar sequence that could base pair with the 18S rRNA is also found in an equivalent region of HIV-2 and SIV<sub>Mac</sub>. The eight first positions (370 to 377) are over ninety percent conserved in amongst the HIV-1 isolates collected at the “HIV sequence database” of Los Alamos. However, this region also encodes a conserved sequence of the Gag polyprotein which could account for the nucleotide conservation. Adenosine rich sequence which have been shown to recruit the 40S subunit and to be involved in atypical IRES activity(Abaeva et al, 2016; Dorokhov et al, 2002; Gilbert et al, 2007; Shirokikh & Spirin, 2008). This led us to investigate Adenosine distribution in the HIV-1 genome (sequence HBX2 Accession Number K03455.1). As previously described(Berkhout & van Hemert, 1994), there is a bias in the nucleotide distribution in both Gag open reading frame and HIV full genome which are about 34% A-rich. This trend is accentuated in the Gag-IRES (from <sub>336</sub>AUG to <sub>759</sub>AUG) which is 39% rich. This is in contrast with the isolated 5' and 3' UTR which are 24% A-rich. We then used a 20 nucleotide sliding window to examine the local abundance in adenosine residues. As can be observed in Figure 6, two regions appear to be particularly A-rich with up to 15 Adenosine residues out of twenty. These A-rich sequences culminate in A<sub>419</sub> and G<sub>667</sub>

which are the centre of Site 1 and Site 2, respectively. In contrast, in the region in between the two sites, there is never more than 8 Adenosine out of 20 residues, and the structured region between U<sub>476</sub> and U<sub>539</sub> is “only” 30% A-rich. This is nevertheless over 25% and could explain how this sequence shows some residual affinity for the 40S subunit. We therefore make the proposition that its high adenosine content endows the Gag-IRES with the capacity of recruiting the ribosome in multiple sites, and especially in two hot spots that we have defined as Site 1 and Site 2. This may be modulated by sequences complementary to the 18S rRNA, and independently by the IRES-Gag structure hindering or exposing A-rich sequences.

### **IRES driven translation and physiological consequences**

The molecular mechanism underlying translation initiation driven by A-rich IRES has been recently reported by Abaeva et al. (Abaeva et al, 2016). In summary, the 40S subunit is recruited in one or several A-rich regions within the IRES, then the ribosome subunit scans the surrounding region for an initiation triplet, not only in the canonical 5' to 3' direction, but also and even preferentially in the 3' to 5' direction. The “backward” scanning itself may only require the 40S subunit, but initiation on such IRES also requires eIF3, eIF2 and the Met- tRNA<sup>i</sup>. In addition, eIF4G/eIF4A favours 5' to 3' scanning, and eIF1 helps to discriminate AUG triplets (Abaeva et al, 2016). Such a mechanism offers a rationale for most of the finding on HIV-1 and HIV-2 gag-IRES reported in this study and in earlier works (Balvay et al, 2007; Buck et al, 2001; Chamond et al, 2010; de Breyne et al, 2012; Herbreteau et al, 2005; Locker et al, 2011; Ricci et al, 2008a; Ricci et al, 2008b; Weill et al, 2010). In such a hypothesis, ribosomal small subunit would be recruited within L3 and/or N2 loop for Site 1, and in the single strand

A-rich sequence between <sub>658</sub>A and <sub>696</sub>A. Ribosomes would then scan either backward or forward until they encounter an initiation triplet. Note that in the absence of the Met-tRNA, 40S progression may be arrested only by stable structures. As a consequence, 40S ribosomal subunit may be trapped within L3 of Site 1 while those recruited on Site 2 are relatively free to diffuse until the 3' terminus. This could explain the relative instability on sucrose gradients of complexes recruited on  $\Delta$  Site 1 constructs and is reminiscent of what was observed for the Halastavi Arva Virus (Abaeva et al, 2016). This also gives a rationale for the presence of footprints and toeprints outside of the recruiting region, and for the efficient translation of transcripts lacking large part of the IRES ((Ricci et al, 2008a) and this work). Such mechanism is compatible with the originally proposed provocative idea of an IRES located downstream the initiation codon (Herbreteau et al, 2005). In this respect, it is interesting to note that Site 1 deletion brings <sub>336</sub>AUG close to Site 2 and eliminates most of the secondary structure between the recruitment site and the initiation codon. This probably accounts for the stimulation of p55 Gag translation observed upon P1-P4 deletion in a bicistronic context where only IRES dependant initiation is observed.

The presence of the Gag-IRES and of the shorter isoforms translation is both conserved amongst primate lentiviruses and in clinical isolates derived from HIV-1 infected patients (Daude et al, 2016; de Breyne et al, 2012; Weill et al, 2010). Here by a combination of biochemical and functional analysis we have characterized two redundant ribosome binding sites driving p40 expression and modulating p55 in some contexts. These results converge with the conservation of p40 initiation sites (Daude et al, 2016) and its requirement for optimal infectivity ((Buck et al, 2001) and C.Swanson pers. com.)

providing compelling evidences of the importance of the Gag IRES. However, translation of the Gag full length polyprotein can clearly be driven by a canonical cap mechanism stimulated by a cellular helicase(Soto-Rifo et al, 2012; Soto-Rifo et al, 2013), or by another IRES located within the 5'UTR. In this context, full length Gag translation through the Gag-IRES may be required at a very specific stage of the viral life cycle, or in specific cell type. Alternatively, the observation of the full length Gag as a result of ribosome entry within the Gag ORF could be a side effect of a physiological phenomenon selected for other reasons. Highly structured viral IRESes such as HCV (type III) or CrPV (type IV) IRES strongly constrain the position of the ribosome on the initiation site, although to a lesser extent, this holds true for type I and type II IRESes. In the case of the “poly-A” type IRES where the ribosome is recruited in many locations and scans both forward and backward, virtually any AUG triplet can be used as an initiation codon(Abaeva et al, 2016). Indeed, translation from illegitimate initiation codons within the Gag-IRES has been observed in several occasions. For instance, translation can initiate from triplets that were introduced by site-directed mutagenesis(Herbreteau et al, 2005), or that are uniquely present in patients isolates(de Breyne et al, 2012), or which are non-AUG initiation codons(Ricci et al, 2008a; Weill et al, 2010), or downstream <sub>749</sub>AUG (see figure 4B). All these initiation events could be observed because they result from initiation at triplets in phase with the Gag ORF thus yielding relatively long proteins. More importantly, this raises the possibility that numerous significant translation events happen from non-in frame initiation codons yielding peptides unrelated to Gag. Interestingly enough, one such peptide which translation is initiated at <sub>797</sub>AUG has been shown to be immunogenic and activated T-cells against this peptide have been found in patients' blood(Bansal et al, 2010; Berger et al, 2010;

Cardinaud et al, 2011; Cardinaud et al, 2004). Initiation at  $_{797}$ AUG can undoubtedly be driven from the Gag-IRES, since we recurrently observe the translation of shorter isoforms initiated downstream from  $_{759}$ AUG. In addition to the production of p40, translation of such peptide could be the main role of HIV Gag-IRES, and explain the conservation of this mechanism amongst primate lentiviruses.

### **ACKNOWLEDGMENTS:**

The authors wish to thank C. Swanson for sharing unpublished data. The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HIV-1 p24 Monoclonal Antibody (183-H12-5C) from Dr. Bruce Chesebro and Kathy Wehrly. Research in B.S. and T.O laboratory was financed by ANRS (AO 2012-2 – AO2014-2), M.A is a recipient of a PhD fellowship from the French minister for research and education (Allocation MRE). A.S is a recipient of a PhD fellowship from the “Fondation pour la Recherche Medicale”. Many thanks to G. De Bisschop for help with figure 6.

**TABLES:**

RNA	Nucleotides	Kd (nM) Mean $\pm$ s.e.m.
Gag-IRES	A <sub>336</sub> -U <sub>851</sub>	25.1 $\pm$ 4.8
	A <sub>336</sub> -C <sub>540</sub>	23.2 $\pm$ 2.4
	A <sub>336</sub> -C <sub>477</sub>	27.3 $\pm$ 3.9
	A <sub>336</sub> -A <sub>458</sub>	49.3 $\pm$ 5.6
	A <sub>336</sub> -A <sub>433</sub>	N.D (> 100)
	A <sub>336</sub> -G <sub>409</sub>	N.D (> 100)
	A <sub>368</sub> -C <sub>540</sub>	30.1 $\pm$ 5.6
	U <sub>396</sub> -C <sub>540</sub>	N.D (>100)
	A <sub>431</sub> -C <sub>540</sub>	N.D (>200)
	G <sub>459</sub> -C <sub>540</sub>	N.D (>200)
	A <sub>475</sub> -C <sub>540</sub>	N.D (>200)

**Table 1:** 40S binding affinities of Gag-IRES Site1 deletion mutants as evaluated by filter binding assay. The results are the mean and standard deviation of at least three independent experiments. N.D : not determined (saturation was not reached).



RNA	Nucleotides	Kd (nM) Mean $\pm$ s.e.m.
Gag-IRES	A <sub>336</sub> -U <sub>851</sub>	25.1 $\pm$ 4.8
	A <sub>475</sub> -U <sub>851</sub>	55.6 $\pm$ 6.8
	A <sub>475</sub> -G <sub>752</sub>	70.6 $\pm$ 10.5
	A <sub>475</sub> -G <sub>656</sub>	N.D (>200)
	A <sub>475</sub> -C <sub>634</sub>	N.D >200
	G <sub>501</sub> -U <sub>851</sub>	48.8 $\pm$ 3.7
	G <sub>547</sub> -U <sub>851</sub>	59.7 $\pm$ 11.8
	G <sub>630</sub> -U <sub>851</sub>	61.4 $\pm$ 8.5
	U <sub>728</sub> -U <sub>851</sub>	N.D (>400)

**Table 2:** 40S binding affinities of Gag-IRES Site2 deletion mutants as evaluated by filter binding assay. The results are the mean and standard deviation of at least three independent experiments. N.D : not determined (saturation was not reached).

RNA	Nucleotides	Kd (nM)
		Mean ± s.e.m.
IRES-Gag	A <sub>336</sub> -U <sub>851</sub>	25.1 ± 4.8
Δ Site1	A <sub>336</sub> -U <sub>881</sub> Δ363-485	41.8 ± 14.6
Δ Site2	A <sub>336</sub> -U <sub>881</sub> Δ632-756	22.0 ± 6.6
Δ Site1 Δ Site2	A <sub>336</sub> -U <sub>881</sub> Δ363-485Δ632- 756	N.D (>300)
Site1	G <sub>367</sub> -U <sub>486</sub>	18.6 ± 5.3
Site2	G <sub>630</sub> -A <sub>751</sub>	28.2 ± 8.7
Intersite region	A <sub>500</sub> -G <sub>633</sub>	ND (>300)

**Table 3:** 40S binding affinities of the two sites as evaluated by filter binding assay. The results are the mean and standard deviation of at least three independent experiments. N.D: not determined (saturation was not reached).

## FIGURE LEGENDS:

Figure 1: **RNA secondary structure model of the HIV1 Gag-IRES and 40S ribosome subunit footprints and toeprints. (A) Schematic representation of the secondary structure model of Gag-IRES.** Nucleotides are coloured according to their reactivity toward 1M7 as indicated in the box. Red triangles and blue dots represent RNase T1 and V1 cleavages. Pairings numbering ( $P_n$  and  $N_n$ ) are as described in the text. Nucleotide numbering is from the +1 of transcription (First nucleotide of TAR). Experimental values result from the mean of three independent experiments (see material and methods and supplementary material). **(B) Footprints and toeprints of the 40S ribosomal subunit on Gag-IRES.** As indicated in the box, the colours and signs compare the reactivities without and with saturating concentration of 40S ribosomal subunit. Nucleotides in red are more reactive to 1M7, while nucleotides in blue are less. Red and Blue triangles indicate nucleotides exposed or protected to RNase T1 respectively. Red and Blue dots indicate nucleotides protected or exposed to RNase V1 respectively. Black Arrows indicate premature RT stops observed in the presence of 40S ribosomal subunits. Small and big arrows are respectively moderate and strong toeprints. Experimental values result from the mean of three independent experiments (see material and methods and supplementary material).

Figure 2: **Mapping the 40S ribosomal subunit on HIV1 Gag-IRES using truncated fragment of the IRES. (A)** Schematic representation of the fragments used in this study. Nucleotides numbering is from the +1 of transcription (first nucleotide of Tar). **(B) to E)** Binding curves of  $^{32}\text{P}$ -labelled fragments to purified 40S ribosomal subunits as measured by filter binding assays (see Materials and Methods). Fragments were

obtained by successive deletions from the 3' (**B**) and the 5' (**C**) of Gag-IRES, or from the 5' of A<sub>336</sub>-C<sub>540</sub> fragment (**D**) or by the 3' of the C<sub>477</sub>-U<sub>851</sub> fragment. The results are the mean of at least three independent experiments ± standard deviation.

**Figure 3: Two sites within the Gag IRES can independently recruit the 40S ribosomal subunit.** (**A**) Binding curves of <sup>32</sup>P-labelled Site 1 (■, G<sub>367</sub>-U<sub>486</sub>), Site 2 (▲, G<sub>630</sub>-A<sub>751</sub>), full length HIV-1 Gag-IRES (●, A<sub>336</sub>-U<sub>851</sub>), the intersite region (▼), and the globin transcript (○) with purified 40s ribosomal subunits. (**B**) Binding curves for HIV-1 Gag-IRES deleted of site 1 (Δ Site 1: A<sub>336</sub>-U<sub>881</sub>Δ363-485), of site 2 (Δ Site 2: A<sub>336</sub>-U<sub>881</sub>Δ632-756), or of both sites (Δ Site 1 Δ Site 2: A<sub>336</sub>-U<sub>881</sub>Δ363-485Δ632-756) as compared to the full length HIV-1 Gag-IRES (A<sub>336</sub>-U<sub>851</sub>), and to the globin transcript. (**C**) Determination of the stoichiometry of the 40S ribosomal subunit and the Gag-IRES RNA. A constant concentration of RNA corresponding to the whole IRES ■ or to the isolated site 1 (G<sub>367</sub>-U<sub>486</sub>) ●, tenfold over the K<sub>d</sub> of the Gag-IRES/40S complex (250 nM) was incubated with increasing concentration of 40S ribosome. The proportion of RNA in complex in the mix was determined by filter binding assay. The % of complex was plotted as a function of [RNA] equivalent of 40S, the linear part of the curve was calculated drawn, and extrapolated to 100% binding. The values are the mean of at least three independent experiments ± standard deviation. The two curves are statistically different ( $P < 10^{-4}$ ). (**D**) Analysis of the complexes formed by the full length HIV-1 Gag-IRES deleted of site 1 (Δ Site 1: A<sub>336</sub>-U<sub>881</sub>Δ363-485), of site 2 (Δ Site 2: A<sub>336</sub>-U<sub>881</sub>Δ632-756), or of both sites (Δ Site 1 Δ Site 2: A<sub>336</sub>-U<sub>881</sub>Δ363-485Δ632-756) and the full length HIV-1 Gag-IRES (A<sub>336</sub>-U<sub>851</sub>) with the 40S ribosomal subunit. <sup>32</sup>P-labelled RNA

transcripts were separated on 10-30% sucrose gradients . Peaks were identified by comparison with UV profiles obtained with purified 40S ribosomal subunits.

Figure 4: **The presence of the ribosome binding sites influence the translation of HIV1 gag short isoform *in vitro* (p40).** (A) Schematic representation of the fragments of the Gag-IRES inserted in a bicistronic reporter for *in vitro* translation assays (nucleotides numbering are from the +1 of transcription). (B) The RRL was programmed for 45 min at 30°C with bicistronic mRNAs that contain the intergenic regions the described in panel A (lane 2 to 13), or with water (lane 1). Protein products were resolved on a 12 % SDS-PAGE and quantified using a PhosphorImager. Expression initiated at both  $_{336}$ AUG and  $_{759}$ AUG codons was normalized to the expression level observed with the full length IRES, and to the expression of the first cistron. Data are representative of three independent experiments error bars are  $\pm$  the standard deviation. (C) Capped and polyadenylated monocistronic mRNAs (200fmol) coding for the 5'UTR followed by the Gag-IRES (WT,  $\Delta$ site 1,  $\Delta$ site2 and  $\Delta$ site 1/ $\Delta$  site2) and the *Renilla* reporter gene were translated in RRL. After 45 min incubation at 30°C, protein products of three independent experiments were analyzed and quantified as above. (D) The HIV-1 5'UTR-GAG (WT,  $\Delta$ site 1,  $\Delta$ site2 and  $\Delta$ site 1/ $\Delta$  site2) mRNAs were used to program RRL and products of translation were run on a 12% SDS-PAGE and quantified. (B), (C) and (D) black and grey squares of the histograms represent the quantification of translation initiated at  $_{336}$ AUG (p55) and  $_{759}$ AUG (p40) respectively.

Figure 5: **The presence of the ribosome binding sites influence the translation of HIV1 gag short isoform *in cellulo*.** Jurkat T-cells were transfected with (A) a series of construct in which HIV-1 5'UTR and Gag-IRES have been fused with the HA-tagged

renilla luciferase ORF in <sub>759</sub>AUG. WT-Ren (Full length Gag-IRES),  $\Delta$ S1-Ren ( $\Delta$  Site 1: A<sub>336</sub>-U<sub>881</sub> $\Delta$ 363-485),  $\Delta$ S2-Ren( $\Delta$  Site 2: A<sub>336</sub>-U<sub>881</sub> $\Delta$ 632-756) and  $\Delta$ S1+2-Ren ( $\Delta$  Site 1  $\Delta$  Site 2: A<sub>336</sub>-U<sub>881</sub> $\Delta$ 363-485 $\Delta$ 632-756) mRNAs (lanes 2-5), or **(B)** with a series of construct carrying HIV-1 5'UTR and Gag ORF. WT-GAG,  $\Delta$ S1-GAG ( $\Delta$  Site 1: A<sub>336</sub>-U<sub>881</sub> $\Delta$ 363-485) ,  $\Delta$ S2-GAG ( $\Delta$  Site 2: A<sub>336</sub>-U<sub>881</sub> $\Delta$ 632-756) and  $\Delta$ S1+2-GAG ( $\Delta$  Site 1  $\Delta$  Site 2: A<sub>336</sub>-U<sub>881</sub> $\Delta$ 363-485 $\Delta$ 632-756) mRNAs (lanes 2-5), or water (lane 1) and then incubated for 90 min. Cellular extracts were sonicated and equal amount of proteins were loaded on 12% SDS-PAGE. Proteins of interest were revealed by western blotting. Data are representative of three independent experiments.

**Figure 6: Adenosine distribution in HIV-1 Gag open reading frame.** For each nucleotide within the 5' UTR and the whole Gag ORF of the PNL4.3, the number of Adenosine within the 20 surrounding positions (9 in 5' and 10 in 3') was plotted. Two dotted lines indicate the level of 25% Adenosine (As expected in a random distribution) and 35% (as the mean observed in Gag ORF).The nucleotides corresponding to the two major peaks (A<sub>416</sub> and A<sub>669</sub>) are indicated.

## REFERENCES:

Abaeva IS, Pestova TV, Hellen CU (2016) Attachment of ribosomal complexes and retrograde scanning during initiation on the Halastavi arva virus IRES. *Nucleic Acids Res* **44**: 2362-2377

Ali IK, McKendrick L, Morley SJ, Jackson RJ (2001) Activity of the hepatitis A virus IRES requires association between the cap-binding translation initiation factor (eIF4E) and eIF4G. *J Virol* **75**: 7854-7863.

Amorim R, Costa SM, Cavaleiro NP, da Silva EE, da Costa LJ (2014) HIV-1 transcripts use IRES-initiation under conditions where Cap-dependent translation is restricted by poliovirus 2A protease. *PLoS One* **9**: e88619

Angulo J, Ulryck N, Deforges J, Chamond N, Lopez-Lastra M, Masquida B, Sargueil B (2016) LOOP IIIId of the HCV IRES is essential for the structural rearrangement of the 40S-HCV IRES complex. *Nucleic Acids Res* **44**: 1309-1325

Balvay L, Lastra ML, Sargueil B, Darlix JL, Ohlmann T (2007) Translational control of retroviruses. *Nat Rev Microbiol* **5**: 128-140

Bansal A, Carlson J, Yan J, Akinsiku OT, Schaefer M, Sabbaj S, Bet A, Levy DN, Heath S, Tang J, Kaslow RA, Walker BD, Ndung'u T, Goulder PJ, Heckerman D, Hunter E, Goepfert PA (2010) CD8 T cell response and evolutionary pressure to HIV-1 cryptic epitopes derived from antisense transcription. *The Journal of experimental medicine* **207**: 51-59

Baudin F, Marquet R, Isel C, Darlix JL, Ehresmann B, Ehresmann C (1993) Functional sites in the 5' region of human immunodeficiency virus type 1 RNA form defined structural domains. *J Mol Biol* **229**: 382-397

Berger CT, Carlson JM, Brumme CJ, Hartman KL, Brumme ZL, Henry LM, Rosato PC, Piechocka-Trocha A, Brockman MA, Harrigan PR, Heckerman D, Kaufmann DE, Brander C (2010) Viral adaptation to immune selection pressure by HLA class I-restricted CTL responses targeting epitopes in HIV frameshift sequences. *The Journal of experimental medicine* **207**: 61-75

Berkhout B (1996) Structure and function of the human immunodeficiency virus leader RNA. *Prog Nucleic Acid Res Mol Biol* **54**: 1-34

Berkhout B, van Hemert FJ (1994) The unusual nucleotide content of the HIV RNA genome results in a biased amino acid composition of HIV proteins. *Nucleic Acids Res* **22**: 1705-1711

Bolinger C, Sharma A, Singh D, Yu L, Boris-Lawrie K (2010) RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. *Nucleic Acids Res* **38**: 1686-1696

Borman AM, Michel YM, Kean KM (2001) Detailed analysis of the requirements of hepatitis A virus internal ribosome entry segment for the eukaryotic initiation factor complex eIF4F. *J Virol* **75**: 7864-7871

Brasey A, Lopez-Lastra M, Ohlmann T, Beerens N, Berkhout B, Darlix JL, Sonenberg N (2003) The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. *J Virol* **77**: 3939-3949

Buck CB, Shen X, Egan MA, Pierson TC, Walker CM, Siliciano RF (2001) The human immunodeficiency virus type 1 gag gene encodes an internal ribosome entry site. *J Virol* **75**: 181-191.

Caceres CJ, Contreras N, Angulo J, Vera-Otarola J, Pino-Ajenjo C, Llorian M, Ameur M, Lisboa F, Pino K, Lowy F, Sargueil B, Lopez-Lastra M (2016) Polypyrimidine tract-binding protein binds to the 5'untranslated region of the mouse mammary tumor virus mRNA and stimulates cap-independent translation initiation. *The FEBS journal*

Cardinaud S, Consiglieri G, Bouziat R, Urrutia A, Graff-Dubois S, Fourati S, Malet I, Guernon J, Guihot A, Katlama C, Autran B, van Endert P, Lemonnier FA, Appay V, Schwartz O, Kloetzel PM, Moris A (2011) CTL escape mediated by proteasomal destruction of an HIV-1 cryptic epitope. *PLoS pathogens* **7**: e1002049

Cardinaud S, Moris A, Fevrier M, Rohrlich PS, Weiss L, Langlade-Demoyen P, Lemonnier FA, Schwartz O, Habel A (2004) Identification of cryptic MHC I-restricted epitopes encoded by HIV-1 alternative reading frames. *The Journal of experimental medicine* **199**: 1053-1063

Carvajal F, Vallejos M, Walters B, Contreras N, Hertz MI, Olivares E, Caceres CJ, Pino K, Letelier A, Thompson SR, Lopez-Lastra M (2016) Structural domains within the HIV-1 mRNA and the ribosomal protein S25 influence cap-independent translation initiation. *The FEBS journal* **283**: 2508-2527



Chamond N, Deforges J, Ulryck N, Sargueil B (2014) 40S recruitment in the absence of eIF4G/4A by EMCV IRES refines the model for translation initiation on the archetype of Type II IRESs. *Nucleic Acids Res* **42**: 10373-10384

Chamond N, Locker N, Sargueil B (2010) The different pathways of HIV genomic RNA translation. *Biochem Soc Trans* **38**: 1548-1552

Chesebro B, Wehrly K, Nishio J, Perryman S (1992) Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell-tropic isolates: definition of critical amino acids involved in cell tropism. *J Virol* **66**: 6547-6554

Costantino DA, Pfingsten JS, Rambo RP, Kieft JS (2008) tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat Struct Mol Biol* **15**: 57-64

Daude C, Decimo D, Trabaud MA, Andre P, Ohlmann T, de Breyne S (2016) HIV-1 sequences isolated from patients promote expression of shorter isoforms of the Gag polyprotein. *Arch Virol*

de Breyne S, Chamond N, Decimo D, Trabaud MA, Andre P, Sargueil B, Ohlmann T (2012) In vitro studies reveal that different modes of initiation on HIV-1 mRNA have different levels of requirement for eIF4F. *The FEBS journal* **279**: 3098-3111

de Breyne S, Soto-Rifo R, Lopez-Lastra M, Ohlmann T (2013) Translation initiation is driven by different mechanisms on the HIV-1 and HIV-2 genomic RNAs. *Virus Res* **171**: 366-381

de Breyne S, Yu Y, Unbehaun A, Pestova TV, Hellen CU (2009) Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proc Natl Acad Sci U S A* **106**: 9197-9202

Deforges J, Chamond N, Sargueil B (2012) Structural investigation of HIV-1 genomic RNA dimerization process reveals a role for the Major Splice-site stem loop. *Biochimie* **94**: 1481-1489

Dirac AM, Huthoff H, Kjemis J, Berkhout B (2002) Regulated HIV-2 RNA dimerization by means of alternative RNA conformations. *Nucleic Acids Res* **30**: 2647-2655

Dmitriev SE, Pisarev AV, Rubtsova MP, Dunaevsky YE, Shatsky IN (2003) Conversion of 48S translation preinitiation complexes into 80S initiation complexes as revealed by toeprinting. *FEBS Lett* **533**: 99-104

Dorokhov YL, Skulachev MV, Ivanov PA, Zvereva SD, Tjulkina LG, Merits A, Gleba YY, Hohn T, Atabekov JG (2002) Polypurine (A)-rich sequences promote cross-kingdom conservation of internal ribosome entry. *Proc Natl Acad Sci U S A* **99**: 5301-5306

Fernandez IS, Bai XC, Murshudov G, Scheres SH, Ramakrishnan V (2014) Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* **157**: 823-831

Gendron K, Ferbeyre G, Heveker N, Brakier-Gingras L (2011) The activity of the HIV-1 IRES is stimulated by oxidative stress and controlled by a negative regulatory element. *Nucleic Acids Res* **39**: 902-912

Gilbert WV, Zhou K, Butler TK, Doudna JA (2007) Cap-independent translation is required for starvation-induced differentiation in yeast. *Science* **317**: 1224-1227

Guerrero S, Batisse J, Libre C, Bernacchi S, Marquet R, Paillart JC (2015) HIV-1 replication and the cellular eukaryotic translation apparatus. *Viruses* **7**: 199-218

Hashem Y, des Georges A, Dhote V, Langlois R, Liao HY, Grassucci RA, Pestova TV, Hellen CU, Frank J (2013) Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **503**: 539-543

Hellman LM, Fried MG (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc* **2**: 1849-1861

Herbreteau CH, Weill L, Decimo D, Prevot D, Darlix JL, Sargueil B, Ohlmann T (2005) HIV-2 genomic RNA contains a novel type of IRES located downstream of its initiation codon. *Nat Struct Mol Biol* **12**: 1001-1007

Hofacker IL (2009) RNA secondary structure analysis using the Vienna RNA package. *Current protocols in bioinformatics / editorial board, Andreas D Baxevanis [et al]* **Chapter 12**: Unit12 12

Jackson RJ, Hellen CU, Pestova TV (2009) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**: 113-127

James L, Sargueil B (2008) RNA secondary structure of the feline immunodeficiency virus 5'UTR and Gag coding region. *Nucleic Acids Res* **36**: 4653-4666

Jang SK, Krausslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E (1988) A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J Virol* **62**: 2636-2643

Karabiber F, McGinnis JL, Favorov OV, Weeks KM (2013) QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *Rna* **19**: 63-73

Kenyon JC, Ghazawi A, Cheung WK, Phillip PS, Rizvi TA, Lever AM (2008) The secondary structure of the 5' end of the FIV genome reveals a long-range interaction between R/U5 and gag sequences, and a large, stable stem-loop. *Rna*

Koh CS, Brilot AF, Grigorieff N, Korostelev AA (2014) Taura syndrome virus IRES initiates translation by binding its tRNA-mRNA-like structural element in the ribosomal decoding center. *Proc Natl Acad Sci U S A* **111**: 9139-9144

Kolupaeva VG, Lomakin IB, Pestova TV, Hellen CU (2003) Eukaryotic Initiation Factors 4G and 4A Mediate Conformational Changes Downstream of the Initiation Codon of the Encephalomyocarditis Virus Internal Ribosomal Entry Site. *Mol Cell Biol* **23**: 687-698

Kozak M (1998) Primer extension analysis of eukaryotic ribosome-mRNA complexes. *Nucleic Acids Res* **26**: 4853-4859

Locker N, Chamond N, Sargueil B (2011) A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3. *Nucleic Acids Res* **39**: 2367-2377

Lomakin IB, Hellen CU, Pestova TV (2000) Physical association of eukaryotic initiation factor 4G (eIF4G) with eIF4A strongly enhances binding of eIF4G to the internal ribosomal entry site of encephalomyocarditis virus and is required for internal initiation of translation. *Mol Cell Biol* **20**: 6019-6029.

Lowman HB, Draper DE (1986) On the recognition of helical RNA by cobra venom V1 nuclease. *J Biol Chem* **261**: 5396-5403

Miele G, Mouland A, Harrison GP, Cohen E, Lever AM (1996) The human immunodeficiency virus type 1 5' packaging signal structure affects translation but does not function as an internal ribosome entry site structure. *J Virol* **70**: 944-951

Monette A, Valiente-Echeverria F, Rivero M, Cohen EA, Lopez-Lastra M, Mouland AJ (2013) Dual mechanisms of translation initiation of the full-length HIV-1 mRNA contribute to gag synthesis. *PLoS One* **8**: e68108

Mortimer SA, Weeks KM (2007) A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J Am Chem Soc* **129**: 4144-4145

Nicholson MG, Rue SM, Clements JE, Barber SA (2006) An internal ribosome entry site promotes translation of a novel SIV Pr55(Gag) isoform. *Virology*

Olivares E, Landry DM, Caceres CJ, Pino K, Rossi F, Navarrete C, Huidobro-Toro JP, Thompson SR, Lopez-Lastra M (2014) The 5' untranslated region of the human T-cell lymphotropic virus type 1 mRNA enables cap-independent translation initiation. *J Virol* **88**: 5936-5955

Othman Z, Sulaiman MK, Willcocks MM, Ulryck N, Blackbourn DJ, Sargueil B, Roberts LO, Locker N (2014) Functional analysis of Kaposi's sarcoma-associated herpesvirus vFLIP expression reveals a new mode of IRES-mediated translation. *Rna* **20**: 1803-1814

Parkin NT, Cohen EA, Darveau A, Rosen C, Haseltine W, Sonenberg N (1988) Mutational analysis of the 5' non-coding region of human immunodeficiency virus type 1: effects of secondary structure on translation. *EMBO J* **7**: 2831-2837

Pelletier J, Sonenberg N (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**: 320-325

Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU (1998) A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev* **12**: 67-83

Pestova TV, Shatsky IN, Hellen CU (1996) Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol Cell Biol* **16**: 6870-6878

Plank TD, Whitehurst JT, Kieft JS (2013) Cell type specificity and structural determinants of IRES activity from the 5' leaders of different HIV-1 transcripts. *Nucleic Acids Res* **41**: 6698-6714

Pyronnet S, Sonenberg N (2001) Cell-cycle-dependent translational control. *Curr Opin Genet Dev* **11**: 13-18

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

Ricci EP, Herbreteau CH, Decimo D, Schaupp A, Datta SA, Rein A, Darlix JL, Ohlmann T (2008a) In vitro expression of the HIV-2 genomic RNA is controlled by three distinct internal ribosome entry segments that are regulated by the HIV protease and the Gag polyprotein. *Rna* **14**: 1443-1455

Ricci EP, Soto Rifo R, Herbreteau CH, Decimo D, Ohlmann T (2008b) Lentiviral RNAs can use different mechanisms for translation initiation. *Biochem Soc Trans* **36**: 690-693

Rojas-Araya B, Ohlmann T, Soto-Rifo R (2015) Translational Control of the HIV Unspliced Genomic RNA. *Viruses* **7**: 4326-4351

Ryder SP, Recht MI, Williamson JR (2008) Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods Mol Biol* **488**: 99-115

Sharma A, Yilmaz A, Marsh K, Cochrane A, Boris-Lawrie K (2012) Thriving under stress: selective translation of HIV-1 structural protein mRNA during Vpr-mediated impairment of eIF4E translation activity. *PLoS pathogens* **8**: e1002612

Shirokikh NE, Spirin AS (2008) Poly(A) leader of eukaryotic mRNA bypasses the dependence of translation on initiation factors. *Proc Natl Acad Sci U S A* **105**: 10738-10743

Siegfried NA, Busan S, Rice GM, Nelson JA, Weeks KM (2014) RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nature methods* **11**: 959-965

Smirnova VV, Terenin IM, Khutornenko AA, Andreev DE, Dmitriev SE, Shatsky IN (2016) Does HIV-1 mRNA 5'-untranslated region bear an internal ribosome entry site? *Biochimie* **121**: 228-237

Sobczak K, Michlewski G, de Mezer M, Krol J, Krzyzosiak WJ (2010) Trinucleotide repeat system for sequence specificity analysis of RNA structure probing reagents. *Anal Biochem* **402**: 40-46

Soto-Rifo R, Rubilar PS, Limousin T, de Breyne S, Decimo D, Ohlmann T (2012) DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. *Embo J* **31**: 3745-3756

Soto-Rifo R, Rubilar PS, Ohlmann T (2013) The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA. *Nucleic Acids Res* **41**: 6286-6299

Steen KA, Rice GM, Weeks KM (2012) Fingerprinting noncanonical and tertiary RNA structures by differential SHAPE reactivity. *J Am Chem Soc* **134**: 13160-13163

Svitkin YV, Pause A, Sonenberg N (1994) La autoantigen alleviates translational repression by the 5' leader sequence of the human immunodeficiency virus type 1 mRNA. *J Virol* **68**: 7001-7007

Sweeney TR, Abaeva IS, Pestova TV, Hellen CU (2014) The mechanism of translation initiation on Type 1 picornavirus IRESs. *Embo J* **33**: 76-92

Terenin IM, Dmitriev SE, Andreev DE, Royall E, Belsham GJ, Roberts LO, Shatsky IN (2005) A cross-kingdom internal ribosome entry site reveals a simplified mode of internal ribosome entry. *Mol Cell Biol* **25**: 7879-7888

Toohey K, Wehrly K, Nishio J, Perryman S, Chesebro B (1995) Human immunodeficiency virus envelope V1 and V2 regions influence replication efficiency in macrophages by affecting virus spread. *Virology* **213**: 70-79

Vallejos M, Deforges J, Plank TD, Letelier A, Ramdohr P, Abraham CG, Valiente-Echeverria F, Kieft JS, Sargueil B, Lopez-Lastra M (2011) Activity of the human immunodeficiency virus type 1 cell cycle-dependent internal ribosomal entry site is modulated by IRES trans-acting factors. *Nucleic Acids Res* **39**: 6186-6200

Vallejos M, Ramdohr P, Valiente-Echeverria F, Tapia K, Rodriguez FE, Lowy F, Huidobro-Toro JP, Dangerfield JA, Lopez-Lastra M (2010) The 5'-untranslated region of the mouse mammary tumor virus mRNA exhibits cap-independent translation initiation. *Nucleic Acids Res* **38**: 618-632

Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW, Jr., Swanstrom R, Burch CL, Weeks KM (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**: 711-716

Wehrly K, Chesebro B (1997) p24 antigen capture assay for quantification of human immunodeficiency virus using readily available inexpensive reagents. *Methods* **12**: 288-293

Weill L, James L, Ulryck N, Chamond N, Herbreteau CH, Ohlmann T, Sargueil B (2010) A new type of IRES within gag coding region recruits three initiation complexes on HIV-2 genomic RNA. *Nucleic Acids Res* **38**: 1367-1381

Wilkinson KA, Gorelick RJ, Vasa SM, Guex N, Rein A, Mathews DH, Giddings MC, Weeks KM (2008) High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol* **6**: e96

Wilkinson KA, Merino EJ, Weeks KM (2006) Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat Protoc* **1**: 1610-1616

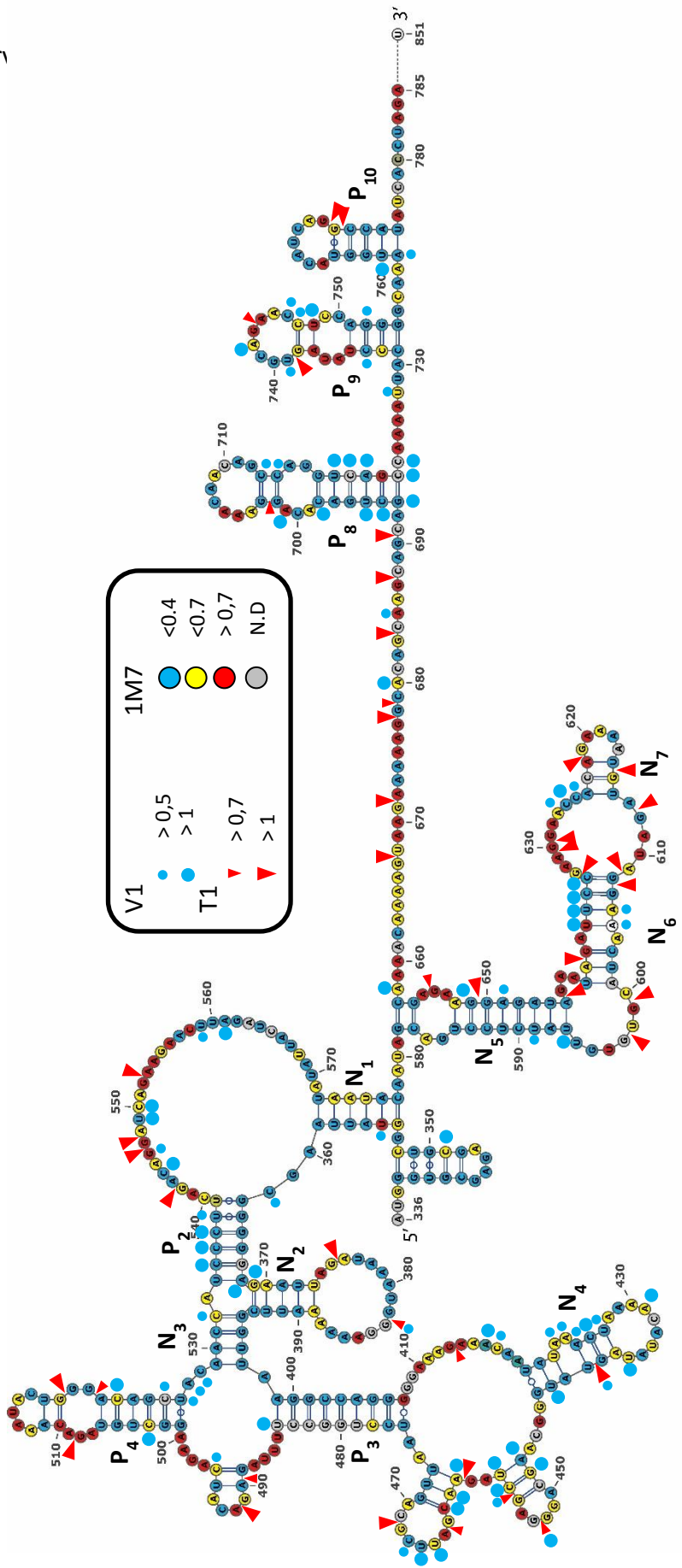
Wilson JE, Pestova TV, Hellen CU, Sarnow P (2000) Initiation of protein synthesis from the A site of the ribosome. *Cell* **102**: 511-520

Xu ZZ, Mathews DH (2016) Secondary Structure Prediction of Single Sequences Using RNAstructure. *Methods Mol Biol* **1490**: 15-34

Yamamoto H, Unbehaun A, Loerke J, Behrmann E, Collier M, Burger J, Mielke T, Spahn CM (2014) Structure of the mammalian 80S initiation complex with initiation factor 5B on HCV-IRES RNA. *Nat Struct Mol Biol* **21**: 721-727

Yilmaz A, Bolinger C, Boris-Lawrie K (2006) Retrovirus translation initiation: Issues and hypotheses derived from study of HIV-1. *Curr HIV Res* **4**: 131-139

Zuker M (2003) Mfold web server  
*Acids* 31: 3406-3415  
**Figure 1A**



:leic



Figure 1B

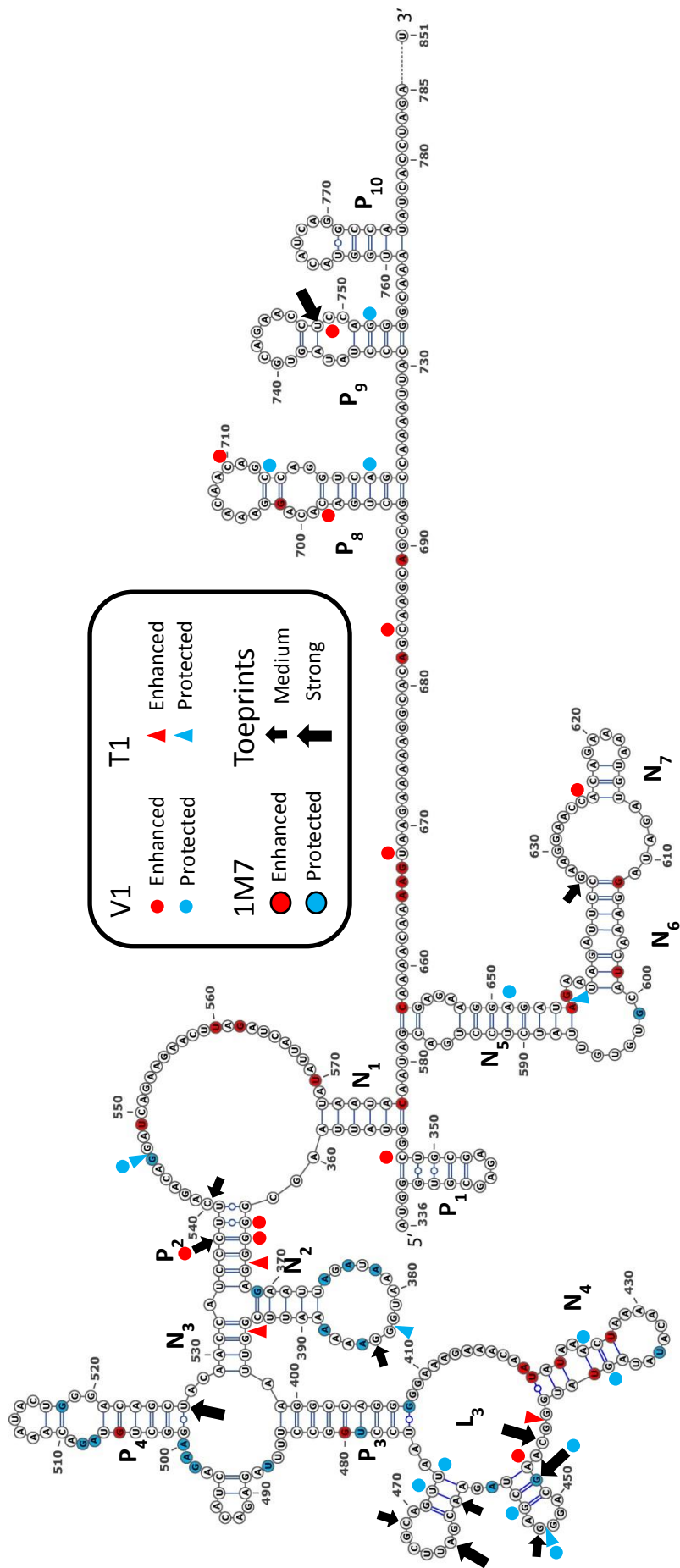
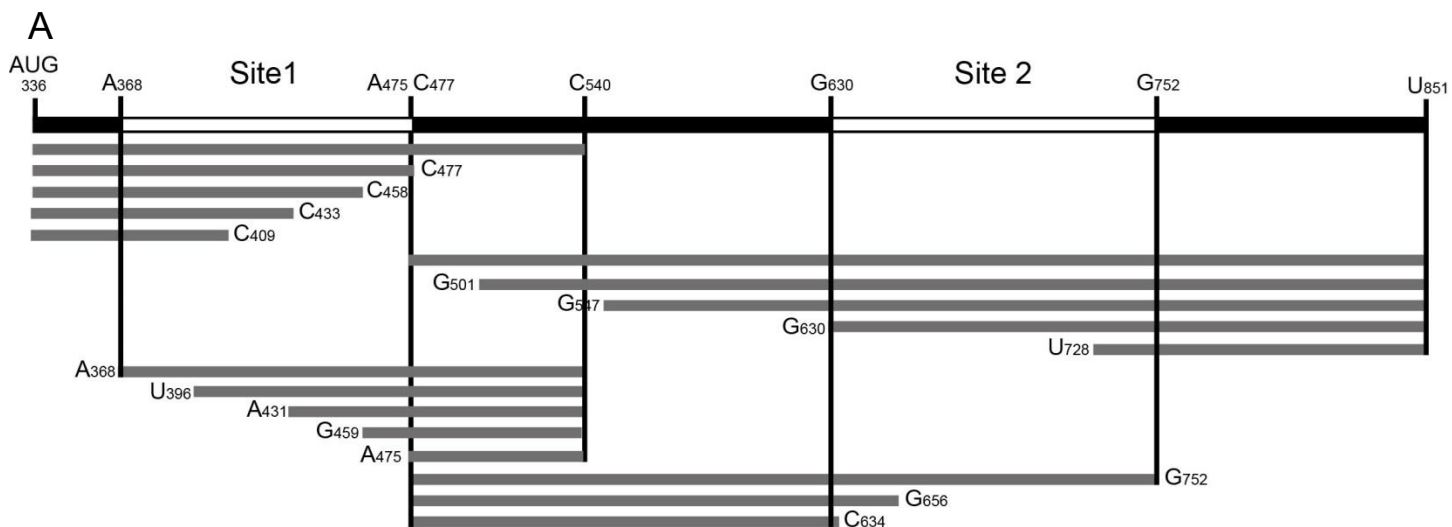
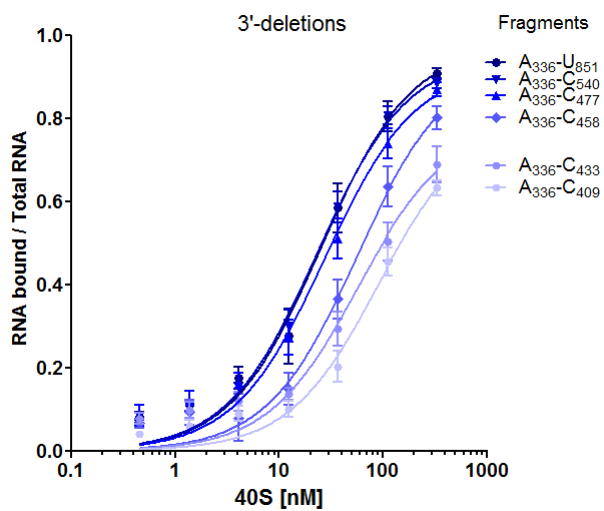


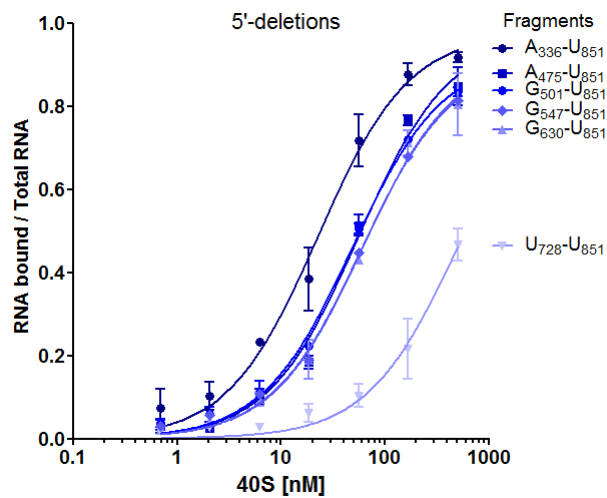
Figure 2



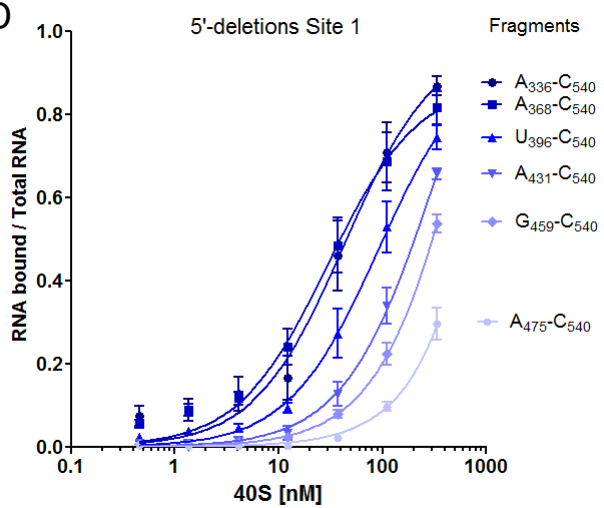
**B**



**C**



**D**



**E**

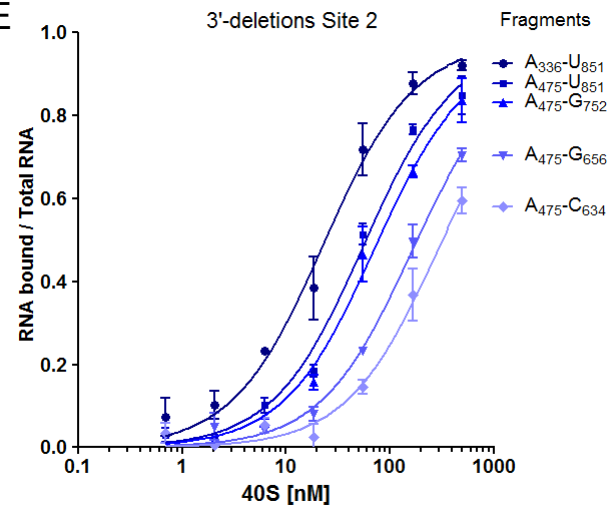
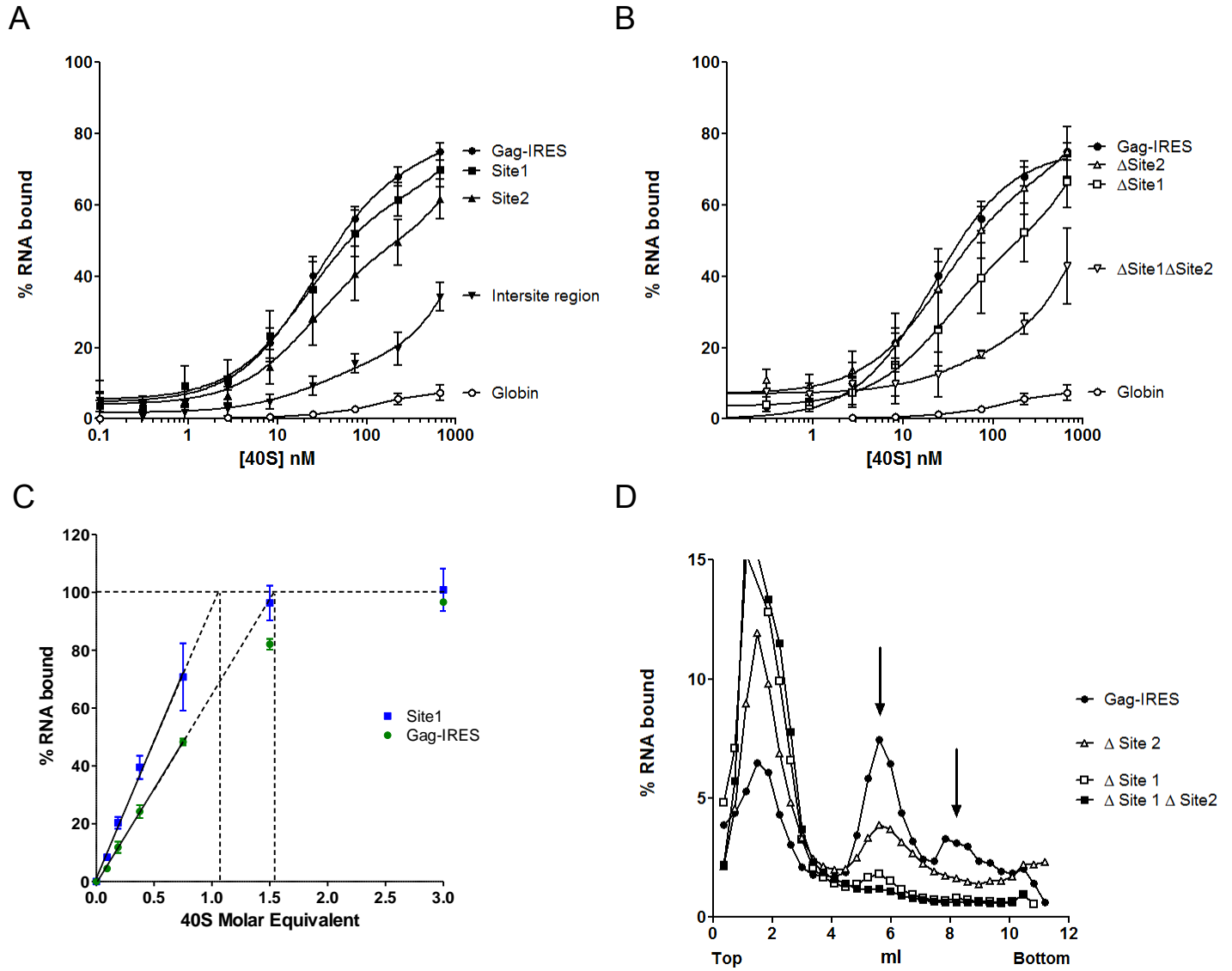
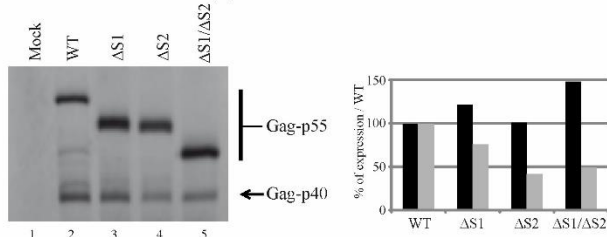
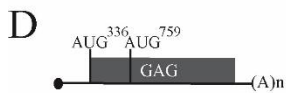
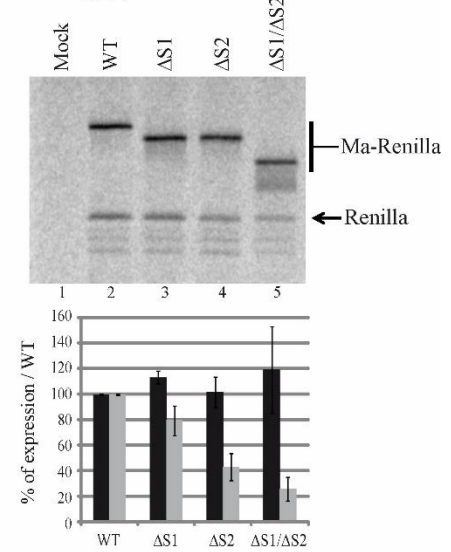
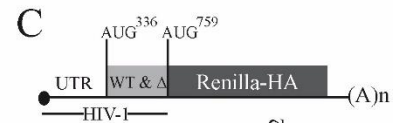
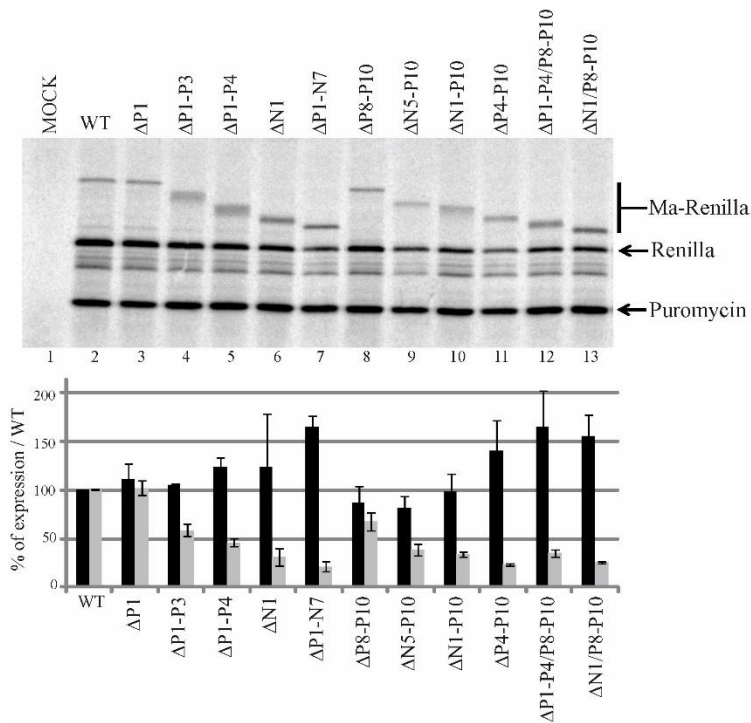
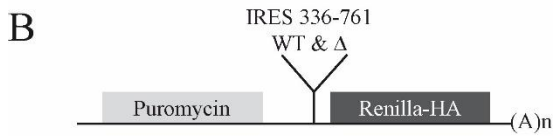
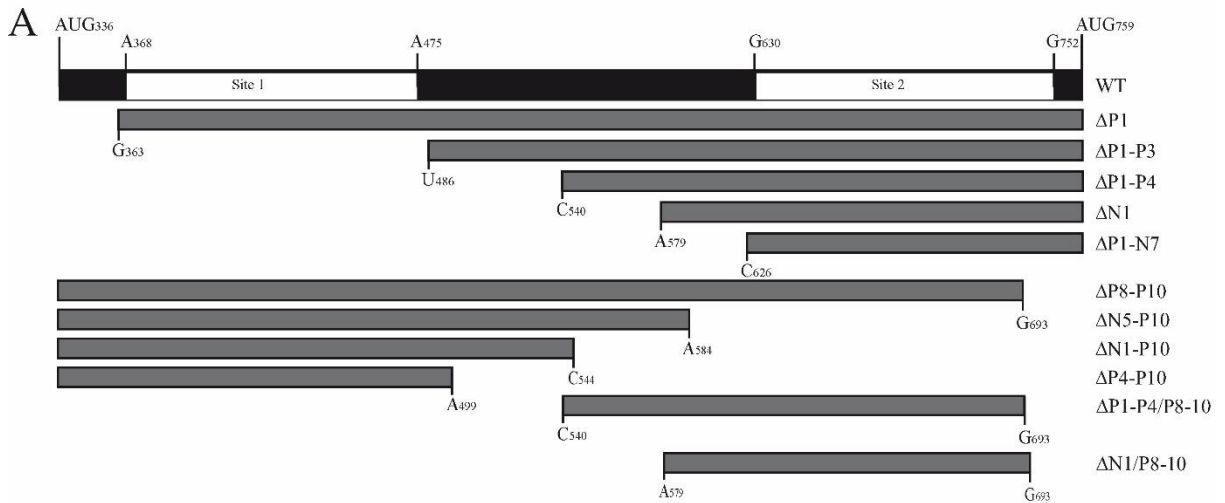


Figure 3

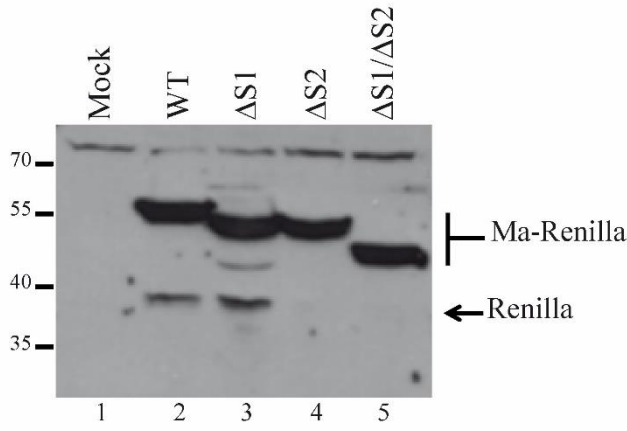
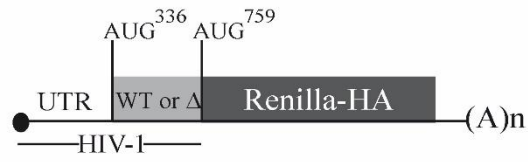




■ expression initiated at AUG(336) codon  
 □ expression initiated at AUG(759) codon

Figure 4

A



B

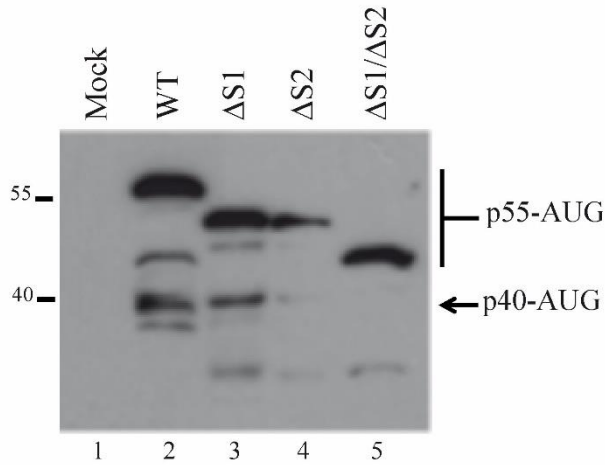
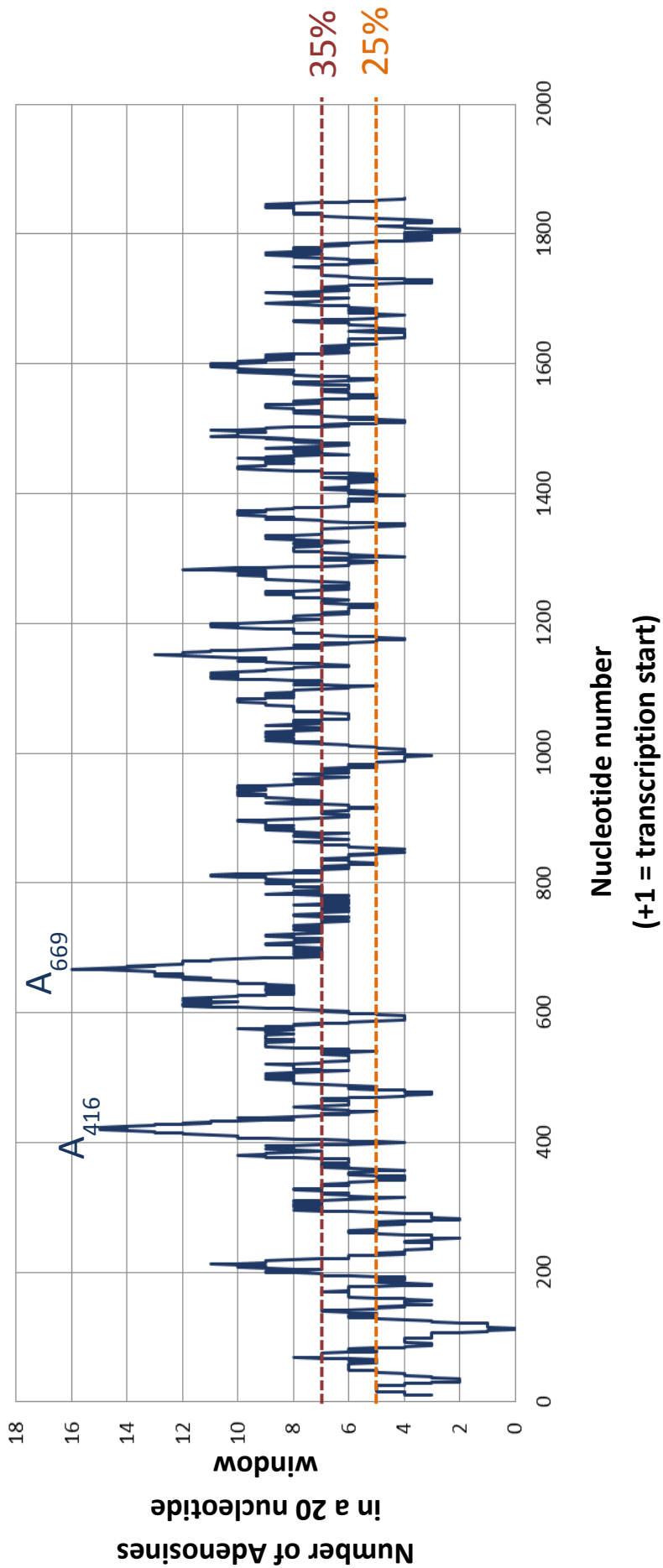


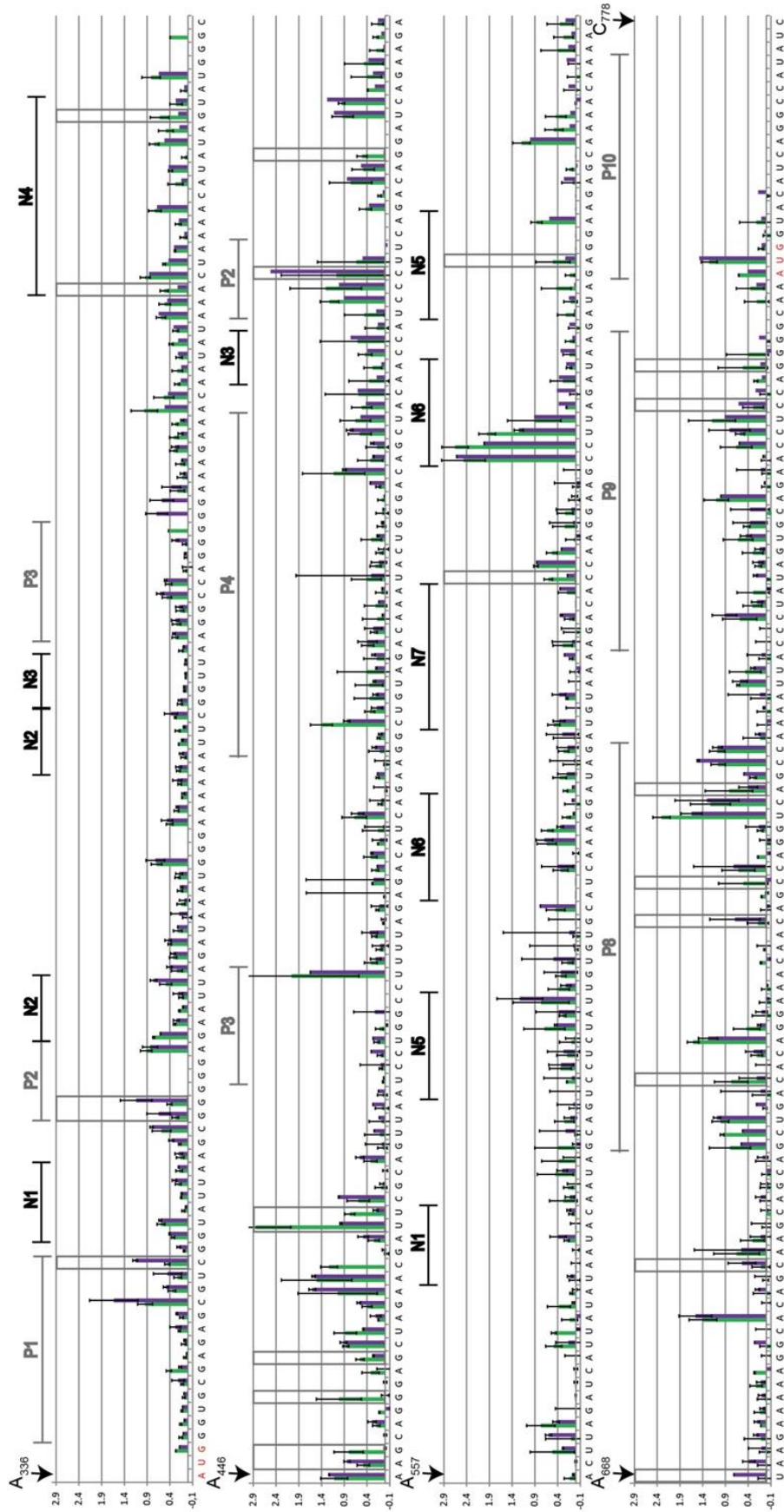
Figure 5

Figure 6





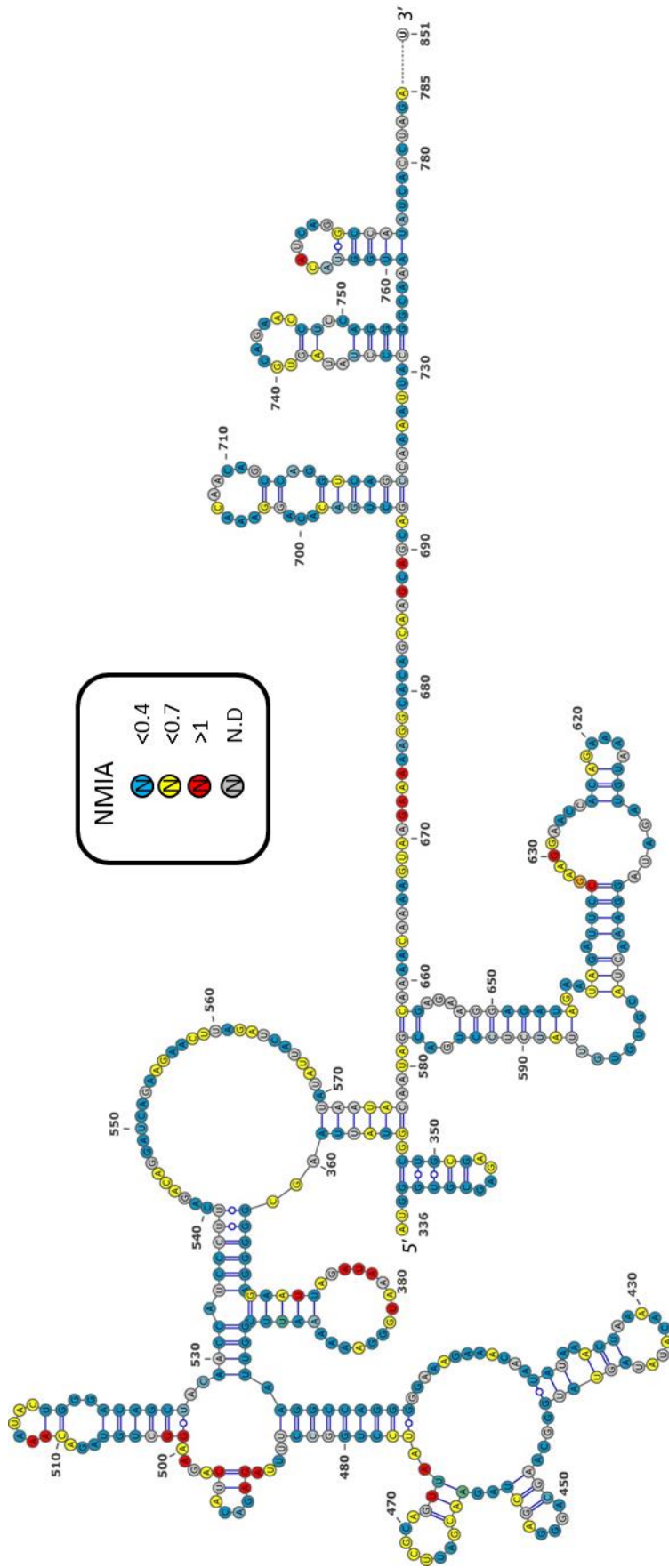




Supplemental Figure 2: V1 probing of Gag-IREs in the presence of ribosomal 40S. The V1 «reactivity» obtained in the absence (green) or presence (purple) of ribosomal 40S. Previously (grey) or newly (black) identified domains are indicated in the Figure as well as relevant nucleotide positions. AUG<sub>336</sub> and AUG<sub>759</sub> are indicated in red. Boxed regions highlight significant differences as measured by the bilateral Student Test with an acceptance value of 0.05. Error bars are the s.e.m. for at least three independent experiments.

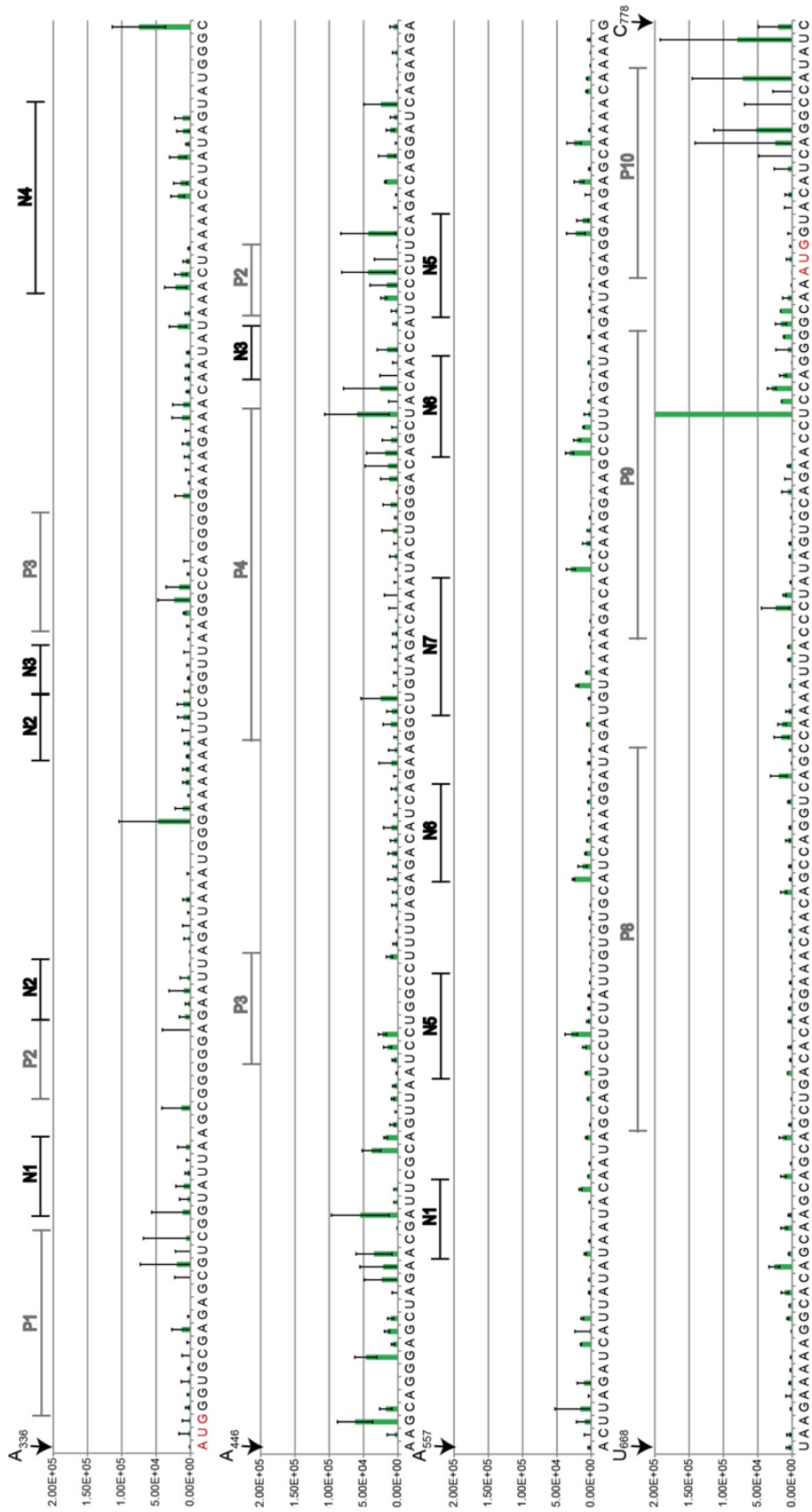






Supplemental figure 3B : NMIA reactivity colour coded on the Gag-HIV1 IRES secondary structure model





Supplemental Figure 5: Toe-print analysis of Gag-IRES in the presence of ribosomal 40S. Toe-print was performed by capillary electrophoresis and histograms represent the differential peak intensity (with-without 40S) for each position. Previously (grey) and newly (black) identified domains are indicated in the Figure as well as relevant nucleotide positions. AUG<sub>336</sub> and AUG<sub>778</sub> are indicated in red. Error bars are the s.e.m. for at least three independent experiments.



## II. Etude du rôle de DDX3 dans l'initiation de la traduction du VIH-1

Comme nous l'avons vu précédemment, la traduction de la polyprotéine Gag s'appuie sur trois mécanismes distincts ce qui indique que l'expression de Gag est finement régulée. L'initiation peut se faire selon un mécanisme dépendant de la présence de la coiffe en 5' ou par entrée interne du ribosome *via* la présence de deux IRES. La 5'-UTR de l'ARNg du VIH-1 est longue et contient de nombreuses structures qui sont essentielles à la coordination des différentes étapes du cycle viral. La présence de ces signaux structuraux interfère clairement avec la traduction et affecte l'expression de Gag en empêchant le balayage du ribosome à partir de la coiffe.

Récemment, il a été montré que DDX3 favorise la traduction de Gag *in cellulo*. DDX3 est une hélicase à boîte DEAD dont les propriétés biochimiques incluent une activité ATPase dépendante de l'ARN et une capacité à dérouler des duplex d'ARN. Ces activités permettent à l'enzyme de lier et de remodeler l'ARN ainsi que des complexes ARN-protéines en présence d'ATP. Néanmoins, et contrairement à son homologue chez la levure, il a été montré que DDX3 1/ présente une activité ATPase indépendante de l'ARN et 2/ est capable de dérouler des duplex de 50 paires de bases.

Ainsi, avant de disséquer le mécanisme moléculaire sous-tendant le rôle de DDX3 dans l'initiation de la traduction de Gag, nous avons produit la protéine recombinante, nous avons caractérisé ses activités biochimiques en utilisant des approches cinétiques et évalué son rôle sur la traduction de Gag dans notre système *in vitro*.

Nos résultats nous permettent d'invalider les propriétés inhabituelles de DDX3 et de proposer une hypothèse explicative aux études publiées sur le sujet. DDX3 présentent donc des activités biochimiques proches de celle de son homologue chez la levure. De plus, la protéine ainsi produite stimule la traduction de Gag *in vitro*. Ces résultats ouvrent la voie à une dissection moléculaire du rôle de DDX3 dans l'initiation de la traduction de Gag et plus généralement à une dissection moléculaire des rôles physiologiques de cette protéine.

A. Obtention d'une protéine recombinante soluble et active.

### Production de rDDX3

Dans un premier temps, nous avons essayé d'obtenir une forme soluble de DDX3 à partir du plasmide pET21a-DDX3-WT (disponible au laboratoire, voir matériels et méthodes). Pour cela, l'expression de la protéine DDX3 sauvage (DDX3-WT) en système hétérologue chez *Escherichia Coli* (*E.coli*)-BL21 (DE3) est induite par ajout d'IPTG. Le culot bactérien est ensuite lysé par action mécanique puis centrifugé afin de séparer la fraction soluble (surnageant, SUP) de la fraction insoluble (débris cellulaires, culot P). La solubilité de la protéine est ensuite, vérifiée par électrophorèse sur gel d'acrylamide. Les résultats montrent que la protéine est essentiellement retrouvée dans les fractions insolubles (indiquées par un P sur la Figure 46). La présence de la protéine dans la fraction insoluble peut s'expliquer par un repliement incorrect de la protéine ou bien encore par la structure de DDX3 qui contient des domaines N-et C-terminaux non structurés, ce qui favoriserait sa présence dans les fractions insolubles. Nous avons donc essayé de moduler le niveau d'expression de DDX3, de favoriser son repliement et de modifier le site d'insertion de l'étiquette His<sub>6</sub>.

Pour cela, nous avons utilisé différentes souches bactériennes et fait varier les paramètres physico-chimiques tels que la température d'induction, le temps d'induction, les tampons utilisés pour la lyse bactérienne, ainsi que la technique utilisée pour la lyse. Nous avons également essayé de Co-exprimer DDX3 avec les protéines chaperonnes GroES/L et enfin, changé la position de l'étiquette histidine [en C-terminal avec le plasmide pET21-a et en N-terminale avec le plasmide pET28-a]. Les résultats obtenus indiquent que quelles que soient les conditions que nous avons testées, la protéine est dans la fraction insoluble. La Figure46-A est donnée à titre d'exemple.

Nous avons donc décidé de sous-cloner la séquence codante de DDX3 dans le vecteur pETM-41 contenant la protéine MBP (Maltose Binding Protein) capable d'aider au repliement et à la solubilisation protéique. Les résultats (Figure 46) du test de solubilité de la protéine montrent que la protéine de fusion MBP-DDX3 est exprimée majoritairement dans les fractions solubles, notamment, dans la condition d'expression 2h à 37°C après ajout de 0.1mM d'IPTG. Pour la suite des expériences nous avons gardé cette condition.

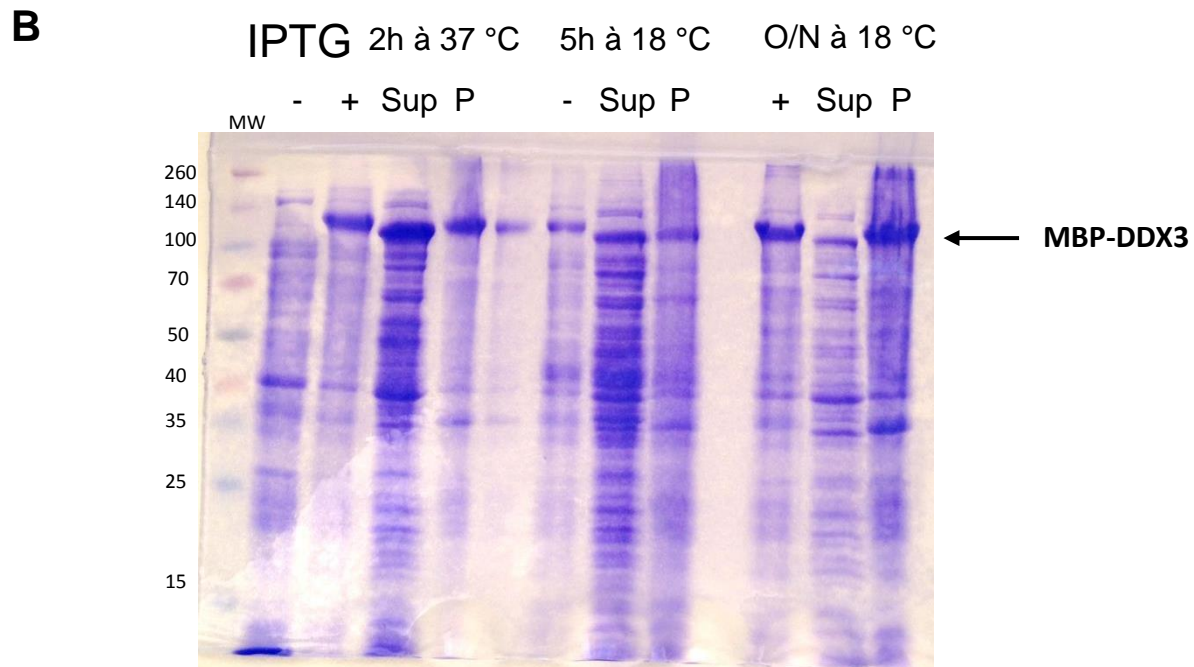
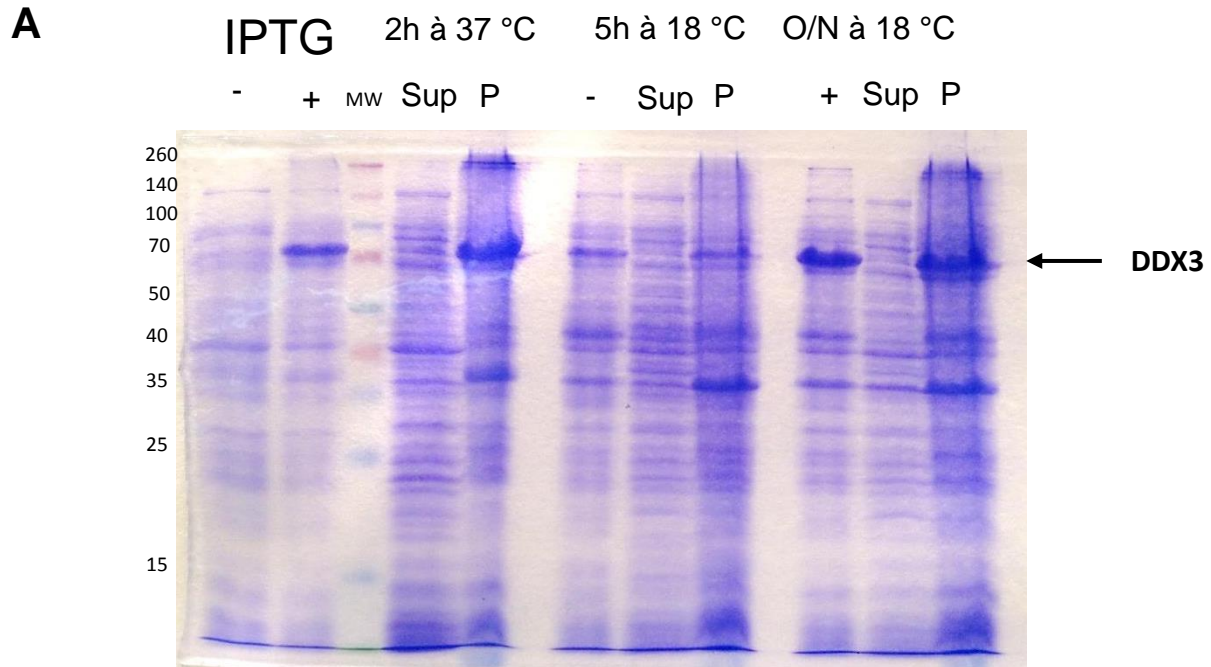


Figure 46 : Expression et solubilité de DDX3

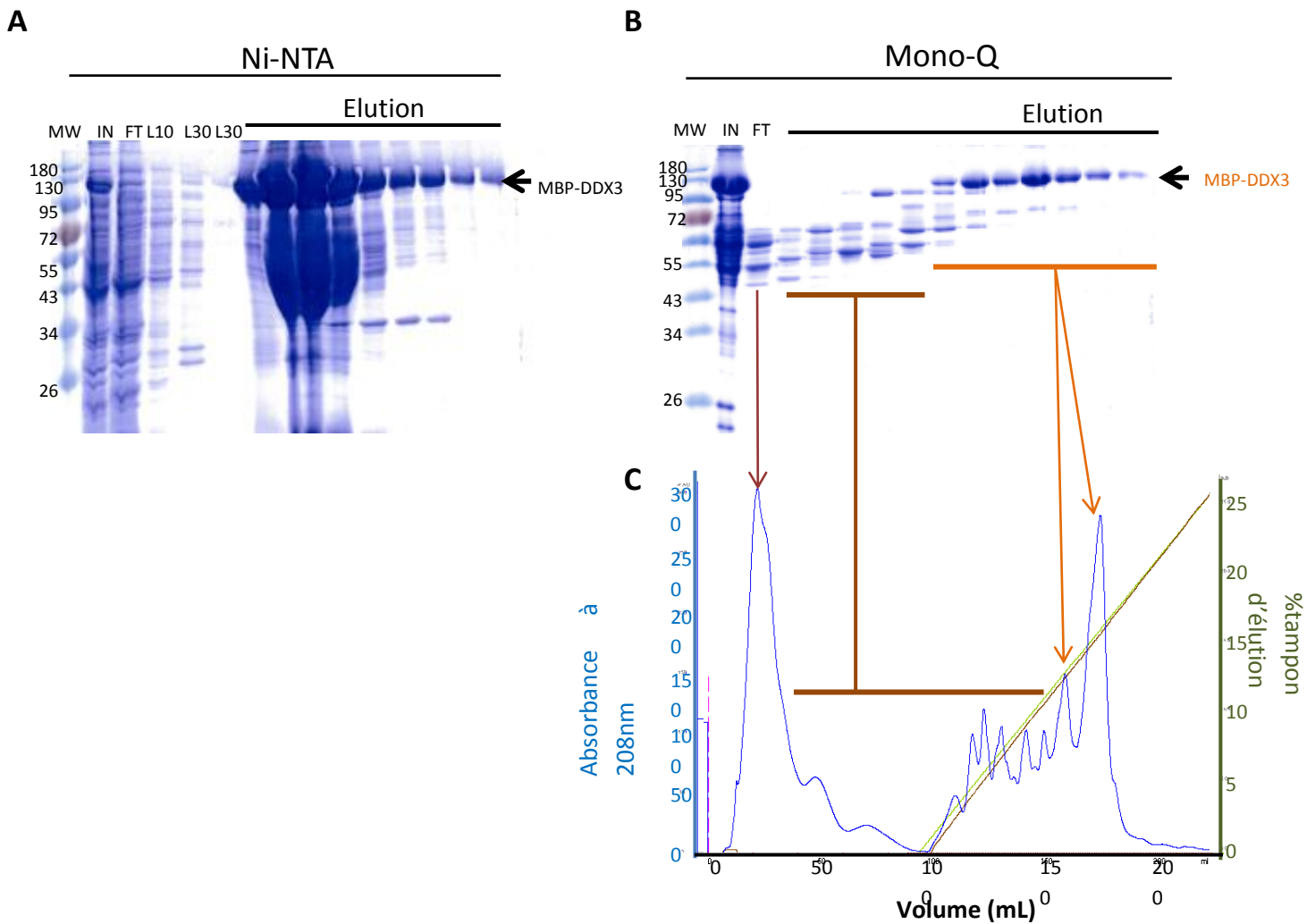
Analyse par électrophorèse sur gel de polyacrylamide 12% de l'expression et la solubilité de DDX3 après ajout (+) d'IPTG. Sup: Surnageant, P: culot après centrifugation du lysat bactérien. MW : marqueur de poids moléculaire, la taille des protéines du marqueur est indiquée en kDa. En A, expression de His6-DDX3 et en B, expression de MBP-DDX3.

## Purification de MBP-DDX3

MBP-DDX3 est ensuite purifiée en deux étapes. D'abord, le lysat bactérien est filtré sur une chromatographie d'affinité au nickel. La protéine MBP-DDX3 étant dotée d'une étiquette histidine à l'extrémité N-terminale de la MBP, est retenue sur la résine de nickel alors que les contaminants ne le sont pas. Après lavage de la résine, la protéine de fusion MBP-DDX3 est éluée en présence de 500 mM d'imidazole (Figure 47-A). Puis, les fractions d'intérêts sont rassemblées et dialysées afin de retirer l'imidazole avant la dernière étape de purification sur une colonne échangeuse d'anions (Figure 47-B et C)

Suite aux deux étapes de purification, la protéine obtenue est relativement pure. D'après l'analyse par électrophorèse sur gel d'acrylamide des différentes fractions (Figure 47-B), et le profil d'éluion (Figure 47-C), la protéine est éluée à une concentration moyenne de 230 mM en KCl en deux pics majoritaires. Ce qui suggère qu'elle présente différentes conformations ou différents états oligomériques. La protéine purifiée a ensuite été analysée en spectroscopie de masse et le résultat confirme que seule la protéine MBP-DDX3 est retrouvée dans notre fraction purifiée. Afin d'analyser l'importance du motif DQAD sur les propriétés de DDX3, nous avons construit le plasmide correspondant. La protéine mutante DDX3-DQAD a ensuite été purifiée dans les mêmes conditions que la protéine sauvage (DDX3-WT) et présente le même rendement et le même niveau de pureté.





**Figure 47 : Purification de MBP-DDX3**

*Purification de DDX3-MBP recombinante. A : analyse par électrophorèse sur gel de polyacrylamide des produits de purification par la chromatographie d'affinité au nickel. B : analyse par électrophorèse sur gel de polyacrylamide des fractions de la chromatographie par échange d'anions. Les fractions déposées sur gel sont indiquées sur le profil d'éluion. C : la courbe d'absorbance à 280nm est représentée en bleu, le pourcentage de tampon d'éluion en vert. IN : input; FT : fraction non retenue ; L10, L30 : lavage avec 10mM ou 30mM d'imidazole. Elution en A (Ni-NTA) se fait avec 500mM imidazole. Éluion en B (MonoQ) se fait avec un gradient de sels allant de 80mM à 1M en KCl*

### **Clivage de l'étiquette MBP**

Afin de poursuivre nos expériences sans la protéine MBP, nous avons cherché à déterminer si la protéine DDX3 restait stable après clivage de la MBP. La protéine de

fusion MBP-DDX3 possède un site de clivage spécifique de la protéase TEV (*Tobacco Etch Virus protease*), qui permet de retirer la MBP et libérer DDX3.

Nous avons, dans un premier temps, clivé la protéine de fusion en présence de la TEV pendant 4h à 20°C et vérifié le produit de clivage par électrophorèse en gel d'acrylamide (Figure 48 et 49). Le résultat montre que le clivage est total et que la protéine DDX3 clivée est soluble

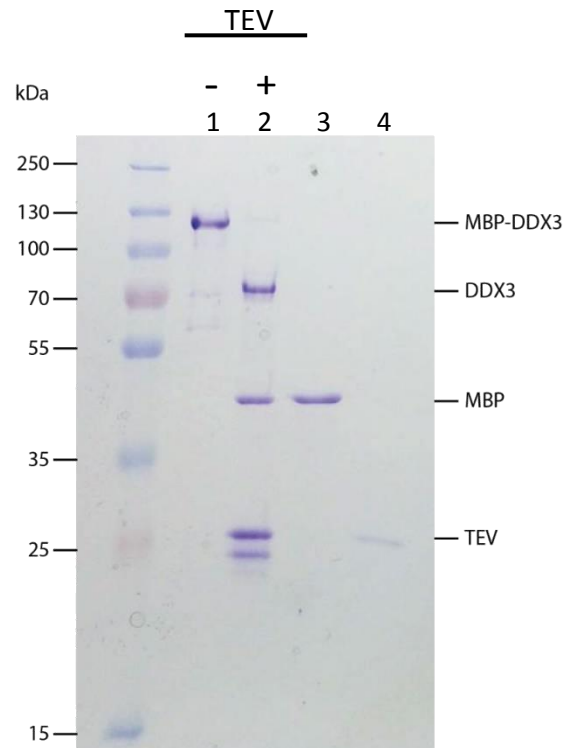


Figure 48 : Clivage de la protéine de fusion MBP-DDX3 WT par la TEV

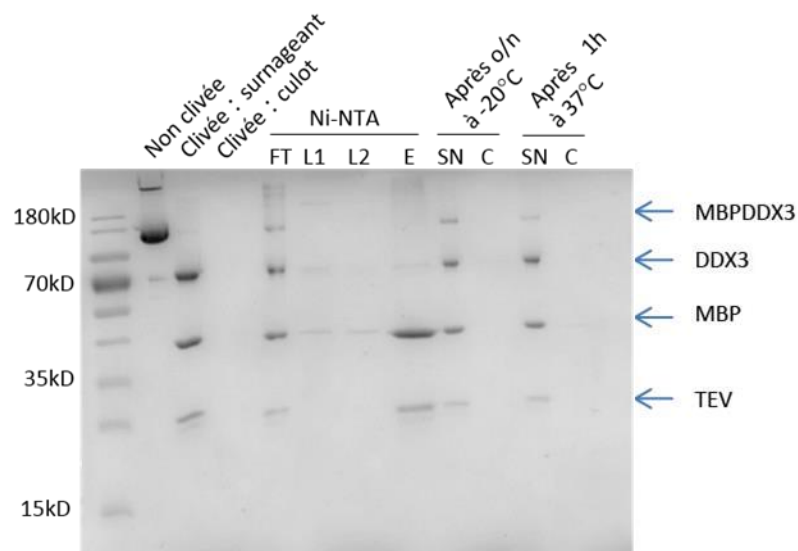
Analyse par électrophorèse sur gel d'acrylamide 12% du produit de clivage (+) de la protéine de fusion MBP-DDX3 par la protéase TEV. La position des différents produits de clivage est indiquée. Piste 1 : la protéine de fusion MBP-DDX3, piste 2 la protéine de fusion MBP-DDX3 après clivage, Piste 3 la protéine MBP, Piste 4 la protéase du virus Tobacco Etch : TEV.

Dans un second temps, nous avons procédé à la purification du produit de clivage sur matrice de nickel. Les protéines MBP et TEV présentent toutes deux une étiquette histidine, ainsi lors de la purification sur la matrice de nickel, elles sont retenues sur la colonne par affinité vis-à-vis de la matrice alors que DDX3 clivée n'est pas retenue. On peut ainsi la récupérer.

Puis, pour évaluer la stabilité de DDX3 après clivage, nous avons incubée la protéine à 37°C (température à laquelle sont conduits les tests d'activité) pendant une heure ou à -20°C (température à laquelle la protéine est conservée) pendant une nuit.

La fraction soluble (surnageant) a été séparée de la fraction insoluble (culot) par centrifugation à 13000rpm pendant 10 minutes. Ces différentes fractions sont analysées par électrophorèse sur gel de polyacrylamide (Figure 49).

On constate que la purification sur la matrice de nickel n'est pas optimale. En effet, bien que DDX3 clivée soit présente dans la fraction non retenue nous retrouvons également les protéines MBP et TEV censées être accrochées sur la matrice. Ce résultat peut s'expliquer par plusieurs hypothèses ; le temps d'incubation entre le produit de clivage et la matrice de nickel pourrait être trop court et donc pas suffisant pour fixer les protéines. Le tampon de clivage contient du DTT (Dithiothréitol) nécessaire à l'activité de la TEV, nous pouvons penser que la présence du DTT réduit le Nickel de la matrice et par conséquent diminue l'affinité de l'histidine à la matrice.



**Figure 49 : Clivage et stabilité de la protéine DDX3**

*Analyse du clivage et de la stabilité de DDX3 par électrophorèse sur gel d'acrylamide. DDX3-MBP (« non clivée ») est clivée par la TEV, puis les produits de clivage sont centrifugés et séparés en surnageant (« Clivée : surnageant ») et culot (« Clivée : culot »). La fraction soluble est purifiée sur colonne de nickel. FT : fraction non retenue sur matrice de nickel ; L 1 et 2 : fractions de lavage ; E : fraction d'élution. La fraction non retenue sur colonne de nickel est incubée à -20°C sur la nuit, ou à 37°C pendant 1h, et les fractions solubles et insolubles sont séparées par centrifugation. SN : surnageant ; C : culot*

Bien que la purification ne soit pas optimale, les résultats indiquent que la protéine DDX3 clivée est retrouvée dans les fractions solubles (Figure 49, SN) et cela aussi bien après une incubation à 37°C qu'après un cycle de congélation/décongélation à -20°C.

## Evaluation de l'activité ATPase des protéines recombinantes

Afin d'évaluer les différentes préparations, nous avons procédé à un test d'activité ATPase. Pour cela, les différentes protéines sont incubées en présence d'ATP- $\gamma P^{32}$  en présence ou non d'ARN. Puis, le pourcentage de phosphate inorganique ( $P_i$ ) relargué est quantifié. Deux contrôles négatifs sont réalisés : ATP- $\gamma P^{32}$  seul afin de quantifier le pourcentage d'auto-hydrolyse de l'ATP, et l'ATP en présence d'ARN uniquement afin de valider que l'ajout d'ARN ne provoque pas l'auto-hydrolyse de l'ATP- $\gamma P^{32}$ .

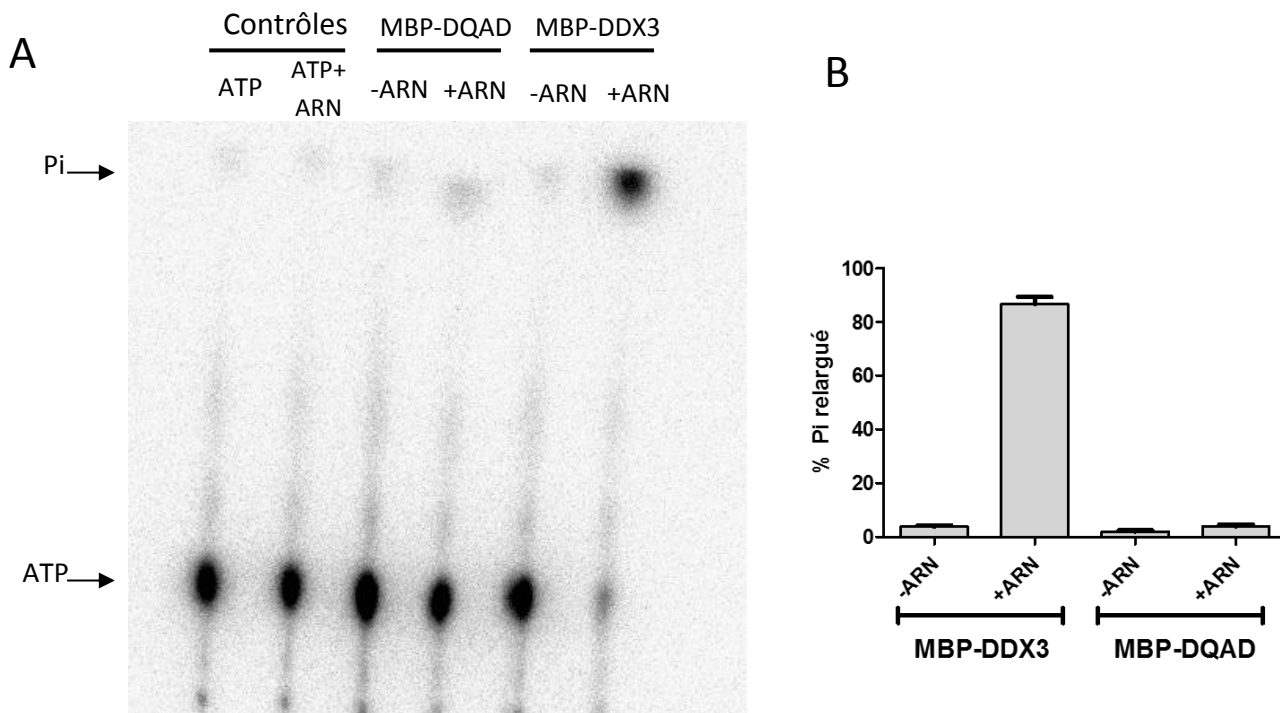


Figure 50 : Mesure de l'activité ATPase

*A : autoradiographie des produits des réactions migrés sur une plaque PEI-cellulose et exposés à un écran sensible à la radioactivité. Les positions de l'ATP et du  $P_i$  sont indiquées. ATP seul et ATP+ARN représentent les contrôles négatifs de l'expérience B : Histogramme représentant le pourcentage de phosphate ( $P_i$ ) relargué au cours de l'activité ATPase. La protéine MBP-DDX3 WT est active en présence d'ARN et le mutant DQAD ne l'est pas.*

Les résultats (Figure 50), montrent que la protéine MBP-DDX3 WT est active et que cette activité est dépendante de la présence de l'ARN. De plus cette activité est spécifique de DDX3 car le mutant DQAD purifié dans les mêmes conditions ne présente pas d'activité ATPase. Il est intéressant de noter ici que ni DDX3-WT ni le mutant DQAD ne présentent d'activité ATPase intrinsèque.

Nous avons ensuite évalué l'activité ATPase de la protéine DDX3 clivée comparée à celle de la protéine de fusion MBP-DDX3. Pour cela, nous avons conduit des cinétiques enzymatiques avec les différentes protéines en présence de 100 nM d'ARN ainsi qu'en présence de contrôles appropriés (le mutant DQAD, la protéine MBP et la protéase TEV).

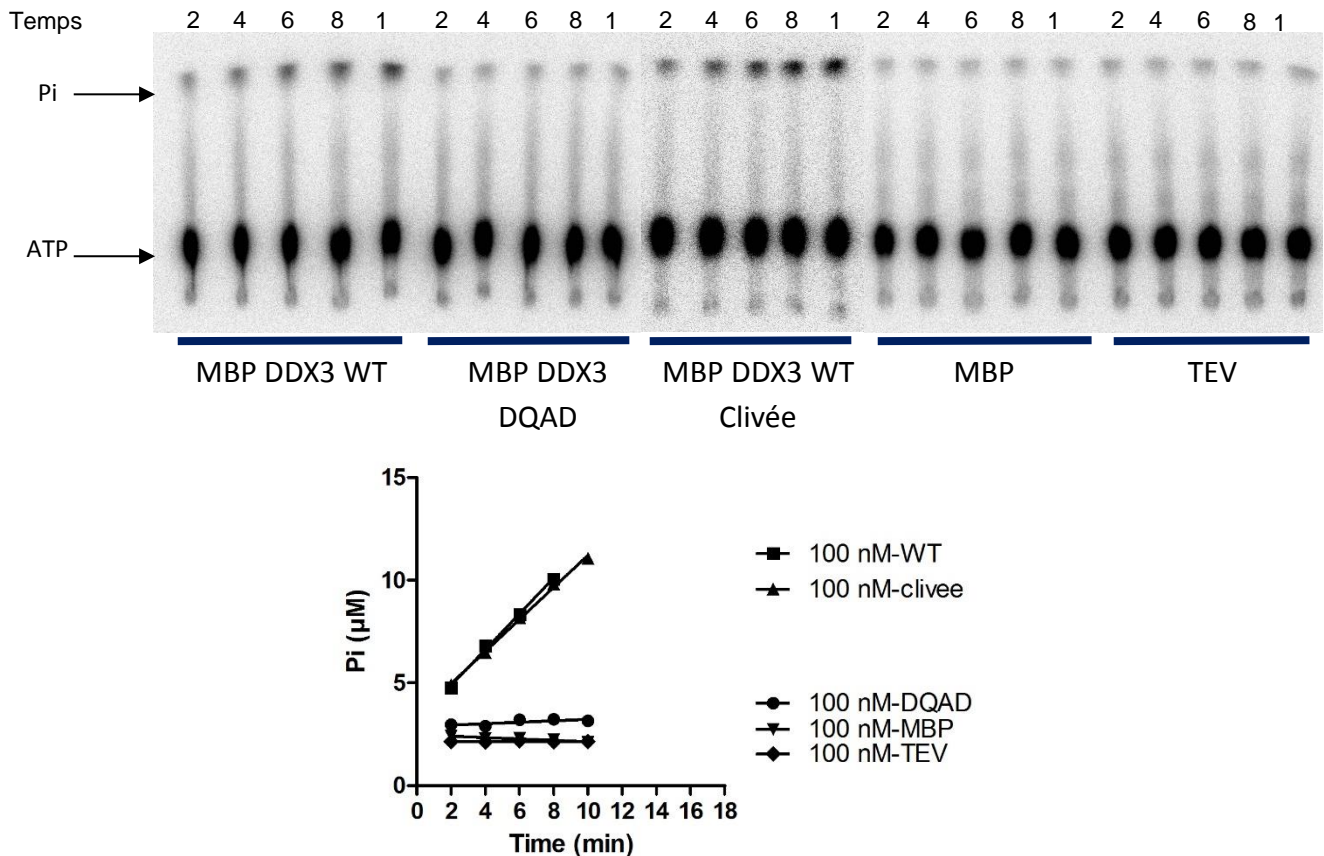


Figure 51 : Test d'activité ATPase de la protéine clivée vs la protéine non clivée

*Autoradiographie des produits des réactions après migration sur une plaque PEI-cellulose dans du  $KH_2PO_4$  et exposition à un film sensible à la radioactivité. Le signal quantifié est rapporté sous forme de graphique représentant la quantité de Pi relarguée au cours du temps. La protéine clivée (triangles) présente la même activité ATPase que protéine non clivée WT (carrés) et ce contrairement au mutant DQAD (ronds), à la protéine MBP (triangles inversés ou encore à la TEV (losanges)).*

Les résultats obtenus (Figure 51) montrent que les formes non clivée et clivée présentent une activité ATPase identique. Aussi, les expériences ultérieures ont été réalisées avec la protéine MBP-DDX3.

## B. Caractérisation de l'activité ATPase

### Détermination de l'influence de la nature de l'acide nucléique sur la stimulation de l'activité ATPase.

DDX3 a été décrite comme une protéine ayant une activité ATPase stimulée aussi bien par l'ADN que par l'ARN. Nous avons alors testé l'activité de la protéine MBP-DDX3 en présence de substrat de nature et de longueur variables (ADN, ARN, hétéroduplexes ADN-ARN). Contrairement à ce qui a été décrit, nos résultats révèlent que la protéine MBP-DDX3 WT n'est active qu'en présence d'ARN (Figure 52, MBP DDX3 WT conditions 12, 13, 14, 15) avec une activité maximale en présence de l'ARN

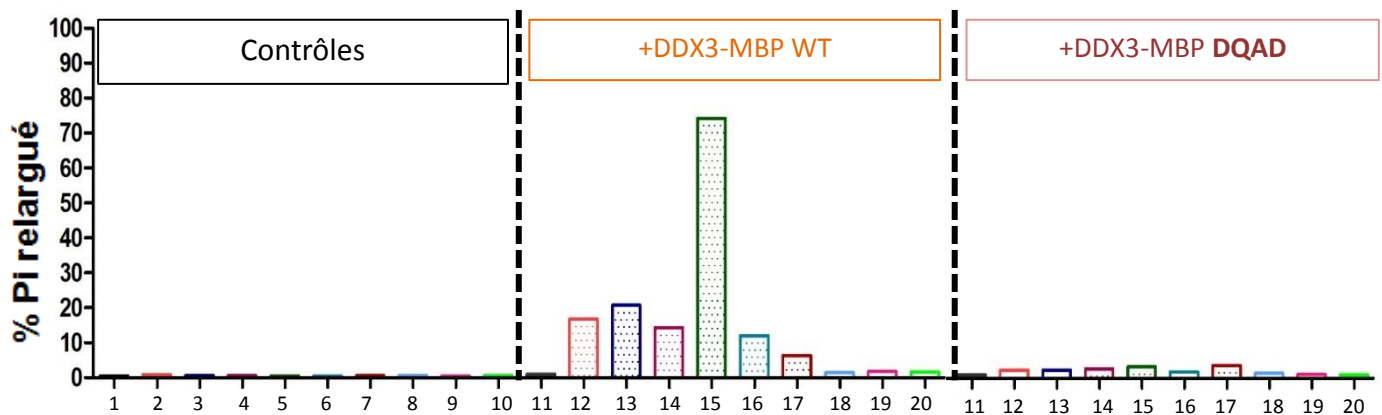


Figure 52 : activité ATPase de MBP-DDX3 en fonction de la nature du substrat

Histogrammes représentant le pourcentage de phosphate (Pi) relargué. L'activité ATPase de la protéine MBP-DDX3 WT ou de MBP-DDX3 DQAD a été testée en absence d'acide nucléique (1,11), en présence d'un ARN simple brin de 64 nt (2,12), en présence d'un ARN simple brin de 90 nt (3,13), en présence d'un ARN de 56 nt contenant la tige boucle TAR du VIH-1 (4,14), en présence d'un ARN du VIH-1 (nt 1-1635, 5,15), en présence d'un hétéroduplexe de 26bp (6/16) ou encore d'un ADN simple brin de 64 nt (7/17), d'un ADN double brin de 64 nt (8/18), d'un ADN contenant la tige boucle TAR (56bp, 9/19) et enfin d'un ADN double brin correspondant au nt 1-1635 du génome du VIH-1 (10/20). nt : nucléotides, bp : paire de bases.

Une faible activité est observée en présence de l'hétéroduplexe ADN-ARN (Figure 52, MBP DDX3 WT conditions 16). Aucune activité n'est observée en présence d'ADN. Le mutant DQAD ne présente pas d'activité ATPase et ceux dans toutes les conditions de l'expérience



## Détermination de l'influence de la séquence et/ou de la structure de l'ARN sur l'activité ATPase

Afin de tester si DDX3 est activée spécifiquement par l'ARN du VIH-1, nous avons procédé à des tests d'activité ATPase en présence de concentration croissante de l'ARN pNL4 VIH-1 (5'UTR VIH-1-région codante de Gag) ou d'un ARN de même longueur (5'UTR globine-région codante  $\beta$ -galactosidase). Les résultats obtenus (Figure 53) montrent que l'activité ATPase est d'autant plus importante quand la concentration en ARN augmente cependant, il n'existe pas de spécificité vis-à-vis de l'ARN du VIH-1. La protéine est activée autant par l'ARN du VIH-1 que par un ARN de même longueur. On peut en conclure que l'activité ATPase de DDX3 est activée par une alternance entre les motifs double-brin et simple brin plutôt que par une séquence spécifique.

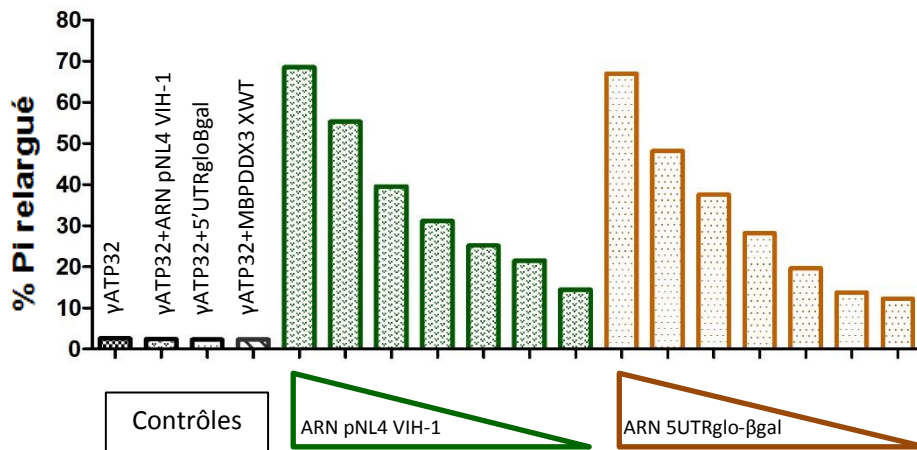


Figure 53 : spécificité de l'activité ATPase de MBP-DDX3

Histogrammes représentant le pourcentage de phosphate ( $P_i$ ) relargué au cours de l'activité ATPase en fonction de concentrations croissantes de l'ARN du VIH-1 (histogrammes en vert), ou d'ARN 5'UTRglo- $\beta$ gal (histogrammes en orange). Les contrôles utilisés pour l'expérience sont indiqués.

## Détermination des paramètres cinétiques de la protéine MBP-DDX3

Pour déterminer les paramètres cinétiques de la protéine MBP-DDX3 nous avons conduit des cinétiques enzymatiques. Pour cela, la protéine est incubée en présence d'ARN et de concentrations différentes en ATP- $\gamma P^{32}$  puis la vitesse initiale est déterminée en utilisant la régression linéaire du  $P_i$  relargué en fonction du temps. Enfin les courbes de  $V_i$  en fonction de la concentration en ATP permettent de déterminer les paramètres cinétiques apparents de l'enzyme. Les résultats (Figure 54-B) indiquent que la protéine MBP-DDX3 arbore une cinétique de type Mickaelis-Menten avec une  $V_{max}$  de 2,4  $\mu M$  d'ATP hydrolysé par minute et un  $K_M$  apparent de 170  $\mu M$ .

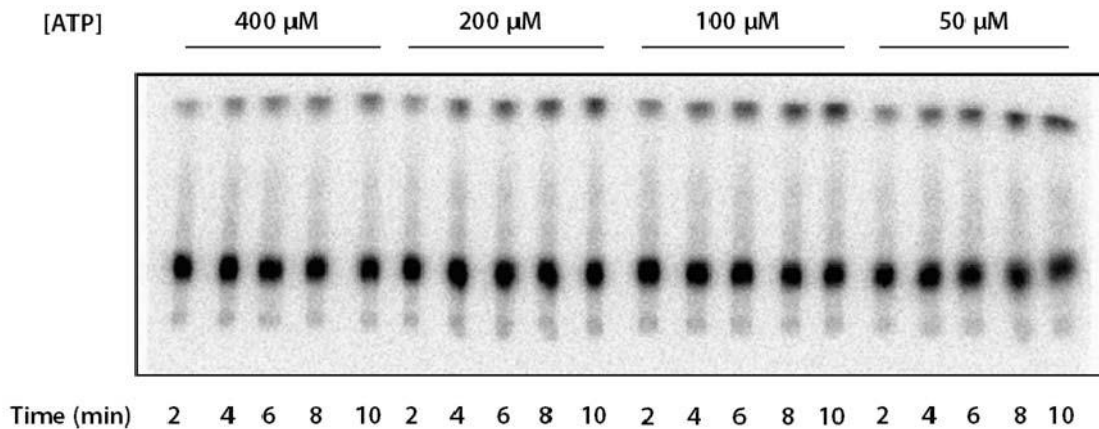
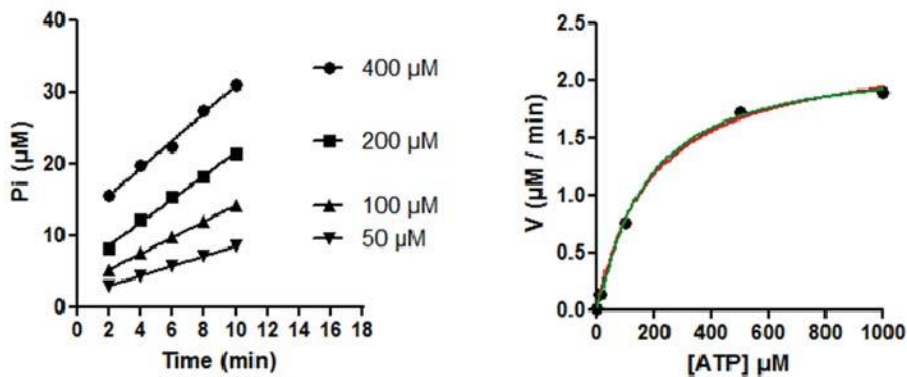
**A****B**

Figure 54 : Détermination des paramètres cinétiques de DDX3

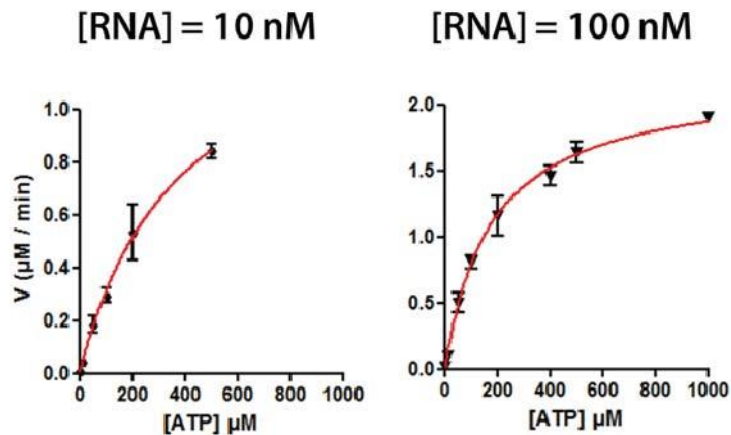
(A) Autoradiographie des produits des réactions après migration sur une plaque PEI-cellulose dans du  $\text{KH}_2\text{PO}_4$  et exposition à un film sensible à la radioactivité. Le signal quantifié est rapporté sous forme de graphique représentant la quantité de  $\text{P}_i$  relargué en fonction du temps et ce, à différentes concentrations en ATP. Puis la régression linéaire du  $\text{P}_i$  relargué en fonction du temps (B) permet de déterminer la vitesse initiale de la réaction (V) pour différentes concentrations en ATP. MBP-DDX3 présente une cinétique de type Michaelis-Menten avec des paramètres apparents:  $V_{\text{max}}$  :  $2,4 \mu\text{M} \cdot \text{min}^{-1}$  et  $K_M$   $170 \mu\text{M}$ .



La stimulation de l'activité ATPase par l'ARN a été mesurée pour différentes concentrations en ARN et en ATP (Figure 55-A). Les résultats Figure 55-B indiquent que la protéine MBP-DDX3 s'oligomérisé. En effet, à de faibles concentration en ARN, la courbe de la vitesse en fonction de la concentration en enzyme est sigmoïdale cela traduit une coopérativité d'oligomérisation

Nos résultats, contrairement à ce qui a été décrit indiquent que les paramètres cinétiques de DDX3 sont similaires à ceux décrit pour son homologue de levure Ded1p.

**A**



**B**

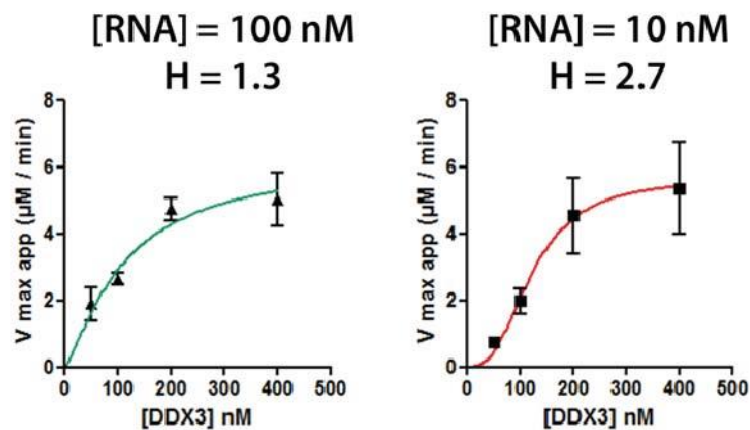


Figure 55 : Evaluation de la capacité d'oligomérisation de DDX3

(A) vitesse initiale de la réaction en fonction de la concentration en ATP pour différentes concentrations en ARN. (B) Les constantes de vitesse de l'activité ATPase montrent une dépendance sigmoïdale de la concentration de l'enzyme à de faible concentration en ARN (10 nM) ce qui est en accord avec une oligomérisation coopérative

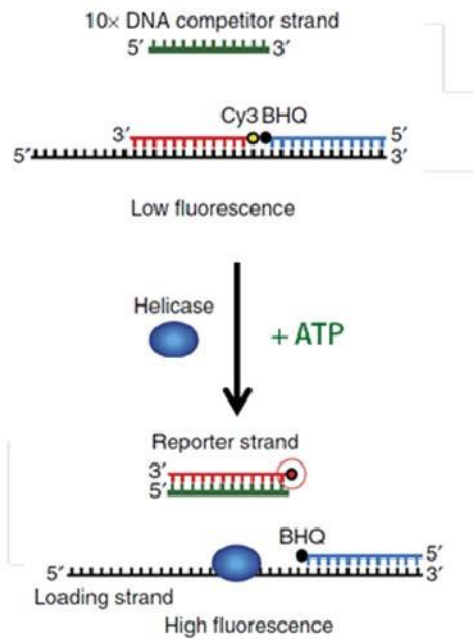
### C. Caractérisation de l'activité hélicase

De nombreuses fonctions biologiques assurées par les hélicases sont dépendante de la capacité de ces enzymes à dérouler des duplex d'acides nucléiques de manière ATP-dépendante. Il est donc crucial de déterminer l'activité hélicase de DDX3 afin de comprendre le fonctionnement de cette protéine.

Récemment, une nouvelle technique basée sur un la fluorescence a été mise au point (Özeş *et al.*, 2014), et permet de suivre en temps réel l'activité hélicase (Figure 56-A). En effet, deux oligonucléotides courts d'ARN sont modifiés par fluorescence, l'un est couplé à un fluorophore (la cyanine 3, présenté en brin rouge sur le schéma Figure 56-A) et l'autre à un quencheur (BHQ pour Black Hole Quencher, présenté en brin bleu sur le schéma Figure 56-A). Ces deux ARNs sont ensuite hybridés à un brin non modifié d'ARN (présenté en brin noir sur le schéma Figure 56-A), de telle sorte que le fluorophore et le quencheur portés par les deux ARNs courts se retrouvent à proximité l'un de l'autre et que la fluorescence puisse être éteinte. Lorsque l'hélicase est active, elle déroule le duplex. Ainsi, l'ARN portant le fluorophore (cy3) est libéré et la fluorescence est détectée. Pour prévenir la réassociation de l'ARN portant le fluorophore à l'ARN non modifié, un compétiteur ADN (présenté en brin vert sur le schéma Figure 56-A) se lie à l'ARN portant le fluorophore est présent en excès dans le mélange réactionnel.

Nous avons utilisé cette technique pour tester l'activité hélicase de MBP-DDX3. D'abord, la fluorescence du duplex hybridé est mesurée afin de valider que la fluorescence est éteinte (pas ou peu détectable) et par conséquent que l'hybridation correct des deux ARNs modifiés sur l'ARN non modifié a eu lieu. Puis l'enzyme est incubée en présence du duplex et la réaction hélicase est déclenchée par ajout d'ATP au mélange réactionnel. Les résultats présentées sur la Figure 56-B valident que la protéine MBP-DDX3-WT présente une activité hélicase puisqu'elle est capable de débobiner le duplex d'ARN. En effet, on mesure une augmentation de la fluorescence dû à la libération de l'ARN portant le fluorophore. Cette activité est dépendante de l'hydrolyse de l'ATP. Le mutant de perte d'activité ATPase (DQAD) ne présente pas d'activité hélicase. La quantité de duplex déroulé est ensuite rapportée en déterminant la régression linéaire dans les conditions de vitesse initiale pour les différentes concentrations en substrats (Figure 56-C).

A



B

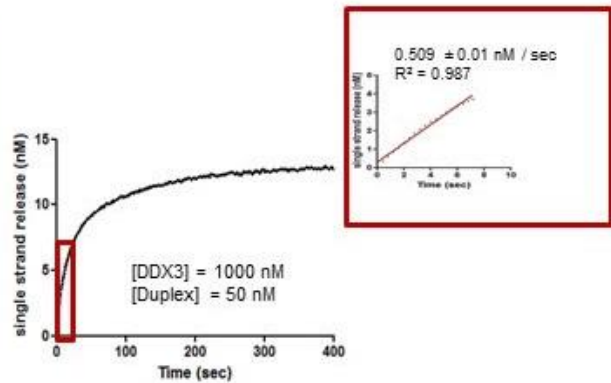
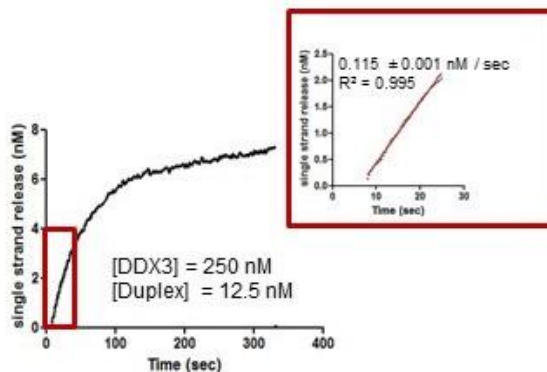
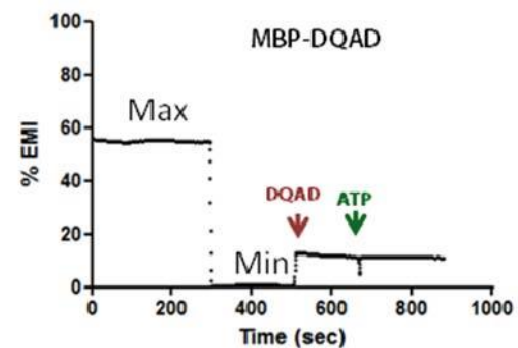
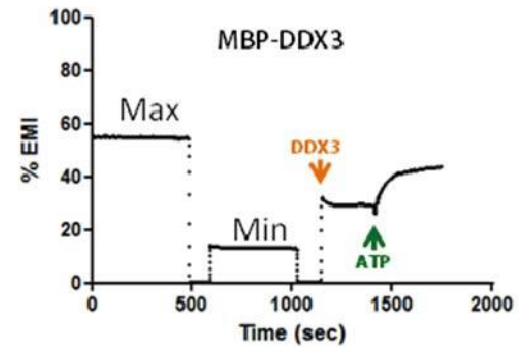
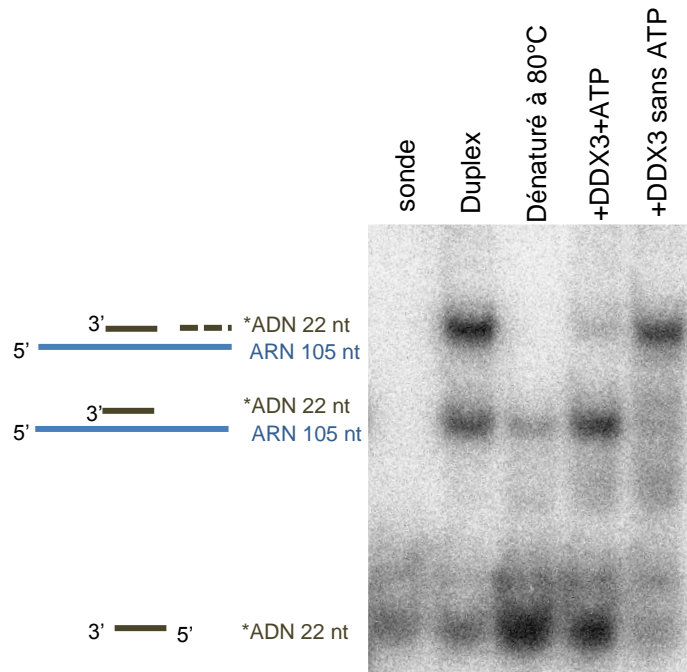


Figure 56 : Activité hélicase de DDX3 sur un duplex ARN/ARN

A : schéma illustrant le principe du test hélicase utilisé (d'après Özeş et al., 2014) . B : l'activité hélicase de MBP-DDX3 WT (en haut) et de MBP DDX3 DQAD (en bas) est testée en présence d'un duplex d'ARN. La réaction est déclenchée par ajout d'ATP (flèche verte) lorsque l'enzyme est active, une augmentation de la fluorescence est mesurée en temps réel. C : mesure de l'activité hélicase pour deux concentrations en enzyme et en substrats. L'encadrée rouge est agrandi, il correspond à la régression linéaire de la quantité d'ARN fluorescent libéré en fonction du temps (Vi).

Nous avons également testé l'activité hélicase de la protéine sur un hétéroduplex ADN/ARN. Pour cela, 1  $\mu$ M d'enzyme est incubée avec 150 nM d'hétéroduplexe ADN-ARN dont le substrat ADN est un oligonucléotide radiomarqué au P<sup>32</sup>, en présence ou non de 1 mM d'ATP. Puis, les produits de la réaction sont déposés sur un gel d'acrylamide 7% en conditions natives. Les résultats montrent que la protéine MBP-DDX3 est capable de dérouler des hétéroduplexe ADN-ARN d'une manière ATP dépendante (Figure 57, comparer les pistes 4 et 5). Néanmoins, cette activité doit être confirmée. En effet, les conditions utilisées pour ce test n'étaient pas optimales puisque l'ADN radiomarqué utilisé comme sonde, peut également s'hybrider partiellement à une autre région de l'ARN simple brin. Par conséquent, le duplex est formé d'une population hétérogène (un duplex avec l'ARN et l'ADN hybridé à une seule des deux régions uniquement, ou bien un duplex avec l'ARN et deux ADN hybridés).

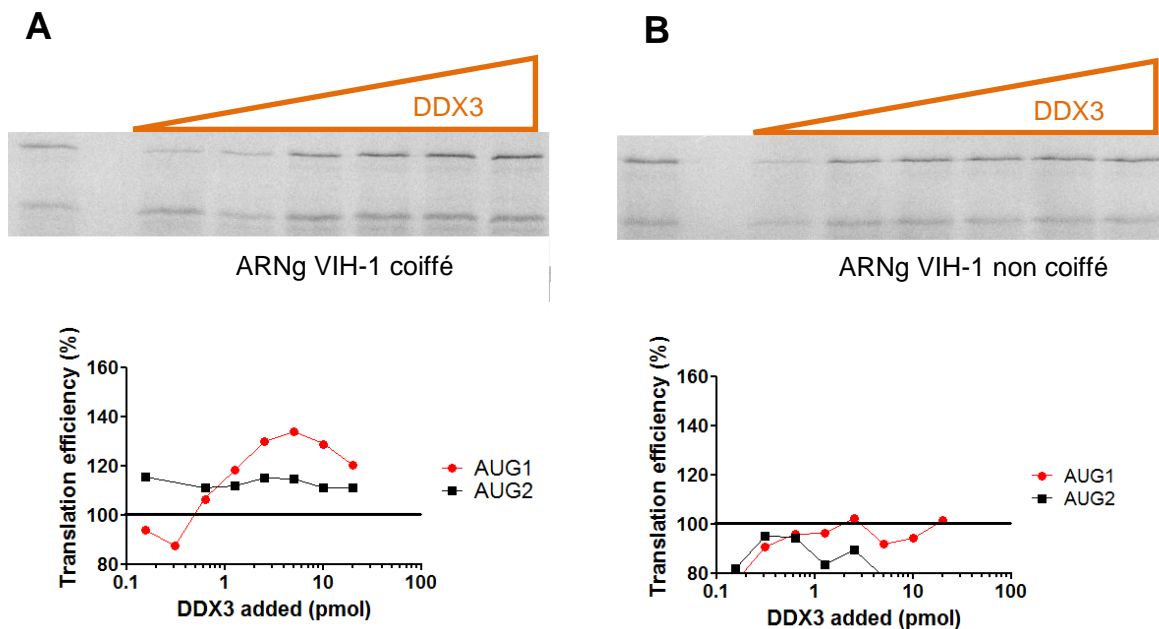


**Figure 57 : Test d'activité hélicase de DDX3 sur un hétéroduplexe**

*Séparation des produits de la réaction d'hélicase sur gel d'acrylamide en conditions natives. La piste 1 : sonde seule (ADN radiomarqué), piste 2 : le duplex ADN/ARN, piste 3 : le duplex dénaturé à 80°C pendant 5min, piste 4 : le duplex incubé avec l'enzyme et l'ATP et piste 5 : le duplex incubé avec l'enzyme mais sans ATP. MBP-DDX3 présente une activité hélicase ATP dépendante. A gauche, les duplex ainsi que la sonde sont représentés schématiquement. Hybridation totale de la sonde (trait plein). Hybridation partielle de la sonde (trait discontinu)*

#### D. Etude fonctionnelle de DDX3 sur l'initiation de la traduction virale

D'un point de vue fonctionnel, afin d'étudier le rôle de DDX3 sur l'initiation de la traduction du VIH-1, nous avons conduit une première série d'expérience *in vitro* de traduction de la protéine Gag à partir de l'ARNg du VIH-1 coiffé ou non en présence de quantité croissante (de 0 à 5 pmol) de la protéine MBP-DDX3. Les deux isoformes de Gag (P55 et P40) sont séparés sur gel SDS-PAGE. Puis quantifiées en exposant le gel à un écran sensible à la radioactivité. Les résultats montrent que DDX3 est capable de stimuler la production uniquement de la première isoforme de Gag (P55) et seulement à partir de l'ARN génomique coiffé (Figure 58-A). DDX3 semble donc impliquée dans le mécanisme de traduction dépendant de la coiffe.



**Figure 58 : La traduction *in vitro* de la polyprotéine Gag du VIH-1 en présence de DDX3**

*Traduction in vitro dans du RRL de l'ARNg du VIH coiffé (A) ou non (B) en présence de concentrations croissantes de la protéine DDX3. L'efficacité de traduction est mesurée pour les deux isoformes de Gag (P55 et P41). DDX3 stimule la production de la première isoforme de Gag*

Nous avons, donc obtenu avec succès une forme recombinante, soluble et active de DDX3 présentant des caractéristiques biochimiques similaires à celles décrites pour son homologue Ded1p. D'un point de vue fonctionnel DDX3 semble stimuler la production protéique dépendante de la coiffe de l'isoforme Gag p55 de du VIH-1.

## Discussion et perspectives

## I. Un nouveau rôle pour l'IRES Gag ?

Comme nous l'avons vu dans l'introduction, l'ARNg du VIH peut être traduit selon divers mécanismes et ce, malgré la présence de structures stables dans la région 5' non traduite. A ce jour, l'existence de chacun de ces mécanismes reste controversée et particulièrement, l'existence d'un IRES dans la région codante qui permettrait la traduction à partir d'un codon d'initiation amont apparaissait spécialement provocante. L'étude des mécanismes moléculaires gouvernant cet IRES nous a permis de formuler une nouvelle hypothèse quant au rôle physiologique de cet IRES. Ainsi, nous postulons que ce mécanisme pourrait certes être utilisé dans certaines conditions pour produire la protéine p55 Gag pleine longueur, ainsi que l'isoforme tronquée p40, mais pourrait également permettre la traduction de petits peptides codés dans des phases de lecture alternatives. Outre l'importance de l'isoforme tronquée décrite par Buck ([Buck et al., 2001](#)) et par l'équipe du Dr Swanson (communication personnelle), il se pourrait que la conservation de la propriété IRES dans la région codante de Gag trouve une explication dans la production de ces petits peptides.

Toutefois, ces conclusions s'appuient essentiellement sur une analogie frappante entre l'IRES Gag et celui présent dans le virus de la carpe *Halasti arva* ([Abaeva et al., 2016](#)). Il faudra donc au laboratoire démontrer qu'il ne s'agit pas seulement d'hypothèses, mais que l'IRES Gag fonctionne réellement ainsi, c'est-à-dire recrute le ribosome par la présence de séquence très riche en A, et que ceux-ci scannent ensuite dans les deux directions à la recherche d'un codon d'initiation. Pour ceci, à la manière de l'étude réalisée par Abaeva et al., il pourrait être judicieux de reconstituer l'initiation sur l'IRES Gag, en suivant les mouvements du ribosome, par toeprint et footprint.

Par ailleurs, il faudra également démontrer que les petits peptides codés sur des phases alternatives sont effectivement produits dans des cellules infectées, voire chez des patients, et comprendre le rôle de ces peptides. Ces travaux pourraient être réalisés en collaboration avec l'équipe du Dr Moris (INSERM UMR-S945, Hôpital de la pitié Salpêtrière).

Enfin, par une analyse bio-informatique, il serait intéressant d'évaluer si cette propriété n'est pas plus répandue au sein des virus et si elle ne pourrait pas être un mécanisme commun pour leurrer le système immunitaire de l'hôte.



## II. Caractérisation biochimique de DDX3

Nous avons caractérisé l'activité ATPase de DDX3 et contrairement à ce qui a été décrit nous avons montré que cette activité est ARN-dépendante et qu'elle présente des constantes cinétiques similaires à celles décrites pour son homologue Ded1p. Il est à noter ici que nous avons réalisé des essais d'activité ATPase après la première étape de purification (purification sur colonne de Nickel uniquement) et que la protéine ainsi semi-purifiée présente une activité ATPase intrinsèque de même que le mutant DQAD. Dans ce contexte, nous pouvons conclure que les résultats publiés précédemment étaient un artefact probablement dû à la co-purification d'une protéine d'*E. coli* présentant une activité ATPase intrinsèque. Lors de ces mêmes études, il avait été montré que DDX3 était capable de lier et d'hydrolyser tous les NTP (Franca *et al.*, 2007). Il serait ainsi peut-être nécessaire de réévaluer la spécificité de DDX3 face à ses substrats.

Concernant l'activité hélicase, il est nécessaire d'effectuer d'autres expériences afin de mieux la caractériser. Pour cela, les tests d'activité doivent être reproduits en présence de duplex ADN, ARN, et hétéroduplexe ADN/ARN, afin de conclure sur la spécificité de substrat. De même, des expériences en présence de duplex de longueurs variables présentant des extrémités libres ou franches permettront de statuer sur la polarité de l'activité hélicase et sur la processivité de l'enzyme.

Bien que nous ayons démontré que les propriétés biochimiques de MBP-DDX3 et de la protéine clivée étaient similaires, voire identique, il pourrait être pertinent d'évaluer les activités biochimiques de la protéine "naturelle". Pour cela, il serait intéressant d'utiliser une lignée cellulaire exprimant "fortement" DDX3 (telle que la lignée Jurkat) et d'effectuer des expériences de fractionnement cellulaire et de purification successives sur des colonnes de type MonoS et MonoQ.

Une étude suggère que DDX3 peut se substituer à eIF4E et se lier à la coiffe (Soto-Rifo *et al.*, 2013a), or pour nos expériences nous avons utilisé des un transcrit non coiffé. Il serait donc judicieux de reproduire ces expériences en utilisant un ARN coiffé et de comparer les résultats obtenus afin d'évaluer l'effet de la coiffe sur les activités ATPase et hélicase de DDX3.



### III. Rôle de DDX3 dans l'initiation de la traduction dépendante de la coiffe de l'ARNg du VIH-1

#### A. Caractérisation du mécanisme d'initiation de la traduction virale impliquant DDX3

Des travaux menés par Soto-Rifo et al, montrent que DDX3 est au cœur de la régulation de la traduction du VIH-1 (Soto-Rifo *et al.*, 2012c, 2013a). Nos résultats préliminaires *in vitro* sont en accord avec ceux observés sur la stimulation de la traduction virale *in cellulo*. Cependant, ces travaux ont été réalisés en utilisant l'ARN du VIH-1 entier et rendent compte de l'effet de DDX3 sur la traduction virale globale. Afin de préciser le mécanisme moléculaire impliquant DDX3, il faudrait explorer le rôle potentiel de DDX3 dans l'initiation de traduction du VIH-1 aussi bien dépendant de la coiffe que des séquences IRES. Pour ce faire, la traduction à partir des séquences IRES (IRES<sub>5'UTR</sub> et IRES<sub>gag</sub>) seront étudiées en contexte monocistronique et bicistronique. C'est la comparaison de l'expression protéique dans ces deux contextes qui permettra de déterminer l'implication de DDX3 sur le mécanisme d'initiation coiffe et/ou IRES dépendant. Ces expériences seront dans un premier temps menées *in vitro* en utilisant la protéine recombinante sauvage et mutée (les mutants sont disponibles au laboratoires, voir matériels et méthodes), ainsi que différents segments de l'ARNg coiffé ou non : la 5'UTR du VIH-1 suivie d'un gène rapporteur, la région codant pour Gag, la 5' UTR du VIH-1 suivie de la région codant pour Gag ou encore la 5'UTR contenant des mutations/délétions, notamment au niveau de TAR et des motifs impliqués dans l'activité IRES. La dépendance vis-à-vis du facteur eIF4E sera également évaluée en utilisant des analogues de la coiffe qui titre eIF4E. Puis, *ex vivo*, ces mêmes constructions seront utilisées dans différents types cellulaires déplétée en DDX3.

Ces différentes expériences permettront non seulement de préciser dans quel mécanisme d'initiation de la traduction est impliquée DDX3 mais également d'évaluer si cette implication est dépendante des activités ATPase et hélicase de la protéine.

#### B. L'interaction DDX3 – ARN du VIH-1

##### Impact de DDX3 sur la structure de la 5'UTR de l'ARN viral

Il existe plusieurs modèles de structure secondaire de la 5'UTR du VIH-1 et il est vraisemblable que celle-ci évolue au cours du cycle viral en fonction de la concentration cellulaire en ARNg (états monomère-dimère) mais également sous l'influence de facteurs cellulaires et viraux. Soto-Rifo et al, proposent trois sites d'interaction entre l'hélicase cellulaire DDX3 et la 5'UTR du VIH-1 : au niveau de la tige boucle TAR (interaction majeure) et au niveau de régions simple brins situés sur le site de liaison du ARN<sub>t</sub><sup>lys3</sup> et entre le site majeur donneur d'épissage et le motif Psi  $\psi$  d'encapsidation (Soto-Rifo et al, 2012). Il serait intéressant de compléter cette étude en déterminant l'impact de DDX3 sur la structure globale de la 5'UTR, en utilisant la technique SHAPE.

Afin d'évaluer l'impact de DDX3 sur la structure de la 5'UTR aussi bien sous sa forme monomère que dimère, la protéine sauvage peut être utilisée en présence et en absence d'ATP, mais également en présence d'un analogue non hydrolysable de l'ATP comme l'AMPPNP, ou d'un mutant de DDX3 capable de fixer l'ATP mais incapable de l'hydrolyser (mutant DQAD disponible au laboratoire) de telle sorte à définir l'empreinte de la protéine sur l'ARN. Il serait également intéressant de reproduire ces expériences en utilisant un ARN coiffé, pour déterminer si l'effet de DDX3 sur la 5'UTR est dépendant de la présence de la coiffe.

#### Détermination du site de fixation de DDX3 sur l'ARN viral

L'affinité de la protéine pour l'ARN viral pourra être déterminée en utilisant la technique de rétention sur filtre. Pour cela des concentrations croissantes de la protéine sont incubés en présence de l'ARN viral. L'implication de la coiffe peut être évaluée en comparant l'affinité de la protéine pour un ARN coiffé et un ARN non coiffé.

Le ou les sites de liaison pourraient être précisés en utilisant cette technique et en suivant la même approche de délétion par PCR (mutants de délétions déjà disponible au laboratoire) que celle utilisée pour la définition du site de liaison entre l'ARN viral et la sous-unité ribosomale 40S.

#### C. Interactions DDX3 avec les facteurs d'initiation de la traduction

Afin de caractériser le mécanisme moléculaire impliquant DDX3, l'une des premières étapes consiste à déterminer la composition des complexes d'initiation de la traduction des ARNm dépendant de DDX3. Pour cela, les complexes d'initiation formés dans un lysat cellulaire sur un ARN cible de DDX3 seront purifiées par chromatographie d'affinité et analysées par Western Blot et / ou spectroscopie de masse.

Une première série de pull down réalisée par un autre membre du laboratoire, montre une interaction de DDX3 avec eIF4G (le fragment p100, délétée du domaine N-terminal de liaison à eIF4E) et confirme les précédents résultats obtenus *in vitro* et *in cellulo* dans le contexte de lymphocytes CD4+ infectées (Soto-Rifo et al., 2012c, 2013a). De même une interaction avec eIF3 a été observée et confirme les résultats d'une étude similaire (Lee et al., 2008b). Aucune interaction avec eIF4E n'a été observée lors des *pull-down* ceci est en accord avec les résultats observés *in cellulo* (Soto-Rifo et al., 2013a) mais en contradiction avec les résultats obtenus par Shih et al. En effet, ces derniers ont identifié un site d'interaction entre DDX3 et eIF4E au niveau du domaine N-terminale de DDX3, où l'on retrouve la séquence consensus YxxxL $\phi$  de liaison à eIF4E (Shih et al., 2008). L'absence d'interaction entre DDX3 et eIF4E tend plutôt à favoriser le modèle selon lequel DDX3 se substituerait à eIF4E pour permettre l'assemblage du complexe d'initiation au niveau de la coiffe.

Enfin, l'interaction observée entre DDX3 et DHX9 (RHA) conforte l'hypothèse d'un possible rôle de DDX3 dans la traduction d'ARNm dont la 5'UTR présente des structures secondaires stables. Toutefois, des expériences complémentaires s'avèrent nécessaires pour confirmer les résultats obtenus et pour distinguer les interactions directes protéines-protéines des interactions indirectes médiée par exemple par l'ARN. Pour cela, ces expériences seront reproduites en présence de RNase pour comparer les protéines éluées avec DDX3 aux résultats précédents.

Il est également possible de recourir à des méthodes biophysiques de caractérisation d'interactions protéiques telles que la mesure de fluorescence ou encore la microcalorimétrie.

Il semble également très intéressant de tester l'effet des partenaires protéiques de DDX3 sur ces activités biochimiques.

#### D. Validation par reconstruction

Afin de disséquer la séquence des évènements requise pour former un complexe d'initiation sur l'ARNg du VIH-1, il faudrait reproduire entièrement l'étape d'initiation *in vitro*. Un tel système existe et est utilisé en routine au laboratoire. Les facteurs purifiés sont ajoutés un par un et à chaque étape, les complexes formés sont caractérisés par toeprint, footprint et gradient de densité. DDX3 pourra être ajouté à ce système

#### E. Recherche d'ARNm cibles de DDX3

Pour déterminer les ARNm cellulaires dont la traduction est dépendante de DDX3, des expériences de CLIP (Cross Linking and Immunoprécipitation) et de SHAPE-map seront conduites dans des cellules exprimant ou non DDX3. Une fois les cibles ARN déterminés, il serait intéressant de déterminer si cette interaction est directe et d'évaluer l'affinité de la protéine pour ces ARNs ainsi que l'impact de la liaison de DDX3 sur la structure de ces ARNs.

## Matériels et méthodes

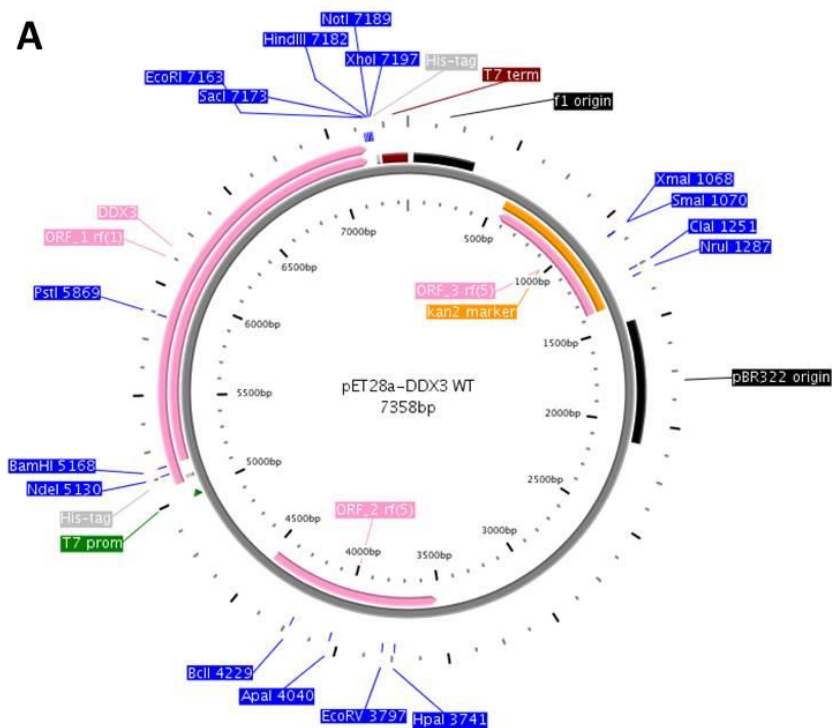
## I. Plasmides

Les plasmides utilisés pour l'expression de la protéine DDX3 ont été obtenus en sous-clonant la séquence codante de DDX3 sauvage (DDX3-WT) dans les vecteurs :

pET28-a entre les sites de restriction BamHI et EcoRI

pET21-a entre les sites de restriction NdeI et Xho I (déjà disponible au laboratoire)

pETM-41 entre les sites de restriction NcoI et EcoRI



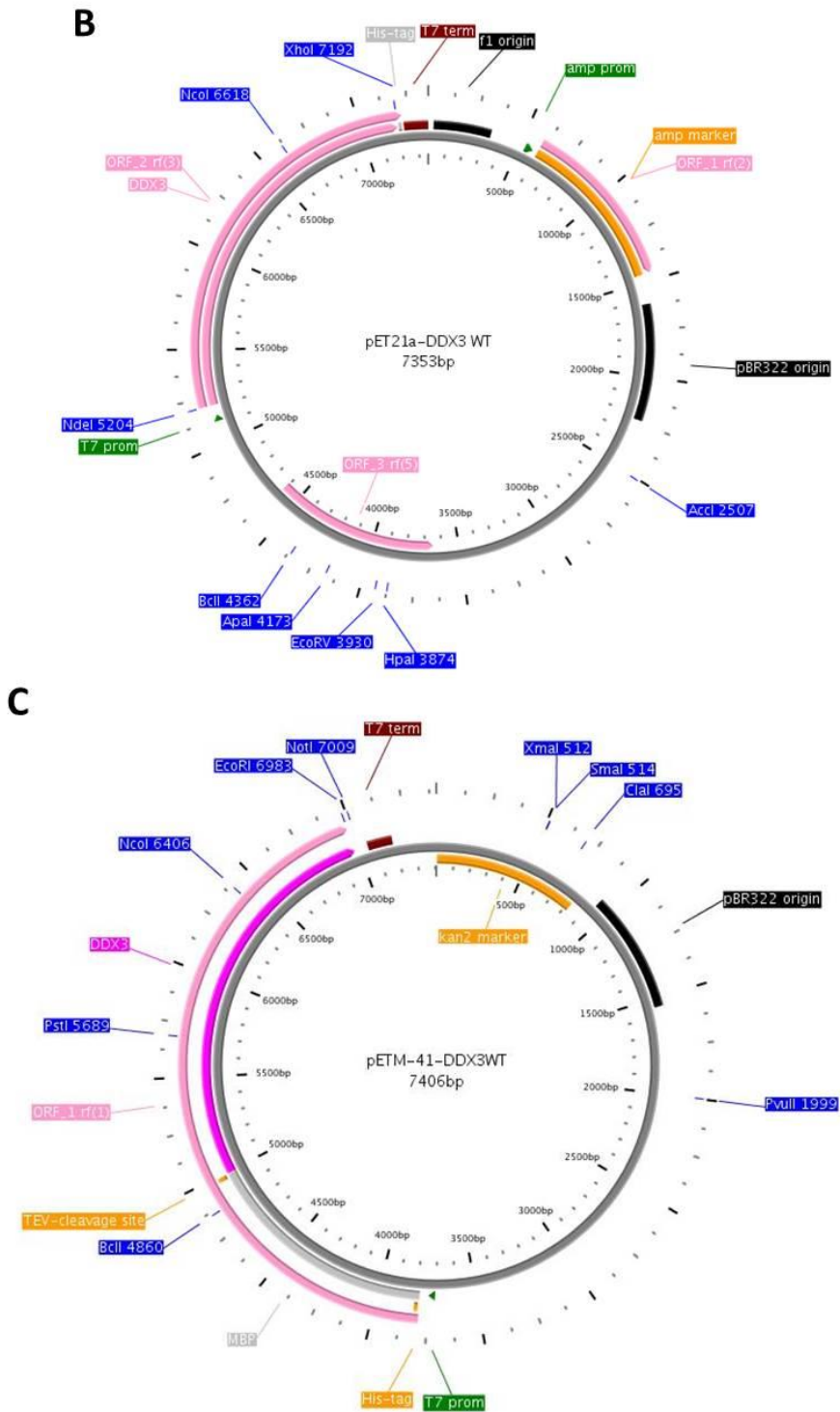


Figure 59: Plasmides utilisés pour l'expression protéiques

A : pET28-a DDX3WT, clonée entre les sites de restriction BamHI et EcoRI, B : pET21-a DDX3WT, clonée entre les sites de restriction NdeI et Xho I C: pETM-41 DDX3WT, clonée entre les sites de restriction NcoI et EcoRI



## II. Les mutants de DDX3

Les mutants de DDX3 ont été sous clonés dans le vecteur pETM-41 (les nucléotides mutés sont encadrés en noir dans la séquence de la phase codante de DDX3 ci-dessous, le code couleur est respecté)

- **Y38A/L43A** : mutation au niveau du domaine N-terminal, mutation du  $_{112}\text{TAT}_{114}$  en GCT, et mutation du  $_{127}\text{TTA}_{129}$  en GCA, altération de la liaison à eIF4E
- **K230E** : mutation au niveau du motif I, mutation du  $_{688}\text{AAA}_{690}$  en GAG, abolit l'activité ATPase
- **DQAD** : mutation au niveau du motif II (boîte DEAD), mutation du  $_{1042}\text{GAA}_{1044}$  en CAA, abolit l'activité ATPase
- **S382L** : mutant au niveau du motif III, mutation du  $_{1144}\text{AGT}_{1146}$  en CTA, abolit partiellement l'activité ATPase

## III. Séquence de la phase codante de DDX3-WT

```
5'ATGAGTCATGTGGCAGTGGAAAATGCGCTCGGGCTGGACCAGCAGTTTGCTGGCCTAG
ACCTGAACTCTTCAGATAATCAGAGTGGAGGAAGTACAGCCAGCAAAGGGCGCTATAATT
CCTCCTCATTTAAAGGAACCGAGAAGCTACTAAAGGTTTCTACGATAAAGACAGTTTCAGG
GTGGAGTTCTAGCAAAGATAAGGATGCGTATAGCAGTTTGGATCTCGTAGTGATTCAAG
AGGGAAGTCTAGCTTCTCAGTGATCGTGAAGTGGATCAAGGGGAAGGTTTGATGATC
GTGGACGGAGTGATTACGATGGCATTGGCAGCCGTGGTGACAGAAGTGGCTTTGGCAAA
TTTGAACGTGGTGGAAACAGTCGCTGGTGTGACAAAATCAGATGAAGATGATTGGTCAAA
ACCACTCCCACCAAGTGAACGCTTGGAAACAGGAACTCTTTCTGGAGGCAACACTGGGA
TTAATTTTGAGAAAATACGATGACATTCCAGTTGAGGCAACAGGCAACAACACTGTCCTCCAC
ATATTGAAAGTTTCAGTGATGTTGAGATGGGAGAAATTATCATGGGAAACATTGAGCTT
ACTCGTTATACTCGCCAACTCCAGTGCAAAAGCATGCTATTCTATTATCAAAGAGAAA
AGAGACTTGATGGCTTGTGCCAAACAGGGTCTGGAATAAAGCTGCAGCATTCTGTGGCC
ATCTTGAGTCAGATTTATTCAGATGGTCCAGGCGAGGCTTTGAGGGCCATGAAGGAAAA
TGGAAGGTATGGGCGCCGCAACAATACCAATCTCCTTGGTATTAGCACCAACGAGAG
AGTTGGCAGTACAGATCTACGAGGAAGCCAGAAAATTTTCATACCGATCTAGAGTTCGTC
CTTGCGTGGTTTATGGTGGTGCCGATATTGGTTCAGCAGATTCGAGACTTGGAAACGTGGAT
GCCATTGTTAGTAGCCACTCCAGGACGTCTAGTGGATATGATGGAAAGAGGAAAGATT
GGATTAGACTTTTGCAAATACTTGGTGTAGATGAACTGATCGGATGTTGGATATGGGG
TTTGAGCCTCAGATTCGTAGAATAGTCAACAAGATACTATGCCTCCAAAGGGTGTCCGC
CACACTATGATGTTTGTCTACTTTTCTAAGGAAATACAGATGCTGGCTCGTGATTTCT
TAGATGAATATACTTCTTGGCTGTAGGAAGAGTTGGCTCTACCTCTGAAAACATCACAC
AGAAAGTAGTTTGGGTGGAAAGATCAGACAAACGGTCATTTCTGCTTGACCTCCTAAATG
CAACAGGCAAGGATTCAGTACCTTAGTGTGTTGTGGAGACCAAAAAGGGTGCAGATTCT
CTGGAGGATTTCTATACCATGAAGGATACGCATGTACCAGCATCCATGGAGACCGTTCT
CAGAGGGATAGAGAAGAGGCCCTTACCAGTCCGCTCAGGAAAAAGCCCAATTTAGT
GGCTACAGCAGTAGCAGCAAGAGGACTGGACATTTCAAATGTGAAACATGTTATCAATT
TTGACTTGCCAAGTGATATTGAAGAATATGTACATCGTATTGGTCTGACGGGACGTGTAG
GAAACCTTGGCCTGGCAACCTCATTCTTTAACGAGAGGAACATAAATATTACTAAGGATT
TGTGGATCTTCTTGTGAAGCTAAACAAGAAGTGCCGCTTGGTTAGAAAACATGGCTT
ATGAACACCACTACAAGGGTAGCAGTCGTGGACGTTCTAAGAGTAGCAGATTTAGTGGAA
GGGTTTGGTGCCAGAGACTACCGACAAAGTAGCGGTGCCAGCAGTTCCAGCTTCCAGCAG
CAGCCGCGCAAGCAGCAGCCGCAAGTGGCGGAGGTGGCCACGGTAGCAGCAGAGGATTT
GGTGGAGGTGGCTATGGAGGCTTTTACAACAGTGATGGATATGGAGGAAATTATAACTC
CCAGGGGGTTGACTGGTGGGGTAAC3'
```

#### IV. Souches bactériennes

Pour la préparation des plasmides la souche utilisée est XL1-blue *d'E.coli*

Pour l'expression des protéines recombinantes la souche utilisée est BL21(DE3) expriment la polymérase du phage T7 sous contrôle du promoteur de l'opéron lactose

#### V. Expression protéique

100 µl de bactéries BL21 (DE3) chimio-compétentes sont transformées par choc thermique en présence de 100 ng de plasmide puis étalées sur une boîte de milieu LB agar en présence d'antibiotique pour la sélection des clones (ampicilline à 100 mg/ml avec le plasmide pET21a-DDX3 et kanamycine à 50 mg/ml avec le plasmide pET28a-DDX3 et pETM41-DDX3) et incubée à 37°C pendant 16h (toute la nuit). À partir d'un clone, une préculture de 10 ml de milieu LB avec ampicilline ou kanamycine (LBA/LBK) estensemencée et incubée 16h (toute la nuit) sous agitation à 37°C.

Un clone estensemencé en pré-culture de 20 ml de milieu LB contenant l'antibiotique de sélection adéquat, puis placée sous agitation à 37°C pendant 16h (toute la nuit). Une culture de 1L de milieu LB avec antibiotique adéquat est inoculée à une densité optique à 600 nm (D.O.600) de 0,05 et placée sous agitation à 37°C. Lorsque la culture atteint une D.O.600 de 0,5 (phase exponentielle de croissance) l'expression protéique est induite par l'ajout d'isopropyl β-D-galactopyranoside (IPTG) à une concentration finale de 0,1 mM. La culture est ensuite placée sous agitation à la température et au temps indiqués. Les cultures sont ensuite centrifugées à 4°C à 3200 rpm pendant 15 minutes. Le surnageant est retiré. Le culot est conservée à -80°C.

#### VI. Purification

Le culot est repris dans 10 mL de tampon contenant 50 mM de phosphate de sodium à pH8, 300 mM de NaCl, 10 mM d'imidazole et 1 mM de PMSF (inhibiteur de protéases). Les cellules sont ensuite lysées au moyen d'une presse d'Eaton. La fraction soluble est récupérée par centrifugation à 4°C à 24000 rpm pendant 30 minutes, puis filtrée sur une colonne de nickel (HisTrap 5 mL, GE Healthcare). Les protéines retenues sont éluées avec 500 mM d'imidazole et résolues sur un gel SDS-PAGE. Les fractions contenant la protéine DDX3 sont rassemblées et dialysées contre le tampon A : 20 mM de Tris pH8, 80 mM de KCl, 1 mM de DTT et 4 % de glycérol. La deuxième étape de purification consiste en une colonne échangeuse d'anion (MonoQ, GE Healthcare). Le produit de purification par la colonne d'affinité (HisTrap) dialysé, est déposé sur la colonne de MonoQ préalablement équilibrée dans le tampon A. les protéines sont ensuite éluées en fraction de 1 mL avec un gradient allant de 0 à 30% en tampon B : 20 mM de Tris pH8, 1 M de KCl, 1 mM de DTT et 4 % de glycérol.



## VII. Clivage et purification de la protéine clivée

Pour les tests de stabilité de la protéine clivée, 500 µg de protéine de fusion DDX3-MBP sont incubés sur roue en présence de 1mM de DTT et de 1 U de TEV (produite au laboratoire) pendant 4 h à 20°C. 100 µL de matrice de nickel (ThermoFisher Scientific) préalablement équilibrée en tampon A (Tris pH8 20 mM, KCl 260 mM, imidazole 10 mM) sont ajoutée. L'incubation est poursuivie pendant 30 minutes. Deux lavages en tampon A, sont réalisés. L'élution se fait en tampon A complété de 500 mM d'imidazole. Un aliquot est incubé pendant une nuit à -20°C, un autre pendant 1 h à 37°C. Les fractions solubles et insolubles résultantes sont séparées par centrifugation à 13000 rpm pendant 10 minutes. Les différentes fractions sont analysées par électrophorèse sur gel de polyacrylamide.

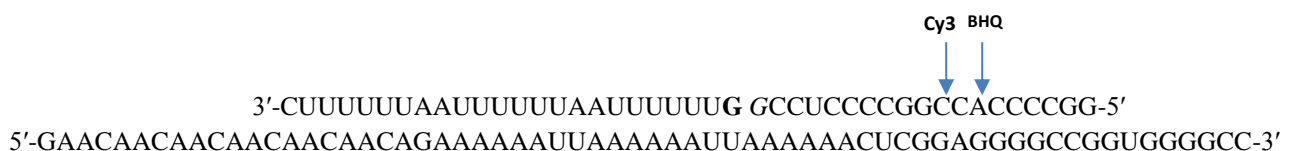
## VIII. Test d'activité ATPase

Les cinétiques enzymatiques sont réalisées en présence de 100 nM d'enzyme incubée durant 10 min à 37°C en présence de 100 nM d'ARN à des concentrations différentes d'ATP-γP<sup>32</sup> (50µM, 100µM, 200µM, 400µM). Puis, les produits de la cinétique sont déposés sur une plaque de PEI cellulose et migrés dans du KH<sub>2</sub>PO<sub>4</sub>. Le phosphate inorganique (P<sub>i</sub>) ayant une plus forte affinité pour le solvant (phase mobile) que pour la phase solide, va migrer jusqu'au front de migration. L'ATP lui possède une plus forte affinité pour la phase solide est reste donc au niveau du point de dépôt. Après exposition de la plaque à un écran sensible à la radioactivité, le signal est quantifié et rapporté sous forme de graphe représentant le P<sub>i</sub> relargué en fonction du temps. Ensuite, la régression linéaire du P<sub>i</sub> en fonction du temps permet de déterminer la vitesse initiale (V<sub>i</sub>) pour chaque concentration en ATP. Les valeurs de V<sub>i</sub> sont rapportées sous forme de courbe V<sub>i</sub> en fonction de la concentration en ATP et afin de mesurer V<sub>max</sub> et K<sub>M</sub> apparents.

## IX. Test d'activité hélicase

Nous avons utilisé pour le test d'activité hélicase, une nouvelle technique basée sur un la fluorescence (Özeş *et al.*, 2014).

Pour cela, deux ARN courts sont modifiés par fluorescence, l'un est couplé à un fluorophore (la cyanine 3) et l'autre à un quencheur (BHQ pour Black Hole Quencher). Ces deux ARNs sont ensuite hybridés à un brin non modifié d'ARN transcrit *in vitro*.



Ces trois ARNs sont mélangés à une concentration finale de 500 nM de chaque (volume dépendant des besoins) dans un tampon contenant 20 mM de tris-acétate à pH 7,5, 100 mM de KCl, 2 mM de MgCl<sub>2</sub> et 1 mM de DTT. L'hybridation se fait en plaçant d'abord le tube contenant ces ARNs à 80°C dans un bain sec pendant 2min puis laisser redescendre à température ambiante (25°C) recouvert d'aluminium pour éviter le photobleaching. Puis placés sur la glace 10min. Un ADN compétiteur de l'ARN portant le fluorophore (Cy3) est ajoutée à une concentration 10 fois plus importante, afin de piéger les ARNs libres lors de la réaction et éviter leur ré-hybridation sur l'ARN non marqué.

Un contrôle positif est préparé en mélangeant tous les constituants de la réaction à l'exception de l'ARN portant le quencher BHQ. Ce contrôle servira à mesurer le maximum de fluorescence de la réaction. Le contrôle négatif consiste à mesurer la fluorescence du duplex après hybridation.

Des concentrations variables en duplex (12 nM à 250 nM) sont ensuite incubées en présence de 250 nM à 1 µM d'enzyme (MBP-DDX3). La réaction d'hélicase est déclenchée par l'ajout de 2mM d'ATP. La fluorescence émise est mesurée sur le SAFAS-monaco. Puis la fraction de duplex déroulé est rapportée en fonction du temps et la vitesse initiale est calculée en mesurant la régression linéaire du duplex déroulé.

## X. PCR

Le mélange réactionnel comprend, dans un volume final de 100 µl, 4 ng d'ADN matrice, 0,2 µM d'amorces sens et anti-sens, 200 µM de chaque dNTP, et 1U de Taq polymérase utilisée avec le tampon 1X recommandé par le fournisseur. L'amorce s'hybridant en 5' comporte une séquence s'appariant parfaitement avec la matrice, précédée d'une séquence correspondant à la séquence du promoteur de l'ARN polymérase T7.

Une réaction de PCR de 30 cycles est ensuite réalisée avec le programme de PCR suivant:

- dénaturation initiale: 94°C pendant 5 minutes.

- 30 cycles composés chacun :

  - D'une étape de dénaturation à 94°C pendant 30 secondes.

  - D'une étape d'hybridation à 55°C pendant 45 secondes (température d'hybridation variable et dépendante des amorces).

  - D'une étape d'élongation à 72°C pendant 1 minute (temps d'élongation variable et dépendant de la taille du produit à amplifier. 1min /Kb).

- élongation terminale: 72°C 10 minutes.

Après électrophorèse sur gel d'agarose 1% en présence de sybersafe, les amplicons sont visualisés sous UV afin de vérifier que leur taille correspond à celle attendue, puis purifiés à l'aide du kit GenJet PCR purification kit (Thermo Fisher Scientific)

## **XI. Amorces PCR**

La séquence ADN correspondant à la 5'UTR suivie de la région codante de GAG DU VIH-1 (5'UTR\_Gag), obtenue à partir du clone proviral pNL4.3, amplifiée par PCR en utilisant:

L'amorce sens : T7 5'UTR HIV-1

5'-TAATACGACTCACTATAGGTCTCTCTGGTTAGACCAGATCT-3',

Et l'amorce antisens: Hiv1-3'UTR reverse

5'-AGGACACCAAATGAAAGATTGACTGAGAGACAGGCTAATTTTTT-3'

Les amorces utilisées pour la délétion des sites de liaisons sont:

Pour la deletion du site 1 (del363-435)

L'amorce sens : HIV1\_del363-435

5'-AGAGCGTCGGTATTAAGCTTAGAGACATCAGAAGGC-3'

Et l'amorce antisens : HIV1\_del363-435

5'-GCCTTCTGATGTCTCTAAGCTTAATACCGACGCTCT-3'

Pour la deletion du site 2 (del632-736)

L'amorce sens : HIV1\_del632-736

5'-GGATAGATGTAAAAGACACCAAGGAAGTGCAGAACCTCC-3'

Et l'amorce antisens HIV1\_del632-736

5'-GGAGGTTCTGCACTTCCTTGGTGTCTTTTACATCTATCC-3'

Les amorces utilisées pour les sites de liaisons seuls :

Pour le site 1

L'amorce sens : T7-HIV-1 J2-3

5'- TAATACGACTCACTATAGGAGAATTAGATAAATGGGAAAAAATTCGG-3'

Et l'amorce antisens : HIV-1 P3 rev

5'- AAAAGGCCAGGATTAAGTGC GAATC-3'

Pour le site 2

L'amorce sens : T7-HIV1\_P7

5'- TAATACGACTCACTATAGGAAGCCTTAGATAAGATAGAGGAAGAG-3'

Et l'amorce antisens HIV1\_P9 rev

5'-CTGGAGGTTCTGCACTATAGGGTAA -3'

## **XII. Transcription *in vitro***

Les ARNs monocistroniques 5'UTR\_Gag du VIH-1 WT et mutants sont directement transcrits à partir du produit de PCR purifié contenant le promoteur T7 : 5'-TAATACGACTCACTATA-3'.

La transcription est réalisée dans un volume final de 100 µl avec au moins 2 µg de matrice ADN (le produit de PCR purifié) transcrit en présence de tampon 1X (contenant : 5 mM de DTT, 40 mM de Tris pH 8, 1 mM de spermidine) de 4 mM de chaque NTP, 22 mM de MgCl<sub>2</sub>, 1U d'inhibiteur de RNases(Promega), et 1U d'ARN polymérase T7 purifiée au laboratoire. Le mélange réactionnel est incubé à 37°C, 1U de pyrophosphatase est ajouté après apparition d'un précipité. Après 3 heures de transcription, 1U de DNase RQ1(Promega), est ajoutée, pour dégrader la matrice ADN. L'ARN néo-synthétisé est précipité par addition de 2,5 M de LiCl. Après lavage à l'éthanol 70% et séchage, le culot est repris dans 50 µl d'eau et purifié par gel filtration (colonnes de sephadex G50). Les ARN sont dosés par mesure de l'absorbance à 260 nm. (Biospec Nano). Puis vérifiés par électrophorèse sur gel d'agarose

Pour chaque ARN, la coiffe et la queue polyA sont ajoutés en utilisant respectivement les kits M2080 et M0276 tous les deux de New England Biolabs et purifiés par chromatographie d'exclusion de taille.

Les transcriptions radioactives sont réalisées suivant le même protocole, mais avec 1 mM d'UTP non marqué et 1 µl de 3000 mCi / mmol d'UTPα-<sup>32</sup>P.

### XIII. Traduction *in vitro*

Des dilutions en série d'une 1 pmol d'ARNs transcrits *in vitro* sont préparés et traduits pendant 60 min à 30 ° C dans un volume réactionnel de 10 ul en présence de 60% (v / v) d'un lysat de réticulocytes de lapin traité (Flexi®-RRL; Promega), supplémenté avec 0,5 mM d'acétate de magnésium, 100 mM de chlorure de potassium et 0,02 mM d'acides aminés sans méthionine et 0,6 mCi / ml de méthionine-<sup>35</sup>S.

Pour la variation en concentration de protéines, 0,25 pmol d'ARN est traduit en présence de 0 à 5 pmol de protéines (DDX3). Le protocole suivi est le même que celui décrit plus haut.

La réaction est ensuite stoppée avec 90 ul de tampon de charge (10% de SDS, 50 mM de Tris pH 6,8, 10% de glycérol, 100 mM de DTT et 0,1% de bleu de bromophénol). 10 ul du mélange final est chargé et résolu sur gel SDS-PAGE (12%). Le gel est fixé pendant 20 minutes dans une solution de 30% d'éthanol et 10% d'acide acétique, et séché pendant 60 minutes dans un système de séchage sous vide à 75 ° C. Les produits radio-marqués sont visualisés en utilisant PhosphorImager Storm (GE Healthcare) et quantifiés en utilisant le logiciel MultiGauge.

Chaque expérience est répétée trois fois de façon indépendante à l'aide de différentes préparations d'ARN.

### XIV. La rétention sur filtre

L'ARN radiomarqué est dénaturé 2 minutes à 80°C puis renaturé par retour à température ambiante durant 10 minutes. 10 nM d'ARN sont incubés en présence de différentes concentrations en ligand (la sous-unité 40S du ribosome, purifiés à partir de cellules Hela) pour permettre la formation des complexes, dans un tampon contenant 20 mM de Tris pH 7.5, 200 mM de chlorure de potassium (KCl), 100 mM d'acétate de sodium (Koac), 2.5 mM de MgCl<sub>2</sub> et 1 mM de DTT. Le mélange est ensuite filtré sur une membrane de nitrocellulose, retenant les protéines et donc les complexes formés, et une membrane de nylon retenant les ARN libres. Les membranes sont au préalable incubées dans le tampon réactionnel. Puis les membranes sont exposées à un écran sensible à la radioactivité, révélé sur à l'aide du phosphorImager (GE Healthcare) et le signal est quantifié à l'aide du logiciel « multigauge ».

Pour chaque concentration, la proportion de complexes formés est calculée en rapportant le signal quantifié sur la membrane de nitrocellulose au signal total (nitrocellulose plus nylon). La constante de dissociation  $K_d$  apparent est déterminé en utilisant l'isotherme de Langmuir dont l'équation est la suivante :  $\Phi = [P] / [P] + K_d$  où  $[P]$  est la concentration en eIF3,  $\Phi$  est le pourcentage d'ARN lié et  $K_d$  correspond à la

concentration en ligand pour laquelle on observe deux fois moins de complexes formés qu'à saturation en ligand

## **XV. Gradient de densité**

Les complexes ribonucléoprotéiques sont assemblés sur l'ARN marqué au  $^{32}\text{P}$ . 2 pmol d'ARN (10  $\mu\text{l}$ ) sont dénaturées à 85 ° C pendant 2 min, suivie d'une renaturation à température ambiante pendant 10 min. La formation de complexes avec la sous-unité ribosomale 40S purifiée à partir de cellules HeLa ou RRL, est réalisée par incubation de 20 nM d'ARN avec 400 nM de 40S pendant 15 min à 37 ° C dans un tampon contenant 20 mM de Tris pH 7,5 100 mM de KOAc, 200 mM de KCl, 1 mM de DTT, 2,5 mM de  $\text{MgCl}_2$ . Les réactions sont arrêtées sur de la glace et déposées à la surface d'un gradient de 10 à 30% de succrose (25 mM de Tris pH 7,6, 6 mM de  $\text{MgCl}_2$ , 75 mM de KCl) puis sédimentés par ultracentrifugation à 39 000 rpm pendant 4 heures à 4 ° C avec le rotor SW40 Ti (Beekman Coulter). Des fractions de 300  $\mu\text{l}$  sont récoltées. 100  $\mu\text{l}$  de chaque fraction sont déposés sur une membrane en nylon (Hybond N+, Amersham), à l'aide d'un appareil à dot blot. La membrane est exposée à un écran sensible à la radioactivité, puis révélé au PhosphorImager Storm (GE Healthcare) et quantifiée en utilisant le logiciel MultiGauge. La quantité d'ARN dans chaque fraction est déterminée et exprimée en pourcentage des comptages totaux. Puis rapporté sous forme de graphe de pourcentage d'ARN lié à la sous-unité ribosomale 40S en fonction de la position dans le gradient.

# Annexe

Facteurs d'initiation	Fonctions
<b>eIF1, eIF1A</b>	Se lient directement au ribosome et induisent une conformation «ouverte» du tunnel d'entrée de l'ARNm au niveau du ribosome, ce qui est crucial pour l'étape de balayage. Ils assurent également la fidélité de reconnaissance du codon d'initiation et stimulent le recrutement du complexe ternaire (eIF2-GTP- ARNt <sub>i</sub> <sup>Met</sup> ) à la sous-unité 40S.
<b>eIF2</b>	Elle est formée de trois sous-unités ( $\alpha, \beta, \gamma$ ) et assure la délivrance de l'ARNt <sub>i</sub> <sup>Met</sup> au niveau du site P. eIF2 forme un complexe ternaire, avec l'ARNt <sub>i</sub> <sup>Met</sup> et une molécule de GTP, qui se lie au ribosome.
<b>eIF2B</b>	Facteur d'échange du GDP en GTP. Le GDP d'eIF2-GDP est échangé par eIF2B en GTP afin de former un eIF2-GTP actif.
<b>eIF3</b>	Facteur multi-protéique composée de 13 sous-unités impliqué dans différentes étapes de l'initiation. Ainsi, eIF3 se lie directement à la sous-unité ribosomale 40S et empêche l'association de la sous-unité 60S, il stimule le recrutement du complexe ternaire, l'assemblage et la stabilisation du complexe de préinitiation 43S. eIF3 est également impliqué dans la formation du complexe 48S du fait de son interaction directe avec le facteur eIF4G, permettant ainsi de faire le lien entre le complexe eIF4F et la sous-unité ribosomale 40S. Il est aussi impliqué dans le recrutement de l'ARNm par l'intermédiaire de domaines spécifiques de liaison à l'ARN. Il joue aussi un rôle dans le balayage de la 5'UTR et dans la reconnaissance du codon initiateur par l'action de sa sous-unité eIF3c qui interagit avec eIF5 et eIF1. eIF3 agit également comme un facteur de ré-initiation et de recyclage des ribosomes en se liant à la sous-unité 40S dès la fin de la terminaison.
<b>eIF4A</b>	eIF4A est une hélicase à ARN ATP- dépendante appartenant à la famille des hélicases à boîte « DEAD », son interaction avec le domaine central HEAT d'eIF4G stimule son activité hélicase ce qui favorise la liaison et le balayage du ribosome en déstabilisant les structures secondaires présentes dans la région 5'UTR. Forme avec eIF4E et eIF4G le complexe eIF4F.
<b>eIF4B, eIF4H</b>	cofacteurs d'eIF4A stimulant son activité hélicase.
<b>eIF4E</b>	Reconnaît et se lie spécifiquement à la coiffe (5'-m <sup>7</sup> GpppN-5'), située à l'extrémité 5' de l'ARNm. Forme avec eIF4A et eIF4G le complexe eIF4F.



<b>eIF4G</b>	Protéine dite d'échafaudage. Se lie à l'ARNm et à différents facteurs impliqués dans la formation du complexe d'initiation (eIF4E, eIF4A dont elle stimule l'activité hélicase, eIF3 ce qui permet de recruter le complexe 43S au niveau de l'extrémité 5' de l'ARNm et la PABP ce qui permet de rapprocher les extrémités 3' et 5' de l'ARNm et de favoriser l'initiation). Forme avec eIF4A et eIF4E le complexe eIF4F.
<b>eIF5</b>	eIF5 est une GAP (GTPase Activating Protein) qui conduit à l'augmentation du taux de GTP lié à eIF2 hydrolysé au sein du complexe de préinitiation.
<b>PABP</b>	Se lie au niveau de la queue polyA et interagit ensuite avec le complexe multi-protéique eIF4F (eIF4A+eIF4E+eIF4G)
<b>Ribosome 80S</b>	Le ribosome eucaryote 80S est une particule ribonucléoprotéique composé de deux sous unités : la petite sous-unité 40S et la grande sous-unité 60S La petite sous-unité 40S se compose de 33 protéines et de l'ARNr 18S. Elle est impliquée dans la fidélité de la traduction en permettant via l'interaction du codon de l'ARNm avec l'anticodon de l'ARNt de décoder l'information génétique portée par les ARNm. La grande sous-unité 60S est, elle, constituée de 49 protéines et des ARNr 5S, 28S et 5,8S. Elle contient le centre de peptidyl-transférase qui catalyse la formation de liaisons peptidiques entre les acides aminés incorporés lors de l'étape de l'élongation

Tableau 1: Les facteurs d'initiations eucaryotes et leurs fonctions

## Bibliographie

- Abaeva, I. S., Pestova, T. V. & 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.
- Abrahamyan, L. G., Chatel-Chaix, L., Ajamian, L., Milev, M. P., Monette, A., Clément, J.-F., Song, R., Lehmann, M., DesGroseillers, L. & other authors. (2010).** Novel Staufen1 ribonucleoproteins prevent formation of stress granules but favour encapsidation of HIV-1 genomic RNA. *J Cell Sci***123**, 369–383.
- Aitken, C. E. & Lorsch, J. R.(2012).** A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol***19**, 568–576.
- Ali, I. K., McKendrick, L., Morley, S. J. & Jackson, R. J.(2001).** Activity of the Hepatitis A Virus IRES Requires Association between the Cap-Binding Translation Initiation Factor (eIF4E) and eIF4G. *J Virol***75**, 7854–7863.
- Alvarez, E., Castelló, A., Menéndez-Arias, L. & Carrasco, L.(2006).** HIV protease cleaves poly(A)-binding protein. *Biochem J***396**, 219–226.
- Anderson, E. C. & Lever, A. M. L.(2006).** Human immunodeficiency virus type 1 Gag polyprotein modulates its own translation. *J Virol***80**, 10478–10486.
- Angulo, J., Ulryck, N., Deforges, J., Chamond, N., Lopez-Lastra, M., Masquida, B. & Sargueil, B.(2016).** LOOP IIIId of the HCV IRES is essential for the structural rearrangement of the 40S-HCV IRES complex. *Nucleic Acids Res***44**, 1309–1325.
- Ariumi, Y.(2014).** Multiple functions of DDX3 RNA helicase in gene regulation, tumorigenesis, and viral infection. *Front Genet***5**.
- Ariumi, Y., Kuroki, M., Abe, K., Dansako, H., Ikeda, M., Wakita, T. & Kato, N.(2007).** DDX3 DEAD-Box RNA Helicase Is Required for Hepatitis C Virus RNA Replication. *J Virol***81**, 13922–13926.

- Arrigo, S. J. & Chen, I. S.(1991).** Rev is necessary for translation but not cytoplasmic accumulation of HIV-1 vif, vpr, and env/vpu 2 RNAs. *Genes Dev***5**, 808–819.
- Baer, B. W. & Kornberg, R. D.(1980).** Repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *Proc Natl Acad Sci U S A***77**, 1890–1892.
- Balvay, L., Soto Rifo, R., Ricci, E. P., Decimo, D. & Ohlmann, T.(2009).** Structural and functional diversity of viral IRESes. *Biochim Biophys Acta***1789**, 542–557.
- Banroques, J., Cordin, O., Doère, M., Linder, P. & Tanner, N. K.(2011).**Analyses of the Functional Regions of DEAD-Box RNA ‘Helicases’ with Deletion and Chimera Constructs Tested In Vivo and In vitro . *Journal of Molecular Biology***413**, 451–472.
- Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F. & other authors. (1983).** Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science***220**, 868–871.
- Belsham, G. J., McInerney, G. M. & Ross-Smith, N.(2000).** Foot-and-mouth disease virus 3C protease induces cleavage of translation initiation factors eIF4A and eIF4G within infected cells. *J Viro***74**, 272–280.
- Berkhout, B. & van Wamel, J. L.(2000).** The leader of the HIV-1 RNA genome forms a compactly folded tertiary structure. *RNA***6**, 282–295.
- Berkhout, B., Arts, K. & Abbink, T. E. M.(2011).** Ribosomal scanning on the 5'-untranslated region of the human immunodeficiency virus RNA genome. *Nucleic Acids Res***39**, 5232–5244.
- Berry, K. E., Waghray, S. & Doudna, J. A.(2010).** The HCV IRES pseudoknot positions the initiation codon on the 40S ribosomal subunit. *RNA***16**, 1559–1569.

**Boehringer, D., Thermann, R., Ostareck-Lederer, A., Lewis, J. D. & Stark, H.(2005).**

Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. *Structure***13**, 1695–1706.

**Boeras, I., Song, Z., Moran, A., Franklin, J., Brown, W. C., Johnson, M., Boris-**

**Lawrie, K. & Heng, X.(2016).** DHX9/RHA Binding to the PBS-Segment of the Genomic RNA during HIV-1 Assembly Bolsters Virion Infectivity. *J Mol Biol***428**, 2418–2429.

**Bolinger, C., Sharma, A., Singh, D., Yu, L. & Boris-Lawrie, K.(2010).**

RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. *Nucleic Acids Res***38**, 1686–1696.

**Bonnal, S., Pileur, F., Orsini, C., Parker, F., Pujol, F., Prats, A.-C. & Vagner,**

**S.(2005).**Heterogeneous nuclear ribonucleoprotein A1 is a novel internal ribosome entry site trans-acting factor that modulates alternative initiation of translation of the fibroblast growth factor 2 mRNA. *J Biol Chem***280**, 4144–4153.

**Borden, K. L. B., Culjkovic-Kraljacic, B. & Volpon, L.(2011).**

EIF4E (eukaryotic translation initiation factor 4E) Atlas of Genetics and Cytogenetics in Oncology and Haematology Volume 15, Number 2.

**Borman, A. M., Michel, Y. M. & Kean, K. M.(2000).**

Biochemical characterisation of cap-poly(A) synergy in rabbit reticulocyte lysates: the eIF4G-PABP interaction increases the functional affinity of eIF4E for the capped mRNA 5'-end. *Nucleic Acids Res***28**, 4068–4075.

**Borman, A. M., Michel, Y. M. & Kean, K. M.(2001).**

Detailed analysis of the requirements of hepatitis A virus internal ribosome entry segment for the eukaryotic initiation factor complex eIF4F. *J Viro***75**, 7864–7871.

- Bossi, P., Martinez, V., Strady, C. & Bricaire, F.(2001).** Actualités sur les traitements antirétroviraux. *La Revue de Médecine Interne***22**, 42–52.
- Both, G. W., Furuichi, Y., Muthukrishnan, S. & Shatkin, A. J.(1975).** Ribosome binding to reovirus mRNA in protein synthesis requires 5' terminal 7-methylguanosine. *Cell***6**, 185–195.
- Bouvet, M., Ferron, F., Imbert, I., Gluais, L., Selisko, B., Coutard, B., Canard, B. & Decroly, E.(2012).** [Capping strategies in RNA viruses]. *Med Sci (Paris)***28**, 423–429.
- Bowers, H. A., Maroney, P. A., Fairman, M. E., Kastner, B., Lührmann, R., Nilsen, T. W. & Jankowsky, E.(2006).** Discriminatory RNP remodeling by the DEAD-box protein DED1. *RNA***12**, 903–912.
- Braddock, M., Thorburn, A. M., Chambers, A., Elliott, G. D., Anderson, G. J., Kingsman, A. J. & Kingsman, S. M.(1990).** A nuclear translational block imposed by the HIV-1 U3 region is relieved by the Tat-TAR interaction. *Cell***62**, 1123–1133.
- Braddock, M., Powell, R., Blanchard, A. D., Kingsman, A. J. & Kingsman, S. M.(1993).** HIV-1 TAR RNA-binding proteins control TAT activation of translation in *Xenopus* oocytes. *FASEB J***7**, 214–222.
- Brai, A., Fazi, R., Tintori, C., Zamperini, C., Bugli, F., Sanguinetti, M., Stigliano, E., Esté, J., Badia, R. & other authors. (2016).** Human DDX3 protein is a valuable target to develop broad spectrum antiviral agents. *Proc Natl Acad Sci U S A***113**, 5388–5393.
- Brasey, A., Lopez-Lastra, M., Ohlmann, T., Beerens, N., Berkhout, B., Darlix, J.-L. & Sonenberg, N.(2003).** The leader of human immunodeficiency virus type 1

genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. *J Virol* **77**, 3939–3949.

**de Breyne, S., Chamond, N., Décimo, D., Traub, M.-A., André, P., Sargueil, B. & Ohlmann, T.(2012).** In vitro studies reveal that different modes of initiation on HIV-1 mRNA have different levels of requirement for eukaryotic initiation factor 4F. *FEBS J* **279**, 3098–3111.

**de Breyne, S., Soto-Rifo, R., López-Lastra, M. & Ohlmann, T.(2013).** Translation initiation is driven by different mechanisms on the HIV-1 and HIV-2 genomic RNAs. *Virus Res* **171**, 366–381.

**Briggs, J. A. G. & Kräusslich, H.-G.(2011).** The Molecular Architecture of HIV. *Journal of Molecular Biology, Structural and Molecular Biology of HIV* **410**, 491–500.

**Bruchfeld, J., Correia-Neves, M. & Källenius, G.(2015).** Tuberculosis and HIV Coinfection. *Cold Spring Harb Perspect Med* **5**, a017871.

**Buchkovich, N. J., Yu, Y., Zampieri, C. A. & Alwine, J. C.(2008).** The TORrid affairs of viruses: effects of mammalian DNA viruses on the PI3K–Akt–mTOR signalling pathway. *Nat Rev Micro* **6**, 266–275.

**Buck, C. B., Shen, X., Egan, M. A., Pierson, T. C., Walker, C. M. & Siliciano, R. F.(2001).** The human immunodeficiency virus type 1 gag gene encodes an internal ribosome entry site. *J Virol* **75**, 181–191.

**Bushell, M., Stoneley, M., Kong, Y. W., Hamilton, T. L., Spriggs, K. A., Dobbyn, H. C., Qin, X., Sarnow, P. & Willis, A. E.(2006).** Polypyrimidine tract binding protein regulates IRES-mediated gene expression during apoptosis. *Mol Cell* **23**, 401–412.

- Camerini, V., Decimo, D., Balvay, L., Pistello, M., Bendinelli, M., Darlix, J.-L. & Ohlmann, T.(2008).** A dormant internal ribosome entry site controls translation of feline immunodeficiency virus. *J Virol***82**, 3574–3583.
- Cammas, A., Pileur, F., Bonnal, S., Lewis, S. M., Lévêque, N., Holcik, M. & Vagner, S.(2007).** Cytoplasmic relocalization of heterogeneous nuclear ribonucleoprotein A1 controls translation initiation of specific mRNAs. *Mol Biol Cell***18**, 5048–5059.
- Campbell, E. M. & Hope, T. J.(2015).** HIV-1 capsid: the multifaceted key player in HIV-1 infection. *Nature Reviews Microbiology***13**, 471–483.
- Caruthers, J. M., Johnson, E. R. & McKay, D. B.(2000).** Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc Natl Acad Sci USA***97**, 13080–13085.
- Carvajal, F., Vallejos, M., Walters, B. A., Contreras, N., Hertz, M. I., Olivares, E., Cáceres, C. J., Pino, K., Letelier, A. & other authors. (2016).** Structural domains within the HIV-1 mRNA and the ribosomal protein S25 influence cap-independent translation initiation. *FEBS J*.
- Cavener, D. R. & Ray, S. C.(1991).** Eukaryotic start and stop translation sites. *Nucleic Acids Res***19**, 3185–3192.
- Chahar, H. S., Chen, S. & Manjunath, N.(2013).** P-body components LSM1, GW182, DDX3, DDX6 and XRN1 are recruited to WNV replication sites and positively regulate viral replication. *Virology***436**, 1–7.
- Chamond, N., Locker, N. & Sargueil, B.(2010).** The different pathways of HIV genomic RNA translation. *Biochem Soc Trans***38**, 1548–1552.



- Chamond, N., Deforges, J., Ulryck, N. & Sargueil, B.(2014).** 40S recruitment in the absence of eIF4G/4A by EMCV IRES refines the model for translation initiation on the archetype of Type II IRESs. *Nucleic Acids Res***42**, 10373–10384.
- Chappell, S. A., Edelman, G. M. & Mauro, V. P.(2000).** A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. *Proc Natl Acad Sci USA***97**, 1536–1541.
- Charbonneau, J., Gendron, K., Ferbeyre, G. & Brakier-Gingras, L.(2012).**The 5' UTR of HIV-1 full-length mRNA and the Tat viral protein modulate the programmed -1 ribosomal frameshift that generates HIV-1 enzymes. *RNA***18**, 519–529.
- Charnay, N., Ivanyi-Nagy, R., Soto-Rifo, R., Ohlmann, T., López-Lastra, M. & Darlix, J.-L.(2009).** Mechanism of HIV-1 Tat RNA translation and its activation by the Tat protein. *Retrovirology***6**, 74.
- Chatel-Chaix, L., Boulay, K., Mouland, A. J. & Desgroseillers, L.(2008).**The host protein Staufen1 interacts with the Pr55Gag zinc fingers and regulates HIV-1 assembly via its N-terminus. *Retrovirology***5**, 41.
- Chatterjee, S. & Pal, J. K.(2009).** Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biol Cell***101**, 251–262.
- Cho, S. & Hoffman, D. W.(2002).** Structure of the beta subunit of translation initiation factor 2 from the archaeon *Methanococcus jannaschii*: a representative of the eIF2beta/eIF5 family of proteins. *Biochemistry***41**, 5730–5742.
- Christensen, H. S., Daher, A., Soye, K. J., Frankel, L. B., Alexander, M. R., Lainé, S., Bannwarth, S., Ong, C. L., Chung, S. W. L. & other authors. (2007).** Small interfering RNAs against the TAR RNA binding protein, TRBP, a Dicer cofactor,

inhibit human immunodeficiency virus type 1 long terminal repeat expression and viral production. *J Virol* **81**, 5121–5131.

**Chuang, R. Y., Weaver, P. L., Liu, Z. & Chang, T. H.(1997).** Requirement of the DEAD-Box protein ded1p for messenger RNA translation. *Science* **275**, 1468–1471.

**Clavel, F., Guyader, M., Guétard, D., Sallé, M., Montagnier, L. & Alizon, M.(1986).** Molecular cloning and polymorphism of the human immune deficiency virus type 2. *Nature* **324**, 691–695.

**Clerzius, G., Gélinas, J.-F. & Gatignol, A.(2011).** Multiple levels of PKR inhibition during HIV-1 replication. *Rev Med Virol* **21**, 42–53.

**Cobbold, L. C., Wilson, L. A., Sawicka, K., King, H. A., Kondrashov, A. V., Spriggs, K. A., Bushell, M. & Willis, A. E.(2010).** Upregulated c-myc expression in multiple myeloma by internal ribosome entry results from increased interactions with and expression of PTB-1 and YB-1. *Oncogene* **29**, 2884–2891.

**Cochrane, A. W., Jones, K. S., Beidas, S., Dillon, P. J., Skalka, A. M. & Rosen, C. A.(1991).** Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression. *J Virol* **65**, 5305–5313.

**Cordes, S., Kusov, Y., Heise, T. & Gauss-Müller, V.(2008).** La autoantigen suppresses IRES-dependent translation of the hepatitis A virus. *Biochem Biophys Res Commun* **368**, 1014–1019.

**Cordin, O., Tanner, N. K., Doère, M., Linder, P. & Banroques, J.(2004).** The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J* **23**, 2478–2487.

- Cordin, O., Banroques, J., Tanner, N. K. & Linder, P.(2006).**The DEAD-box protein family of RNA helicases. *Gene***367**, 17–37.
- Costa-Mattioli, M., Svitkin, Y. & 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 . *Mol Cell Biol***24**, 6861–6870.
- Costantino, D. A., Pfingsten, J. S., Rambo, R. P. & Kieft, J. S.(2008).** tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat Struct Mol Bio***15**, 57–64.
- Craig, A. W., Svitkin, Y. V., Lee, H. S., Belsham, G. J. & Sonenberg, N.(1997).** The La autoantigen contains a dimerization domain that is essential for enhancing translation. *Mol Cell Bio***17**, 163–169.
- Craig, A. W., Haghghat, A., Yu, A. T. & Sonenberg, N.(1998).** Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature***392**, 520–523.
- Craigie, R. & Bushman, F. D.(2012).** HIV DNA integration. *Cold Spring Harb Perspect Med***2**, a006890.
- 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. *J Cell Bio***175**, 415–426.
- 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**, 65–69.

- D'Agostino, D. M., Felber, B. K., Harrison, J. E. & Pavlakis, G. N.(1992).** The Rev protein of human immunodeficiency virus type 1 promotes polysomal association and translation of gag/pol and vpu/env mRNAs. *Mol Cell Biol***12**, 1375–1386.
- Daher, A., Longuet, M., Dorin, D., Bois, F., Segeral, E., Bannwarth, S., Battisti, P. L., Purcell, D. F., Benarous, R. & other authors. (2001).** Two dimerization domains in the trans-activation response RNA-binding protein (TRBP) individually reverse the protein kinase R inhibition of HIV-1 long terminal repeat expression. *J Biol Chem***276**, 33899–33905.
- De Leys, R., Vanderborght, B., Vanden Haesevelde, M., Heyndrickx, L., van Geel, A., Wauters, C., Bernaerts, R., Saman, E., Nijs, P. & Willems, B.(1990).** Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin. *J Virol***64**, 1207–1216.
- Decroly, E., Ferron, F., Lescar, J. & Canard, B.(2012).** Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol***10**, 51–65.
- Deeks, S. G.(2006).** Antiretroviral treatment of HIV infected adults. *BMJ: British Medical Journal***332**, 1489.
- Deforges, J., Locker, N. & Sargueil, B.(2015).** mRNAs that specifically interact with eukaryotic ribosomal subunits. *Biochimie***114**, 48–57.
- Devaney, M. A., Vakharia, V. N., Lloyd, R. E., Ehrenfeld, E. & Grubman, M. J.(1988).** Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. *J Virol***62**, 4407–4409.

- Dias, A., Bouvier, D., Crépin, T., McCarthy, A. A., Hart, D. J., Baudin, F., Cusack, S. & Ruigrok, R. W. H.(2009).** The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature***458**, 914–918.
- Ditton, H. J., Zimmer, J., Kamp, C., Rajpert-De Meyts, E. & Vogt, P. H.(2004).**The AZFa gene DBY (DDX3Y) is widely transcribed but the protein is limited to the male germ cells by translation control. *Hum Mol Genet***13**, 2333–2341.
- 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. *J Biol Chem***285**, 26779–26787.
- Donahue, T. F., Cigan, A. M., Pabich, E. K. & Valavicius, B. C.(1988).** Mutations at a Zn(II) finger motif in the yeast eIF-2 beta gene alter ribosomal start-site selection during the scanning process. *Cell***54**, 621–632.
- Donnelly, N., Gorman, A. M., Gupta, S. & Samali, A.(2013).** The eIF2 $\alpha$  kinases: their structures and functions. *Cell Mol Life Sci***70**, 3493–3511.
- Dorin, D., Bonnet, M. C., Bannwarth, S., Gatignol, A., Meurs, E. F. & Vaquero, C.(2003).** The TAR RNA-binding protein, TRBP, stimulates the expression of TAR-containing RNAs in vitro and in vivo independently of its ability to inhibit the dsRNA-dependent kinase PKR. *J Biol Chem***278**, 4440–4448.
- Doudna, J. A., Sarnow, P. (2007).** Chapter 5 *Translation initiation by viral internal ribosome entry sites*, from the book *Translational Control in Biology and medicine* by CSHL Press, Cold Spring Harbor, NY. Mathews, N. Sonenberg, and J. W. B. Hershey.

- Dugré-Brisson, S., Elvira, G., Boulay, K., Chatel-Chaix, L., Mouland, A. J. & DesGroseillers, L. (2005).** Interaction of Staufen1 with the 5' end of mRNA facilitates translation of these RNAs. *Nucleic Acids Res***33**, 4797–4812.
- Dulude, D., Baril, M. & Brakier-Gingras, L.(2002).** Characterization of the frameshift stimulatory signal controlling a programmed -1 ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic Acids Res***30**, 5094–5102.
- Ederly, I., Petryshyn, R. & Sonenberg, N.(1989).** Activation of double-stranded RNA-dependent kinase (dsl) by the TAR region of HIV-1 mRNA: a novel translational control mechanism. *Cell***56**, 303–312.
- Engelman, A. & Cherepanov, P.(2012).** The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat Rev Micro***10**, 279–290.
- Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C. & Dumas, P.(1999).** The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure***7**, 1439–1449.
- Ennifar, E., Walter, P., Ehresmann, B., Ehresmann, C. & Dumas, P.(2001).** Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. *Nat Struct Biol***8**, 1064–1068.
- Fairman, M. E., Maroney, P. A., Wang, W., Bowers, H. A., Gollnick, P., Nilsen, T. W. & Jankowsky, E.(2004).** Protein displacement by DEXH/D 'RNA helicases' without duplex unwinding. *Science***304**, 730–734.
- Fairman-Williams, M. E., Guenther, U.-P. & Jankowsky, E.(2010).** SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol***20**, 313–324.

- Farruggio, D., Chaudhuri, J., Maitra, U. & RajBhandary, U. L.(1996).**The A1 x U72 base pair conserved in eukaryotic initiator tRNAs is important specifically for binding to the eukaryotic translation initiation factor eIF2. *Mol Cell Bio***16**, 4248–4256.
- Ferlin, A., Moro, E., Rossi, A., Dallapiccola, B. & Foresta, C.(2003).** The human Y chromosome's azoospermia factor b (AZFb) region: sequence, structure, and deletion analysis in infertile men. *J Med Genet***40**, 18–24.
- Fernández, I. S., Bai, X.-C., Murshudov, G., Scheres, S. H. W. & Ramakrishnan, V.(2014).** Initiation of Translation by Cricket Paralysis Virus IRES Requires Its Translocation in the Ribosome. *Cell***157**, 823–831.
- Filbin, M. E., Vollmar, B. S., Shi, D., Gonen, T. & Kieft, J. S.(2013).** HCV IRES manipulates the ribosome to promote the switch from translation initiation to elongation. *Nat Struct Mol Biol***20**, 150–158.
- Fischer, P. M.(2009).** Cap in hand: targeting eIF4E. *Cell Cycle***8**, 2535–2541.
- Floor, S. N., Condon, K. J., Sharma, D., Jankowsky, E. & Doudna, J. A.(2016).** Autoinhibitory Interdomain Interactions and Subfamily-specific Extensions Redefine the Catalytic Core of the Human DEAD-box Protein DDX3. *J Biol Chem***291**, 2412–2421.
- Foresta, C., Ferlin, A. & Moro, E.(2000).** Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility. *Hum Mol Gene***9**, 1161–1169.
- Franca, R., Belfiore, A., Spadari, S. & Maga, G.(2007).** Human DEAD-box ATPase DDX3 shows a relaxed nucleoside substrate specificity. *Proteins***67**, 1128–1137.

- Frankel, A. D. & Young, J. A.(1998).** HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem***67**, 1–25.
- Fraser, C. S. & Doudna, J. A.(2007).** Structural and mechanistic insights into hepatitis C viral translation initiation. *Nat Rev Micro***5**, 29–38.
- Freed, E. O.(2001).** HIV-1 replication. *Somat Cell Mol Genet***26**, 13–33.
- Frisby, D., Eaton, M. & Fellner, P.(1976).** Absence of 5' terminal capping in encephalomyocarditis virus RNA. *Nucleic Acids Res***3**, 2771–2787.
- Fröhlich, A., Rojas-Araya, B., Pereira-Montecinos, C., Dellarossa, A., Toro-Ascuy, D., Prades-Pérez, Y., García-de-Gracia, F., Garcés-Alday, A., Rubilar, P. S. & other authors. (2016).** DEAD-box RNA helicase DDX3 connects CRM1-dependent nuclear export and translation of the HIV-1 unspliced mRNA through its N-terminal domain. *Biochim Biophys Acta***1859**, 719–730.
- Gallo, R. C., Sarin, P. S., Gelmann, E. P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V. S., Mann, D., Sidhu, G. D., Stahl, R. E. & other authors. (1983).** Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science***220**, 865–867.
- Ganser-Pornillos, B. K., Yeager, M. & Pornillos, O.(2012).** Assembly and architecture of HIV. *Adv Exp Med Bio***726**, 441–465.
- Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., Cummins, L. B., Arthur, L. O., Peeters, M. & other authors. (1999).** Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature***397**, 436–441.
- Garbelli, A., Radi, M., Falchi, F., Beermann, S., Zanolli, S., Manetti, F., Dietrich, U., Botta, M. & Maga, G.(2011a).** Targeting the human DEAD-box polypeptide 3



(DDX3) RNA helicase as a novel strategy to inhibit viral replication. *Curr Med Chem***18**, 3015–3027.

**Garbelli, A., Beermann, S., Cicco, G. D., Dietrich, U. & Maga, G.(2011b).** A Motif Unique to the Human Dead-Box Protein DDX3 Is Important for Nucleic Acid Binding, ATP Hydrolysis, RNA/DNA Unwinding and HIV-1 Replication. *PLOS ONE***6**, e19810.

**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. *J Virol***84**, 1124–1138.

**Gatignol, A., Buckler-White, A., Berkhout, B. & Jeang, K. T.(1991).** Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science***251**, 1597–1600.

**Gatignol, A., Buckler, C. & Jeang, K. T.(1993).** Relatedness of an RNA-binding motif in human immunodeficiency virus type 1 TAR RNA-binding protein TRBP to human P1/dsl kinase and *Drosophila* staufen. *Mol Cell Bio***13**, 2193–2202.

**Geissler, R., Golbik, R. P. & Behrens, S.-E.(2012).** The DEAD-box helicase DDX3 supports the assembly of functional 80S ribosomes. *Nucleic Acids Res***40**, 4998–5011.

**Gelderblom, H. R., Hausmann, E. H., Ozel, M., Pauli, G. & Koch, M. A.(1987).** Fine structure of human immunodeficiency virus (HIV) and immunolocalization of structural proteins. *Virology***156**, 171–176.

**Gendron, K., Ferbeyre, G., Heveker, N. & Brakier-Gingras, L.(2011).** The activity of the HIV-1 IRES is stimulated by oxidative stress and controlled by a negative regulatory element. *Nucleic Acids Res***39**, 902–912.

- des 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**, 491–495.
- Giraud, S., Greco, A., Brink, M., Diaz, J. J. & Delafontaine, P.(2001).** Translation initiation of the insulin-like growth factor I receptor mRNA is mediated by an internal ribosome entry site. *J Biol Chem***276**, 5668–5675.
- Goldman, E.(2008).** Transfer RNA. In *Encyclopedia of Life Sciences*. Edited by John Wiley & Sons, Ltd. Chichester, UK: John Wiley & Sons, Ltd.
- Gorbalenya, A. E. & Koonin, E. V.(1993).** Helicases: amino acid sequence comparisons and structure-function relationships. *Current Opinion in Structural Biology***3**, 419–429.
- Gray, N. K. & Hentze, M. W.(1994).** Regulation of protein synthesis by mRNA structure. *Mol Biol Rep***19**, 195–200.
- Grohman, J. K., Del Campo, M., Bhaskaran, H., Tijerina, P., Lambowitz, A. M. & Russell, R.(2007).** Probing the mechanisms of DEAD-box proteins as general RNA chaperones: the C-terminal domain of CYT-19 mediates general recognition of RNA. *Biochemistry***46**, 3013–3022.
- Groom, H. C. T., Anderson, E. C., Dangerfield, J. A. & Lever, A. M. L.(2009).** Rev regulates translation of human immunodeficiency virus type 1 RNAs. *J Gen Viro***90**, 1141–1147.
- Groppo, R. & Richter, J. D.(2009).** Translational control from head to tail. *Curr Opin Cell Biol***21**, 444–451.
- Gross, J. D., Moerke, N. J., von der Haar, T., Lugovskoy, A. A., Sachs, A. B., McCarthy, J. E. G. & Wagner, G.(2003).** Ribosome loading onto the mRNA cap

is driven by conformational coupling between eIF4G and eIF4E. *Cell***115**, 739–750.

**Grossman, Z., Meier-Schellersheim, M., Paul, W. E. & Picker, L. J.(2006).**

Pathogenesis of HIV infection: what the virus spares is as important as what it destroys. *Nat Med***12**, 289–295.

**Grundhoff, A. & Ganem, D.(2001).** Mechanisms Governing Expression of the v-FLIP

Gene of Kaposi's Sarcoma-Associated Herpesvirus. *J Viro***75**, 1857–1863.

**von der Haar, T., Gross, J. D., Wagner, G. & McCarthy, J. E. G.(2004).** The mRNA

cap-binding protein eIF4E in post-transcriptional gene expression. *Nat Struct Mol Bio***11**, 503–511.

**Hartman, T. R., Qian, S., Bolinger, C., Fernandez, S., Schoenberg, D. R. & Boris-**

**Lawrie, K.(2006).** RNA helicase A is necessary for translation of selected messenger RNAs. *Nat Struct Mol Bio***13**, 509–516.

**Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H. Y., Grassucci, R. A.,**

**Pestova, T. V., Hellen, C. U. T. & Frank, J.(2013).** HCV-like IRESs displace eIF3 to gain access to the 40S subunit. *Nature***503**, 539–543.

**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**, 558–570.

**Herbreteau, C. H., Weill, L., Décimo, D., Prévôt, D., Darlix, J.-L., Sargueil, B. &**

**Ohlmann, T.(2005).** HIV-2 genomic RNA contains a novel type of IRES located downstream of its initiation codon. *Nat Struct Mol Bio***12**, 1001–1007.

**Hernández, G.(2008).** Was the initiation of translation in early eukaryotes IRES-driven?

*Trends Biochem Sci***33**, 58–64.

- Hilliker, A., Gao, Z., Jankowsky, E. & Parker, R.(2011).** The DEAD-box protein Ded1 modulates translation by the formation and resolution of an eIF4F-mRNA complex. *Mol Cell***43**, 962–972.
- Hinnebusch, A. G.(2006).** eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem Sci* **31**, 553–562.
- Hinnebusch, A. G., Dever,T.E., Asano,K. (2007).** Chapter 9 *Mechanism of Translation Initiation in the Yeast Saccharomyces cerevisiae*, from the book *Translational Control in Biology and medicine* by CSHL Press, Cold Spring Harbor, NY. Mathews, N. Sonenberg, and J. W. B. Hershey.
- Hinnebusch, A. G. (2011).** Molecular Mechanism of Scanning and Start Codon Selection in Eukaryotes. *Microbiol Mol Biol Rev***75**, 434–467.
- Hinnebusch, A. G.(2014).** The scanning mechanism of eukaryotic translation initiation. *Annu Rev Biochem***83**, 779–812.
- Hinnebusch, A. G., Ivanov, I. P. & Sonenberg, N.(2016).** Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science***352**, 1413–1416.
- Hirsch, V. M., Olmsted, R. A., Murphey-Corb, M., Purcell, R. H. & Johnson, P. R.(1989).** An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature***339**, 389–392.
- Ho, D. D. & Bieniasz, P. D.(2008).**HIV-1 at 25. *Cell***133**, 561–565.
- Högbom, M., Collins, R., van den Berg, S., Jenvert, R.-M., Karlberg, T., Kotenyova, T., Flores, A., Karlsson Hedestam, G. B. & Schiavone, L. H.(2007).** Crystal structure of conserved domains 1 and 2 of the human DEAD-box helicase DDX3X in complex with the mononucleotide AMP. *J Mol Bio***372**, 150–159.

- Hoshino, S., Imai, M., Kobayashi, T., Uchida, N. & Katada, T.(1999).** The eukaryotic polypeptide chain releasing factor (eRF3/GSPT) carrying the translation termination signal to the 3'-Poly(A) tail of mRNA. Direct association of erf3/GSPT with polyadenylate-binding protein. *J Biol Chem***274**, 16677–16680.
- Hu, M. C., Tranque, P., Edelman, G. M. & Mauro, V. P.(1999).** rRNA-complementarity in the 5' untranslated region of mRNA specifying the Gtx homeodomain protein: evidence that base- pairing to 18S rRNA affects translational efficiency. *Proc Natl Acad Sci USA***96**, 1339–1344.
- Hua, S., Lécuroux, C., Sáez-Ciri3n, A., Pancino, G., Girault, I., Versmisse, P., Boufassa, F., Taulera, O., Sinet, M. & other authors. (2014).** Potential Role for HIV-Specific CD38 - /HLA-DR + CD8 + T Cells in Viral Suppression and Cytotoxicity in HIV Controllers. *PLOS ONE***9**, e101920.
- Imataka, H. & Sonenberg, N.(1997).** Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol Cell Bio***17**, 6940–6947.
- Iost, I., Dreyfus, M. & Linder, P.(1999).** Ded1p, a DEAD-box protein required for translation initiation in *Saccharomyces cerevisiae*, is an RNA helicase. *J Biol Chem***274**, 17677–17683.
- Jackson, R. J., Hellen, C. U. T. & Pestova, T. V.(2010).** The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Bio***11**, 113–127.
- Jan, E., Thompson, S. R., Wilson, J. E., Pestova, T. V., Hellen, C. U. & Sarnow, P.(2001).** Initiator Met-tRNA-independent translation mediated by an internal

- ribosome entry site element in cricket paralysis virus-like insect viruses. *Cold Spring Harb Symp Quant Biol* **66**, 285–292.
- Jan, E. & Sarnow, P.(2002).** Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J Mol Biol* **324**, 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. *J Viro* **62**, 2636–2643.
- Jang, S. K.(2006).** Internal initiation: IRES elements of picornaviruses and hepatitis c virus. *Virus Res* **119**, 2–15.
- Jarmoskaite, I. & Russell, R.(2014).** RNA helicase proteins as chaperones and remodelers. *Annu Rev Biochem* **83**, 697–725.
- Jennings, M. D. & Pavitt, G. D.(2010).** eIF5 has GDI activity necessary for translational control by eIF2 phosphorylation. *Nature* **465**, 378–381.
- Joshi, C. P., Zhou, H., Huang, X. & Chiang, V. L.(1997).** Context sequences of translation initiation codon in plants. *Plant Mol Biol* **35**, 993–1001.
- Kaminski, A., Howell, M. T. & Jackson, R. J.(1990).** Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning mechanism. *EMBO J* **9**, 3753–3759.
- Kapp, L. D. & Lorsch, J. R.(2004a).** The molecular mechanics of eukaryotic translation. *Annu Rev Biochem* **73**, 657–704.
- Kapp, L. D. & Lorsch, J. R.(2004b).** GTP-dependent Recognition of the Methionine Moiety on Initiator tRNA by Translation Factor eIF2. *Journal of Molecular Biology* **335**, 923–936.

- Kapp, L. D., Kolitz, S. E. & Lorsch, J. R.(2006).** Yeast initiator tRNA identity elements cooperate to influence multiple steps of translation initiation. *RNA***12**, 751–764.
- Karim, M. M., Svitkin, Y. V., Kahvejian, A., De Crescenzo, G., Costa-Mattioli, M. & Sonenberg, N.(2006).** A mechanism of translational repression by competition of Paip2 with eIF4G for poly(A) binding protein (PABP) binding. *Proc Natl Acad Sci USA***103**, 9494–9499.
- Kenyon, J. C., Prestwood, L. J., Le Grice, S. F. J. & Lever, A. M. L.(2013).** In-gel probing of individual RNA conformers within a mixed population reveals a dimerization structural switch in the HIV-1 leader. *Nucleic Acids Res***41**, e174.
- Khatter, H., Myasnikov, A. G., Natchiar, S. K. & Klaholz, B. P.(2015).** Structure of the human 80S ribosome. *Nature***520**, 640–645.
- Kieft, J. S., Zhou, K., Jubin, R. & Doudna, J. A.(2001).** Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA***7**, 194–206.
- Kieft, J. S.(2008).** Viral IRES RNA structures and ribosome interactions. *Trends Biochem Sci***33**, 274–283.
- Kim, J. H., Park, S. M., Park, J. H., Keum, S. J. & Jang, S. K.(2011).** eIF2A mediates translation of hepatitis C viral mRNA under stress conditions. *EMBO J***30**, 2454–2464.
- Kim, Y. S., Lee, S. G., Park, S. H. & Song, K.(2001).** Gene structure of the human DDX3 and chromosome mapping of its related sequences. *Mol Cells***12**, 209–214.
- Kimball, S. R.(1999).** Eukaryotic initiation factor eIF2. *The International Journal of Biochemistry & Cell Biology***31**, 25–29.
- Klann, E. & Dever, T. E.(2004).** Biochemical mechanisms for translational regulation in synaptic plasticity. *Nat Rev Neurosci***5**, 931–942.

- Koh, D. C.-Y., Liu, D. X. & Wong, S.-M.(2002).** A six-nucleotide segment within the 3' untranslated region of hibiscus chlorotic ringspot virus plays an essential role in translational enhancement. *J Viro***76**, 1144–1153.
- Kolupaeva, V. G., Unbehaun, A., Lomakin, I. B., Hellen, C. U. T. & Pestova, T. V.(2005).** Binding of eukaryotic initiation factor 3 to ribosomal 40S subunits and its role in ribosomal dissociation and anti-association. *RNA***11**, 470–486.
- Kozak, M.(1980).** Influence of mRNA secondary structure on binding and migration of 40S ribosomal subunits. *Cell***19**, 79–90.
- Kozak, M.(1986).** Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell***44**, 283–292.
- Kozak, M.(1987a).** An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res***15**, 8125–8148.
- Kozak, M.(1987b).** At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol***196**, 947–950.
- Kozak, M.(1991).** A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expr***1**, 111–115.
- Kozak, M.(1997).** Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. *EMBO J***16**, 2482–2492.
- Krol, J., Loedige, I. & Filipowicz, W.(2010).** The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet***11**, 597–610.
- Kuyumcu-Martinez, N. M., Joachims, M. & Lloyd, R. E.(2002).** Efficient Cleavage of Ribosome-Associated Poly(A)-Binding Protein by Enterovirus 3C Protease. *J Viro***76**, 2062–2074.



- Kuyumcu-Martinez, N. M., Van Eden, M. E., Younan, P. & Lloyd, R. E.(2004).** Cleavage of poly(A)-binding protein by poliovirus 3C protease inhibits host cell translation: a novel mechanism for host translation shutoff. *Mol Cell Biol***24**, 1779–1790.
- Lai, M.-C., Chang, W.-C., Shieh, S.-Y. & Tarn, W.-Y.(2010).** DDX3 regulates cell growth through translational control of cyclin E1. *Mol Cell Biol***30**, 5444–5453.
- Lai, M.-C., Wang, S.-W., Cheng, L., Tarn, W.-Y., Tsai, S.-J. & Sun, H. S.(2013).** Human DDX3 interacts with the HIV-1 Tat protein to facilitate viral mRNA translation. *PLoS ONE***8**, e68665.
- Laletina, E., Graifer, D., Malygin, A., Ivanov, A., Shatsky, I. & Karpova, G.(2006).** Proteins surrounding hairpin IIIe of the hepatitis C virus internal ribosome entry site on the human 40S ribosomal subunit. *Nucleic Acids Res***34**, 2027–2036.
- Lambotte, O., Boufassa, F., Madec, Y., Nguyen, A., Goujard, C., Meyer, L., Rouzioux, C., Venet, A., Delfraissy, J.-F. & SEROCO-HEMOCO Study Group. (2005).** HIV controllers: a homogeneous group of HIV-1-infected patients with spontaneous control of viral replication. *Clin Infect Dis***41**, 1053–1056.
- Larder, B. A. & Kemp, S. D.(1989).** Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science***246**, 1155–1158.
- Lee, C.-S., Dias, A. P., Jedrychowski, M., Patel, A. H., Hsu, J. L. & Reed, R.(2008a).** Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. *Nucleic Acids Res***36**, 4708–4718.
- Lee, C.-S., Dias, A. P., Jedrychowski, M., Patel, A. H., Hsu, J. L. & Reed, R.(2008b).** Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. *Nucleic Acids Research***36**, 4708–4718.

- LeFebvre, A. K., Korneeva, N. L., Trutschl, M., Cvek, U., Duzan, R. D., Bradley, C. A., Hershey, J. W. B. & Rhoads, R. E.(2006).** Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J Biol Chem***281**, 22917–22932.
- Li, C., Ge, L., Li, P., Wang, Y., Dai, J., Sun, M., Huang, L., Shen, Z., Hu, X. & other authors. (2014).** Cellular DDX3 regulates Japanese encephalitis virus replication by interacting with viral un-translated regions. *Virology***449**, 70–81.
- Li, Q., Pène, V., Krishnamurthy, S., Cha, H. & Liang, T. J.(2013).** Hepatitis C virus infection activates an innate pathway involving IKK- $\alpha$  in lipogenesis and viral assembly. *Nat Med***19**, 722–729.
- Linder, P., Lasko, P. F., Ashburner, M., Leroy, P., Nielsen, P. J., Nishi, K., Schnier, J. & Slonimski, P. P.(1989).** Birth of the D-E-A-D box. *Nature***337**, 121–122.
- Linder, P.(2006).** Dead-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic Acids Res***34**, 4168–4180.
- Linder, P. & Jankowsky, E.(2011).** From unwinding to clamping — the DEAD box RNA helicase family. *Nat Rev Mol Cell Bio***12**, 505–516.
- Liu, F., Putnam, A. & Jankowsky, E.(2008).** ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *Proc Natl Acad Sci USA***105**, 20209–20214.
- Liu, F., Putnam, A. A. & Jankowsky, E.(2014).** DEAD-box helicases form nucleotide-dependent, long-lived complexes with RNA. *Biochemistry***53**, 423–433.
- Llácer, J. L., Hussain, T., Marler, L., Aitken, C. E., Thakur, A., Lorsch, J. R., Hinnebusch, A. G. & Ramakrishnan, V.(2015).** Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. *Mol Cell***59**, 399–412.

- Locker, N., Easton, L. E. & Lukavsky, P. J.(2007).** HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO J***26**, 795–805.
- Locker, N., Chamond, N. & Sargueil, B.(2011).** A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3. *Nucleic Acids Res***39**, 2367–2377.
- Lorsch, J. R. & Dever, T. E.(2010).** Molecular view of 43 S complex formation and start site selection in eukaryotic translation initiation. *J Biol Chem***285**, 21203–21207.
- Lozano, G., Fernandez, N. & Martinez-Salas, E.(2016).** Modeling Three-Dimensional Structural Motifs of Viral IRES. *Journal of Molecular Biology, Challenges in RNA Structural Modeling and Design* **428**, 767–776.
- Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E. L., Kharytonchyk, S., Dorjsuren, B., Kulandaivel, G., Jones, S. & other authors. (2011).** NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging. *Science***334**, 242–245.
- Lucas, S. & Nelson, A. M.(2015).** HIV and the spectrum of human disease. *J Pathol***235**, 229–241.
- Lyons, A. J. & Robertson, H. D.(2003).** Detection of tRNA-like structure through RNase P cleavage of viral internal ribosome entry site RNAs near the AUG start triplet. *J Biol Chem***278**, 26844–26850.
- Maag, D., Algire, M. A. & Lorsch, J. R.(2006).** Communication between eukaryotic translation initiation factors 5 and 1A within the ribosomal pre-initiation complex plays a role in start site selection. *J Mol Biol***356**, 724–737.
- Macejak, D. G. & Sarnow, P.(1991).** Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature***353**, 90–94.

- Majumdar, R., Bandyopadhyay, A. & Maitra, U.(2003).** Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40S preinitiation complex. *J Biol Chem***278**, 6580–6587.
- Maldarelli, F., Martin, M. A. & Strebel, K.(1991).** Identification of posttranscriptionally active inhibitory sequences in human immunodeficiency virus type 1 RNA: novel level of gene regulation. *J Viro***65**, 5732–5743.
- 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 Res***41**, 8706–8714.
- Mangus, D. A., Evans, M. C. & Jacobson, A.(2003).** Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol***4**, 223.
- Maquat, L. E., Tarn, W.-Y. & Isken, O.(2010).** The pioneer round of translation: features and functions. *Cell***142**, 368–374.
- Marcotrigiano, J., Gingras, A. C., Sonenberg, N. & Burley, S. K.(1999).** Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell***3**, 707–716.
- Marissen, W. E. & Lloyd, R. E.(1998).** Eukaryotic translation initiation factor 4G is targeted for proteolytic cleavage by caspase 3 during inhibition of translation in apoptotic cells. *Mol Cell Biol***18**, 7565–7574.
- Marquet, R., Paillart, J. C., Skripkin, E., Ehresmann, C. & Ehresmann, B.(1994).** Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. *Nucleic Acids Res***22**, 145–151.

**Martin Stoltzfus, C.(2009).** Chapter 1 Regulation of HIV-1 Alternative RNA Splicing and Its Role in Virus Replication, pp. 1–40. Edited by B.-A. in V. Research. Academic Press.

**Martineau, Y., Derry, M. C., Wang, X., Yanagiya, A., Berlanga, J. J., Shyu, A.-B., Imataka, H., Gehring, K. & Sonenberg, N.(2008).** Poly(A)-binding protein-interacting protein 1 binds to eukaryotic translation initiation factor 3 to stimulate translation. *Mol Cell Biol***28**, 6658–6667.

**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. *Proc Natl Acad Sci USA***111**, 15385–15389.

**McMillan, N. A., Chun, R. F., Siderovski, D. P., Galabru, J., Toone, W. M., Samuel, C. E., Mak, T. W., Hovanessian, A. G., Jeang, K. T. & Williams, B. R.(1995).** HIV-1 Tat directly interacts with the interferon-induced, double-stranded RNA-dependent kinase, PKR. *Virology***213**, 413–424.

**Medina, M., Domingo, E., Brangwyn, J. K. & Belsham, G. J.(1993).** The two species of the foot-and-mouth disease virus leader protein, expressed individually, exhibit the same activities. *Virology***194**, 355–359.

**Meng, Z., Jackson, N. L., Shcherbakov, O. D., Choi, H. & Blume, S. W.(2010).** The human IGF1R IRES likely operates through a Shine-Dalgarno-like interaction with the G961 loop (E-site) of the 18S rRNA and is kinetically modulated by a naturally polymorphic polyU loop. *J Cell Biochem***110**, 531–544.

**Michel, Y. M., Poncet, D., Piron, M., Kean, K. M. & Borman, A. M.(2000).** Cap-Poly(A) synergy in mammalian cell-free extracts. Investigation of the requirements for

poly(A)-mediated stimulation of translation initiation. *J Biol Chem***275**, 32268–32276.

**Millevoi, S. & Vagner, S.(2010).** Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res***38**, 2757–2774.

**Monette, A., Ajamian, L., López-Lastra, M. & Mouland, A. J.(2009).** Human immunodeficiency virus type 1 (HIV-1) induces the cytoplasmic retention of heterogeneous nuclear ribonucleoprotein A1 by disrupting nuclear import: implications for HIV-1 gene expression. *J Biol Chem***284**, 31350–31362.

**Monette, A., Valiente-Echeverría, F., Rivero, M., Cohen, É. A., Lopez-Lastra, M. & Mouland, A. J.(2013).** Dual mechanisms of translation initiation of the full-length HIV-1 mRNA contribute to gag synthesis. *PLoS ONE***8**, e68108.

**Morlat, Philippe.** prise en charge médicale des personnes vivants avec le vih\_actualisation2014 - rapport Morlat.

**Morris, D. R. & Geballe, A. P.(2000).** Upstream open reading frames as regulators of mRNA translation. *Mol Cell Biol***20**, 8635–8642.

**Mouland, A. J., Mercier, J., Luo, M., Bernier, L., DesGroseillers, L. & Cohen, E. A.(2000).** The double-stranded RNA-binding protein Staufen is incorporated in human immunodeficiency virus type 1: evidence for a role in genomic RNA encapsidation. *J Viro***74**, 5441–5451.

**Nisole, S. & Saïb, A.(2004).** Early steps of retrovirus replicative cycle. *Retrovirology***1**, 9.

**Nomoto, A., Lee, Y. F. & Wimmer, E.(1976).** The 5' end of poliovirus mRNA is not capped with m7G(5')ppp(5')Np. *Proc Natl Acad Sci USA***73**, 375–380.

- Novoa, I. & Carrasco, L.(1999).** Cleavage of eukaryotic translation initiation factor 4G by exogenously added hybrid proteins containing poliovirus 2Apro in HeLa cells: effects on gene expression. *Mol Cell Biol***19**, 2445–2454.
- Oda, S.-I., Schröder, M. & Khan, A. R.(2009).** Structural basis for targeting of human RNA helicase DDX3 by poxvirus protein K7. *Structure***17**, 1528–1537.
- Ohlmann, T., Lopez-Lastra, M. & Darlix, J. L.(2000).** An internal ribosome entry segment promotes translation of the simian immunodeficiency virus genomic RNA. *J Biol Chem***275**, 11899–11906.
- Ohlmann, T., Prévôt, D., Décimo, D., Roux, F., Garin, J., Morley, S. J. & Darlix, J.-L.(2002).** In vitro cleavage of eIF4GI but not eIF4GII by HIV-1 protease and its effects on translation in the rabbit reticulocyte lysate system. *J Mol Biol***318**, 9–20.
- Ohlmann, T., Mengardi, C. & López-Lastra, M.(2014).** Translation initiation of the HIV-1 mRNA. *Translation (Austin)***2**, e960242.
- Ong, C. L., Thorpe, J. C., Gorry, P. R., Bannwarth, S., Jaworowski, A., Howard, J. L., Chung, S., Campbell, S., Christensen, H. S. & other authors. (2005).** Low TRBP levels support an innate human immunodeficiency virus type 1 resistance in astrocytes by enhancing the PKR antiviral response. *J Virol***79**, 12763–12772.
- Orna Elroy-Stein ; William C. Merrick. (2007).***Translation Initiation Via Cellular Internal Ribosome Entry Sites*, CSHL Press, Cold Spring Harbor, NY. Mathews, N. Sonenberg, and J. W. B. Hershey,.
- Othman, Z., Sulaiman, M. K., Willcocks, M. M., Ulryck, N., Blackbourn, D. J., Sargueil, B., Roberts, L. O. & Locker, N.(2014).** Functional analysis of Kaposi's sarcoma-associated herpesvirus vFLIP expression reveals a new mode of IRES-mediated translation. *RNA***20**, 1803–1814.

- Otto, G. A., Lukavsky, P. J., Lancaster, A. M., Sarnow, P. & Puglisi, J. D.(2002).** Ribosomal proteins mediate the hepatitis C virus IRES-HeLa 40S interaction. *RNA***8**, 913–923.
- Özeş, A. R., Feoktistova, K., Avanzino, B. C. & Fraser, C. S.(2011).**Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B. *J Mol Biol***412**, 674–687.
- Özeş, A. R., Feoktistova, K., Avanzino, B. C., Baldwin, E. P. & Fraser, C. S.(2014).** Real-time fluorescence assays to monitor duplex unwinding and ATPase activities of helicases. *Nat Protocols***9**, 1645–1661.
- Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C. & Marquet, R.(1996).** A loop-loop ‘kissing’ complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc Natl Acad Sci USA***93**, 5572–5577.
- Paillart, J.-C., Shehu-Xhilaga, M., Marquet, R. & Mak, J.(2004).** Dimerization of retroviral RNA genomes: an inseparable pair. *Nat Rev Microbiol***2**, 461–472.
- Park, H., Davies, M. V., Langland, J. O., Chang, H. W., Nam, Y. S., Tartaglia, J., Paoletti, E., Jacobs, B. L., Kaufman, R. J. & Venkatesan, S.(1994).** TAR RNA-binding protein is an inhibitor of the interferon-induced protein kinase PKR. *Proc Natl Acad Sci USA***91**, 4713–4717.
- Park, S. H., Lee, S. G., Kim, Y. & Song, K.(1998).** Assignment of a human putative RNA helicase gene, DDX3, to human X chromosome bands p11.3-->p11.23. *Cytogenet Cell Genet***81**, 178–179.
- Patel, S. S. & Picha, K. M.(2000).** Structure and function of hexameric helicases. *Annu Rev Biochem***69**, 651–697.



- Pelletier, J. & Sonenberg, N.(1985).** Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell***40**, 515–526.
- Pelletier, J. & Sonenberg, N.(1988).** Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature***334**, 320–325.
- Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J. & Hellen, C. U.(1998).** A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev***12**, 67–83.
- Pestova, T. V., Lorsch, J. R., and Hellen, C. U.(2007).** *The mechanism of translation initiation in eukaryotes.*, CSHL Press, Cold Spring Harbor, NY. Mathews, N. Sonenberg, and J. W. B. Hershey,.
- Pestova, T. V. & Kolupaeva, V. G.(2002).** The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev***16**, 2906–2922.
- Pestova, T. V., de Breyne, S., Pisarev, A. V., Abaeva, I. S. & Hellen, C. U. T.(2008).** eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: a common role of domain II. *EMBO J***27**, 1060–1072.
- Pfingsten, J. S., Costantino, D. A. & Kieft, J. S.(2006).** Structural basis for ribosome recruitment and manipulation by a viral IRES RNA. *Science***314**, 1450–1454.
- Pickering, B. M., Mitchell, S. A., Spriggs, K. A., Stoneley, M. & Willis, A. E.(2004).** Bag-1 internal ribosome entry segment activity is promoted by structural changes

mediated by poly(rC) binding protein 1 and recruitment of polypyrimidine tract binding protein 1. *Mol Cell Biol***24**, 5595–5605.

**Pisarev, A. V., Kolupaeva, V. G., Pisareva, V. P., Merrick, W. C., Hellen, C. U. T. & Pestova, T. V.(2006).** Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev***20**, 624–636.

**Pisarev, A. V., Hellen, C. U. T. & Pestova, T. V.(2007).** Recycling of eukaryotic posttermination ribosomal complexes. *Cell***131**, 286–299.

**Plank, T.-D. M., Whitehurst, J. T. & Kieft, J. S.(2013).** Cell type specificity and structural determinants of IRES activity from the 5' leaders of different HIV-1 transcripts. *Nucleic Acids Res***41**, 6698–6714.

**Plantier, J.-C., Leoz, M., Dickerson, J. E., De Oliveira, F., Cordonnier, F., Lemée, V., Damond, F., Robertson, D. L. & Simon, F.(2009).**A new human immunodeficiency virus derived from gorillas. *Nat Med***15**, 871–872.

**Proud, C. G.(2005).** eIF2 and the control of cell physiology. *Semin Cell Dev Biol***16**, 3–12.

**Putnam, A. A. & Jankowsky, E.(2013).** AMP sensing by DEAD-box RNA helicases. *J Mol Biol***425**, 3839–3845.

**Putnam, A. A., Gao, Z., Liu, F., Jia, H., Yang, Q. & Jankowsky, E.(2015).** Division of Labor in an Oligomer of the DEAD-Box RNA Helicase Ded1p. *Molecular Cell***59**, 541–552.

**Quade, N., Boehringer, D., Leibundgut, M., van den Heuvel, J. & Ban, N.(2015).** Cryo-EM structure of Hepatitis C virus IRES bound to the human ribosome at 3.9-Å resolution. *Nat Commun***6**.

- Radi, M., Falchi, F., Garbelli, A., Samuele, A., Bernardo, V., Paolucci, S., Baldanti, F., Schenone, S., Manetti, F. & other authors. (2012).** Discovery of the first small molecule inhibitor of human DDX3 specifically designed to target the RNA binding site: towards the next generation HIV-1 inhibitors. *Bioorg Med Chem Lett***22**, 2094–2098.
- Rao, P., Yuan, W. & Krug, R. M.(2003).** Crucial role of CA cleavage sites in the capsid-snatching mechanism for initiating viral mRNA synthesis. *EMBO J***22**, 1188–1198.
- Raught, B. & Gingras, A. C.(1999).** eIF4E activity is regulated at multiple levels. *Int J Biochem Cell Biol***31**, 43–57.
- Rauschendorf, M.-A., Zimmer, J., Hanstein, R., Dickemann, C. & Vogt, P. H.(2011).** Complex transcriptional control of the AZFa gene DDX3Y in human testis. *Int J Androl***34**, 84–96.
- Ricci, E. P., Soto Rifo, R., Herbreteau, C. H., Decimo, D. & Ohlmann, T.(2008).** Lentiviral RNAs can use different mechanisms for translation initiation. *Biochem Soc Trans***36**, 690–693.
- Richman, D. D.(1993).** Resistance of clinical isolates of human immunodeficiency virus to antiretroviral agents. *Antimicrob Agents Chemother***37**, 1207–1213.
- Rivas-Aravena, A., Ramdohr, P., Vallejos, M., Valiente-Echeverría, F., Dormoy-Raclet, V., Rodríguez, F., Pino, K., Holzmann, C., Huidobro-Toro, J. P. & other authors. (2009).** The Elav-like protein HuR exerts translational control of viral internal ribosome entry sites. *Virology***392**, 178–185.
- Roberts, L. O., Seamons, R. A. & Belsham, G. J.(1998).** Recognition of picornavirus internal ribosome entry sites within cells; influence of cellular and viral proteins. *RNA***4**, 520–529.

- Rocak, S. & Linder, P.(2004).** DEAD-box proteins: the driving forces behind RNA metabolism. *Nat Rev Mol Cell Bio*5, 232–241.
- Rojas-Araya, B., Ohlmann, T. & Soto-Rifo, R.(2015).** Translational Control of the HIV Unspliced Genomic RNA. *Viruses*7, 4326–4351.
- 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. *Mol Cell Bio*10, 1134–1144.
- Sáez-Cirión, A., Pancino, G., Sinet, M., Venet, A., Lambotte, O. & ANRS EP36 HIV CONTROLLERS study group. (2007).** HIV controllers: how do they tame the virus? *Trends Immuno*28, 532–540.
- Sasaki, J. & Nakashima, N.(1999).** Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus in vitro . *J Viro*73, 1219–1226.
- Schmitt, E., Naveau, M. & Mechulam, Y.(2010).** Eukaryotic and archaeal translation initiation factor 2: A heterotrimeric tRNA carrier. *FEBS Letters*, Transfer RNA **584**, 405–412.
- Schneider, R., Campbell, M., Nasioulas, G., Felber, B. K. & Pavlakis, G. N.(1997).** Inactivation of the human immunodeficiency virus type 1 inhibitory elements allows Rev-independent expression of Gag and Gag/protease and particle formation. *J Viro*71, 4892–4903.
- Schneider, R. J. & Mohr, I.(2003).** Translation initiation and viral tricks. *Trends Biochem Sci*28, 130–136.
- Schoenberg, D. R. & Maquat, L. E.(2012).** Regulation of cytoplasmic mRNA decay. *Nat Rev Genet*13, 246–259.

- Schröder, M.(2011).** Viruses and the human DEAD-box helicase DDX3: inhibition or exploitation? *Biochem Soc Trans***39**, 679–683.
- Schüler, M., Connell, S. R., Lescoute, A., Giesebrecht, J., Dabrowski, M., Schroeer, B., Mielke, T., Penczek, P. A., Westhof, E. & Spahn, C. M. T.(2006).** Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nat Struct Mol Biol***13**, 1092–1096.
- Schwartz, S., Felber, B. K., Fenyö, E. M. & Pavlakis, G. N.(1990).** Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. *J Viro***64**, 5448–5456.
- Schwartz, S., Felber, B. K. & Pavlakis, G. N.(1992a).** Mechanism of translation of monocistronic and multicistronic human immunodeficiency virus type 1 mRNAs. *Mol Cell Bio***12**, 207–219.
- Schwartz, S., Felber, B. K. & Pavlakis, G. N.(1992b).** Distinct RNA sequences in the gag region of human immunodeficiency virus type 1 decrease RNA stability and inhibit expression in the absence of Rev protein. *J Viro***66**, 150–159.
- SenGupta, D. N. & Silverman, R. H.(1989).** Activation of interferon-regulated, dsRNA-dependent enzymes by human immunodeficiency virus-1 leader RNA. *Nucleic Acids Res***17**, 969–978.
- SenGupta, D. N., Berkhout, B., Gagnol, A., Zhou, A. M. & Silverman, R. H.(1990).** Direct evidence for translational regulation by leader RNA and Tat protein of human immunodeficiency virus type 1. *Proc Natl Acad Sci USA***87**, 7492–7496.
- Sharma, A., Yilmaz, A., Marsh, K., Cochrane, A. & Boris-Lawrie, K.(2012).** Thriving under stress: selective translation of HIV-1 structural protein mRNA during Vpr-mediated impairment of eIF4E translation activity. *PLoS Pathog***8**, e1002612.

- Sharma, D. & Jankowsky, E.(2014).** The Ded1/DDX3 subfamily of DEAD-box RNA helicases. *Crit Rev Biochem Mol Biol***49**, 343–360.
- Shen, Z.-J. & Malter, J. S.(2015).** Regulation of AU-Rich Element RNA Binding Proteins by Phosphorylation and the Prolyl Isomerase Pin1. *Biomolecules***5**, 412–434.
- Shih, J.-W., Wang, W.-T., Tsai, T.-Y., Kuo, C.-Y., Li, H.-K. & Wu Lee, Y.-H.(2012).** Critical roles of RNA helicase DDX3 and its interactions with eIF4E/PABP1 in stress granule assembly and stress response. *Biochemical Journal***441**, 119–129.
- Shih, J.-W., Tsai, T.-Y., Chao, C.-H. & Wu Lee, Y.-H.(2008).** Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. *Oncogene***27**, 700–714.
- Shin, B.-S., Kim, J.-R., Walker, S. E., Dong, J., Lorsch, J. R. & Dever, T. E.(2011).** Initiation factor eIF2 $\gamma$  promotes eIF2-GTP-Met-tRNA<sup>i</sup>(Met) ternary complex binding to the 40S ribosome. *Nat Struct Mol Biol***18**, 1227–1234.
- Shine, J. & Dalgarno, L.(1974).** The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci USA***71**, 1342–1346.
- Shuman, S.(2000).** Structure, mechanism, and evolution of the mRNA capping apparatus. *Prog Nucleic Acid Res Mol Biol***66**, 1–40.
- Sierra, S., Kupfer, B. & Kaiser, R.(2005).** Basics of the virology of HIV-1 and its replication. *Journal of Clinical Virology, Focus on HIV* **34**, 233–244.
- Simon, F., Maucière, P., Roques, P., Loussert-Ajaka, I., Müller-Trutwin, M. C., Saragosti, S., Georges-Courbot, M. C., Barré-Sinoussi, F. & Brun-Vézinet, F.(1998).** Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med***4**, 1032–1037.

- Singleton, M. R., Dillingham, M. S. & Wigley, D. B.(2007).** Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem***76**, 23–50.
- Siridechadilok, B., Fraser, C. S., Hall, R. J., Doudna, J. A. & Nogales, E.(2005).** Structural Roles for Human Translation Factor eIF3 in Initiation of Protein Synthesis. *Science***310**, 1513–1515.
- Skabkin, M. A., Skabkina, O. V., Dhote, V., Komar, A. A., Hellen, C. U. T. & Pestova, T. V.(2010).** Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes Dev***24**, 1787–1801.
- Sonenberg, N. & Dever, T. E.(2003).** Eukaryotic translation initiation factors and regulators. *Current Opinion in Structural Biology***13**, 56–63.
- Sonenberg, N. & Hinnebusch, A. G.(2009).** Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell***136**, 731–745.
- Soto-Rifo, R., Rubilar, P. S. & Ohlmann, T.(2013a).** The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA. *Nucleic Acids Research***41**, 6286–6299.
- Soto-Rifo, R. & Ohlmann, T.(2013).** The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism. *Wiley Interdiscip Rev RNA***4**, 369–385.
- Soto-Rifo, R., Limousin, T., Rubilar, P. S., Ricci, E. P., Décimo, D., Moncorgé, O., Traub, M.-A., André, P., Cimarelli, A. & Ohlmann, T.(2012a).** Different effects of the TAR structure on HIV-1 and HIV-2 genomic RNA translation. *Nucleic Acids Res***40**, 2653–2667.

**Soto-Rifo, R., Rubilar, P. S., Limousin, T., de Breyne, S., Décimo, D. & Ohlmann, T.(2012b).** DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. *EMBO J***31**, 3745–3756.

**Soto-Rifo, R., Rubilar, P. S., Limousin, T., de Breyne, S., Décimo, D. & Ohlmann, T.(2012c).** DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs: Translation initiation mediated by DDX3. *The EMBO Journal***31**, 3745–3756.

**Soto-Rifo, R., Rubilar, P. S. & Ohlmann, T.(2013b).** The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA. *Nucleic Acids Res***41**, 6286–6299.

**Spahn, C. M., Kieft, J. S., Grassucci, R. A., Penczek, P. A., Zhou, K., Doudna, J. A. & Frank, J.(2001).** Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science***291**, 1959–1962.

**Spahn, C. M. T., Jan, E., Mulder, A., Grassucci, R. A., Sarnow, P. & Frank, J.(2004).** Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell***118**, 465–475.

**Stoneley, M., Subkhankulova, T., Le Quesne, J. P., Coldwell, M. J., Jopling, C. L., Belsham, G. J. & Willis, A. E.(2000).** Analysis of the c-myc IRES; a potential role for cell-type specific trans-acting factors and the nuclear compartment. *Nucleic Acids Res***28**, 687–694.

**Story, R. M. & Steitz, T. A.(1992).** Structure of the recA protein-ADP complex. *Nature***355**, 374–376.



- Story, R. M., Li, H. & Abelson, J. N.(2001).** Crystal structure of a DEAD box protein from the hyperthermophile *Methanococcus jannaschii*. *Proc Natl Acad Sci USA***98**, 1465–1470.
- Svitkin, Y. V., Pause, A. & Sonenberg, N.(1994).** La autoantigen alleviates translational repression by the 5' leader sequence of the human immunodeficiency virus type 1 mRNA. *J Viro***68**, 7001–7007.
- Swanson, C. M., Sherer, N. M. & Malim, M. H.(2010).** SRp40 and SRp55 promote the translation of unspliced human immunodeficiency virus type 1 RNA. *J Viro***84**, 6748–6759.
- Tarn, W.-Y. & Chang, T.-H.(2009).** The current understanding of Ded1p/DDX3 homologs from yeast to human. *RNA Biol***6**, 17–20.
- Tazi, J., Bakkour, N., Marchand, V., Ayadi, L., Aboufirassi, A. & Branlant, C.(2010).** Alternative splicing: regulation of HIV-1 multiplication as a target for therapeutic action. *FEBS J***277**, 867–876.
- Terenin, I. M., Dmitriev, S. E., Andreev, D. E. & Shatsky, I. N.(2008).** Eukaryotic translation initiation machinery can operate in a bacterial-like mode without eIF2. *Nat Struct Mol Biol***15**, 836–841.
- Topisirovic, I., Svitkin, Y. V., Sonenberg, N. & Shatkin, A. J.(2011).** Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA***2**, 277–298.
- Vagner, S., Galy, B. & Pyronnet, S.(2001).** Irresistible IRES. Attracting the translation machinery to internal ribosome entry sites. *EMBO Rep***2**, 893–898.
- Valiente-Echeverría, F., Vallejos, M., Monette, A., Pino, K., Letelier, A., Huidobro-Toro, J. P., Mouland, A. J. & López-Lastra, M.(2013).** A cis-acting element

present within the Gag open reading frame negatively impacts on the activity of the HIV-1 IRES. *PLoS ONE*8, e56962.

**Valiente-Echeverría, F., Hermoso, M. A. & Soto-Rifo, R.(2015).** RNA helicase DDX3: at the crossroad of viral replication and antiviral immunity. *Rev Med Virol*25, 286–299.

**Vallejos, M., Deforges, J., Plank, T.-D. M., Letelier, A., Ramdohr, P., Abraham, C. G., Valiente-Echeverría, F., Kieft, J. S., Sargueil, B. & López-Lastra, M.(2011).** Activity of the human immunodeficiency virus type 1 cell cycle-dependent internal ribosomal entry site is modulated by IRES trans-acting factors. *Nucleic Acids Res*39, 6186–6200.

**Vallejos, M., Carvajal, F., Pino, K., Navarrete, C., Ferres, M., Huidobro-Toro, J. P., Sargueil, B. & López-Lastra, M.(2012).** Functional and structural analysis of the internal ribosome entry site present in the mRNA of natural variants of the HIV-1. *PLoS ONE*7, e35031.

**Ventoso, I., MacMillan, S. E., Hershey, J. W. & Carrasco, L.(1998).** Poliovirus 2A proteinase cleaves directly the eIF-4G subunit of eIF-4F complex. *FEBS Lett*435, 79–83.

**Ventoso, I., Blanco, R., Perales, C. & Carrasco, L.(2001).** HIV-1 protease cleaves eukaryotic initiation factor 4G and inhibits cap-dependent translation. *Proc Natl Acad Sci USA*98, 12966–12971.

**Waysbort, A., Bonnal, S., Audigier, S., Estève, J. P. & Prats, A. C.(2001).** Pyrimidine tract binding protein and La autoantigen interact differently with the 5' untranslated regions of lentiviruses and oncoretrovirus mRNAs. *FEBS Lett*490, 54–58.

- Wei, C. M. & Moss, B.(1975).** Methylated nucleotides block 5'-terminus of vaccinia virus messenger RNA. *Proc Natl Acad Sci USA***72**, 318–322.
- Weill, L., James, L., Ulryck, N., Chamond, N., Herbreteau, C. H., Ohlmann, T. & Sargueil, B.(2010).** A new type of IRES within gag coding region recruits three initiation complexes on HIV-2 genomic RNA. *Nucleic Acids Research***38**, 1367–1381.
- Weingarten-Gabbay, S., Elias-Kirma, S., Nir, R., Gritsenko, A. A., Stern-Ginossar, N., Yakhini, Z., Weinberger, A. & Segal, E.(2016).** Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science***351**.
- Wilen, C. B., Tilton, J. C. & Doms, R. W.(2012).** HIV: cell binding and entry. *Cold Spring Harb Perspect Med***2**.
- Wilson, J. E., Powell, M. J., Hoover, S. E. & Sarnow, P.(2000a).** Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol Cell Biol***20**, 4990–4999.
- Wilson, J. E., Pestova, T. V., Hellen, C. U. T. & Sarnow, P.(2000b).** Initiation of Protein Synthesis from the A Site of the Ribosome. *Cell***102**, 511–520.
- Woolaway, K., Asai, K., Emili, A. & Cochrane, A.(2007).** hnRNP E1 and E2 have distinct roles in modulating HIV-1 gene expression. *Retrovirology***4**, 28.
- Yanagiya, A., Svitkin, Y. V., Shibata, S., Mikami, S., Imataka, H. & Sonenberg, N.(2009).** Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4GI (eIF4GI) for efficient interaction of eIF4E with the mRNA cap. *Mol Cell Biol***29**, 1661–1669.

- Yang, Q. & Jankowsky, E.(2005).** ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry***44**, 13591–13601.
- Yang, Q. & Jankowsky, E.(2006).** The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nat Struct Mol Bio***13**, 981–986.
- Yang, Q., Del Campo, M., Lambowitz, A. M. & Jankowsky, E.(2007).** DEAD-box proteins unwind duplexes by local strand separation. *Mol Cell***28**, 253–263.
- Yedavalli, V. S. R. K., Neuveut, C., Chi, Y.-H., Kleiman, L. & Jeang, K.-T.(2004).** Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. *Cell***119**, 381–392.
- Le sida en chiffres 2015 | ONUSIDA. . .**