Université de Paris

École doctorale Médicament, Toxicologie, Chimie, Imagerie

Laboratoire Cibles Thérapeutiques et Conception du Médicament

Caractérisation biochimique et enzymatique de l'hélicase DDX3 : influence sur la dimérisation de l'ARN génomique du VIH

Par Grégoire de Bisschop

Thèse de doctorat de Biologie moléculaire

Dirigée par Bruno Sargueil

Présentée et soutenue publiquement le 4 octobre 2019

Devant un jury composé de :

Jean-Christophe Paillart : Rapporteur

Laurence Briant : Rapporteure

Ricardo Soto-Rifo : Examinateur

Francesca Fiorini : Examinatrice

Nathalie Chamond : Examinatrice

Except where otherwise noted, this is work licensed under
https://creativecommons.org/licenses/by-nc-nd/3.0/fr/

Table des matières

Liste des abréviations ..... **iii**

Résumé ..... **v**

Introduction **1**

Introduction générale 2

L'ARN génomique du VIH-1 3

1) Le VIH-1, un rétrovirus 3

2) Formes du génome viral au cours du cycle 4

3) *Traduction de l'ARN génomique* 12

4) *Dimérisation de l'ARN génom*ique du VIH-1 20

5) *Encapsidation de l'ARN génomique* 30

6) Conclusion 33

DDX3, une hélicase cellulaire impliquée dans le cycle réplicatif du VIH-1 35

1) Classification des hélicases 35

2) Relation structure – *fonction du cœur catalyt*ique et des extensions terminales de DDX3 36

3) *DDX3 joue de nombreux rôles dans le contrôle de l'expression génétique* 52

4) *Le VIH détourne les fonctions cellulaires de DDX3 lors de l'infection* 63

5) Conclusion 67

Objectifs **69**

Résultats **71**

Etude structurale de l'ARN génomique du VIH-1 73

1) Les dimères formés peuvent être de différentes natures 73

2) *Interdépendance du repliement de la région codante et de la 5'UTR* 84

3) Effet de la méthylation des adénosines en N6 sur la dimérisation 88

Etude de l'interaction de DDX3 avec l'ARN génomique du VIH 113
1) Caractérisation biochimique de DDX3 113
2) *Caractérisation de l'interaction de DDX3 à l'ARN génomique* 127
DDX3 déstabilise des dimères d'ARN du VIH-1 131
1) DDX3 déstabilise *des dimères d'ARN du VIH*-1 131
2) Analyse cinétique des paramètres de dissociation des dimères 132
3) *Caractérisation du remaniement de l'ARN par SHAPE résolu en temps* . 136
Rôle de DDX3 dans la traduction de l'ARN génomique du VIH-1 141
1) *Recherche des interactions de DDX3 avec des facteurs de l'initiation de la* traduction 141
2) *Evaluation de l'effet de DDX3 sur l*a traduction in vitro 146
Discussion 149
Relations structure – fonction de l'ARNg du VIH-1 151
1) Structures secondaire et tertiaire du PBS 151
2) Différents sous-types, différents mécanismes de dimérisation? 152
3) Régulation de la dimérisation par la méthylation 157
4) *Interdépendance du repliement de la 5'UTR et de la région codante de Gag* 157
DDX3 régule l'expression de l'ARNg 161
1) Effet de DDX3 sur la *traduction de l'ARN génomique* 161
2) *Autres effets possibles de DDX3 sur l'expression de l'ARN génomique* .. 162
Perspectives 165
Annexes 167
1) La technique de SHAPE 167
2) Analyse de la dimérisation par électrophorèse en conditions natives 168
Références bibliographiques 169

Liste des abréviations

1M7 : 1-Methyl-7-nitroisatoic anhydride – Anhydride 1-méthyl-7-nitroisatoïque

4E-BP : eIF4E Binding Protein – Protéine de liaison à eIF4E

5'/3'UTR : 5'/3' UnTranslated Region - Région non traduite 5'/3'

BMH : Branched Multiple Hairpins – Multiples tige-boucles connectées

BzCN : Benzoyl Cyanide – Cyanure de benzoyle

CBC : Cap Binding Complex – Complexe de liaison à la coiffe

CRM-1 : Chromosomal Maintenance 1 – Maintenance du Chromosome 1

DIS : Dimerization Initiation Site – Site d'initiation de la dimérisation

DLS : Dimer Linkage Structure – Structure de liaison des dimères

eIF : eukaryotic Initiation Factor - Facteur d'initiation eucaryote

EJC : Exon Jonction Complex – Complexe de jonction d'exon

Env : viral Enveloppe protein – Protéine de l'enveloppe virale

Gag : Group Specific Antigen – Antigène spécifique du groupe

IFN : Interferon

IKK : IκB kinase – Kinase IκB

IRES : Ribosome Entry Site – Site d'entrée interne du ribosome

LDI : Long Distance Interactions – Interactions de longue distance

LTR : Long Terminal Repeat – Longue répétition terminale

LysRS : Lysyl-$_t$RNA Synthase – Lysyl-ARN$_t$ synthétase

m6A : N6-Methyl-Adenosine

m7G : N7-Methyl-Guanosine

MAVS : Mitochondrial Antiviral Signaling protein – Protéine mitochondriale de signalisation antivirale

MSD : Major Splice Donor site – Site majeur donneur d'épissage

MBP : Maltose Binding Protein – Protéine de liaison au maltose

NC : NucleoCapsid – Nucléocapside

Nef : Negative regulatory factor – Facteur régulateur négatif

NES : Nuclear Export Signal – Signal d'export nucléaire

NELF : Negative Elongation Factor – Facteur négatif d'élongation

NFκB : Nuclear Factor κ B – Facteur nucléaire κ B

NLS : Nuclear Localization Signal – Signal de localisation nucléaire

NXF1 : Nuclear eXport Factor 1 – Facteur d'export nucléaire 1

ORF : Open Reading Frame – Cadre de lecture ouvert

PABP : PolyA Binding Protein – Protéine de liaison à la polyA

PBS : Primer Binding Site – Site de liaison à l'amorce

Pol : viral DNA Polymerase – Polymérase d'ADN virale

P-TEFb : Positive Transcription Elongation Factor b – Facteur positif d'élongation de la transcription b

RecA : Recombinase A

Rev : Regulator of Expression of Virion proteins – Régulateur de l'expression des protéines de virion

RRE : Rev Responsive Element – Elément de réponse à Rev

RRL : Rabbit Reticulocytes Lysate – Lysta de réticulocytes de lapin

SHAPE : Selective Hydroxyl-acylation Analyzed by Primer Extension – Hydroxyl-acylation sélective analysée par extension d'amorce

TAP : Tip Associated Protein – Protéine associée à l'embout

TAR : Trans-Activation Response element – Elément de réponse à la transactivation

Tat : Tran-Activator of Transcription – Trans-activateur de la transcription

TLE : $_{t}$RNA-Like Element – Elément similaire à l'ARN$_{t}$

Vif : Viral Infectivity Factor – Facteur d'infectivité virale

VHC : Heptitis C Virus – Virus de l'Hépatite C

Vpr : Viral Protein R – Protéine virale R

Vpu : Viral Protein U – Protéine virale U

Résumé

DDX3 est une hélicase à boîte DEAD humaine impliquée dans de nombreuses étapes du métabolisme de l'ARN. Elle joue un rôle clef pour la réplication du Virus de l'Immunodéficience Humaine de type 1 (VIH-1). En particulier, elle intervient dans l'export nucléaire de l'ARN génomique viral et dans sa traduction. Cependant, les bases moléculaires du rôle de cette protéine cellulaire pour le cycle viral ne sont pas connues, les propriétés biochimiques de DDX3 étant elles-mêmes incertaines. J'ai effectué la caractérisation biochimique et enzymatique de DDX3 et j'ai confirmé qu'elle présente une activité hélicase dépendante de l'ATP. J'ai mis en évidence la présence de déterminants de spécificité dans l'ARN génomique du VIH-1 capables de stimuler l'activité de cette hélicase à un niveau encore jamais observé. J'ai recherché des sites d'interaction de DDX3 avec l'ARN génomique et j'ai montré que DDX3 interagit avec des structures impliquées dans la dimérisation de cet ARN. Ces observations m'ont conduit à examiner l'effet de DDX3 sur la dimérisation de l'ARNg du VIH. DDX3 s'avère capable de dissocier des dimères d'ARNg avec une efficacité remarquable pour une hélicase à boîte DEAD. J'ai caractérisé le remaniement structural qu'elle entraîne en utilisant la technique de SHAPE en temps réel. Parallèlement, j'ai mis en évidence l'existence de différents mécanismes de dimérisation de l'ARN génomique en comparant deux sous-types couramment utilisés en laboratoire. Enfin, j'ai montré que la dimérisation pouvait être régulée post-transcriptionnellement par la méthylation des adénosines.

DDX3 is a human DEAD-box helicase involved in many steps of RNA metabolism. It has a key role in type 1 human immunodeficiency virus (HIV-1) cycle. In particular, it is involved in the nuclear export of the viral genomic RNA and in its translation. However, the molecular basis of the role of this cellular protein on the virus cycle are not known. The biochemical properties of DDX3 themselves are still unclear. During my PhD, I performed the biochemical and enzymatic characterization of DDX3 and I confirmed that it has an ATP dependent helicase activity. Moreover, I discovered the presence of specificity determinants within HIV-1 genomic RNA that are able to stimulate DDX3's activity to an unprecedented level for a DEAD-box helicase. I looked for the binding sites of DDX3 on HIV gRNA and I showed that DDX3 interacts with structures involved in the dimerization of this RNA. DDX3 turns out to be able to dissociate gRNA dimer with a high efficiency. I characterized the structural remodeling triggered by DDX3 using time-resolved SHAPE. I have also highlighted the existence of different mechanisms of dimerization by comparing two subtypes commonly used in laboratories. Finally, I showed that dimerization could be post-transcriptionally regulated by adenosine methylation.

Introduction

Introduction générale

La famille des hélicases à boîte DEAD a été définie en 1989 à partir d'un alignement des séquences de sept protéines de divers organismes, homologues de l'hélicase humaine eIF4A (Linder et al., 1989). De nombreuses hélicases à boîte DEAD ont été découvertes dans les trois domaines du vivant ainsi que chez certains virus. Elles assurent des fonctions dans la quasi-totalité des processus biologiques (Linder and Fuller-Pace, 2013).

DDX3 est une hélicase à boîte DEAD humaine découverte en 1998. Une expression anormale de DDX3 a été retrouvée dans de nombreuses situations pathologiques, principalement des cancers et des infections virales (Ariumi, 2014) mais aussi des troubles du développement (Kellaris et al., 2018). Cette hélicase constitue donc une cible thérapeutique d'importance.

En particulier, DDX3 s'avère nécessaire à la réplication du VIH-1 (Yedavalli et al., 2004). Il semble que plusieurs étapes du cycle de réplication du virus soient concernées, notamment les étapes tardives d'export nucléaire et de traduction de l'ARN génomique, mais on ignore les bases moléculaires de l'interaction entre DDX3 et cet ARN viral. L'objectif de ma thèse consiste à établir une caractérisation biochimique de DDX3 et à étudier son interaction avec l'ARN génomique du VIH-1 pour identifier d'éventuels motifs de spécificité dans l'ARN.

L'ARN génomique du VIH-1

1) Le VIH-1, un rétrovirus

Le Virus de l'Immunodéficience Humaine (VIH) est l'agent infectieux pandémique responsable du Syndrome de l'Immunodéficience Acquise (SIDA). L'Organisation Mondiale de la Santé rapporte que le VIH a causé 35 millions de décès à travers le monde, et qu'environ autant de personnes sont actuellement infectées. La pandémie du VIH représente deux enjeux : d'une part, il n'existe pour l'heure aucun traitement permettant de guérir de l'infection ; d'autre part, un quart des personnes potentiellement infectées ignore son statut (http://www.who.int/hiv/fr). On distingue principalement deux types de VIH, le VIH-1 qui est le plus répandu et le VIH-2 qui est moins infectieux et retrouvé principalement en Afrique de l'Ouest.

Le VIH est un lentivirus, un genre rattaché à la famille des rétrovirus. Les rétrovirus sont des virus enveloppés contenant deux transcrits d'ARN de polarité positive de 7 à 12 kb (9,7 kb pour le VIH-1). On distingue trois groupes de VIH-1, M (Main), O (Outlier) et N (Non-M, Non-O), le groupe M étant le plus répandu. Ce groupe se divise lui-même en 9 clades ou sous-types : A-D, F-H, J et K. Le sous-type B est le plus répandu dans le monde occidental, le C est répandu en Chine, en Inde et en Afrique sub-saharienne tandis que les autres sous-types sont retrouvés principalement en Afrique (Karlsson Hedestam et al., 2008). Dans ma thèse, j'ai principalement utilisé le clone moléculaire infectieux NL4-3 correspondant au sous-type B. Il résulte de la recombinaison de l'isolat NY5 (partie 5') et LAI (partie 3'). Ce plasmide est largement utilisé expérimentalement, et il semble être suffisamment représentatif du groupe M (Adachi et al., 1986; Thomson et al., 2002).

Les particules virales matures mesurent 80 à 120 nm de diamètre et sont composées d'une membrane d'origine cellulaire dans laquelle sont insérées les glycoprotéines (gp) virales gp120 et gp41. La membrane est tapissée d'une protéine de matrice (MA). Elle renferme une nucléocapside formée par l'assemblage de protéines de capside (CA). Au sein de la capside se trouvent les deux molécules d'ARN de polarité positive liées par des liaisons hydrogène ainsi que des protéines virales nécessaires à la réplication du virus, des protéines cellulaire et l'ARN de transfert 3 de la lysine (ARN_t^{Lys3}) (Karlsson Hedestam et al., 2008) (Figure 1).

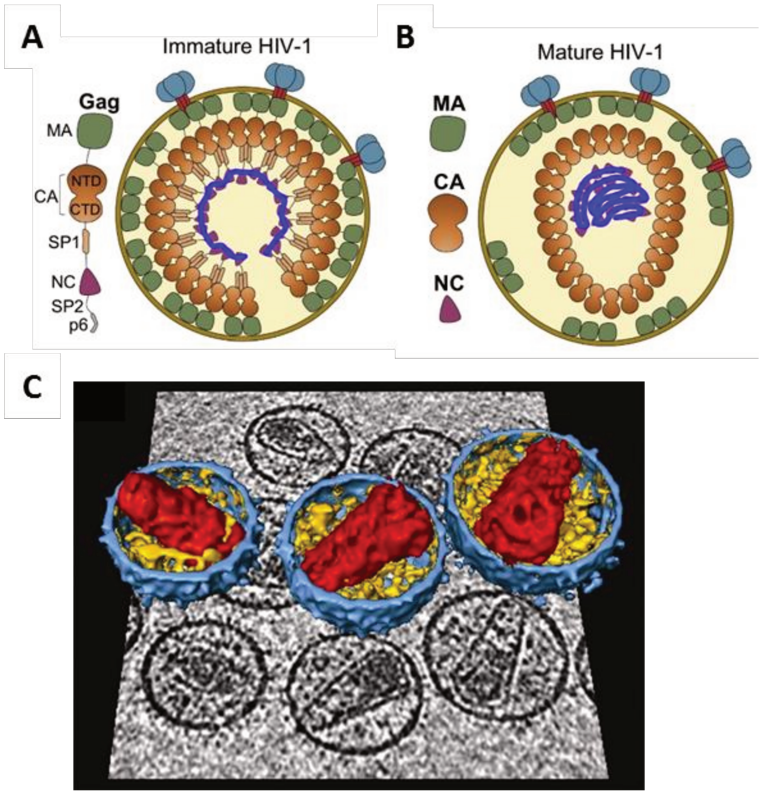


Figure 1 : Structure des particules virales du VIH-1. Schéma des particules virales néoformées, immatures (A), et matures (B). Seuls les protéines structurales Gag, gp120 et gp41 *ainsi que l'ARN génomique (en bleu foncé) sont représentés* (Pornillos and Ganser-Pornillos, 2019). C/ Structure des particules virales matures obtenue en cyro-électrotomographie (Briggs et al., 2006). *La densité correspondant à l'enveloppe, composée d'une membrane et des* protéines de matrice, est représentée en bleu, la capside est en rouge et la densité intermédiaire en jaune.

2) Formes du génome viral au cours du cycle

i. Fusion de la particule virale et transcription inverse de l'ARN génomique

Les particules virales interagissent via leur glycoprotéine membranaire gp120 avec les récepteurs CD4 et CCR5 ou CXCR4, présents majoritairement à la surface des lymphocytes T mais aussi des monocytes, des cellules dendritiques et des macrophages. Cette interaction provoque la fusion de la membrane virale avec la membrane cellulaire, dont résulte la libération de la nucléocapside dans le cytoplasme. La structure de la nucléocapside est remaniée pendant qu'elle interagit avec les microtubules et progresse vers le noyau. Lors de cette phase cytoplasmique, l'ARN est rétrotranscrit en ADN double brin. Ce mécanisme complexe fait intervenir la polymérase virale, qui peut synthétiser de l'ADN à partir d'ARN ou d'ADN, ainsi

que l'ARN_t^{Lys3}. Il est relativement conservé chez tous les rétrovirus. L'ARN génomique viral présente une région répétée R à chaque extrémité, immédiatement suivie d'une région unique U5 en 5' et U3 en 3'. En aval de l'élément U5 se trouve un site de liaison à l'ARN_t^{Lys3}, le PBS (Primer Binding Site). L'ARN_t^{Lys3} est hybridé à l'ARN génomique au cours de l'encapsidation dans la cellule productrice du virus. La polymérase virale initie la rétrotranscription en utilisant l'ARN_t^{Lys3} comme amorce et progresse vers l'extrémité 5' de l'ARN génomique. Elle recopie l'élément U5 puis R en ADN de polarité négative (-) pendant que son module RNAse H dégrade l'ARN génomique à mesure qu'il est rétrotranscrit. Cette amorce ADN (-) est ensuite hybridée à l'élément R situé en 3' de l'ARN génomique et permet la rétrotranscription de la quasi-totalité de l'ARN génomique, jusqu'au PBS. Une seule région, le polypurine tract (PPT) situé en 5' de l'élément U3, échappe à la dégradation par l'activité RNAse H et sert d'amorce à la synthèse du brin d'ADN (+) correspondant aux éléments U3, R et U5. Ce brin peut alors s'hybrider avec l'élément U5 situé en 3' du brin d'ADN (-) et permettre la synthèse du brin d'ADN (+). L'ADN double brin résultant de la rétrotranscription de l'ARN génomique présente donc un élément répété aux deux extrémités, U3 – R – U5 (Basu et al., 2008) (Figure 2).

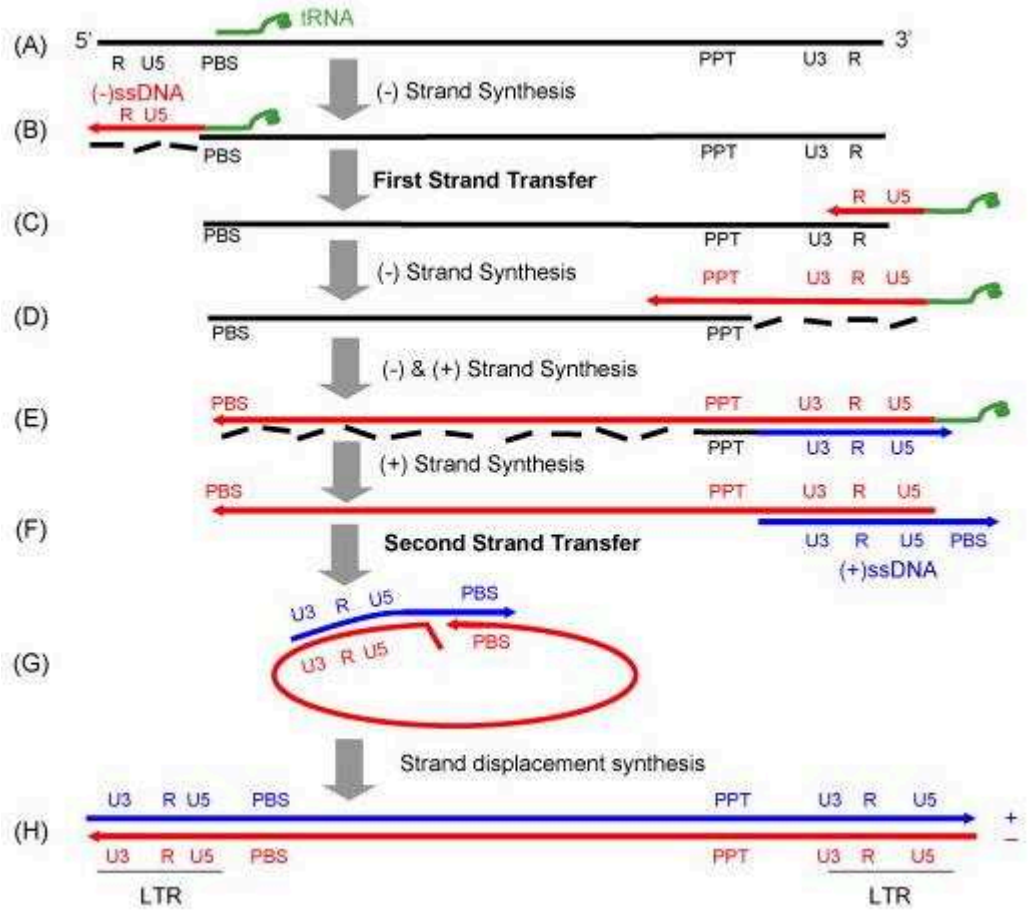


Figure 2 : Mécanisme général de transcription inverse chez les rétrovirus. *L'ARN (de polarité* positive) est repr*ésenté en noir, l'ADN de polarité négative en rouge et l'ADN de polarité* positive en bleu. *L'ARN*$_t^{Lys3}$*, d'origine cellulaire, est en vert.* Tiré de (Basu et al., 2008)

L'ADN double brin est ensuite clivé aux extrémités par l'intégrase virale qui génère des extrémités cohésives. Ce complexe de pré-intégration (PIC, Pre-Integration Complex) pénètre dans le noyau via le complexe de pore nucléaire, et l'intégrase insère l'ADN viral dans une région chromosomique propice. Ce segment d'ADN viral inséré dans le génome de l'hôte se nomme provirus et assure la production des transcrits viraux.

ii. Transcription de l'ARN génomique

La longue répétition terminale (LTR, Long Terminal Repeat) du provirus contient un promoteur de transcription. Ce promoteur présente trois sites de liaison au facteur de transcription Sp1 capable de recruter le facteur général de transcription TFIID au niveau d'une boîte TATA. Le facteur TFIID stimule la formation d'un complexe d'initiation avec la polymérase ARN II (Pol II). Le promoteur du VIH-1 dispose également d'une région stimulatrice (enhancer) qui peut lier le facteur NFκB. NFκB recrute à son tour des

acétyltransférases qui remodèlent la chromatine et augmentent l'efficacité du promoteur (Karn and Stoltzfus, 2012). L'initiation de la transcription débute au niveau de la jonction U3 – R, où trois guanosines successives peuvent chacune être choisies comme premier nucléotide transcrit. Il est possible que le devenir cellulaire de l'ARN viral dépende du nombre de G en 5' (Kharytonchyk et al., 2016). Quoiqu'il en soit, une coiffe est ajoutée en 5' de l'ARN naissant par la machinerie cellulaire après une vingtaine de nucléotides transcrits. Cette coiffe est constituée d'un résidu guanosine méthylé en N7 lié par une liaison 5' – 5' triphosphate au premier nucléotide transcrit. La protéine virale Tat stimule l'activation de la méthylase par une interaction directe via son domaine C-terminal (Chiu et al., 2002).

En absence de Tat, on constate une accumulation de complexes d'initiation au niveau du promoteur (Kao et al., 1987). Cela suggère que Tat est impliqué dans l'élongation de la transcription virale, mais pas dans l'initiation. Une série de travaux a permis d'identifier la cible de Tat et ses partenaires protéiques. Tout d'abord, Tat permet l'expression d'un gène rapporteur sous contrôle du promoteur viral LTR (Sodroski et al., 1985). Par des délétions successives, les auteurs ont montré que l'élément TAR en 3' du LTR est la cible de Tat (Muesing et al., 1987). Tat interagit directement avec la tige-boucle TAR du transcrit en cours d'élongation. La structure de cette tige-boucle est critique pour cette interaction (Berkhout and Jeang, 1989). La structure RMN de Tat lié à TAR révèle que cette interaction se fait au niveau d'un bulge (boucle interne asymétrique) de TAR riche en U à proximité de la boucle apicale. La stimulation de l'élongation de la Pol II par Tat implique également des facteurs cellulaires. En effet, des mutations de la boucle apicale de TAR qui n'affectent pas la liaison de Tat peuvent avoir un effet sur la stimulation de la transcription (Feng and Holland, 1988). Un tel facteur a été identifié, il s'agit du complexe P-TEFb (Positive Transcription Elongation Factor b) (Karn and Stoltzfus, 2012). La structure cristallographique du complexe TAR – Tat – P-TEFb montre que Tat interagit simultanément avec les sous-unités CycT1 et Cdk9 de P-TEFb et que CycT1 lie la boucle apicale de TAR (Tahirov et al., 2010). Le modèle d'activation de la transcription actuellement admis propose qu'une fois l'initiation de la transcription effectuée, la Pol II s'arrête après avoir transcrit TAR sous l'effet du facteur négatif d'élongation NELF (Negative Elongation Factor). TAR recrute Tat, qui avec la boucle apicale de TAR recrutent le complexe P-TEFb. La sous-unité Cdk9 phosphoryle le facteur positif d'élongation Spt5 et hyperphosphoryle le domaine C-terminal de la Pol II. Ceci a pour effet de stimuler la Pol II. En parallèle, Cdk9 phosphoryle NELF et provoque ainsi son relargage (Karn and Stoltzfus, 2012) (Figure 3).

C’est grâce à tous ces éléments activateurs que le promoteur du VIH est susceptible de devenir extrêmement actif lors de la phase virulente (Karn and Stoltzfus, 2012).

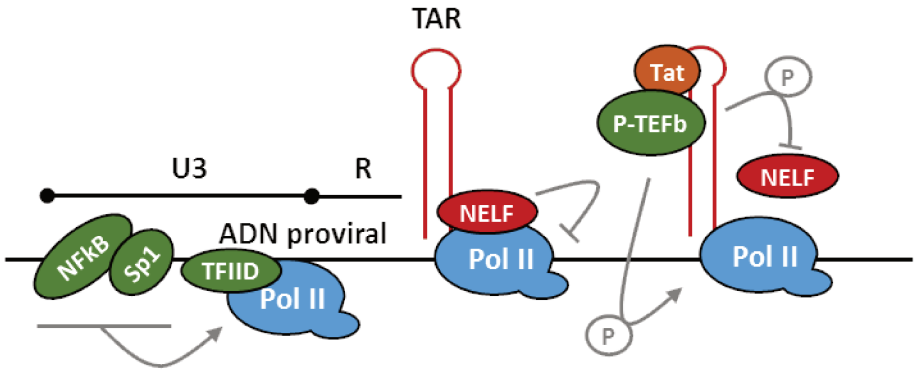


Figure 3 : Mécanisme d'activation de la transcription de l'ARN du VIH-1. *L’ADN proviral est représenté en noir tandis que l’ARN est en rouge. La transcription reprend après l’inhibition de NELF et l’activation de la Pol II.*

iii. Epissage alternatif de l’ARN génomique

Le VIH-1, sous forme provirale, produit par transcription un ARN dit « de pleine longueur » qui présente 4 sites donneurs d’épissage et 8 sites accepteurs (Figure 4). Ces sites sont reconnus par la machinerie d’épissage cellulaire. On retrouve ainsi dans les cellules infectées plus de quarante transcrits différents (Karn and Stoltzfus, 2012) qui assurent la traduction de quinze protéines (Frankel and Young, 1998). On distingue des transcrits totalement épissés, des transcrit partiellement épissés, et des transcrits non épissés qui correspondent à l’ARN génomique (Cullen, 2003). Les ARNm totalement épissés codent les protéines régulatrices Tat, Rev et Nef, tandis que les ARNm partiellement ou pas épissés codent les protéines de structure, notamment Gag et Env. L’écart important entre le nombre de transcrits et le nombre de protéines virales effectivement produites s’explique par la présence de deux exons non codants qui subissent un épissage alternatif (de Breyne and Ohlmann, 2019). La fonction de ces deux petits exons n’est pas connue, et leur présence est limitée à une faible proportion de transcrits (Emery et al., 2017).

L’épissage de l’ARN du VIH-1 est finement régulé, de telle sorte que la production des transcrits viraux évolue au cours du temps. Les transcrits codant Tat, Rev et Nef sont exprimés tôt au cours du cycle viral ; les transcrits codant Env, Vif, Vpr et Vpu apparaissent plus tard ; enfin les transcrits codant Gag et Pol sont exprimés lors des derniers stades du cycle. Cette régulation repose sur des séquences activatrices et inhibitrices présentes dans les exons et dans les introns. Par exemple, le site majeur donneur d’épissage (MSD, Major Splice Donor site),

situé dans la 5'UTR et qui permet de générer tous les transcrits épissés, est stimulé par la liaison du facteur hnRNP H (heterogeneous nuclear Ribonucleoprotein H) (LeBlanc et al., 2013). Les facteurs viraux interviennent également dans la régulation de l'épissage. Ainsi Tat exerce un rétrocontrôle négatif sur la production de ses transcrits et positif sur la production des transcrits de Env (Jablonski et al., 2010). Enfin il est possible que la structure secondaire de l'ARN intervienne dans l'épissage en modulant la liaison de facteurs protéiques régulateurs (Jablonski et al., 2008).

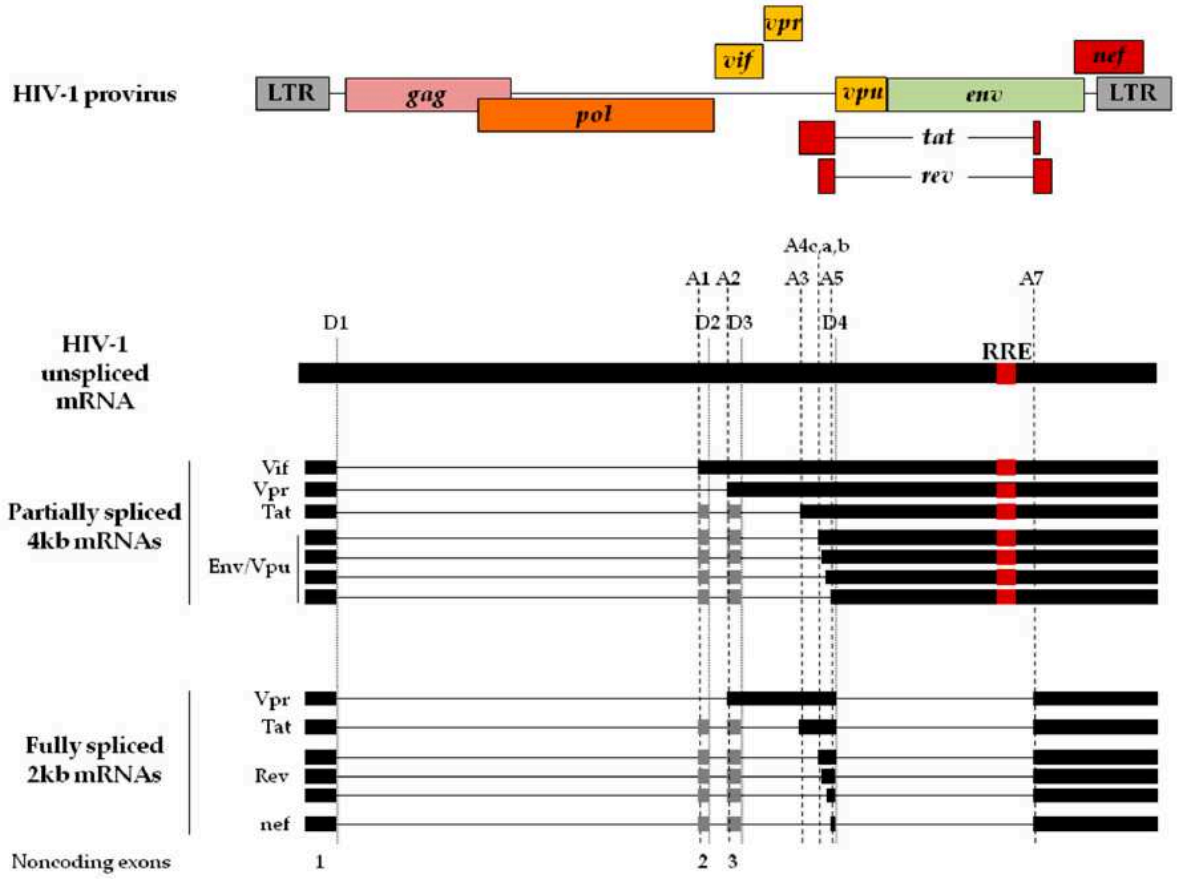


Figure 4 ***: L'é*****pissage alternatif de l'ARN génomique assure la production des différentes protéines virales** (de Breyne et Ohlmann, 2019). Haut : les cadres de lecture sont indiqués en couleur. Bas : Structure des transcrits épissés. Les sites donneurs d'épissage (D) et accepteur (A) sont indiqués par des pointillés. Les exons 1, 2 et 3 sont non codants. Les exons 2 et 3 ne sont pas toujours simultanément présents dans les transcrits. Le RRE, un motif impliqué dans *l'export des transcrits partiellement ou non épissés,* est indiqué en rouge.

iv. Clivage, polyadénylation

La polyadénylation des transcrits cellulaires repose sur la reconnaissance du motif AAUAAA par le complexe CPSF (Cleavage/Polyadenylation Specificity Factor) en amont du site de polyadénylation et d'un motif riche en G et U par le complexe hétérotrimérique CstF

(Cleavage stimulation factor) en aval (Karn and Stoltzfus, 2012). L'ARN génomique du VIH présente un site de polyadénylation dans chaque site R, en 5' et en 3'. Seul le site en 3' est actif car il est stimulé par l'élément U3 qui seul peut recruter le CPSF. Un autre site situé en 5' du gène Nef stimule le recrutement du CPSF via la protéine 9G8.

v. Export de l'ARN génomique

Les transcrits ainsi générés doivent être exportés du noyau afin d'exercer leurs fonctions cytoplasmiques. Néanmoins, la présence d'introns dans une partie des transcrits (ARNg ou transcrits partiellement épissés) empêche l'export nucléaire par la voie classique d'export des ARNm eucaryotes. Le contournement de ce mécanisme de contrôle se fait notamment par la protéine virale Rev. En effet, Rev assure l'export de ces transcrits en se liant à une structure particulière de l'ARN, le RRE (Rev Recognition Element) situé dans le gène Env. L'interaction Rev – RRE s'effectue via un motif riche en arginines qui entraîne l'oligomérisation de Rev autour de RRE (Cullen, 2003). La transposition de ce motif dans un gène dont l'export est initialement indépendant de Rev le rend en effet dépendant de Rev (Malim et al., 1989). Rev présente un signal d'export nucléaire (NES, Nuclear Export Signal) à son extrémité C terminale. L'injection nucléaire d'un large excès du NES de Rev inhibe l'export d'une certaine classe d'ARN, les petits ARN nucléaires riches en U (U snRNA) (Fischer et al., 1995). Ces résultats suggèrent que Rev est exporté suivant le même mécanisme que ces ARN, et non suivant la voie d'export des ARN_t ou celle des ARNm. Ces U-snRNA, de même que les ARN ribosomaux, sont exportés par la voie dépendante de CRM1 (Okamura et al., 2015). Une étude a montré que Rev est capable d'interagir directement avec CRM1 (Neville et al., 1997), ce qui conforte l'hypothèse selon laquelle Rev permet l'export des transcrits partiellement épissés suivant cette voie. Dans le noyau, CRM1 interagit avec Ran liée au GTP (RanGTP). Ce complexe lie les cargos présentant un NES adéquat, puis interagit avec le complexe de pore nucléaire. Une fois dans le cytoplasme, le GTP lié à Ran est hydrolysé ce qui provoque la dissociation du cargo (Okamura et al., 2015) et le recyclage de Rev qui est importée dans le noyau par l'importine β (Karn and Stoltzfus, 2012) (Figure 5). Notons que Rev recrute également une méthyltransférase cellulaire, PIMT (Peroxysome proliferator-activated receptor-Interacting protein with MethylTransferase). En conséquence, une partie des transcrits qui interagissent avec Rev ont une coiffe triméthylée qui stimule leur export et leur expression (Yedavalli and Jeang, 2010).

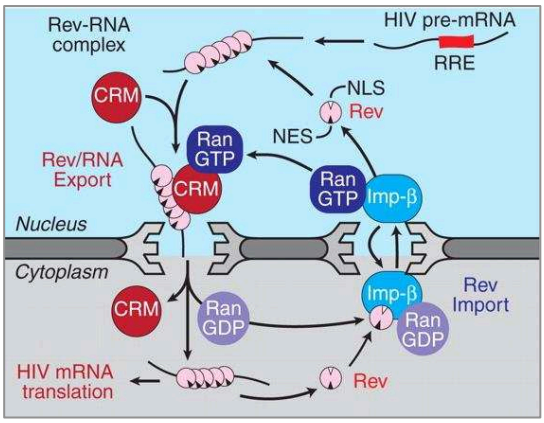


Figure 5 : Export des ARN partiellement épissés par la voie Rev - CRM1 (Karn et Stoltzfus, 2012). La protéine virale Rev lie le RRE sous forme oligomérique. Elle interagit avec le complexe CRM-1 – *RanGTP qui assure son export nucléaire. L'hydrolyse de GTP permet le relargage de l'ARN génomique et le recyclage des facteurs d'exports.*

Les transcrits totalement épissés sont exportés comme la plupart des ARNm cellulaires, suivant le voie TAP/NXF1 (Tip Associated Protein/Nuclear eXport Factor 1). TAP forme un complexe avec NXT1 et le complexe TREX-1. TREX-1 augmente l'affinité de TAP pour l'ARN tandis que NXT1 stimule les interactions du complexe TAP – ARN avec des nucléoporines. TREX-1 est recruté à l'ARN via le complexe de liaison à la coiffe CBC (Cap Binding Complex).

vi. Modifications post-transcriptionnelles de l'ARN viral

Comme tous les ARNm cellulaires, l'ARN génomique est susceptible de recevoir des modifications post-transcriptionnelles. Ces modifications covalentes des nucléotides régulent le métabolisme de l'ARN. Il existe près d'une centaine de modifications post-transcriptionnelles différentes, la méthylation de l'azote 6 de l'adénosine (m6A) étant la plus abondante (Roundtree et al., 2017). Elle est catalysée par un hétérodimère composé de METTL3 et METTL14. Cette modification, quoique stable, est réversible. Le groupement méthyl peut être retiré par les enzymes FTO (Fat mass and Obesity-associated protein) ou AlkBH5. La méthylation des adénosines pourrait modifier le devenir d'un ARN de trois façons différentes. Tout d'abord, les résidus m6A peuvent être spécifiquement reconnus par des protéines, notamment celles qui présentent un domaine « reader » YTH. D'autre part, la méthylation de l'adénosine pourrait engendrer une modification de la structure de l'ARN et être ainsi détectée de façon indirecte. Enfin, la méthylation de l'adénine augmente l'hydrophobicité de la chaîne nucléotidique et pourrait réguler son affinité pour des domaines protéiques hydrophobes (Figure 6).

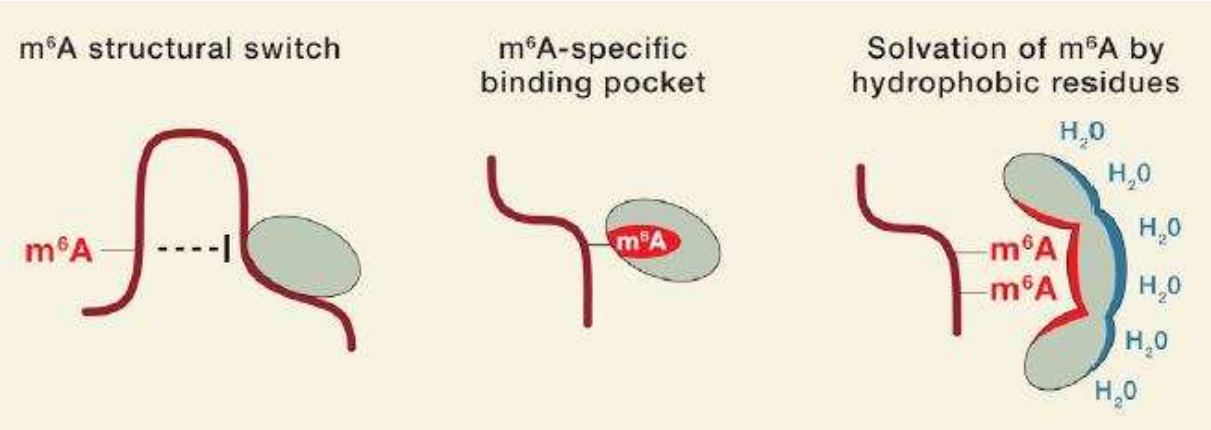


Figure 6 : effet de la méthylation de l'adénosine sur les interactions de l'ARN (Roundtree et al., 2017). *La méthylation de l'adénosine peut (de gauche à droite)* : altérer la structure de *l'ARN et faire disparaître (ou apparaître) des sites de liaison* ; être reconnue spécifiquement par des protéines ; ou favoriser des liaisons hydrophobes.

Des adénosines méthylées ont récemment été identifiées dans l'ARN génomique du VIH-1, au niveau des extrémités 5' et 3' (Kennedy et al., 2016; Tirumuru et al., 2016) ainsi que dans des régions internes comme le RRE (Lichinchi et al., 2016). Ces méthylations sont clairement requises pour assurer l'efficacité de la réplication du virus, mais leur effet exact reste incertain. La méthylation du RRE pourrait augmenter l'affinité de Rev et stimuler l'export nucléaire (Lichinchi et al., 2016), tandis que la méthylation de la région 5' semble inhiber la rétrotranscription (Lu et al., 2018)

3) *Traduction de l'ARN génomique*

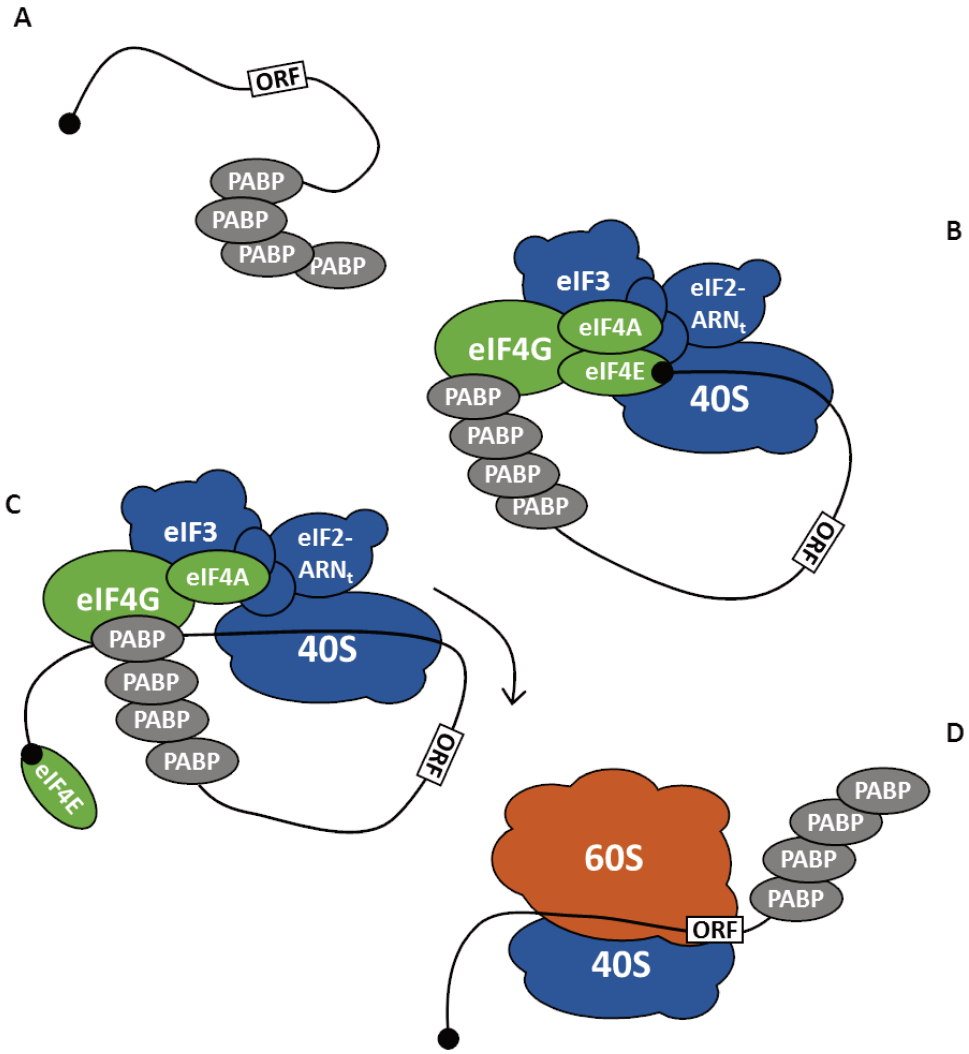
L'initiation de la traduction est un processus moléculaire qui conduit à l'assemblage de ribosomes compétents pour la traduction au niveau de codons initiateurs des ARNm. On distingue deux ensembles de mécanismes d'initiation de la traduction. L'un est dépendant de la coiffe et l'autre se fait par entrée interne du ribosome, indépendamment de la présence de la coiffe (Jackson et al., 2010).

i. Traduction dépendante de la coiffe

Mécanisme général

Dans le premier cas, la petite sous-unité du ribosome est recrutée au niveau de la coiffe par le complexe eIF4F composé de la protéine de liaison à la coiffe eIF4E, de l'hélicase eIF4A et d'une protéine architecturale, eIF4G. Le complexe de pré-initiation 43S, comportant la petite

sous-unité du ribosome, l'ARN_t initiateur et plusieurs facteurs d'initiation, interagit avec eIF4G via le complexe multiprotéique eIF3. L'ARN est circularisé par PABP (PolyA Binding Protein) qui lie la queue polyA et interagit également avec eIF4G. La petite sous-unité balaye alors la région 5' non traduite, assistée par eIF4A qui déstabilise les structures en aval. D'autres hélicases interviennent vraisemblablement pour assister le balayage de la 5'UTR, notamment DHX29 (Dhote et al., 2012). Le complexe s'arrête lorsqu'il rencontre un codon initiateur dans un contexte favorable, et l'assemblage du ribosome se poursuit avec la liaison de la grande sous-unité du ribosome et le recyclage d'une partie des facteurs d'initiation (Jackson et al., 2010) (Figure 7).



Légende page suivante

Figure 7 (page précédente) : L'initiation de la traduction dépendante de la coiffe. *A/ L'ARNm* prêt à être traduit est coiffé et polyadénylé. La queue polyA est tapissée de PABP. B/ Le complexe eIF4F (en vert) lie la coiffe via eIF4E. Le complexe de pré-initiation 43S (en bleu) interagit avec eIF4F via eIF3 et eIF4G et permet le recrutement de la petite sous-unité du *ribosome 40S sur la 5'UTR. C/* Le complexe 48S ainsi formé balaye *la 5'UTR jusqu'à* rencontrer un codon initiateur dans un contexte favorable. D/ La grande sous-unité du ribosome 60S *s'assemble à la petite sous*-unité au niveau du codon initiateur, et une partie des *facteurs d'initiation est recyclée.*

La structure de la région 5' non-traduite de l'ARNg du VIH-1 régule la traduction dépendante de la coiffe

La structure de la région 5' non traduite de l'ARN génomique est fortement structurée. Elle débute notamment par une tige-boucle d'une remarquable stabilité, TAR, dont l'effet sur l'activation de la transcription a été présenté. Concernant la traduction dépendante de la coiffe, on peut supposer que cette structure rende la coiffe inaccessible au complexe eIF4F ou au ribosome. Plusieurs études ont démontré l'effet inhibiteur de TAR, en montrant notamment que l'ajout d'un gène rapporteur en aval de TAR inhibe sa traduction dans des lysats de réticulocytes ainsi que dans des oocytes de xénope (de Breyne and Ohlmann, 2019 pour revue). Par ailleurs, des mutations qui déstabilisent TAR stimulent la traduction (Parkin et al., 1988). Cependant, l'effet inhibiteur de TAR est plus important in vitro que dans des cellules (Soto-Rifo et al., 2012a), ce qui suggère que des facteurs cellulaires stimulent la traduction des transcrits du VIH-1. Parmi ceux-ci, au moins trois hélicases ont été identifiées : La, RHA/DHX9 et DDX3. La est associée aux ARNm du VIH-1 in vivo et interagit directement avec TAR in vitro (Chang et al., 1994). Elle est capable de lever l'inhibition de TAR sur la traduction d'un gène rapporteur dans des RRL (Craig et al., 1997). RHA/DHX9 est une hélicase de la famille DExD/H dont un des rôles semble être de stimuler la traduction d'ARNm cellulaires qui présentent une 5'UTR structurée. Elle s'associe à la région R – U5 de l'ARN du VIH-1 et stimule la traduction de Gag in vivo (Bolinger et al., 2010). Le rôle de DDX3 sur la traduction dépendante de la coiffe sera développé plus loin. L'effet de ces hélicases pourrait être de déstabiliser TAR et d'augmenter l'accessibilité de la coiffe pour favoriser l'initiation dépendante d'eIF4F. D'autres facteurs mis en évidence ne présentent pas d'activité hélicase. Ainsi, TRBP (TAR Binding Protein) est capable de se lier à TAR et stimule la traduction de l'ARN génomique (Dorin et al., 2003). TRBP a au moins deux effets : le premier est d'inhiber la kinase PKR (Protein Kinase R), activée en réponse à la présence d'ARN viral double-brin et qui inhibe le facteur d'initiation eIF2α ; le deuxième est indépendant de PKR et pourrait être de déstabiliser TAR. Enfin Staufen 1 est une protéine de liaison à l'ARN double-brin qui stimule la traduction d'un gène rapporteur

lorsque TAR est placé en 5' de la région 5'UTR (Dugré-Brisson et al., 2005). Staufen 1 interagit avec les deux sous-unités du ribosome et pourrait donc recruter le ribosome et stimuler son positionnement. Un autre rôle envisagé, indirect, serait l'inhibition de PKR par la liaison compétitive à TAR.

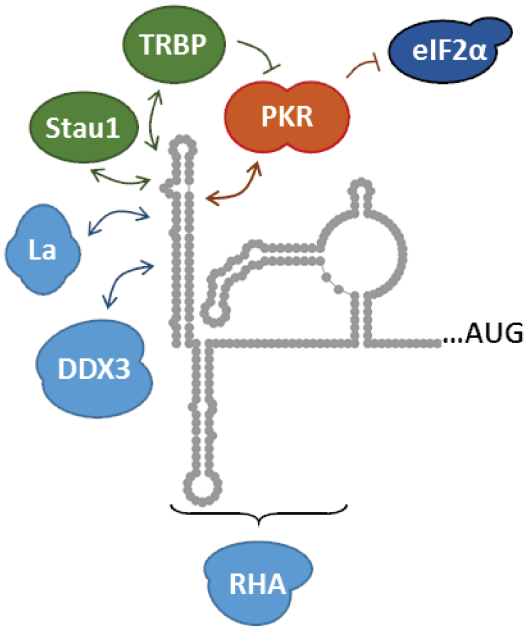


Figure 8 ***:De nombreux facteurs cellulaires régulent l'inhibition de la traduction*** **induite par TAR.** Le début *de la 5'UTR* est indiqué en gris*, les protéines dotées d'une activité hélicase sont* en bleu. Le rôle de ces hélicases est probablement de déstabiliser les structures qui inhibent la traduction dépendante de la coiffe, dont TAR. Les protéines *de liaison à l'ARN Stau*fen 1 (Stau1) *et TRBP recrutent inhibent la kinase PKR et activent indirectement le facteur eIF2α.*

D'autre part, la traduction canonique est inhibée dans les cellules infectées. En effet, la protéine virale Vpr (Viral Protein R) induit l'arrêt du cycle cellulaire en phase G2/M. Au cours de cette phase, la protéine 4E-BP (eIF4E Binding Protein) est phosphorylée et séquestre le facteur eIF4E. Cela se traduit par une baisse importante du niveau de traduction dans les cellules. Cependant, la traduction de l'ARNg continue d'être assurée. De façon remarquable, l'ARN génomique et les ARN viraux partiellement épissés sont associés au complexe nucléaire de liaison à la coiffe CBC. Ce complexe lie la coiffe des transcrits cellulaires dans le noyau, et il permet d'assurer un premier tour de traduction dans le cytoplasme. Il est normalement rapidement recyclé et réimporté dans le noyau. Il se pourrait donc que le VIH contourne l'indisponibilité du facteur eIF4E en ayant empêchant le recyclage du CBC (Sharma et al., 2012). Le CBC est composé de deux protéines, CBP80 et CBP20. La protéine virale Rev et le RRE stimulent la liaison de CBP80 à eIF4AI dans le cytoplasme (Toro-Ascuy et al.). D'autre part CBP80 a une affinité plus grande pour la triméthylguanosine, or certains transcrits non-

épissés du VIH sont effectivement triméthylés (Yedavalli and Jeang, 2010). L'inhibition de la traduction canonique est également induite par la protéase virale qui cible certaines protéines de l'initiation dépendante de la coiffe telle que l'isoforme 1 d'eIF4GI, PABP et eIF3d. L'ajout de PABP exogène à des cellules préalablement traitées par la protéase virale restaure la traduction, ce qui confirme l'effet inhibiteur de cette protéase sur la traduction. La traduction de l'ARN génomique n'est cependant pas affectée par la protéase, ce qui suggère qu'elle se produit suivant un mécanisme non canonique d'initiation par la coiffe, ou bien suivant un mécanisme indépendant de la coiffe (Castelló et al., 2009).

ii. Traduction indépendante de la coiffe

Mécanismes généraux

Dans le deuxième mode d'initiation de la traduction, tout ou partie de la machinerie d'initiation de la traduction est recrutée directement au niveau d'une structure particulière de l'ARNm voire immédiatement au niveau du codon initiateur (Figure 9).

Ce recrutement interne du complexe d'initiation peut résulter d'une structure particulière de l'ARN. On parle alors de site d'entrée interne du ribosome (IRES, Internal Ribosome Entry Site). Ce mécanisme d'initiation est retrouvé fréquemment chez les ARNm viraux, mais il a également été mis en évidence dans des transcrits cellulaires. On distingue quatre types d'IRES. Les IRES de type 1 et 2 requièrent la totalité des facteurs nécessaires à l'initiation dépendante de la coiffe à l'exception d'eIF4E. La liaison à l'ARN est assurée par eIF4G. Ces deux types d'IRES diffèrent entre eux seulement par l'étape de balayage de l'ARN, qui est absente du type 2 où le complexe d'initiation s'assemble directement sur le codon initiateur. Les IRES de type 3 et 4 se caractérisent également par l'absence de balayage ainsi que par la liaison directe de la petite sous-unité du ribosome. L'IRES de type 3 requiert eIF2 qui fournit l'ARN_t initiateur ainsi que les facteurs eIF5 et eIF5B qui stimulent le recrutement de la grande sous-unité du ribosome. Enfin l'IRES de type 4 ne requiert aucun autre facteur que la petite sous-unité du ribosome (Jackson et al., 2010). L'activité des IRES peut être régulée par des facteurs accessoires, les ITAF (IRES Trans-Acting Factors).

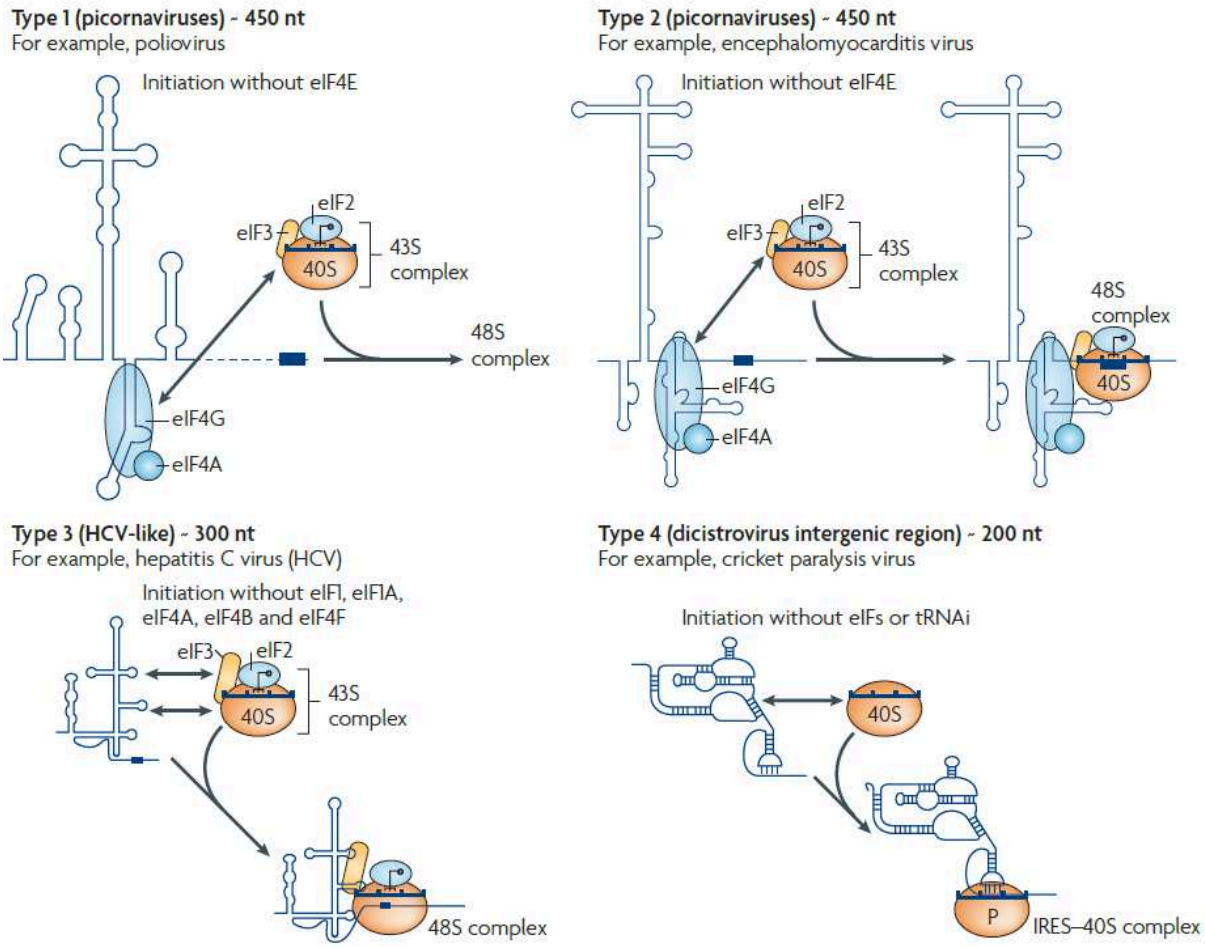


Figure 9 : Les différents types d'IRES (Jackson et al., 2010). Les IRES de type 1 et 2 sont reconnus par le facteur eIF4G tandis que les IRES de type 3 et 4 sont liés directement par la petite sous-*unité du ribosome. Seule l'IRES de type 1 inclut une étape de balayage pour la* reconnaissance du codon initiateur (rectangle bleu).

Enfin, certains ARNm présentant des adénosines méthylées en N6 dans leur région 5' peuvent recruter le ribosome de manière indépendante du complexe eIF4F, et donc indépendante de la coiffe. Ce mécanisme d'engagement du ribosome induit par la N6-méthyladénosine (MIRES, Methyladenosine-Induced Ribosome Enagement Site) rend la traduction de ces ARN dépendante de la méthylase METTL3. Les résidus m6A recruteraient directement le complexe eIF3 pour promouvoir l'initiation de la traduction (Meyer et al., 2015).

Mise en évidence d'IRES dans l'ARN génomique du VIH-1

L'identification d'un IRES peut être effectuée au moyen d'un transcrit bicistronique : la traduction du premier est sous contrôle de la coiffe, tandis que la traduction du deuxième est sous contrôle de l'IRES dont on veut tester l'activité. La comparaison des niveaux d'expressions relatifs des deux gènes rapporteurs permet de quantifier l'activité de l'IRES

(Figure 10). Un premier IRES a été identifié dans la 5'UTR de l'ARN génomique (Brasey et al., 2003). Aucune structure particulière requise pour le fonctionnement de cet IRES n'a pu être identifiée, mais la région minimale requise recouvre les tige-boucles PBS à MSD (**Figure 11**). Cet IRES reste actif après coupure de eIF4G1 et de PABP par une protéase virale in vitro, ce qui révèle que son fonctionnement est indépendant du complexe eIF4F (Monette et al., 2013).

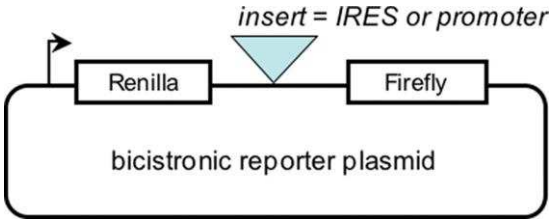


Figure 10 : Le rapporteur bicistronique permet de mettre en évidence une activité IRES dans *une séquence d'ARN donnée*. Après transfection du plasmide, les productions de luciférase *Renilla et Firefly sont comparées pour mesurer le taux relatif d'initiation coiffe dépendante et d'entrée interne du riboso*me. Tiré de (Gilbert, 2010).

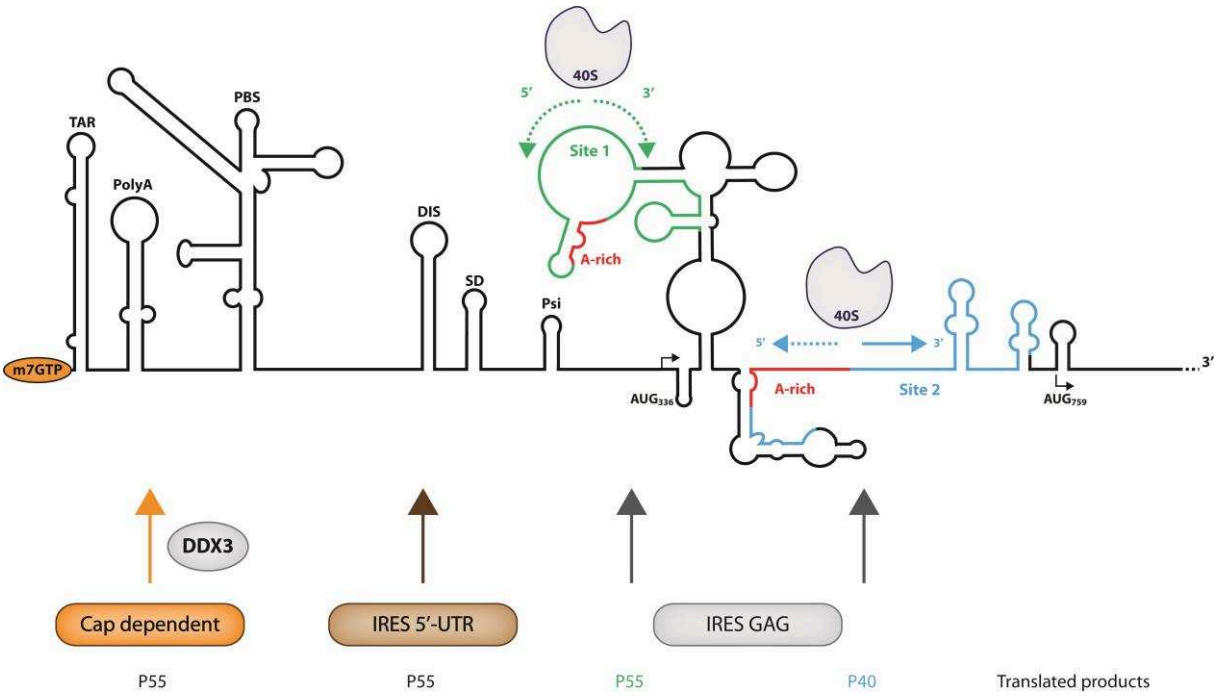
Un autre IRES a été identifié dans la région codante de Gag (Buck et al., 2001). Cet IRES conduit à l'initiation de la traduction au niveau de deux codons, le codon AUG en position 336 et le codon AUG en position 759. Le premier permet la synthèse de l'isoforme de Gag de pleine longueur p55 et le deuxième la synthèse d'une isoforme tronquée, p40. Deux sites de liaison directe à la petite sous-unité du ribosome ont été localisés entre ces deux codons initiateurs. Le mode exact de liaison du ribosome n'est pas connu mais il pourrait être assuré par les adénosines, qui sont particulièrement abondantes dans ces séquences, comme c'est le cas dans d'autres IRES (Abaeva et al., 2016). Le modèle proposé suggère que le ribosome lie l'un ou l'autre de ces deux sites puis balaye l'IRES vers l'extrémité 5' ou 3' jusqu'à rencontrer un codon initiateur (Figure 11). Le rôle des facteurs d'initiation n'est pas clairement établi. Cet IRES est conservé chez le VIH-2 où les complexes d'initiation contiennent eIF2, eIF3, eIF4A, eIF4G et eIF5 en plus de la petite sous-unité du ribosome (Locker et al., 2011; Weill et al., 2010).

Facteurs régulant l'activité de ces IRES

L'activité de l'IRES de la 5'UTR est stimulée au cours de la phase G2/M, lorsque l'initiation dépendante de la coiffe est inhibée. De plus, comme évoqué plus haut, l'ARN génomique n'est que faiblement traduit dans des extraits de réticulocytes. En revanche, l'IRES est actif dans des cellules HeLa et Jurkat (Brasey et al., 2003). Ces observations supposent que des facteurs spécifiques régulent le fonctionnement de l'IRES de la 5'UTR. Trois protéines

cellulaires sont suspectées réguler l'activité de l'IRES : hnRNP A1 (heteronuclear Ribonucleoprotein A1), dont l'expression semble requise pour le fonctionnement de l'IRES ; de même que la protéine ribosomale S25, également impliquée dans l'IRES du virus de l'hépatite C ; et HuR dont le rôle semble être d'inhiber l'IRES. Le rôle exact de ces facteurs n'est cependant pas connu (de Breyne and Ohlmann, 2019). L'activité de l'IRES de la 5'UTR pourrait ainsi également dépendre du type cellulaire. En effet, cet IRES est plus actif dans des cellules représentatives des lymphocytes T comme les cellules de Jurkat que dans des cellules HeLa issues d'un cancer du col de l'utérus (Plank et al., 2013). Enfin, l'état physiologique de la cellule peut influer sur l'activité de l'IRES de la 5'UTR. Cet IRES est en effet plus actif dans des cellules soumises à un stress oxydatif similaire au stress provoqué lors de l'infection (Gendron et al., 2011).

Différents mécanismes permettent donc d'initier la traduction de l'ARN génomique dans un contexte cellulaire globalement défavorable, bien que la contribution exacte de ces différents modes d'initiation ne soit pas connue. Le virus détourne de nombreux facteurs cellulaires pour contourner l'inhibition cellulaire. De plus, certains auteurs proposent que la régulation de la traduction de l'ARNg repose sur la conformation adoptée par ce dernier. En effet, l'ARNg peut former des dimères : seule les monomères seraient efficacement traduits, tandis que les dimères serait sélectionnés pour l'encapsidation (Lu et al., 2011a).



Légende page suivante

Figure 11 (page précédente) ***: Différents mécanismes d'initiation de la traduction de l'ARNg*** (Deforges et al., 2017). *L'initiation dépendante de la coiffe est stimulée par différents facteurs cellulaire, en particulier DDX3. L'IRES situé dans le cadre de lecture de Gag présente deux sites de liaison au ribosome et permet l'*initiation au niveau de deux codons AUG, en positions 336 et 759.

4) *Dimérisation de l'ARN génomique du VIH*-1

L'ARN génomique du VIH peut former des dimères d'une remarquable stabilité (Song et al., 2007). Cette propriété est commune aux rétrovirus. Ce phénomène est d'une importance critique pour le virus (Dubois et al., 2018a) : il est indispensable à l'encapsidation du génome ; il confère une résistance accrue aux dommages de l'ARN lors de la transcription inverse ; et il augmente la diversité génétique des particules virales. Enfin, la dimérisation pourrait jouer un rôle dans la régulation de la traduction de l'ARN génomique et de ce fait, dans la régulation de la production des constituants des particules virales.

i. Mise en évidence des dimères d'ARN génomique

La première observation de la forme dimère de l'ARN génomique dans une particule rétrovirale a été faite chez le virus du sarcome de Rous. Les auteurs ont constaté que l'ARN génomique issu des particules virales présentait deux fois la densité attendue, et que cette propriété disparaissait après dénaturation thermique de l'ARN (Duesberg, 1968; Robinson et al., 1967). Les mêmes observations ont été faites chez le VIH lors de sa découverte, suggérant qu'il s'agit d'une propriété conservée chez les rétrovirus. Par ailleurs, ces observations ont été confirmées par microscopie électronique. Les images obtenues révèlent des dimères d'ARN génomique du VIH-1 présentant deux points de contacts dans la région 5' (Höglund et al., 1997) (Figure 12).

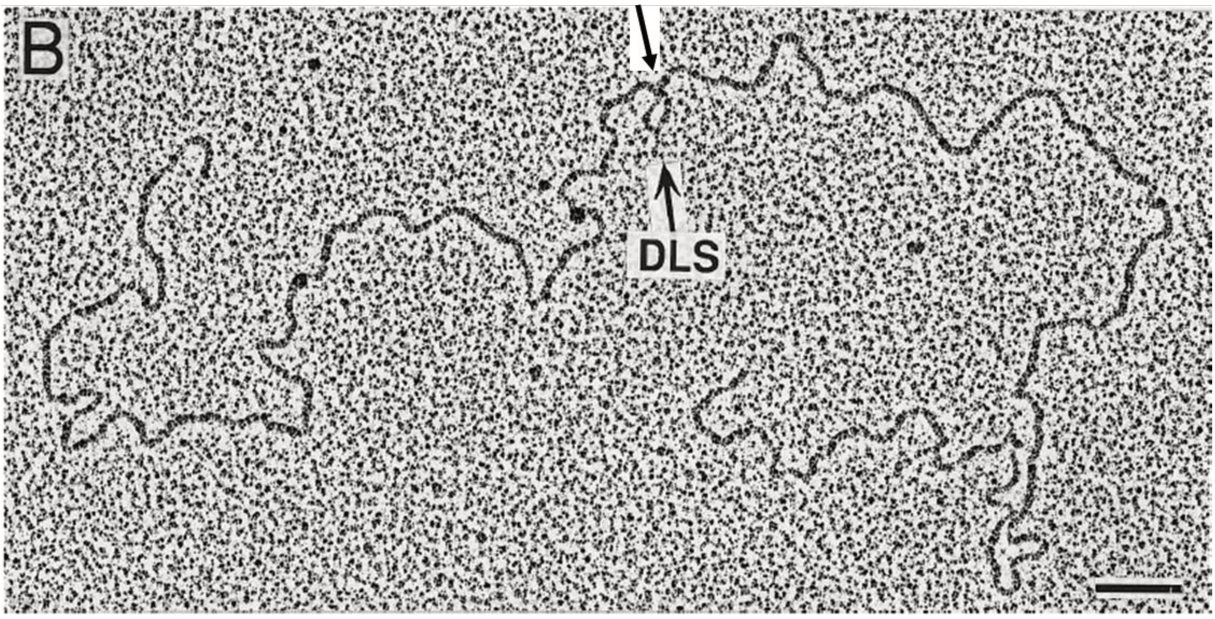


Figure 12 : observation d'un dimère d'ARN génomique du VIH-1 par microscopie électronique (Höglund et al, 1997). *L'ARN est préalablement extrait de particules virales* puis partiellement dénaturé (2,5 M urée, 50% formamide). Les flèches indiquent deux points de contact du dimère. DLS : Dimer Linkage Structure. La barre située au coin inférieur droit représente 100 nm.

Les dimères issus de particules virales se dissocient sous l'effet de la chaleur, ce qui suggère qu'ils sont formés de liaisons hydrogène ou d'autres liaisons faibles. De plus, ils subissent un processus de maturation dépendant de la protéase virale (Fu et al., 1994). En effet, l'ARN génomique issu de particules où la protéase est inactivée (PR-) présente une mobilité électrophorétique plus faible que l'ARN génomique issu de particules sauvages. Par ailleurs, les dimères PR- se dissocient à plus faible température. Ces résultats suggèrent que les dimères d'ARN génomique changent de conformation et sont stabilisés au cours de la maturation des particules virales. La nature de ces changements conformationnels reste à définir.

ii. Le mécanisme de dimérisation de l'ARN génomique du VIH-1

La dimérisation débute par la formation d'un complexe DIS – DIS : le modèle loop – loop kissing complex

De nombreuses études ont progressivement permis d'identifier la structure impliquée dans la liaison des dimères d'ARN génomique (DLS, Dimer Linkage Structure). Cette région a initialement été localisée en aval de la tige-boucle SD, car des transcrits contenant la centaine de nucléotides située en aval de SD ont été observés sous forme de dimères (Darlix et al., 1990; Marquet et al., 1991). Cependant, d'autres études ont rectifié la localisation du site de

dimérisation en amont de SD et en aval du PBS (Laughrea and Jetté, 1994; Marquet et al., 1994; Skripkin et al., 1994). Cette région, susceptible de former une tige-boucle, a donc été nommée DIS, Dimerization Initiation Site. La boucle apicale de DIS présente une séquence palindromique de six nucléotides permettant à deux ARN de s'assembler de façon autocomplémentaire, formant ainsi un complexe kissing. En effet, des mutations qui affectent le palindrome inhibent la dimérisation (Paillart et al., 1994). De plus, le recours à un oligonucléotide complémentaire de cette séquence inhibe la dimérisation (Muriaux et al., 1995). Par ailleurs, la boucle apicale de DIS comprend aussi trois purines qui bordent le palindrome, deux en 5' (AA/G selon les sous-types) et une en 3' (A). La mutation ou la délétion de ces purines inhibe la dimérisation (Clever et al., 1996; Paillart et al., 1997). Enfin, on ne trouve essentiellement que trois séquences auto-complémentaires dans DIS parmi les différents isotypes du VIH-1, CGCGCG, GUGCAC et GUGCGC sur les 64 séquences possibles (Paillart et al., 2004a). La formation du complexe kissing DIS – DIS ne repose donc pas sur la seule complémentarité des séquences, mais elle implique vraisemblablement une structure particulière de DIS.

Transition du complexe kissing de DIS au dimère étendu

In vitro, on peut distinguer par électrophorèse native des dimères stables et des dimères instables. Les dimères instables se dissocient lors d'une électrophorèse en condition semi-native, c'est-à-dire en absence de magnésium et à 27°C. Ils se forment spontanément, tandis qu'un traitement à 60°C est nécessaire pour obtenir des dimères stables (Laughrea and Jetté, 1996). La stabilisation des dimères pourrait correspondre au passage d'un complexe kissing à un dimère étendu mettant en jeu des liaisons entre les nucléotides de la tige de DIS (**Figure 13**).

Plusieurs études structurales ont été menées autour d'une construction de 23 ou 35 nucléotides correspondant à la séquence de DIS. On observe en RMN ainsi qu'en cristallographie des structures correspondant au complexe kissing de DIS (Ennifar et al., 2001; Kieken et al., 2006; Mujeeb et al., 1998), mais aussi à un duplex étendu dans lequel les nucléotides forment exclusivement des appariements inter-brin (Ennifar et al., 1999; Mujeeb et al., 1999; Ulyanov et al., 2006). Deux modèles ont initialement été envisagés pour expliquer la transition d'un complexe kissing à un duplex étendu. Le premier exclut la fusion de la tige-boucle DIS et propose une réaction de trans-estérification semblable à une recombinaison des deux brins d'ARN (Ennifar et al., 2001). Le deuxième propose au contraire que l'appariement entre les séquences palindromiques soit maintenu pendant que les appariements

intramoléculaires se défont puis se reforment de façon intermoléculaire (Skripkin et al., 1994). Ce denier mécanisme a été démontré par des expériences de FRET en molécule unique à partir d'une construction de DIS de 23 nucléotides (Mundigala et al., 2014).

L'existence d'un dimère étendu de DIS est toutefois discutée. En effet, des mutations qui déstabilisent la tige de DIS mais qui n'empêchent pas la formation du complexe kissing n'entraînent pas de modification de la stabilité thermique des dimères (Paillart et al., 1996).

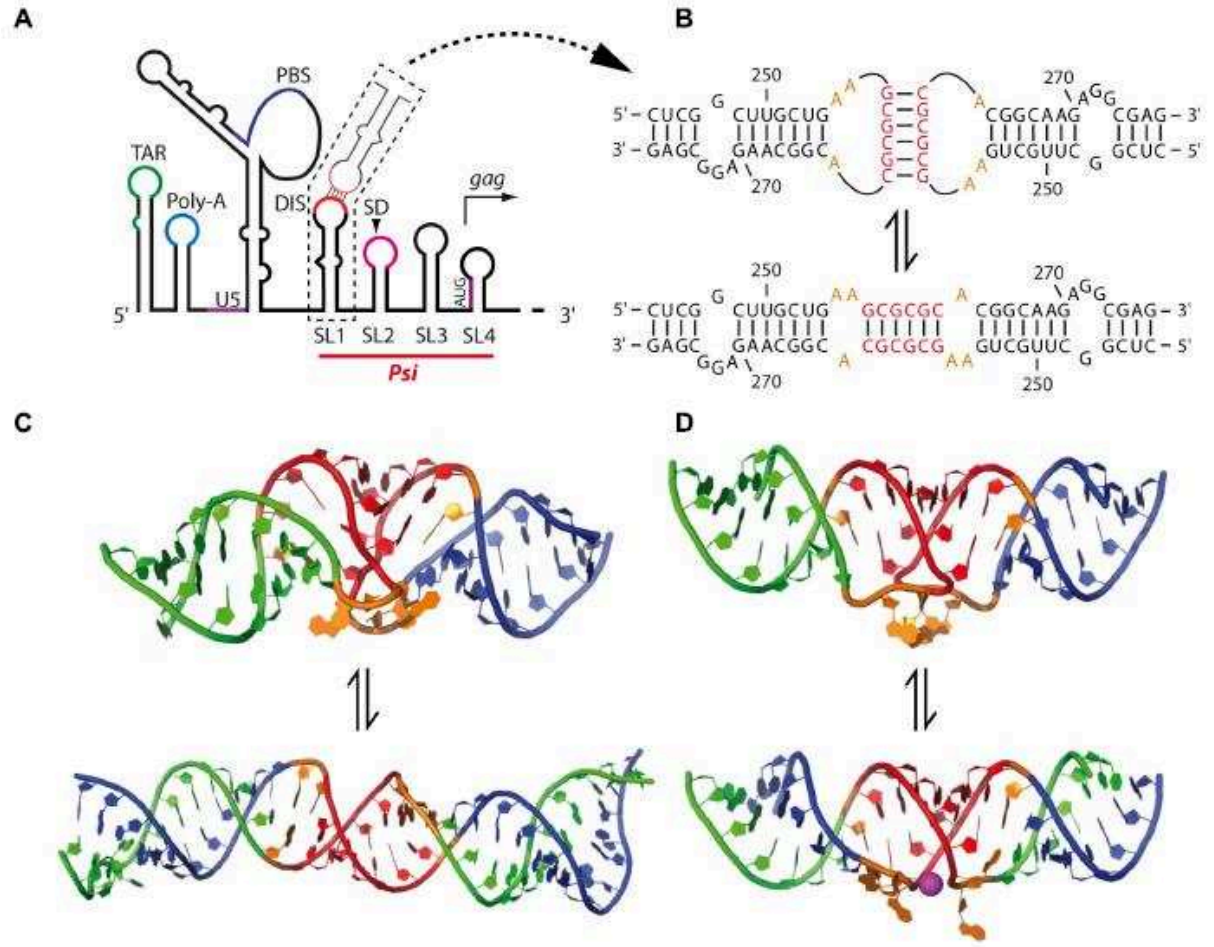


Figure 13 : Modèle de dimérisation médiée par DIS (Dubois et al., 2018a). A/ Structure *secondaire de la 5'UTR. B/ Transition du complexe kissing de DIS au dimère* étendu. C/ Structures en solution du dimère kissing (23 nucléotide) (haut) et du dimère étendu (35nt) (bas) *de l'isolat Lai (sous*-type B) obtenues par RMN D/ Structures cristallographiques (23 nt) du dimère kissing *de l'isolat Lai* (haut) et du dimère étendu *de l'isolat MAL* (bas). Les molécules sont colorées en vert ou en bleu. Le palindrome est en rouge et les purines flanquantes sont en orange. Un des atomes de magnésium observé dans la structure du dimère étendu en cristallographie est représenté en violet.

Les différents modèles de structure monomère et dimère

Si la structure des 23 ou 35 nucléotides de DIS a pu être déterminée par RMN et par cristallographie, la structure de la région de l'ARN génomique impliquée dans la dimérisation est beaucoup plus incertaine. En effet, les fragments plus grands sont difficilement analysables en RMN, et la présence de segments d'ARN flexibles est incompatible avec l'obtention de cristaux. Le recours à des sondes chimiques pour mesurer la flexibilité des nucléotides au sein des structures d'ARN ne permet pas quant à lui de distinguer les liaisons intramoléculaires des liaisons intermoléculaires.

D'importants travaux d'analyse phylogénétique et biochimique ont cherché à établir un modèle de structure de l'ARN monomère et dimère et de proposer un mécanisme de transition entre ces deux structures (Dubois et al., 2018a). L'ARN génomique est destiné à la dimérisation et à l'assemblage des particules virales, mais aussi à la traduction des polyprotéines Gag et Gag-Pol (cet aspect sera développé plus loin). Ces modèles visent donc également à expliquer comment la structure de l'ARN génomique gouverne son devenir dans la cellule.

Un premier modèle propose un mécanisme d'aiguillage moléculaire (molecular switch) permettant d'orienter l'ARN génomique vers la traduction ou vers la dimérisation et l'empaquetage. Dans ce modèle, l'ARN génomique peut adopter une conformation en multiples tige-boucles connectées (BMH, Branched Multiple Hairpins) où la séquence palindromique de DIS est exposée et susceptible de former un complexe kissing avec une autre molécule de façon à initier la dimérisation et permettre l'encapsidation (Figure 14 C). Le codon initiateur de Gag est masqué dans un appariement avec U5, supposé défavorable à la traduction. Cette conformation peut changer pour une conformation avec des interactions à longue distance (LDI, Long-Distance Interactions) où le palindrome de DIS est rendu masqué dans une tige-boucle stable (Figure 14 A). Dans cette conformation, le palindrome de DIS est donc masqué et inapte à former un complexe kissing. A l'inverse, le codon initiateur de Gag est présent dans une boucle interne et supposé plus accessible à la machinerie de traduction. La structure BMH serait donc propice à la dimérisation et la reconnaissance par Gag, tandis que la structure LDI serait favorable à la traduction (Huthoff and Berkhout, 2002). Toutefois, une série d'arguments remet en question l'existence de la conformation LDI (Moore and Hu, 2009). Tout d'abord il s'avère que les deux conformations sont aussi efficacement traduites in vitro (Abbink et al., 2005). D'autre part, des résultats de probing de la structure de l'ARNg dans des particules virales et dans des cellules semblent contredire l'existence de la conformation LDI in vivo (Paillart et al., 2004b).

Un autre modèle de dimère étendu a été proposé par le groupe de M. Summers à partir de données de RMN (Lu et al., 2011b). La forme monomère masquerait le palindrome de DIS par une interaction intramoléculaire DIS – U5 (Figure 14 B). La dimérisation débute par la formation du complexe kissing DIS – DIS et de l'appariement U5 – AUG (Figure 14 C). Ce dimère initial est ensuite stabilisé par la formation d'appariements intermoléculaires entre DIS, SD et U5-AUG (Figure 14 D) (Keane et al., 2015). Ce modèle suppose donc la dissociation et la reformation de dizaines de paires de bases, ce qui représente une contrainte topologique très forte (Dubois et al., 2018a). De plus, seule une partie de la 5'UTR a été analysée pour décrire la forme étendue du dimère : les tiges-boucle TAR et polyA ainsi que la boucle du PBS ont été omises, et le palindrome GCGCGC de DIS a été remplacé par la tétrade GAGA de façon à rendre le signal de RMN analysable. La validité de ce modèle in vivo reste donc très incertaine.

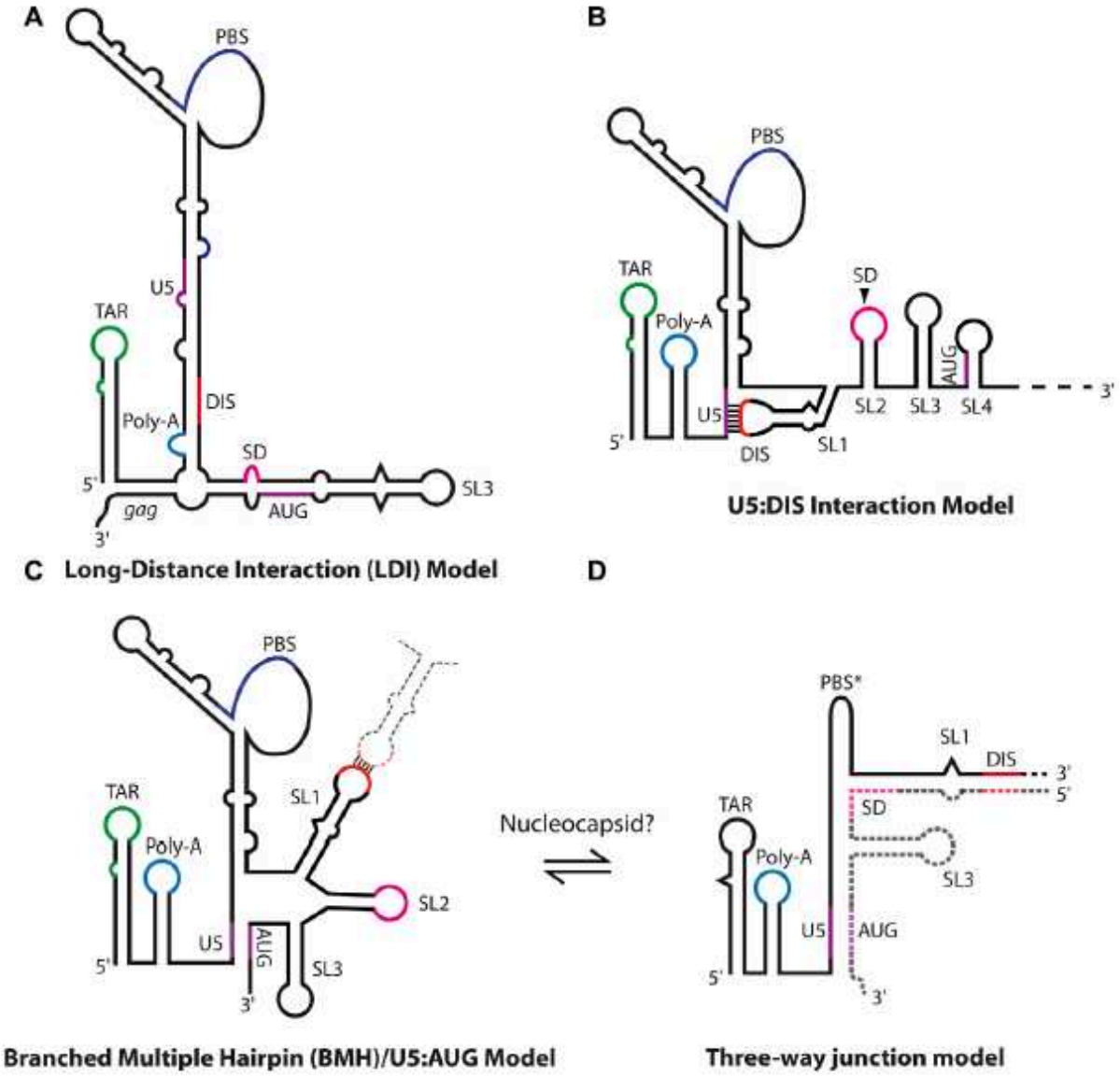


Figure 14 : ***Différents modèles de structure secondaire de la 5'UTR monomère / dimère*** *(Dubois et al., 2018a)*. Les tige-boucles SL1, 2, 3 et 4 sont aussi nommées DIS, (M)SD, Psi et AUG. *La nucléocapside pourrait stimuler la conversion d'un dimère kissing (C) en un dimère étendu (D). Le modèle présenté en D a été obtenu par RMN à partir d'une construction dépourvue d'une partie du PBS.*

Autres sites de dimérisation possibles

Il est possible que DIS ne soit pas indispensable à la dimérisation et qu'il existe d'autres sites capables d'initier la formation de dimères. En effet, il est possible que des molécules d'ARN génomique privées de DIS puissent former des dimères dans les particules virales (Song et al., 2007). Les images de dimères d'ARN viral obtenues en microscopie électronique (Höglund et al., 1997) et en microscopie à force atomique (Andersen et al., 2004; Pallesen,

2011) révèlent deux points de contact entre les molécules d'ARN. Si l'interaction DIS – DIS est vraisemblablement le point de contact le plus en 3', les interprétations diffèrent quant au point de contact en 5'. Les tige-boucles TAR et PolyA ayant chacune un palindrome au niveau apical, il pourrait d'agir d'une interaction TAR – TAR (Andersen et al., 2004; Pallesen, 2011) ou PolyA – PolyA (Höglund et al., 1997). L'hypothèse d'une contribution de TAR dans la formation de dimère est également soutenue par des expériences montrant que la délétion de TAR affecte la dimérisation in virio (Song et al., 2008), par l'utilisation d'oligonucléotides anti-sens (Reyes-Darias et al., 2012) et par des expériences de mutagenèse dirigée (Jalalirad et al., 2012).

L'hypothèse de dimères stabilisés par des quadruplex de G intermoléculaires a également été proposée. Ainsi, des auteurs ont observé la dimérisation du fragment 311 – 612 (numérotation de l'isolat MAL) et avancé l'hypothèse de fragments assemblés par des quadruplex de G (Marquet et al., 1991). Par ailleurs, des expériences de délétions successives ont permis d'identifier in vitro un fragment de la région codante de Gag (732 – 825, numérotation du clone NL4-3) capable de former de dimères spontanément (Awang and Sen, 1993; Sundquist and Heaphy, 1993). Ces dimères sont dépendants de la présence de potassium et sont moins stables en présence d'autres cations (lithium, sodium, césium). Cette dépendance au potassium caractérise les quadruplex de G. Néanmoins l'influence de la concentration sur la dimérisation de ce fragment n'a pas été étudiée, et la forte concentration d'ARN employée dans les tests (1,5 µM) n'argue pas en faveur de l'existence de ces dimères in vivo. D'autre part, une autre étude rapporte que la température de dissociation des dimères issus de particules virales n'est pas affectée par la nature des cations présents, ce qui semble contredire l'existence d'un quadruplex de G dans les dimères natifs (Fu et al., 1994). Cependant, il est possible que la nature des cations n'affecte pas de la même manière la formation de quadruplex de G et leur dissociation. En conclusion, l'existence de quadruplex de G intermoléculaire impliqués dans la dimérisation reste incertaine.

Facteurs influençant la dimérisation

Au cours de la maturation des particules, la protéase virale clive les polyprotéines Gag et Gag-Pol. Ceci s'accompagne d'une modification de la structure de la nucléocapside ainsi que d'un changement de conformation du dimère d'ARN génomique. En effet, il est possible de bloquer la maturation des particules en mutant la protéase virale. Dans les particules formées par de tels mutants, l'ARNg est plus fréquemment présent sous forme monomère, et les dimères

restants sont moins résistants à la dénaturation thermique (Fu et al., 1994). Cette observation suppose qu'il existe au moins deux conformations de dimères, les dimères néo-formés, instables et les dimères matures. Une étude a plus récemment proposé un mécanisme de maturation en trois temps : la dimérisation débute par la formation du complexe kissing, ensuite stabilisé en un dimère étendu DIS – DIS et ultimement par d'autres points de contacts (Song et al., 2007). Un tel mécanisme suppose l'existence de facteurs qui régulent temporellement la dimérisation.

Le principal facteur de maturation des dimères connu est la nucléocapside (NC). Ce domaine de Gag apparaît après des clivages successifs de Gag sous la forme de NCp15 (NC de 15kD) puis NCp9 et enfin NCp7. In vitro, l'incubation de dimères instables avec NCp7 permet en effet d'obtenir des dimères stables. Cette maturation requiert la présence de DIS, ce qui suggère qu'elle transforme des complexes kissing en dimères étendus (Muriaux et al., 1996a). Cette hypothèse a été confortée par une approche par fluorescence (Rist and Marino, 2002). De plus, in vivo, l'introduction d'une mutation dans le site de clivage de Gag qui assure la production de NC empêche la formation de dimères stables dans les virions (Shehu-Xhilaga et al., 2001). NCp7 lie les régions simple brin de DIS, à savoir la boucle interne 272-GG-273 (numérotation NL4-3) et les nucléotides non-appariés de la boucle apicale, grâce à des résidus chargés positivement en N-terminal et à deux motifs en doigt de zinc. Ces structures permettent de déstabiliser l'hélice d'ARN et de promouvoir l'échange intermoléculaire des brins des hélices (Aduri et al., 2013).

L'influence des ions magnésium a été mise en évidence par plusieurs études. Des expériences de RMN ont mis en évidence un site de liaison au magnésium au niveau de la boucle interne de DIS 272-GG-273. Le magnésium semble stabiliser les dimères kissing et rendre la conversion en dimère étendus dépendante de la nucléocapside (Sun et al., 2007). Un modèle similaire a été proposé par une approche de fluorescence (Rist and Marino, 2002). Des expériences de FRET ont montré que le résidu conservé A255 en 5' de la boucle apicale de DIS participe également à la liaison du magnésium. Là aussi, le magnésium semble stabiliser les dimères kissing (Mundigala et al., 2014). Ces résultats suggèrent que plusieurs sites de liaison au magnésium participent à la stabilisation des dimères kissing. Enfin, une étude basée sur l'utilisation de pinces magnétiques a qu'en présence de magnésium, les dimères kissing de DIS formés par les sous-types A et B sont de stabilité équivalente (Li and Tinoco, 2009). Pourtant, la nature de la séquence palindromique prédit un complexe plus stable chez le sous-type B, où la séquence est 255-AAGCGCGCA-263 et qui présente donc 6 paires G-C contre 4 chez le sous-type A, où la séquence de la boucle apicale est 272-AGGUGCACA-280. Ces observations

suggèrent que la contribution du magnésium est différente chez les sous-types A et B. De fait, chez le sous-type A, un site de liaison au magnésium supplémentaire a été proposé par modélisation in silico au niveau des résidus 273-GGUG-278 (Ennifar et al., 2001; Jossinet et al., 1999).

La protéine virale Vif pourrait également jouer un rôle dans la formation des dimères, mais cette hypothèse reste incertaine. En effet, Vif présente les caractéristiques d'une protéine chaperone : elle peut lier l'ARN indépendamment de l'ATP et cette interaction entraîne une structuration de la protéine (Batisse et al., 2012). De plus elle stimule la formation de dimères instables in vitro, mais elle inhibe la maturation des dimères induite par NCp7 (Henriet et al., 2007). Elle pourrait ainsi retarder la formation des dimères matures et réguler l'encapsidation.

Enfin le pH pourrait jouer un rôle sur la dimérisation. Des expériences pionnières sur la dimérisation in vitro ont d'abord montré que la dimérisation du fragment 1-707 de l'isolat MAL (sous-type A) est inhibée aux pH acides (Marquet et al., 1991). A l'inverse, des expériences de RMN ont montré que la protonation de l'adénosine 272 entraîne une déstabilisation locale de la structure et favorise la conversion du dimère kissing au dimère étendu par la NCp7 (Mihailescu and Marino, 2004). Ces résultats ont été obtenus avec une tige-boucle DIS de 23 nucléotides correspondant à l'isolat MAL. En absence de NCp7, Le pH n'a pas d'effet sur la dimérisation de la 5'UTR du clone NL4-3 (sous-type B) (Tran et al., 2015). La pertinence biologique de ces observations tient à un rapport montrant que le pH cytoplasmique diminue lors de l'infection (Makutonina et al., 1996). Cette acidification pourrait donc stimuler - ou inhiber - la dimérisation. L'effet du pH sur la dimérisation pourrait également être étudié au regard de l'acidification des endosomes, premier environnement cellulaire rencontré par les particules virales lors de l'infection. Cependant, il n'est pas certain que l'acidification des endosomes ait un effet sur l'infectivité du VIH (Kubo et al., 2012).

iii. Où se produit la dimérisation ?

Le compartiment cellulaire où s'effectue la dimérisation n'est pas certain. Il est possible que la dimérisation se produise dans le cytoplasme. En effet, la co-expression de deux provirus différents entraîne la formation équivalentes d'homodimères et d'hétérodimères dans les particules virales (Rhodes et al., 2003). Si la dimérisation se produisait dans le noyau, on s'attendrait à un plus fort taux d'homodimères qui seraient susceptibles de se former co-transcriptionnellement ou rapidement après la transcription. Au contraire, une autre étude soutient que la dimérisation se produit dans le noyau. Elle repose sur l'utilisation de provirus

génétiquement modifiés pour favoriser l'export via la voie CRM1 ou NXF1. Les auteurs observent davantage de particules formées de dimères d'ARN exportés suivant la même voie, ce qui suggère que l'ARN dimérise avant son export (revu dans Lu et al., 2011c). Néanmoins il est possible que la voie d'export nucléaire conditionne de façon non négligeable le devenir de l'ARN et rende impossible la formation de dimères exportés suivant des voies différentes dans le cytoplasme. Enfin d'autres études montrent que la dimérisation des virus de la leucémie murine et du sarcome de Rous se produit dans le noyau (Lu et al., 2011c), une propriété qui pourrait donc être conservée chez les rétrovirus. Pour conclure, une récente étude a montré par microscopie de fluorescence que la dimérisation se produit au niveau de la membrane cytoplasmique, où elle est assistée par Gag qui assure l'encapsidation (Chen et al., 2016).

5) *Encapsidation de l'ARN génomique*

L'ARN génomique est sélectionné par la protéine Gag qui assure son encapsidation. Il est généralement admis que seules quelques protéines Gag interagissent avec l'ARN génomique et l'oriente à proximité de la membrane plasmique où d'autres protéines Gag s'ajoutent pour former la capside (Lu et al., 2011c). Etant donné que les virions sont composés d'un dimère d'ARN génomique, et que le domaine NC de Gag stimule la formation de dimères, il est clair que dimérisation et encapsidation son liés. En revanche, on ne sait pas exactement comment sont spécifiquement sélectionnées deux molécules d'ARN génomique parmi tous les autres transcrits viraux, et même parmi tous les transcrits cellulaires (Comas-Garcia et al., 2016).

i. Eléments structuraux nécessaires à l'encapsidation

De nombreuses études ont cherché à identifier la région de l'ARN génomique responsable de l'encapsidation. Initialement, la tige-boucle Psi (SL3) a été considérée comme le principal site de liaison à Gag (Lever et al., 1989). Il semble cependant que la sélection de l'ARN génomique requière la quasi-totalité de la 5'UTR et peut-être aussi en partie la région codante de Gag. De plus, à la différence d'autres rétrovirus, le signal d'encapsidation du VIH-1 est vraisemblablement discontinu (Lu et al., 2011c). Un modèle récemment proposé envisage que la liaison de ce domaine par des molécules de Gag entraîne la structuration du connecteur SP1 bordant le domaine NC de la protéine. Ce changement de conformation se propage au domaine CA qui peut alors déclencher l'oligomérisation (Datta et al., 2016).

Gag est composée de quatre domaines : MA interagit avec la membrane et avec l'ARN, CA assure l'oligomérisation de Gag, NC interagit avec l'ARN et P6 intervient dans le bourgeonnement de la particule virale. Les domaines MA et NC présentent tous deux de nombreuses charges positives et sont susceptibles d'interagir de façon aspécifique avec les acides nucléiques. De fait, des particules virales peuvent se former en absence d'ARNg encapsident alors des ARN cellulaires. In vitro, Berkowitz et al. ont observé qu'un ARN quelconque pouvait entrer en compétition avec l'ARN génomique pour la liaison à Gag (Berkowitz et al., 1993). Ces observations suggèrent que dans la cellule, la spécificité de l'interaction Gag – ARNg n'est pas uniquement assurée par l'affinité de Gag pour le signal d'encapsidation. Comas-Garcia et al. ont analysé les ARN cellulaires encapsidés en absence d'ARNg afin de mettre en évidence des propriétés particulières requises pour l'encapsidation. Ces ARN présentent une 3'UTR particulière longue, à l'instar des ARN rétroviraux. La présence d'une longue région d'ARN dépourvue de ribosomes pourrait donc être une condition nécessaire à d'encapsidation (Comas-Garcia et al., 2016). Plus récemment, Abd El-Wahab et al. ont montré, au moyen d'une forme recombinante de Gag de pleine longueur et de différentes construction d'ARN, que Gag interagit spécifiquement avec DIS et que la boucle interne est critique dans cette interaction. Les régions situées en aval de l'AUG stimulent l'interaction, tandis que celles situées en amont de DIS inhibent l'interaction (Abd El-Wahab et al., 2014). Des auteurs ont comparé les interactions Gag – Psi et Gag – TAR-PolyA. La première est de nature moins électrostatique que la deuxième. Seul le domaine NC intervient dans la liaison à Psi, tandis que les complexes non spécifiques Gag – TAR-PolyA font intervenir les domaines NC et MA (Webb et al., 2013).

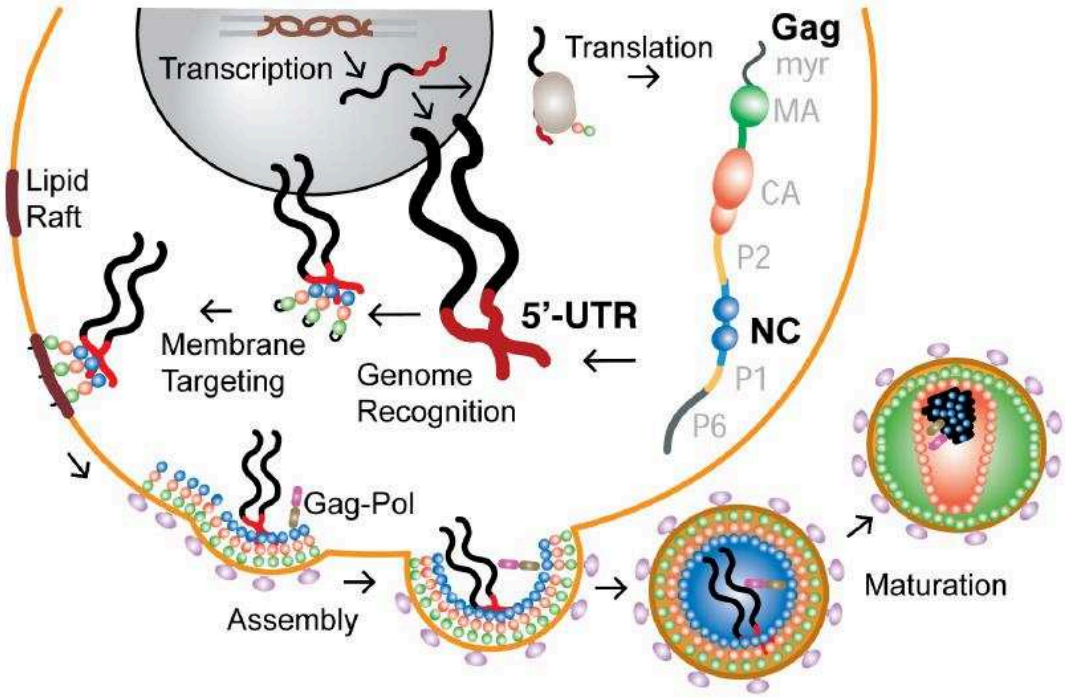


Figure 15 ***: Sélection et encapsidation de l'ARNg par Gag*** (Lu et al., 2011c). La traduction de *l'ARNg conduit majoritairement à la polyprotéine Gag, mais également à la polyprotéine Gag-Pol par un phénomène de changement du cadre de lecture. L'ancrage à la* membrane de ces *polyprotéines est assuré par l'ajout d'un acide myristique au domaine MA. La protéine virale* Env est représentée en violet.

ii. L'ARN génomique est-il sélectionné par Gag sous forme de dimère ?

Plusieurs observations permettent d'avancer que les molécules d'ARN génomique sont sélectionnées sous forme de dimère. En effet, la réplication du DLS au sein d'un même ARN génomique conduit à l'empaquetage d'une seule molécule (Sakuragi et al., 2001). Cela suggère que deux DLS dans un même ARN peuvent s'apparier et former un pseudo-dimère, et que la structure du dimère est spécifiquement reconnue par Gag. D'autre part, si on infecte des cellules avec deux souches virales dont les palindromes de DIS diffèrent, on constate que les virus formés par ces cellules contiennent principalement des dimères provenant d'une seule de ces deux souches (Chin et al., 2005), ce qui laisse penser que l'assemblage des copies d'ARN précède la sélection par Gag.

Au contraire, une étude a montré que si on purifie des particules virales incapables de maturation et ce rapidement après leur bourgeonnement, l'ARN génomique y est présent sous forme de monomère exclusivement (Song et al., 2007). Cela suggère que l'ARN génomique est sélectionné sous forme de monomère et que les dimères se forment progressivement dans les particules au cours de la maturation. Plusieurs études similaires soutiennent cette hypothèse ; il

est cependant difficile d'affirmer que les monomères observés ne résultent pas de la dissociation de dimères instables lors de la purification des particules (revu dans Moore et Hu, 2009).

Ces résultats contradictoires peuvent être éclairés par les observations suivantes. En effet, si on diminue la capacité des virus à former des dimères d'ARN génomique en introduisant des mutations dans DIS, des mutations finissent par apparaître dans Gag et non dans DIS. Ces mutations améliorent la sélection spécifique des transcrits non épissés (Moore and Hu, 2009). En d'autres termes, au cours de l'encapsidation, la dimérisation n'est pas requise en elle-même mais elle est une étape de « contrôle-qualité » contribuant à l'encapsidation spécifique des transcrits non épissés. De façon tout à fait concordante, Abd El-Wahab et al. ont montré que des mutations dans la boucle apicale de DIS qui affectent la dimérisation ont nettement moins d'effet sur la liaison à GagPr55 que la délétion des résidus situés en 3' du site donneur d'épissage (Abd El-Wahab et al., 2014). Comme par ailleurs la présence de Gag est nécessaire à la formation de dimères dans des cellules (Chen et al., 2016), il est vraisemblable que la dimérisation et la sélection par Gag soient deux événements difficilement dissociables dans le temps. Néanmoins, le mécanisme de sélection de l'ARN génomique est encore très incertain.

6) Conclusion

En conclusion, l'ARN génomique du VIH-1 assure de nombreux rôles au cours du cycle de réplication : synthèse de l'ADN proviral, production des transcrits épissés, synthèses des protéines Gag et Gag-Pol et formation des nouveaux virions. Ces fonctions sont régulées spatialement et temporellement par des éléments de structure de l'ARN qui interagissent avec des facteurs cellulaires et viraux. Il est intéressant de remarquer que, si le génome du VIH code des protéines qui ont une activité de chaperonne comme NC et Vif, il ne code aucune hélicase. Ces enzymes sont pourtant fréquemment retrouvées dans les génomes de virus, notamment de virus à ARN de polarité positive (Kadaré and Haenni, 1997). On peut donc facilement supposer que la réplication du VIH implique le détournement de fonctions d'hélicases cellulaire.

DDX3, une hélicase cellulaire impliquée dans le cycle réplicatif du VIH-1

Il existe deux paralogues de la protéine DDX3 : DDX3X et DDX3Y, dont les gènes sont situés sur les chromosomes X et Y, respectivement. DDX3X est retrouvée exprimée dans de nombreux tissus, tandis que DDX3Y n'est exprimée que dans les gonades mâles. Les deux protéines sont très similaires, leurs séquences d'acides aminés présentant 91% d'identité. Cependant il n'est pas certain qu'elles soient fonctionnellement interchangeables. En effet, DDX3X et DDX3Y humaines peuvent toutes deux complémenter une mutation perte de fonction de DDX3X dans une lignée cellulaire de hamster (Sekiguchi et al., 2004). A l'inverse DDX3X ne peut pas complémenter une mutation de DDX3Y, ce qui suggère que ces deux protéines assurent des rôles distincts (Sharma and Jankowsky, 2014). Ces observations pourraient être expliqués par le fait que DDX3X et DDX3Y ne sont pas exprimées de la même manière au cours de la spermatogenèse, DDX3X étant principalement exprimée dans les spermatides à l'issue de la seconde division de méiose et DDX3Y dans les spermatogonies avant la méiose (Rauschendorf et al., 2014). Dans le cadre de ma thèse, je me suis intéressé à DDX3X (notée DDX3 dans ce manuscrit), DDX3Y n'étant pas exprimée dans les cellules infectés par le VIH-1.

1) Classification des hélicases

Les hélicases sont des enzymes qui utilisent l'ATP pour lier et remodeler des acides nucléiques et des complexes protéines – acides nucléiques (Fairman-Williams et al., 2010). Elles sont présentes dans la quasi-totalité des processus biologiques. On les classe en superfamilles suivant leurs propriétés structurales. Les hélicases formant un anneau hexamérique se répartissent entre les superfamilles 3 à 6, tandis que les hélicases non-hexamériques appartiennent aux superfamilles (SF) 1 et 2. La grande majorité des hélicases à ARN appartient aux superfamilles 1 et 2, bien qu'on retrouve également quelques hélicases à ARN virales dans les superfamilles 3 et 4 (Jankowsky and Fairman, 2007).

Les hélicases des superfamilles 1 et 2 présentent un cœur organisé en deux domaines structuraux répétés et relativement conservés, ainsi que des extensions spécifiques en N- et C-terminal (Figure 16). La présence de motifs conservés dans le cœur permet de définir des familles d'hélicases au sein de ces superfamilles. Les hélicases SF1 présentent en plus des

insertions spécifiques dans le cœur. DDX3 appartient à la famille des hélicases à boîte DEAD qui est généralement rattachée à la superfamille 2 (Cordin et al., 2006; Fairman-Williams et al., 2010). Chez l'homme, on identifie 37 membres dans cette famille (Linder and Jankowsky, 2011).

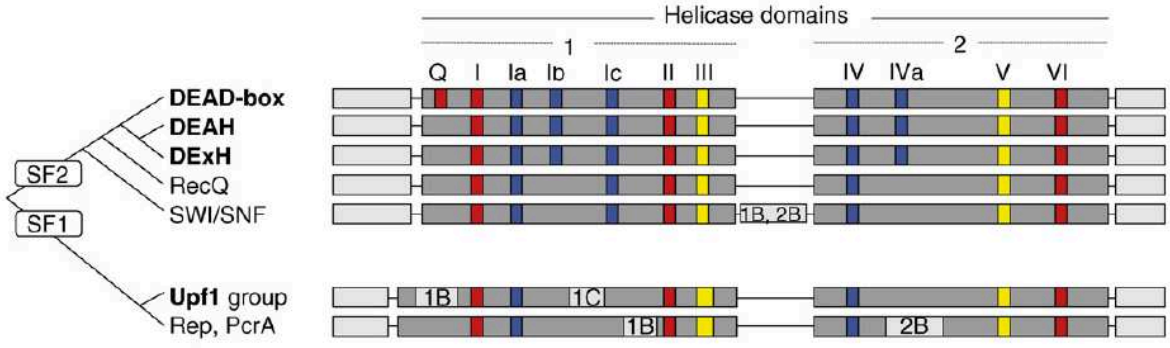


Figure 16 : Organisation phylogéniques des superfamilles SF1 et SF2 (Jankowsky et Fairman 2007). Les rectangles gris indiquent les extensions N- et C-terminales. Les motifs sont colorés suivant leur fonction biochimique *: rouge, liaison et hydrolyse de l'ATP* ; bleu, liaison *à l'ARN* ; jaune, coordination entre les deux activités. Les motifs 1B, 1C et 2B sont spécifiques de la superfamille 2.

Les hélicases à boîte DEAD présentent un cœur catalytique d'environ 350 à 400 acides aminés dont la séquence et la structure sont très conservées, et des extensions amino- et carboxy-terminales peu structurées dont la composition permet de définir des sous-familles. DDX3 et ses orthologues Ded1 et Dbp1 (S. cerevisiae), Belle (D. melanogaster), PL10 (M. musculus) ou An3 (X. laevis) forment ainsi une sous-famille d'hélicases à boîte DEAD.

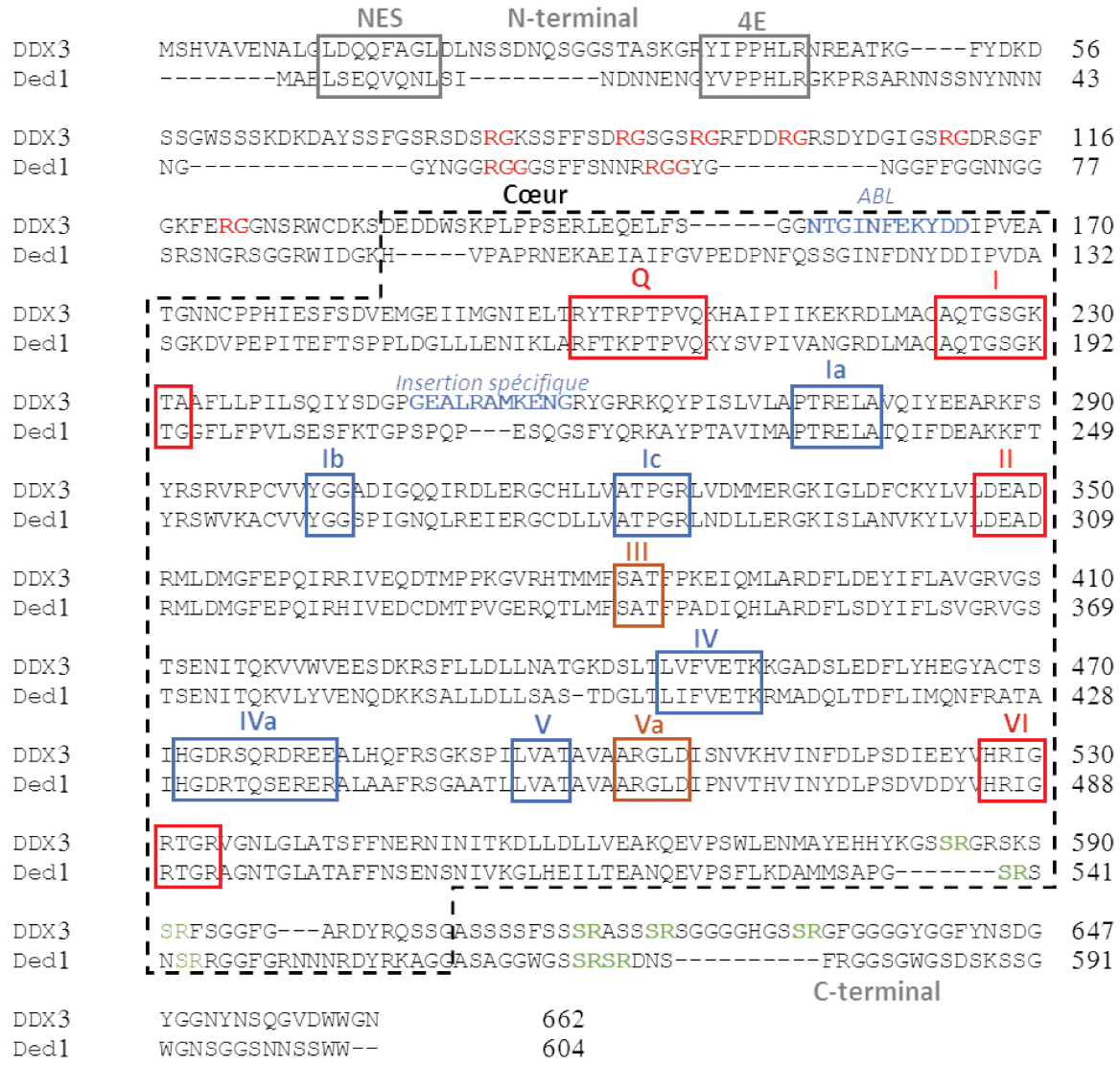
2) Relation structure – *fonction du cœur catal*ytique et des extensions terminales de DDX3

i. Le cœur catalytique et le couplage des activités ATPase et hélicase

Les motifs conservés forment un réseau d'interactions complexe avec les substrats et coordonnent le cycle catalytique

Le cœur catalytique des hélicases à boîte DEAD présente une série de douze motifs conservés (Linder and Jankowsky, 2011). Le motif Q est spécifique de la famille des hélicases à boîte DEAD et intervient dans la liaison à l'adénine (Tanner, 2003). Il pourrait être un régulateur de l'activité ATPase en modulant l'affinité de DDX3 pour l'ATP. Dans les structures disponibles, le motif Q établit des interactions spécifiques avec les azotes 6 et 7 de l'adénine,

ce qui suggère que les hélicases à boîte DEAD lient spécifiquement l'ATP (Cordin et al., 2006). Les motifs I, ou motif Walker A, II, ou Walker B, et VI interviennent dans la liaison au groupement triphosphate et participent à l'hydrolyse de l'ATP. On y trouve des résidus très conservés parmi les différentes familles d'hélicases, notamment ceux impliqués dans la liaison de la molécule d'eau qui intervient dans l'hydrolyse de l'ATP ainsi que dans la liaison du magnésium et des phosphates β et γ. Le motif II comporte en particulier la séquence conservée d'acides aminés aspartate – glutamate – alanine – aspartate (D – E – A – D suivant la nomenclature à une lettre) qui a servi à nommer la famille. Les motifs Ia, Ib, Ic, IIIa/IV, IVa et V lient l'ARN. Les motifs III et Va interagissent avec l'ATP et sont impliqués dans la coordination de la liaison à l'ATP et à l'ARN.



Légende page suivante

Figure 17 (page précédente) : Alignement des séquences de Ded1 et DDX3. Les motifs *impliqués dans la liaison à l'ATP* sont encadrés *en rouge, ceux qui interagissent avec l'ARN* sont encadrés en bleu et les motifs qui assurent la coordination des deux sites de liaison sont encadrés en jaune. Les séquences *d'export nucléaire (NES) et de liaison à eIF4E (4E) sont* indiquées en gris. Les motifs RG répétés en N-terminal sont en rouge, les motifs SR répétés en C-*terminal sont en vert. La région correspondant au cœur catalytique* défini par Floor et al. est encadrée par des tirets.

Structure du cœur catalytique

La structure cristallographique du cœur catalytique de DDX3 a pu être résolue en absence de nucléotide (Rodamilans and Montoya, 2007) et liée à l'AMP (Högbom et al., 2007) ou à l'ADP (Epling et al., 2015). Des structures plus récentes présentent une construction plus étendue du cœur catalytique lié à l'AMP ou à un analogue non hydrolysable de l'ATP (Floor et al., 2015) (Figure 18). Toutes ces structures ont été obtenues en absence d'ARN. Comme pour toutes les hélicases des superfamilles 1 et 2, le cœur catalytique adopte une structure en deux domaines RecA reliés par un connecteur flexible. Ce domaine, qui tire son nom d'une protéine de E. coli (Recombinase A), est retrouvé dans de nombreuses protéines qui lient les nucléotides triphosphate. Il est formé de cinq brins β entourés de cinq hélices α. Le motif Q est présent à l'extrémité N-terminale du premier domaine RecA. Il est compris dans une structure en coiffe, commune aux hélicases à boîte DEAD, composée d'une hélice α et de deux brins β (Högbom et al., 2007). Une séquence susceptible de former une boucle de liaison à l'ATP (ATP Binding Loop, ABL) se trouve une vingtaine de nucléotide en amont du motif Q. Cette boucle s'avère critique dans la stimulation de l'activité ATPase par l'ARN (Epling et al., 2015). Elle n'est pas modélisable en absence d'ARN, mais elle l'est dans la structure de Vasa liée à l'ARN. Il est possible que l'ABL se replie de façon coopérative avec la liaison à l'ARN (Floor et al., 2015). L'interstice entre les deux domaines RecA forme une cavité qui délimite les sites de liaison au nucléotide et à l'ARN. En absence d'ARN, la position relative de ces deux domaines semble être très peu contrainte.

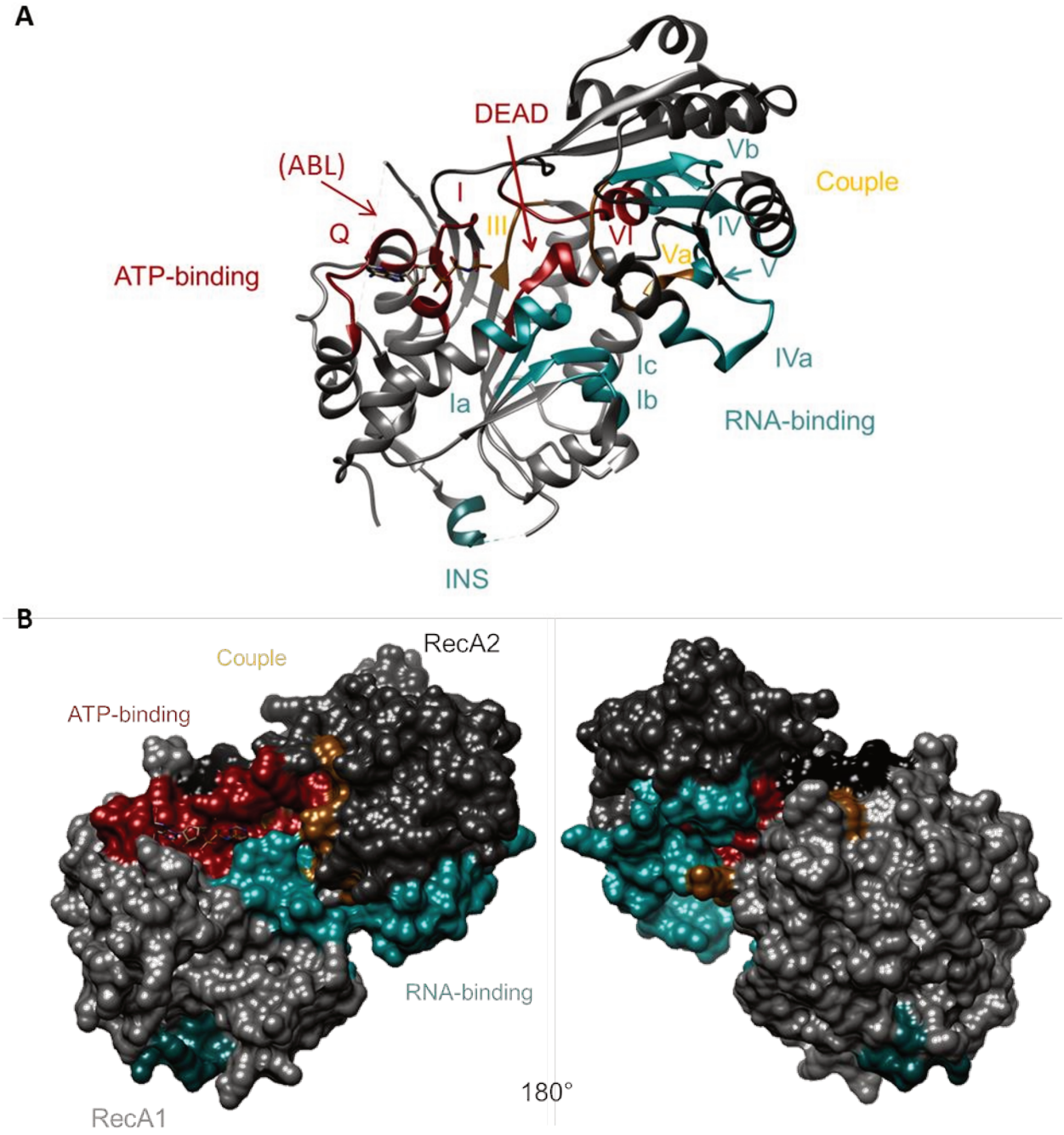


Figure 18 : Structure de DDX3 (PDB 5EM7M, résidus 132 à 607) en complexe avec *l'analogue de l'ATP non*-hydrolysable AMPPNP représentée sous forme de rubans (A) ou de *surface (B). Les régions correspondant au site de liaison à l'ATP, au site de liaison à l'ARN et* au couplage des deux activités sont indiquées en rouge, bleu et jaune respectivement.

De façon étonnante, la structure cristallographique de DDX3 révèle une boucle dans le domaine RecA entre les motifs I et Ia (résidus 250 – 259) qu'on ne retrouve chez aucune autre hélicase à boîte DEAD humaine (Högbom et al., 2007). Cette insertion est partiellement retrouvée chez Ded1 mais sans conservation de séquence (

Figure **17**). La localisation de cette boucle dans la structure de DDX3 laisse penser qu'elle intervient dans la liaison à l'ARN (Figure 19). De fait, la délétion de cette boucle diminue l'affinité de DDX3 pour l'ARN ainsi que son activité ATPase (Garbelli et al., 2011). Elle pourrait aussi permettre des interactions avec d'autres protéines (Sharma and Jankowsky, 2014).

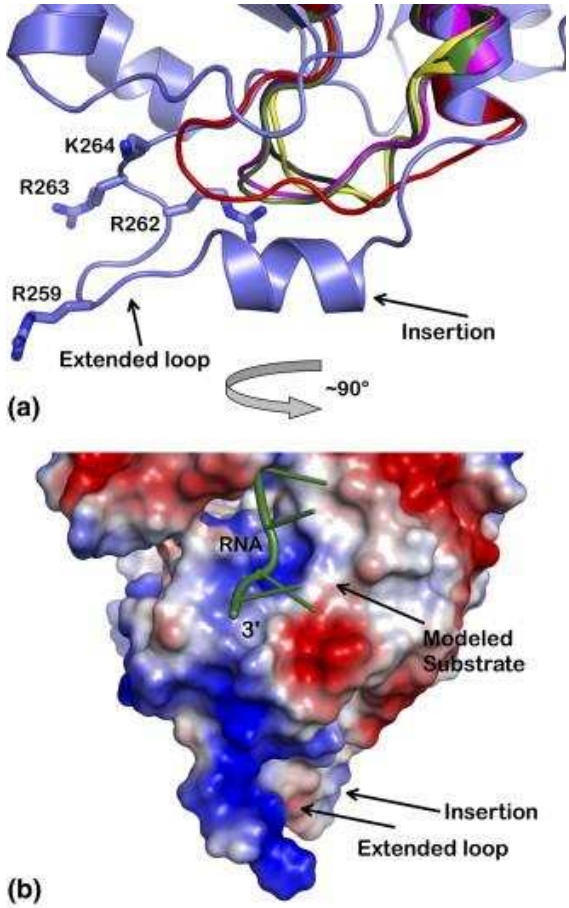


Figure 19 : DDX3 présente une insertion spécifique entre les motifs conservés I et Ia (Högbom et al., 2007). a/ Superposition des structures de DDX3 (en bleu) avec celles de UAP56 *en présence (vert) et en absence d'ADP (magenta),* Dhh1 (jaune) et eIF4A (noir). Les résidus chargés négativement sont indiqués en bâtons. b/ Surface de potentiel électrostatique avec une *molécule d'ARN modélisée à partir de la structure de Vasa. L'extrémité 3' pointe vers l'insert*ion spécifique de DDX3.

Les extensions N- et C-terminales de DDX3 présentent peu de résidus hydrophobes et sont donc vraisemblablement peu structurées. Par conséquent aucune structure cristallographique de DDX3 ne présente ces extensions (Linder and Jankowsky, 2011).

Par ailleurs, aucune structure de DDX3 liée à l'ARN n'est connue, mais on peut déduire la localisation du site de liaison à l'ARN à l'opposé du site de liaison à l'ATP à partir des structures d'autres hélicases à boîte DEAD liées à un fragment d'ARN simple brin, Mss116

(Del Campo and Lambowitz, 2009; Mallam et al., 2014), Vasa (Sengoku et al., 2006), eIF4A-III (Bono et al., 2006; Nielsen et al., 2009) et DDX19 (Collins et al., 2009; von Moeller et al., 2009; Montpetit et al., 2011) (revues dans Rudolph and Klostermeier, 2015). Notons qu'à ce jour, aucune structure d'une hélicase à boîte DEAD en complexe avec de l'ARN double brin n'est connue. Les domaines RecA liés à l'ATP et à l'ARN adoptent une structure fermée. La liaison à l'ARN se fait via la face ribose de quatre ou cinq nucléotides consécutifs, ce qui suggère une faible spécificité de séquence. Elle entraîne une courbure de la chaîne nucléotidique incompatible avec la structure d'une hélice d'ARN. Les structures révèlent de nombreuses interactions de la protéine avec les groupements 2'OH et phosphate.

La séparation de duplex d'acides nucléiques

Les hélicases à boîte DEAD dissocient des duplex d'une taille allant jusque deux tours d'hélice. Cette activité est généralement stimulée par la présence d'une extension d'ARN simple brin, en 3' ou en 5'. (Linder and Jankowsky, 2011).

Le modèle couramment admis est celui de séparation locale des brins (local strand separation) (Yang et al., 2007a). Ce modèle a été proposé en comparant l'affinité de Ded1 et Mss116 pour différents substrats synthétiques. En effet, Ded1 et Mss116 sont capables de séparer un duplex d'ADN comportant des résidus d'ARN au centre ou à l'extrémité d'un des brins (Figure 20). Comme ces hélicases sont incapables de séparer des duplex d'ADN purs, les auteurs proposent que l'hélicase dissocie d'abord la portion ADN/ARN puis que le reste du duplex, alors déstabilisé, se dissocie spontanément. En effet, la dissociation du duplex est légèrement plus lente si la section d'ARN est placée à une extrémité du duplex, car alors la partie du duplex ADN/ADN non dissociée par l'hélicase est plus stable. Ce n'est toutefois vrai que si la section d'ARN est placée en 3', et pas en 5'. La longueur de la section interne peut être réduite à deux résidus, bien que les structures cristallographiques de Vasa et eIF4AIII révèlent que la liaison de l'ARN implique 3 à 4 résidus 2'-OH. Les auteurs proposent le scénario suivant : 1/ l'hélicase lie le duplex d'acide nucléique à n'importe quelle position, 2/ elle sépare quelques paires de bases de façon dépendante de l'ATP puis 3/ la dissociation des paires de bases restantes se produit de façon passive. La présence d'une région simple brin d'au moins cinq nucléotides stimule la dissociation de duplex. Plus spécifiquement, l'ARN simple brin augmente l'affinité de l'hélicase pour le substrat mais pas la constante catalytique. Cette observation conduit les auteurs du modèle à proposer que la liaison à l'ARN simple brin et la dissociation du duplex sont deux étapes distinctes. Il est possible que l'ARN simple brin

permette de recruter l'hélicase à proximité de son substrat et augmente ainsi la concentration locale d'hélicase. Le modèle n'exclut pas qu'après dissociation active de quelques paires de bases et recyclage de l'hélicase, le duplex se reforme. On considère alors qu'il s'agit d'un cycle futile puisque l'état final est identique à l'état initial.

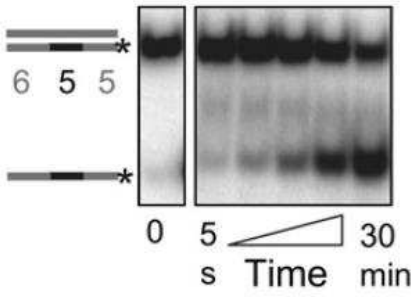


Figure 20 *: Exemple d'expérience utilisée* pour décrire le modèle de séparation locale des brins. *L'ARN est représenté en noir et l'ADN en gris. L'astérisque indique le brin marqué* radioactivement. Ce brin comporte *une insertion d'ARN lié de façon covalente* au reste de la chaîne nucléotidique. Le gel présente la séparation du duplex au cours du temps en présence de Ded1. Tiré de (Yang et al., 2007a)

Ce modèle reste toutefois incomplet car il ne permet pas d'expliquer certaines observations. En particulier, la polarité du substrat peut avoir un effet. Par exemple, un duplex avec une extrémité d'ARN simple brin en 5' est dissocié plus efficacement qu'un duplex avec une extrémité simple brin en 3' (Raj et al., 2019; Yang et al., 2007a). Ceci a également été constaté avec eIF4A lorsque la région simple brin n'excède pas 8 résidus (Rogers et al., 2001). A cet égard, il est intéressant de constater que toutes les structures d'hélicase à boîte DEAD en complexe avec de l'ARN simple-brin révèlent que l'extrémité 5' de l'ARN est toujours positionnée du côté du domaine RecA C-terminal, tandis que la torsion de l'ARN est toujours introduite du côté du domaine RecA N-terminal (Rudolph and Klostermeier, 2015). Les auteurs de la structure de Vasa proposent un mécanisme qui explique la préférence observée pour le 5' libre. L'hélicase lie un brin du duplex au niveau de 3 résidus consécutifs, provoquant la torsion du brin et la séparation de quelques paires de base en 3' ou en 5' suivant leur stabilité relative. Ensuite, la protéine pourrait étendre son interaction au niveau du domaine RecA C-terminal avec deux éventuels résidus simple brin en 5', augmentant ainsi l'efficacité de la réaction (Sengoku et al., 2006).

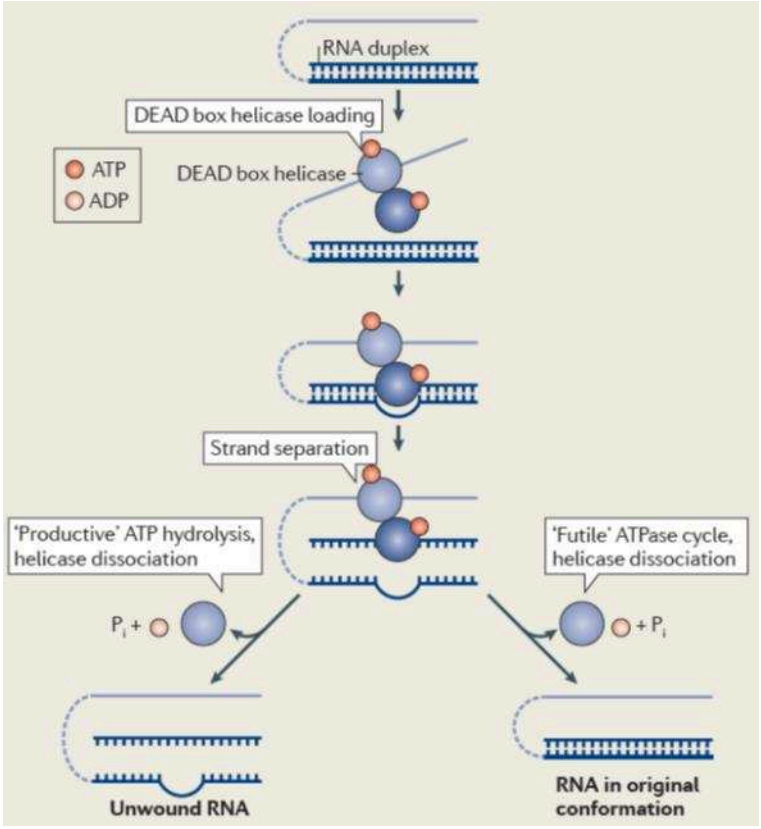


Figure 21 : Mécanisme proposé de séparation locale des brins (Linder and Jankowsky, 2011). Une hélicase lié à l'*ARN simple brin, schématisée en bleu clair, stimule le chargement d'une autre hélicase, en bleu foncé, au niveau d'une région d'ARN double brin.*

Un modèle alternatif au modèle de séparation locale de brins a récemment été proposé (Figure 22). En utilisant un système de pinces optiques, les auteurs confirment que l'activité hélicase de Ded1 est dépendante de l'ATP, non processive et plus prononcée vis-à-vis des duplex présentant une extrémité 5' libre, et que Ded1 a plus d'affinité pour l'ARN simple brin. Les duplex d'ARN sont des structures dynamiques où les paires de bases peuvent se dissocier de façon transitoire. Dans ce modèle, ces états transitoires instables sont stabilisés par Ded1. Le domaine RecA C-terminal lie l'ARN double-brin ou simple brin en 5', et le domaine RecA N-terminal lie uniquement l'ARN simple brin en 3' exposé lors des fluctuations de la structure. Une fois l'hélicase liée, la déstabilisation de la structure se poursuit (Raj et al., 2019).

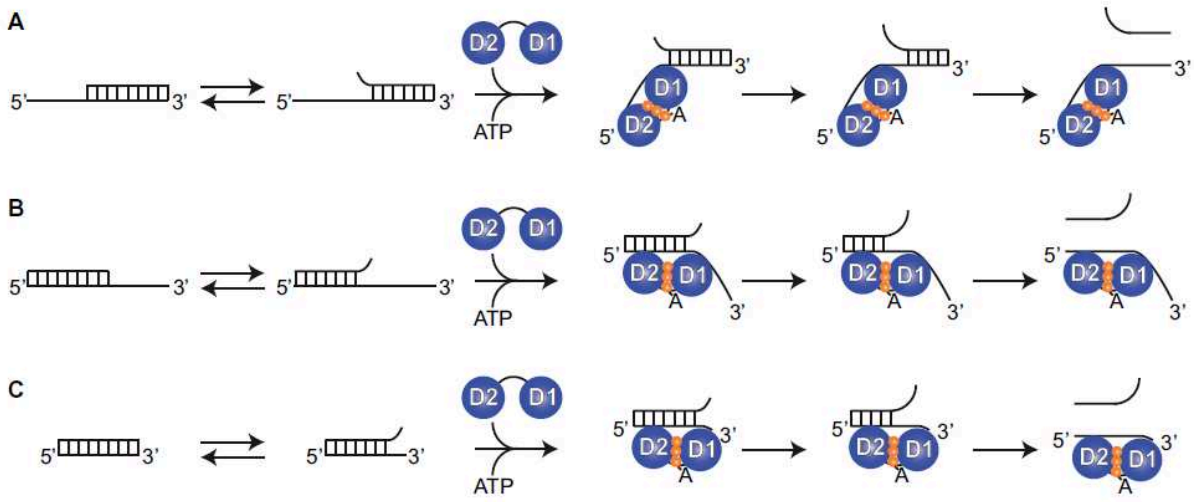


Figure 22 : ***Ded1 dissocie les duplex d'ARN en stabilisant des états transitoires instables***. D1 : domaine RecA N-terminal D2 : domaine Rec C-*terminal. L'ATP est indiqué*e en orange. *A/ Duplex avec une extrémité 5'* simple brin. *B/ Duplex avec une extrémité 3' simple brin.* Le *positionnement de l'hélicase* C/ Duplex à bouts francs. Tiré de (Raj et al., 2019)

La coordination des activités ATPase, de liaison à l'ARN et de séparation des duplex d'ARN est encore incertaine. Quatre réactions peuvent décrire le site de liaison à l'ATP : liaison de l'ATP, hydrolyse de l'ATP, dissociation du phosphate et dissociation de l'ADP. Le site de liaison à l'ARN peut quant à lui être caractérisé par trois réactions, à savoir : liaison de l'ARN double-brin, séparation des brins à l'ARN et dissociation de l'ARN. Ce modèle simplifié suppose donc 20 états biochimiques différents, ce qui en fait un système assez complexe à étudier (Henn et al., 2012). Les hélicases à boîte DEAD présentent peu ou pas d'activité ATPase en absence d'ARN. Il semble que pour la plupart d'entre elles, l'hydrolyse de l'ATP ne soit pas nécessaire à la dissociation du duplex mais seulement pour le recyclage de l'hélicase. Il est cependant possible que le couplage entre liaison à l'ATP, hydrolyse et séparation des brins varie entre les hélicases (Linder and Jankowsky, 2011). Par exemple, Mss116 (S. cerevisiae), Hera (T. thermophilus) et YxiN (B. subtilis) présentent des degrés divers de couplage entre la liaison à l'ATP et à l'ARN (Samatanga and Klostermeier, 2014). Le domaine RecA N-terminal de ces trois hélicases est suffisant pour assurer la liaison de l'ATP, tandis que le domaine RecA C-terminal intervient seulement pour l'hydrolyse. En revanche, la contribution des domaines RecA pour la liaison à l'ARN diffère selon les hélicases. Ainsi le domaine RecA C-terminal de Mss116 lie l'ARN de façon quasi-indépendante de la liaison à l'ATP. A l'inverse, les domaines RecA N- et C-terminaux de YxiN sont tous deux nécessaires à la liaison de l'ARN. YxiN lie séquentiellement l'ATP puis l'ARN. Enfin Hera présente un comportement intermédiaire où chacun des domaines RecA peut lier l'ARN, et la liaison de l'ARN stimule la liaison de l'ATP.

Une fois le complexe ternaire hélicase – ARN – ATP formé, l'hélicase adopte une conformation fermée compatible avec l'hydrolyse de l'ATP et la dissociation des brins d'ARN.

Conversion de structures d'ARN

Certaines hélicases à boîte DEAD peuvent stimuler l'hybridation de brins d'ARN. Le rôle de l'ATP dans cette activité n'est pas très clair, et il semble que le recours à l'ATP varie d'une hélicase à l'autre (Linder and Jankowsky, 2011). Chez Ded1, les activités de séparation de brins et d'hybridation de brins s'équilibrent in vitro suivant les concentrations relatives d'ATP et d'ADP, ainsi que selon la proportion d'ARN double brin. Ainsi, en absence d'ATP, Ded1 stimule l'hybridation de brins d'ARN complémentaires alors qu'elle dissocie des brins hybridés en présence d'un excès d'ATP (Yang and Jankowsky, 2005). Cette activité semble être partagée par les hélicases à boîte DEAD CYT-19 chez N. crassa (Mohr et al., 2002) et Mss116 également chez la levure (Halls et al., 2007) mais elle n'est pas décrite pour DDX3.

En utilisant un système composé de trois brins d'ARN susceptibles de former deux structures de différente stabilité, ces mêmes auteurs ont proposé un modèle pour décrire l'activité de conversion de structure de Ded1 (Figure 23). En absence d'ATP, Ded1 stimule la conversion d'un duplex relativement instable vers un duplex plus stable, c'est-à-dire une réaction thermodynamiquement favorable. Ce mécanisme n'implique pas la séparation complète des brins du duplex initial, mais passe par de multiples intermédiaires où les trois brins forment un seul complexe. Cette propriété est caractéristique des chaperons à ARN. A l'inverse, en présence d'ATP, Ded1 stimule la dissociation complète d'un duplex stable et la reformation d'un duplex moins stable, autrement dit une réaction thermodynamiquement défavorable (Yang et al., 2007b).

Ces résultats paraissent toutefois contradictoire avec le fait que Ded1 ne semble pas lier l'ARN en absence de nucléotide (Yang and Jankowsky, 2006). De plus, ils ne peuvent pas être directement transposés aux conditions physiologiques où la concentration en ATP est maintenue relativement constante. On peut supposer que des facteurs cellulaires interviennent in vivo pour contrôler cette activité de conversion de structure en inhibant la liaison de l'ATP.

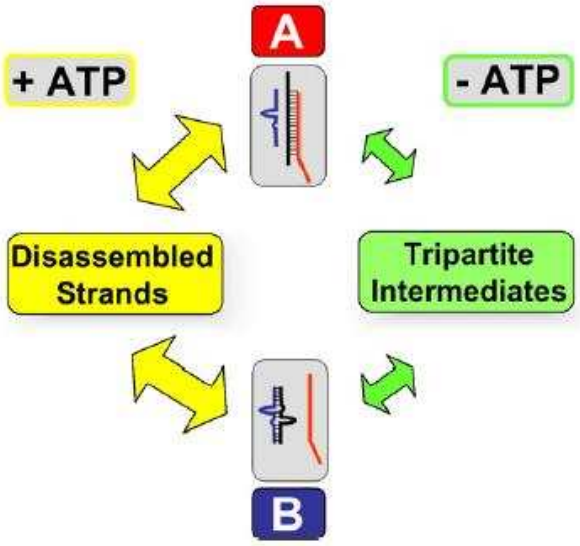


Figure 23 ***: Modèle illustrant l'activité de conversion de structure de Ded1*** (Yang et al., 2007b). Le duplex B est thermodynamiquement moins stable que A du fait de la présence de boucle internes.

Formation et remodelage de complexes ribonucléoprotéiques

Chez la levure, les hélicases Ded1, Mss116 et Sub2 peuvent former des complexes stables pendant plusieurs heures avec l'ARN en présence d'ADP-BeF_x, un analogue non hydrolysable de l'ATP. Notons en revanche que seules Ded1 et Mss116 forment un complexe avec l'ARN en présence d'ADP-AlF_x, un analogue de l'état de transition ADP – Pi, ce qui suggère que les hélicases à boîte DEAD ne suivent pas nécessairement le même couplage entre hydrolyse d'ATP et liaison à l'ARN (Liu et al., 2014). Par ailleurs, in vivo, le complexe MAGOH/Y14 inhibe le relargage des produits d'hydrolyse de l'ATP par l'hélicase eIF4AIII lors de la formation du complexe EJC (Exon Junction Complex) (Nielsen et al., 2009). Un autre exemple concerne le l'export nucléaire chez la levure. Ded1 s'associe à Xpo1 et RanGTP, ce qui a pour effet de réduire son activité ATPase (Hauk and Bowman, 2015). L'hydrolyse de l'ATP semble donc pouvoir être contrôlée pour permettre aux hélicases à boîte DEAD de former des complexes ribonucléoprotéiques stables. Par ailleurs, chez la levure, l'AMP inhibe l'activité de certaines hélicases à boîte DEAD ce qui soulève la possibilité d'une régulation de ces hélicases par l'état métabolique de la cellule, dont dépend la concentration en AMP (Putnam and Jankowsky, 2013).

De plus, Ded1 peut activement dissocier certains complexes ribonucléoprotéiques in vitro à condition qu'une extrémité d'ARN simple brin soit présente. Par exemple, Ded1 dissocie les complexes EJC et U1 snRNP de leur substrat ARN mais pas TRAP ni U1A (Bowers et al.,

2006; Fairman et al., 2004). Les auteurs de ces résultats proposent que la capacité d'une hélicase à dissocier un complexe RNP soit dictée par l'architecture du complexe plutôt que par la différence de stabilité de ces complexes. D'autres exemples de dissociation de complexes RNP par des hélicases à boîte DEAD ont été rapportés, notamment chez la levure avec les hélicases Sub2 et Dbp5 (Linder and Jankowsky, 2011). Cette activité est observable alors même que les complexes sont formés sur de l'ARN non structuré ce qui montre qu'elle est indépendante de l'activité de dissociation de brins. Il est possible que seule la torsion du brin d'ARN sous l'effet de la liaison à l'hélicase suffise à déstabiliser les complexes ARN – protéine (Sengoku et al., 2006).

Interaction avec les quadruplex de G

Des expériences de pull-down ont montré que DDX3 interagit avec les quadruplex de G de la 5'UTR de l'oncogène NRAS (Herdy et al., 2018) et de l'ARN ribosomique (Mestre-Fos et al., 2019). Ces structures non-canoniques particulièrement stables sont formées par des interactions hydrogène et des empilements π – π entre des guanines de quatre brins d'ARN. Chez la levure, Ded1 est capable d'interagir avec les quadruplex de G (ARN ou ADN) et de déstructurer les quadruplex de G ARN (Gao et al., 2019), quoique le mécanisme sous-jacent ne soit pas connu. Ces résultats suggèrent que DDX3 pourrait avoir une fonction similaire chez l'homme.

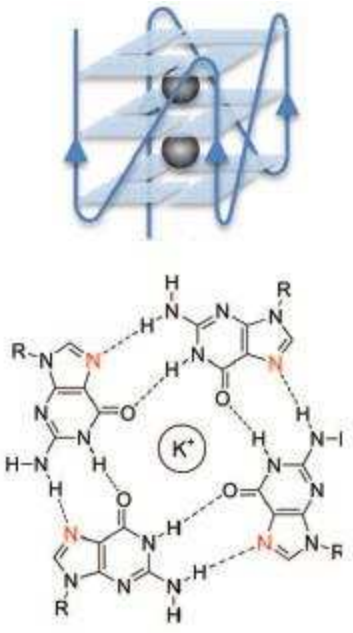


Figure 24 ***: La structure d'un quadruplex de G ARN*** est stabilisée par de nombreuses liaisons hydrogène*, des empilements π – π entre les bases et la coordination de cations. Tiré de* (Guo and Bartel, 2016)

ii. Les extensions C- et N-terminales variables assurent la spécificité

Le génome humain code 37 hélicases à boîte DEAD différentes (Jankowsky et al., 2011) qui présentent toutes un cœur catalytique très conservé. La spécialisation fonctionnelle des hélicases à boîte DEAD provient donc vraisemblablement des extensions N- et C-terminales qui, elles, diffèrent significativement en composition et en taille.

Composition des extensions N- et C-terminales de DDX3

L'extension N-terminale de DDX3 comporte une séquence d'export nucléaire NES ainsi que le motif consensus de liaison au facteur d'initiation de la traduction eIF4E YxxxxLϕ (où x est n'importe quel acide aminé et ϕ un acide aminé hydrophobe) (Sharma and Jankowsky, 2014). De fait, DDX3 interagit in vitro avec eIF4E et la mutation de ce domaine abolit l'interaction (Shih et al., 2008). On trouve de plus dans l'extension N-terminale la répétition de six motifs arginine – glycine (RG), dont trois espacés de moins de quatre résidus. Ce type de répétition du dipeptide RG est fréquemment retrouvé dans les protéines de liaison aux acides nucléiques et constitue lui-même vraisemblablement un site de liaison à l'ARN (revu dans (Thandapani et al., 2013)). En particulier, le motif RG est impliqué dans la liaison de DDX3 aux quadruplex de G (Herdy et al., 2018). En effet, les ARNm auxquels DDX3 s'associe in vivo sont significativement enrichis en quadruplex de G. La mutation de certaines des arginines des motifs RG en alanine diminue la proportion de quadruplex de G dans les ARNm qui s'associent à DDX3. L'effet de cette interaction n'est cependant pas connu. Il pourrait être de stabiliser les quadruplex, comme c'est le cas pour le motif RG de la protéine FMRP. A l'inverse, DDX3 pourrait déstructurer des quadruplex de G comme le fait Ded1 (Gao et al., 2019). Les motifs RG sont également impliqués dans des interactions protéine – protéine. De façon remarquable, des interactions entre des motifs RG et des protéines impliquées dans l'export de l'ARNg du VIH ont été mises en évidence. Par exemple, le motif RG de FMRP interagit avec Ran et celui de hnRNPA1 avec Rev. Enfin, les arginines du motif RG peuvent être méthylées au niveau des deux amines primaires. La méthylation des motifs RG est suceptible de réguler leur fonction. En particulier, elle pourrait avoir un effet sur la localisation cellulaire. C'est par exemple le cas pour protéine Sam68, qui est importée dans le noyau lorsque son motif RG n'est pas méthylé.

Des répétitions du motif sérine – arginine sont retrouvée dans les extensions N- et C-terminales. L'extension C-terminale assure l'interaction de DDX3 avec le facteur d'export nucléaire TAP. Les régions riches en sérine – arginine étant connues pour être de potentiels adaptateurs de TAP, les auteurs de cette étude supposent que la répétition du motif SR en C-

terminal de DDX3 est impliquée dans l'interaction avec TAP (Lai et al., 2008). Notons également qu'il s'agit d'un motif pouvant être phosphorylé, bien qu'aucune phosphorylation n'ait été identifiée au niveau de ces résidus à ce jour.

iii. Coopération des extensions N- et C- terminales avec le cœur catalytique

Plusieurs exemples de coopération entre l'extension C-terminale et cœur catalytique d'hélicases à boîte DEAD n'appartenant pas à la sous-famille de DDX3/Ded1 ont été mises en évidence. Chez N. crassa, l'extrémité C-terminale de CYT-19 présente un accumulation de résidus arginine qui interagissent avec l'ARN double brin ce qui a pour effet d'augmenter l'affinité du cœur catalytique pour l'ARN (Busa et al., 2017). Les résidus arginine sont en effet des résidus basiques susceptibles de former des liaisons hydrogènes avec les groupements phosphate, le ribose ou la base azotée des nucléotides, spécifiquement ou non (Burd and Dreyfuss, 1994). Chez Mss116, l'orthologue de CYT-19 chez la levure, le domaine C-terminal constitue également un second site de liaison à l'ARN qui provoque une deuxième courbure de la partie 5' chaîne nucléotidique (Del Campo and Lambowitz, 2009). Notons que, à la différence de la plupart des hélicases à boîte DEAD, l'extension C-terminale de Mss116 est partiellement structurée et interagit avec le domaine RecA 2. Chez E. coli, l'extension C-terminale de DbpA présente également un module structuré, dont le repliement est indépendant, et qui est capable d'interagir spécifiquement avec la tige-boule 92 de l'ARN ribosomal 23S (Moore et al., 2017). Elle recrute ainsi le cœur hélicase à proximité, ce qui a pour effet de spécifier les régions d'ARN double-brin substrats de DbpA autour de cette tige-boucle.

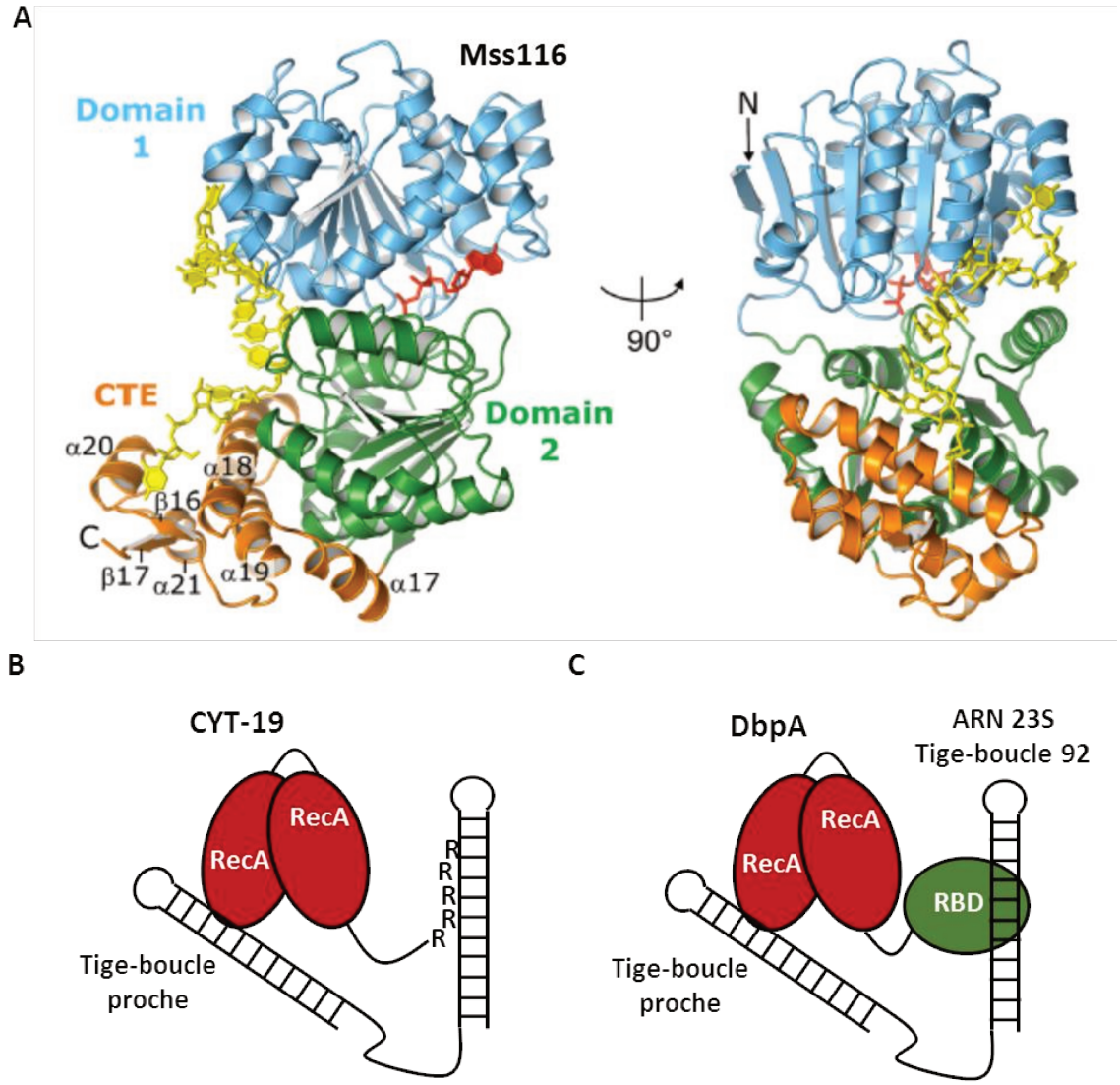


Figure 25 : Coopération des extensions N et C-*terminales avec le cœur catalytique*. A. La structure de Mss116 révèle un domaine additionnel *de liaison à l'ARN* qui entraîne une seconde courbure de la chaîne nucléotidique (Del Campo & Lambovitz, 2009). B. CYT-19 présente une accumulation de résidus arginine dans son extension C-terminale qui augmente son affinité *pour l'ARN. C. DbpA dispose d'un module structuré dans son extension C*-terminale qui ancre *spécifiquement le cœur catalytique* à la tige-*boucle 92 de l'ARN 23S et lui permet de* déstabiliser les structures voisines.

Un modèle biochimique élaboré chez la levure propose une oligomérisation fonctionnelle de Ded1 impliquant l'extension C-terminale (Putnam et al., 2015). In vitro, Ded1 forme des trimères qui s'assemblent en présence d'ARN. Il est possible de former un complexe stable Ded1 –ARN au moyen d'un duplex doté d'une extrémité simple brin et d'un analogue non-hydrolysable d'ATP. L'ajout d'ATP entraîne la libération d'un des brins du duplex, tandis que le complexe Ded1 – ARN reste formé. Ils proposent un modèle dans lequel deux protomères se lient à une région d'ARN simple brin pour recruter un troisième protomère au niveau d'une

région double brin. Les deux protomères « de chargement » nécessitent la présence d'ATP pour lier l'ARN mais ils l'hydrolysent faiblement, de sorte qu'ils restent liés stablement à l'ARN. Le protomère qui sépare les brins d'ARN requiert quant à lui l'hydrolyse de l'ATP pour effectuer plusieurs cycles de façon répétée.

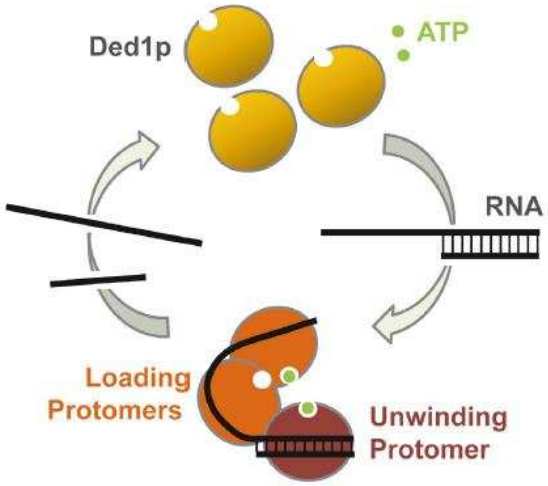


Figure 26 : Modèle d'oligomérisation fonctionnelle de Ded1 (Putnam et al., 2015). Les protomères de chargement lient *l'ARN simple brin de façon stable, tandis que le protomère de* dissociation des brins effectue de multiples cycles catalytiques.

iv. Des modifications post-traductionnelles sont susceptibles de réguler l'activité de DDX3

La base de données PhosphoSite Plus (http://phosphositeplus.org) répertorie 105 sites potentiels de modifications post-traductionnelles diverses réparties sur l'ensemble de la structure de DDX3. Dans le cœur catalytique, les thréonines 204 et 323 peuvent être phosphorylées in vitro par la kinase cycline b/cdc2 (Sekiguchi et al., 2007). L'extension N-terminale est également phosphorylée in vitro par IKKe, au moins sur la sérine 102 (Gu et al., 2013), et TBK1 (Soulat et al., 2008) mais les conséquences moléculaires de ces modifications ne sont pas connues. Aucune autre modification post-traductionnelle n'a à ce jour été vérifiée. Les modifications post-traductionnelles de DDX3 semblent osciller au cours du cycle cellulaire. En effet, durant l'interphase, DDX3 est phosphorylée au cours de la phase de croissance (phase G1) et de la phase précédant la mitose (G2/M), mais pas au cours de la transition vers la phase de duplication des chromosomes (G1/S) (Choi and Lee, 2012).

3) DDX3 joue de nombreux rôles *dans le contrôle de l'expression* génétique

i. DDX3 assure des fonctions tout au long du métabolisme des ARNm

Epissage et export nucléaire / fonctions nucléaires

DDX3 présente plusieurs fonctions nucléaires. Elle intervient tout d'abord dans la régulation de la transcription de certains gènes, dont l'IFN-β (Soulat et al., 2008), $p21^{waf1/cip1}$ (Chao et al., 2006), la E-cadhérine (Botlagunta et al., 2008) et les cytokines IL-2 et IL-6 (Xiang et al., 2016), suivant un mécanisme plus ou moins bien connu. Notons toutefois que DDX3 semble associée à la chromatine. Des expériences d'immunoprécipitation ont en effet montré que DDX3 est liée au promoteur de la E-cadhérine (Botlagunta et al., 2008) et au promoteur de $p21^{waf1/cip1}$ (Chao et al., 2006). D'autre part DDX3 pourrait jouer un rôle dans l'épissage des ARNm. En effet, des expériences de pull-down à partir d'extraits cellulaires ont montré que DDX3 s'associe à des protéines du complexe de jonction d'exons EJC telles que Y14, Magoh, eIF4AIII, Acinus et Aly/REF. L'association de DDX3 avec ces ribonucléoprotéines résiste par ailleurs à l'action conjuguée de la RNAse H et de l'héparine, ce qui suggère une interaction directe et robuste avec l'EJC (Merz et al., 2007). Il n'est pas certain que DDX3 intervienne activement au cours de l'épissage comme c'est le cas pour d'autres hélicases à boîte DEAD. Cependant, des expériences d'interférence ARN chez des cellules de drosophile montrent que le knockdown de DDX3 n'affecte ni la transcription, ni l'épissage (Lee et al., 2008). Ces résultats appellent davantage de clarifications pour comprendre le rôle exact de DDX3. Il est possible que la fonction de DDX3 ne concerne qu'un certain type de transcrits ou un certain état physiologique, ou encore que des mécanismes alternatifs puissent compenser la déplétion de DDX3

L'épissage des ARNm est fortement lié à l'export nucléaire. En effet, la principale protéine d'export des ARNm, TAP, est recrutée par le complexe EJC (Lai et al., 2008). Ce mécanisme assure à la cellule de n'exporter que des ARNm correctement épissés. Des interactions directes entre DDX3 et les facteurs d'export nucléaire TAP et CRM1 ont en été mises en évidence. DDX3 interagit avec TAP via son extension C-terminale riche en arginines et sérines (Lai et al., 2008). DDX3 s'accumule dans le noyau en absence de TAP, ce qui montre qu'elle est au moins partiellement exportée avec des ARNm suivant cette voie. Cependant DDX3 ne semble pas intervenir dans la voie générale d'export des ARNm. En effet, la réduction

de l'expression de DDX3 n'entraîne pas de changement dans la distribution nucléo-cytoplasmique d'ARNm contrôles. D'autre part l'export de l'ARN génomique du VIH-1 dépend de l'interaction DDX3 – CRM1 (Yedavalli et al., 2004). Cet aspect a déjà été développé, mais il suggère de plus que DDX3 puisse être un cofacteur de l'export d'ARNm dépendant de CRM1. L'extension C-terminale de DDX3, et non son NES, a initialement été proposée comme assurant l'interaction de DDX3 - CRM1 (Yedavalli et al., 2004). Une étude plus récente a cependant montré l'inverse, à savoir que le NES et non l'extension C-terminale est requis pour l'export de DDX3 dépendant de CRM1 (Brennan et al., 2018). Ces résultats semblent confortés par le fait que le domaine N-terminal de DDX3 seul, qui contient le NES, peut interagir avec CRM1 (Fröhlich et al., 2016). D'autre part, chez la levure, Ded1 interagit avec l'homologue de CRM1, Xpo1, de façon dépendante de son NES (Hauk and Bowman, 2015). Les activités hélicase et ATPase ne sont pas requises dans l'export nucléaire de DDX3 dépendant de CRM1 (Brennan et al., 2018; Yedavalli et al., 2004). De plus le recours à un inhibiteur de CRM1, la leptomycine B, entraîne une accumulation partielle de DDX3 dans le noyau. Cette observation confirme que DDX3 est un cargo et éventuellement un cofacteur de CRM1, et qu'il existe plusieurs voies d'export nucléaire de DDX3 (Lai et al., 2008).

DDX3 présente également au moins trois séquences de localisation nucléaire (NLS) située dans l'extension N-terminale et dans chacun des deux domaines RecA. Ces NLS, en concert avec le NES, lui permettent de circuler entre le noyau et le cytoplasme. Sa localisation subcellulaire oscille en effet au cours des phases du cycle cellulaire, avec une accumulation nucléaire au cours de la prophase – pro-métaphase (Brennan et al., 2018).

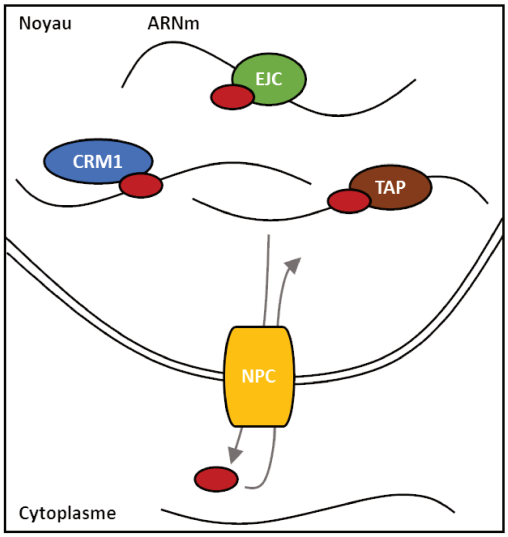


Figure 27: DDX3 (en rouge) ***est associée à la machinerie d'épissaga ainsi qu'à*** **l'export nucléaire des ARN**.

DDX3 et l'initiation de la traduction

L'homologue de DDX3 chez la levure, Ded1, est considéré comme un facteur général de la traduction (Chuang et al., 1997). Le rôle de Ded1 serait d'assister le balayage du ribosome le long de la 5'UTR en déstabilisant les éléments de structure via son activité hélicase (Guenther et al., 2018). DDX3, quant à elle, interagit directement avec plusieurs facteurs d'initiation de la traduction tels que eIF3 (Lee et al., 2008), eIF4A, eIF2a, PABP1 (Lai et al., 2008), eIF4E (Shih et al., 2008) et eIF4G (Soto-Rifo et al., 2013). Cependant, en dépit de l'homologie de Ded1 et DDX3 et de la capacité de DDX3 à complémenter le knockdown de Ded1 chez la levure, il n'est pas certain que DDX3 ait le même rôle que Ded1.

Certaines études soutiennent que DDX3 est un facteur général de la traduction. En effet, l'utilisation d'ARN interférents ciblant DDX3 provoque la diminution de la traduction d'un gène rapporteur exprimant l'ARNm de la β-globine sans affecter l'export (Lee et al., 2008). D'autre part la diminution de l'expression de DDX3 diminue la traduction d'ARN dont la traduction est supposée dépendante de la coiffe et de l'IRES du VHC (Geissler et al., 2012). DDX3 interagit avec eIF3c, la sous-unité 40S et les ribosomes matures au début de l'élongation. Les auteurs proposent un modèle DDX3 s'associe au complexe de pré-initiation 43S pour stimuler l'assemblage de la grande sous-unité du ribosome. Cette fonction est conservée dans l'initiation dépendante de la coiffe et dans certains mécanismes d'entrée interne du ribosome (Geissler et al., 2012). Par ailleurs, DDX3 stimule la traduction de transcrits qui présentent de petits cadres ouverts de lecture (upstream Open Reading Frame, uORF) dans leur région 5' (Chen et al., 2018). Ces uORF inhibent généralement la traduction au niveau de l'ORF principal par un effet compétiteur sur le recrutement du ribosome. Une hypothèse envisageable suggère que DDX3 assiste le ribosome dans la sélection de l'ORF principal. Un tel rôle a d'ailleurs été plus clairement proposé pour Ded1, qui interagit avec le complexe de pré-initiation pour déstabiliser les structures de la 5'UTR des ARNm et favoriser la sélection du codon initiateur de l'ORF principal (Guenther et al., 2018).

La similarité entre le rôle de DDX3 et de Ded1 dans la stimulation de la traduction a cependant été remise en question. En effet, dans un système de traduction in vitro, DDX3 ne stimule pas la formation de complexe 48S comme le fait Ded1. Il semble que la fonction d'assistance au balayage de la 5'UTR décrite pour Ded1 chez la levure soit portée chez l'homme par DHX29, une hélicase à boîte DEAH (Abaeva et al., 2011).

Une autre étude suggère au contraire que DDX3 inhibe la traduction dépendante de la coiffe, mais pas l'entrée interne du ribosome (Shih et al., 2008). En effet, la surexpression de DDX3 dans des cellules HuH7 entraîne la diminution de l'incorporation de méthionine, autrement dit une diminution de la synthèse protéique. In vitro, DDX3 a un effet très similaire à celui de 4EBP1, une protéine qui régule la traduction en inhibant la liaison d'eIF4E à la coiffe. En effet, en présence de DDX3, on observe une diminution de l'expression d'un gène rapporteur sous contrôle de la coiffe, mais pas d'effet sur un gène rapporteur sous contrôle d'une IRES. De plus, DDX3 interagit avec eIF4E via son motif consensus d'interaction. Une mutation dans ce motif abolit simultanément l'interaction de DDX3 avec eIF4E et l'inhibition de la traduction dépendante de la coiffe par DDX3, ce qui suggère que DDX3 joue un rôle similaire à celui de 4EBP (Shih et al., 2008).

La divergence des résultats quant au rôle de DDX3 dans la traduction peut être expliquée de différentes façons. Tout d'abord, les modèles cellulaires utilisés diffèrent (cellules HeLa issues d'un cancer du col de l'utérus et HuH7 issues d'un cancer du foie). Il est possible que le comportement de ces lignées tumorales reflète différents états cellulaires. En particulier, les modifications post-traductionnelles ne sont pas identiques dans toutes les lignées cellulaires (N. Chamond, communication personnelle), or ces modifications sont susceptibles de modifier l'activité de la protéine. Cependant cette explication seule est insuffisante, puisque les cellules Huh7 sont utilisées à la fois pour proposer un rôle inhibiteur (Shih et al., 2008) et activateur (Geissler et al., 2012) de la traduction. Il est également possible que, compte tenu des diverses fonctions assurées par DDX3, une modification de son niveau d'expression ait des conséquences sur la traduction qui ne reflètent pas directement la fonction de DDX3 mais résultent d'un état de stress. Ainsi, les observations qui soutiennent un rôle inhibiteur de DDX3 sur la traduction reposent sur la surexpression de DDX3 (Shih et al., 2008), tandis que les observations qui soutiennent un rôle activateur reposent sur la suppression de l'expression de DDX3 (Chen et al., 2018; Geissler et al., 2012; Lee et al., 2008).

Le dernier modèle pourrait concilier les deux rôles antagonistes précédemment décrits. Dans ce modèle, DDX3 stimule uniquement la traduction de certains ARNm, caractérisés par la présence d'une structure stable à leur extrémité 5'. Son rôle dans la traduction semble se limiter aux granules de stress où elle est associée à des facteurs d'initiation de la traduction (Lai et al., 2008, 2010; Soto-Rifo et al., 2012b). Cette observation est cohérente avec l'inhibition générale de la traduction observée lors de la surexpression de DDX3. Lorsque l'expression de DDX3 est réduite par interférence ARN dans des cellules HeLa, l'expression d'un gène

rapporteur est peu modifiée (Lai et al., 2008), pas plus que l'incorporation de méthionine (Soto-Rifo et al., 2012b). En revanche, l'ajout d'une région 5'UTR structurée en amont du gène rapporteur rend son expression plus dépendante de DDX3. De plus, une mutation affectant l'activité hélicase de DDX3 la rend incapable de stimuler la traduction d'ARNm avec une 5'UTR structurée (Lai et al., 2008). Le même groupe a cherché à identifier les ARNm cibles de DDX3 (Lai et al., 2010). En comparant les ARNm présents dans les complexes d'initiation 40S issus de cellules exprimant ou non DDX3, l'initiation de la traduction d'environ 300 ARNm apparaît dépendante de DDX3. Ces ARNm présentent une 5'UTR sensiblement plus longue et structurée que la moyenne. 10% des ARNm identifiés correspondent à des protéines régulatrices du cycle cellulaire, ce qui confirme le rôle important joué par DDX3 dans le contrôle du cycle cellulaire. D'autres auteurs ont utilisé l'ARNg du VIH comme modèle pour montrer que DDX3 stimule la traduction d'ARN présentant une 5'UTR structurée (Soto-Rifo et al., 2012b).

ii. DDX3 intervient dans plusieurs cascades de signalisation cellulaire

Réponse à l'hypoxie

DDX3 est assez clairement associée à la réponse à l'hypoxie. En effet, dans les cancers du poumon, l'expression de DDX3 est associée à l'expression de marqueurs de l'hypoxie tels que le facteur de transcription HIF-1 α (Bol et al., 2013). Dans des conditions normales, HIF-1 α est ubiquitiné et dégradé par le protéasome tandis qu'il s'associe au transporteur nucléaire ARNT dans des conditions d'hypoxie pour former le facteur de transcription HIF-1. Dans des conditions d'hypoxies, HIF-1 α stimule la transcription du gène DDX3 qui présente trois éléments de réponses à HIF-1 dans son promoteur (Botlagunta et al., 2011).

Régulation du cycle cellulaire et de l'apoptose (Bol et al., 2015a)

DDX3 intervient dans la régulation du cycle cellulaire à plusieurs niveaux. Tout d'abord, elle stimule la transcription de $p21^{waf1/cip1}$, un inhibiteur de kinase impliqué dans le contrôle du cycle cellulaire (Chao et al., 2006). DDX3 est en effet capable d'interagir avec le facteur de transcription Sp1, qui dispose de plusieurs sites spécifiques dans le promoteur de $p21^{waf1/cip1}$. L'utilisation d'un système de gène rapporteur sous contrôle du promoteur de $p21^{waf1/cip1}$ confirme l'effet activateur de DDX3. DDX3 peut aussi réguler la traduction de certains transcrits codant des protéines de contrôle du cycle cellulaire, les cyclines D1 et E1. DDX3 stimule la traduction de la cycline E1, qui est impliquée dans la transition G1 -> S (Lai

et al., 2010). D'autre part, DDX3 régule la transcription de p53, un facteur cellulaire critique dans la régulation de la mitose. La suppression de l'expression de DDX3 entraîne l'activation de méthyltransférases qui hyperméthylent le promoteur de p53 et recrute des histones qui réprime la transcription. De plus, l'utilisation d'un ARNm rapporteur contenant la 5'UTR de p53 révèle que DDX3 stimule la traduction de p53 (Chen et al., 2017). Enfin, DDX3 est associée à la régulation de l'apoptose induite par la voie extrinsèque. Cette voie de signalisation fait intervenir des récepteurs membranaires de la famille du TNF. La stimulation de ces récepteurs par des cytokines entraîne leur trimérisation, qui induit la formation du complexe DISC (Death-Inducing Signaling Complex) et l'activation de caspases entraînant l'apoptose. DDX3 s'associe à GSK3 (glycogen synthase kinase 3) et cIAP1 (cellular inhibitor of apoptosis protein 1) pour former un complexe qui inhibe les quatre principaux récepteurs de mort (Fas, TNF-R, TRAIL-R1 et TRAIL-R2). La stimulation de ces récepteurs de mort entraîne l'inactivation de DDX3 par clivage. Il est possible que certaines cellules cancéreuses échappent au déclenchement de l'apoptose par voie extrinsèque en bloquant la protéolyse de DDX3 (Sun et al., 2008). En conclusion, DDX3 régule le devenir cellulaire en intervenant au niveau de la régulation transcriptionnelle et traductionnelle de gènes, ainsi qu'en intervenant dans des cascades de transduction de signaux extrinsèques.

Contrôle du développement par la voie Wnt

La voie Wnt (Wingless – Int1, du nom des gènes ayant conduit à sa découverte) est une voie de transduction du signal impliquée dans divers processus de développement, y compris cancéreux. Elle est déclenchée par la protéine Wnt qui permet la communication cellulaire paracrine. En absence de Wnt, la β-caténine est recrutée au sein d'un complexe de destruction qui déclenche la polyubiquitination et l'adressage au protéasome de la β-caténine. Lorsque Wnt se lie à son récepteur, elle inhibe l'ubiquitination et la dégradation de la β-caténine. Cette dernière s'accumule et est importée dans le noyau où elle active la transcription de gènes spécifiques. DDX3 interagit avec la kinase CK1ε (*Casein Kinase 1 ε*) dont elle stimule l'activité. En conséquence, CK1ε phosphoryle DVL (Dishevelled), ce qui stimule sa liaison au récepteur FZD et contribue à la dissociation du complexe de destruction. DDX3 est donc requise pour inhiber la dégradation de la β-caténine (Cruciat et al., 2013). L'effet de DDX3 sur CK1ε est indépendant de l'hydrolyse de l'ATP. DDX3 contrôle également Rac1, qui stabilise aussi la β-caténine. A la différence, l'effet de DDX3 sur Rac1 s'établit au niveau de la traduction. En effet, en absence de DDX3, la production de Rac1 est diminuée sans que l'abondance de son transcrit soit modifiée (Chen et al., 2015).

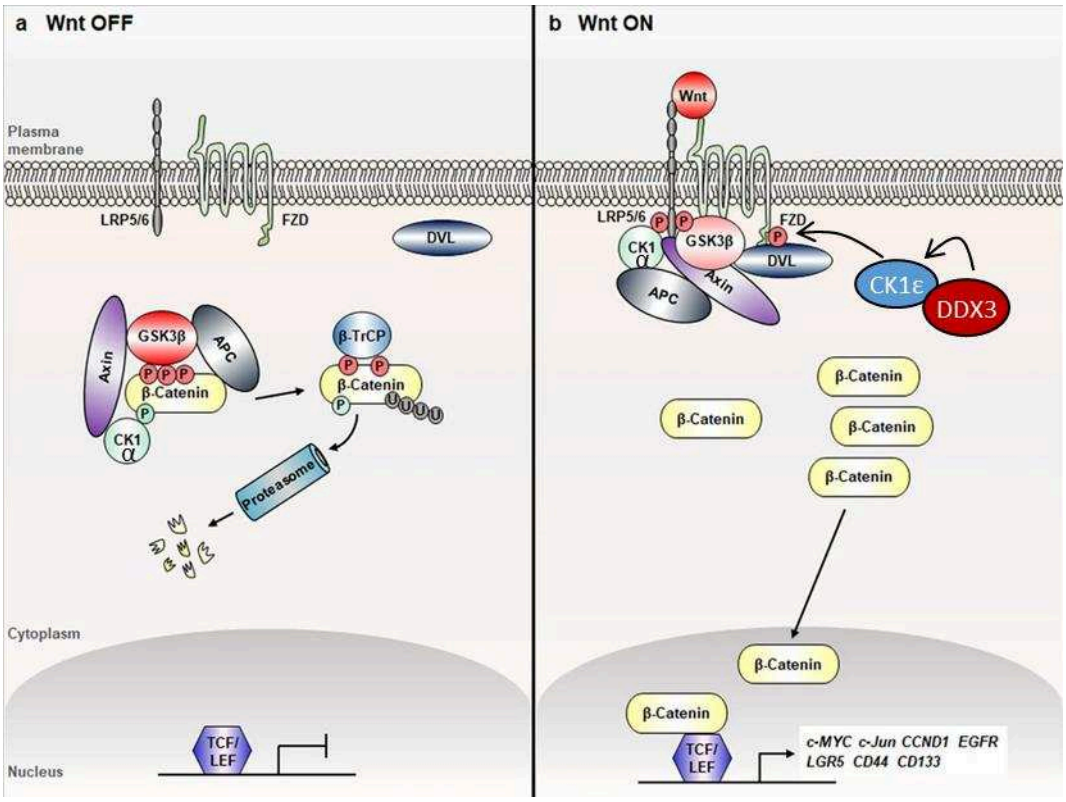


Figure 28 ***: La voie Wnt / β*-caténine canonique.** *En absence de Wnt, la β*-caténine est recrutée *au sein du complexe de destruction formé par APC, l'axine et les kinases GSK3β et CK1α. La phosphorylation de la β*-caténine recrute la ligase *β-TrCP qui ubiquitine la β*-caténine et *l'adresse ainsi au protéasome. En présence de Wnt, le récepteur LRP5/6 et DVL sont* phosphorylés recrutent la dissociation du complexe de destruction. DDX3 stimule la *phosphorylation de DVL en activant la kinase CK1ε. La β*-caténine est alors importée dans le *noyau où elle stimule l'expression de gènes en s'associant à TCF ou LEF. Adapté de* (Jeong et al., 2018)

Immunité innée

Un des mécanismes de l'immunité innée des cellules eucaryotes implique la reconnaissance de motifs moléculaires associés aux pathogènes (PAMP), tels que de l'ADN cytoplasmique, des ARNm non coiffés ou des ARN double-brin (Wu and Chen, 2014). La reconnaissance de ces motifs par des récepteurs de reconnaissance des motifs (PRR) induit une réponse interféron (INF), c'est-à-dire une cascade d'évènements moléculaires induisant la sécrétion d'INF de type I (α ou β) à l'origine d'une réponse antivirale. Les ARN viraux sont détectés par la famille des récepteurs RIG-I-like (RLR), à savoir RIG-I, MDA5 et LGP2 qui sont des hélicases ARN de la famille DExD/H (Yoneyama, Fujita 2007). Les ARN reconnus par RIG-I sont des ARN double-brin comportant une extrémité 5'-tri-, di- ou monophosphate (non coiffé), 5'-hydroxyle ou 3'-monophosphate, ainsi que des ARN circulaires. Les substrats de MDA5 sont moins bien connus, mais il se pourrait qu'ils présentent de longs segments

d'ARN double-brin. La fonction de LGP2 pourrait être de réguler l'activité de RIG-I et MDA5. La reconnaissance d'un PAMP par le domaine C-terminal d'un RLR active son domaine CARD (Caspase Recruitment Domain) en N-terminal qui interagit alors avec le domaine CARD de MAVS (Mitochondrial Antiviral Signaling protein), une protéine mitochondriale, qui polymérise à la surface des mitochondries. Les polymères de MAVS recrutent des facteurs associés au récepteur du TNF (TRAF, Tumor Necrosis Factor Receptor Associated Factors), qui sont des ligases E3 et qui synthétisent en réponse de chaînes d'ubiquitine. Ces chaînes d'ubiquitines sont reconnues par la protéine NEMO (NF-*kβ Essential Modulator*) qui recrute les kinases IKKα/β/ε (*IκB Kinase*) et TBK-1 (TANK Binding Kinase 1) qui à leur tour activent par phosphorylation le facteur régulateur 3 ou 7 de l'IFN (IRF3/7) et le facteur NF-kB (Nuclear Factor k B). Ces facteurs de transcription activent l'expression des gènes de la réponse IFN-I et de cytokines (Figure 29).

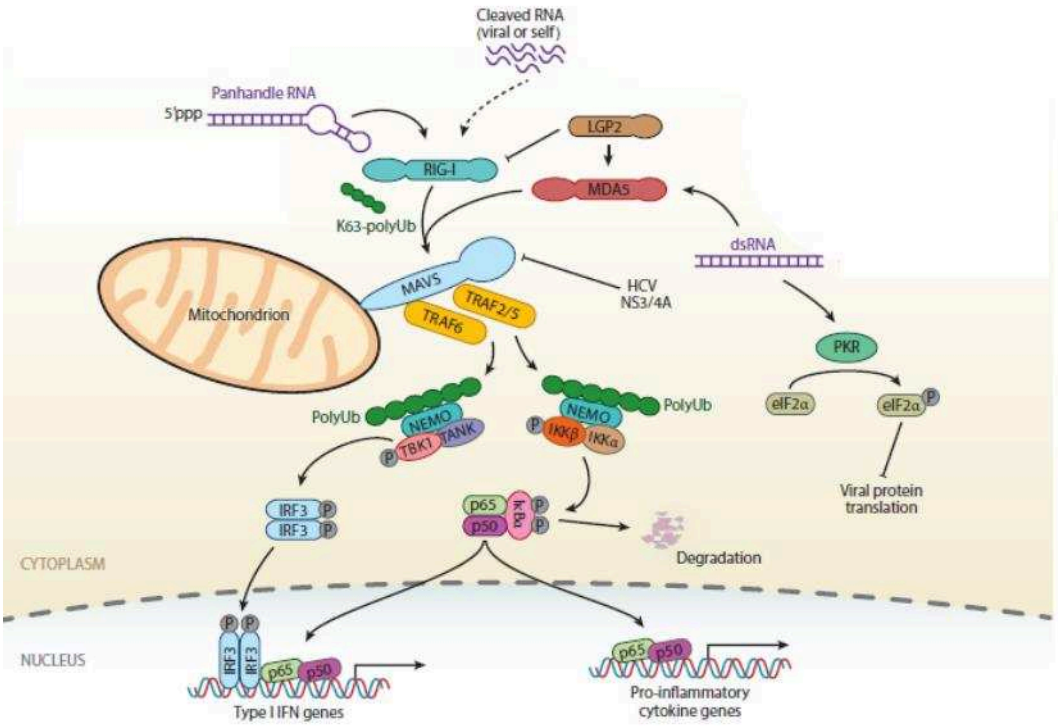


Figure 29 : Reconnaissance des ARN de pathogènes dans le cytoplasme par les récepteurs de la famille RIG. *La reconnaissance d'un PAMP par un PRR entraîne la polymérisation de MAVS à la surface des mitochondries et l'activation des ligases TRAF2/5 et TRAF6 qui* synthétisent des chaînes de polyubiquitine. Ce signal *entraîne la phosphorylation d'*IRF3 et *IκBα. IκBα est dégradé et libère le facteur* NFkB (p65-p50). IRF3 et NFkB sont alors importés dans le noyau où ils activent des gènes de la réponse IFN-I et des cytokines pro-inflammatoires. Adapté de (Wu and Chen, 2014).

DDX3 intervient dans la réponse IFN- I. En effet, DDX3 interagit avec IKKε (Gu et al., 2013), TBK-I (Soulat et al., 2008) et MAVS (Oshiumi et al., 2010) et entraîne ainsi l'activation du promoteur de l'IFN-I. L'effet de DDX3 sur la production d'IFN-I est indépendant de son activité ATPase (Soulat et al., 2008). Il est possible que DDX3 ait un rôle similaire à celui de RIG-I, en amont de IKKε/TBK-I (Oshiumi et al., 2010). Il est aussi possible que l'interaction de DDX3 avec IKKε ou TBK-I entraîne la phosphorylation de DDX3 au niveau de sérines en N-terminal. Cette phosphorylation pourrait permettre à DDX3 de recruter IRF3 (Gu et al., 2013), voire de stimuler directement le promoteur de l'INF-I (Soulat et al., 2008).

Ces résultats semblent toutefois contredits par un autre rapport montrant que DDX3 interagit avec la sous-unité p65 du facteur NFkB et réprime ainsi son activité transcriptionnelle. L'effet de cette répression transcriptionnelle se limite cependant à la production des transcrits IL-6 et IL-8 (Xiang et al., 2016).

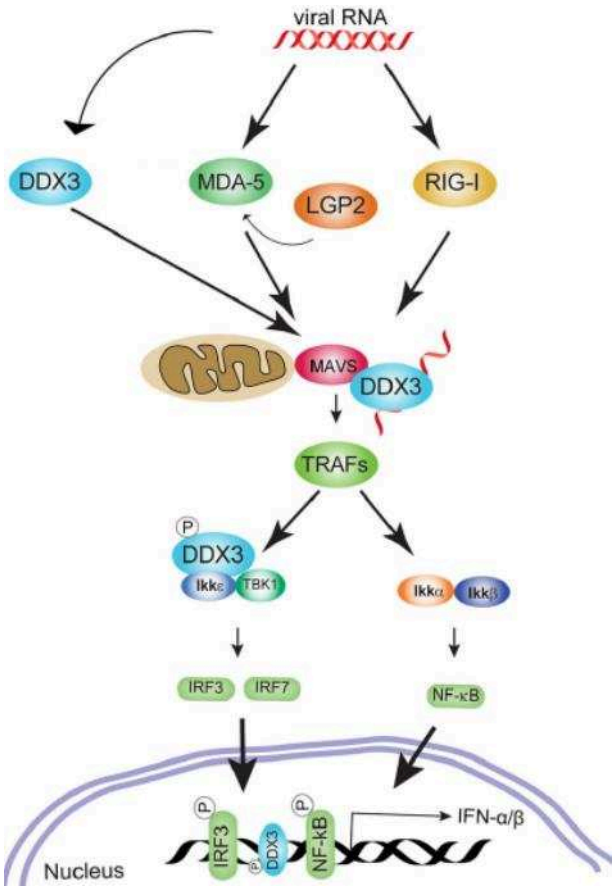


Figure 30 : Rôles possibles *de DDX3 dans l'immunité anti*virale. DDX3 pourrait être un récepteur *d'ARN viraux* et déclencher la cascade de signalisation anti-*virale en s'associant à* MAVS, *être phosphorylée par IKKε* et *recruter l'IRF3*, ou encore promouvoir directement le *promoteur de l'INF*-I. Adapté de (Valiente-Echeverría et al., 2015)

Un autre rôle possible de DDX3 dans l'immunité innée est lié au phénomène d''interférence ARN. Au cours de cette réponse immunitaire innée, les ARN double brin sont clivés pour produire des petits ARN interférents d'une vingtaine de nucléotides sous l'action de la protéine Dicer. Ces ARN interférents s'hybrident avec l'ARNm comportant la séquence complémentaire et forment une structure reconnue par le complexe RISC qui clive l'ARNm, conduisant à sa dégradation complète. Ainsi les ARN viraux sont utilisés pour s'autodétruire. La suppression de l'expression de DDX3 entraîne une diminution de la réponse par interférence ARN. De plus, dans des cellules HeLa, DDX3 est colocalisée avec la protéine Ago2 du complexe RISC (RNA-Induced Silencing Complex) (Kasim et al., 2013).

iii. DDX3 et le développement de cancers

Un nombre croissant d'étude met en lumière les divers rôles joués par DDX3 au cours de processus tumoraux. DDX3 est décrit, suivant les systèmes, comme un suppresseur de tumeur ou comme un oncogène (Zhao et al., 2016).

Plusieurs rapports décrivent la surexpression ou la mutation de DDX3 comme des facteurs carcinogènes. DDX3 est surexprimée dans certains cancers du poumon, des tumeurs colorectales ainsi que des carcinomes hépatocellulaires (HCC) (Bol et al., 2015a). Par ailleurs, dans un modèle de HCC, l'expression ectopique de DDX3 entraîne la capacité des cellules à proliférer au contact d'une matrice extracellulaire, une propriété caractéristique des tumeurs métastatiques. De façon cohérente, DDX3 inhibe l'expression de la E-cadhérine chez un modèle cellulaire de cancer du sein, et elle est requise pour l'expression de Snail, un facteur nécessaire à la transition épithélium – mésenchyme qui caractérise le processus tumoral. De plus, comme présenté plus haut, DDX3 stimule la traduction de la cycline E1 et elle favorise ainsi la transition G1 -> S. Son effet anti-apoptotique du à son association aux récepteurs de mort a également été vu (revu dans (Fuller-Pace, 2013)). Enfin, DDX3 est surexprimée dans des cancers du foie. Le recours à un inhibiteur de DDX3 a un effet anti-tumoral sur un modèle murin de cancer du foie, possiblement en inhibant la voie Wnt/β-caténine (Bol et al., 2015b).

D'autres rapports, au contraire, illustrent l'effet anti-tumoral de DDX3. Chez des patients infectés par le virus de l'hépatite C et présentant un HCC, l'expression de DDX3 est réduite. Une étude clinique montre également qu'une faible expression de DDX3 est associée à un mauvais pronostic de cancer colorectal (Su et al., 2015). De plus, la suppression de l'expression de DDX3 dans une lignée cellulaire non transformée favorise la prolifération cellulaire. L'expression de DDX3 est corrélée à celle de p53 dans un autre type de cancer du

poumon (Wu et al., 2014). Ces résultats sont cohérents avec la stimulation de la transcription et de la traduction de p53 par DDX3 décrite précédemment. Une autre étude rapporte que DDX3 inhibe la prolifération cellulaire de manière dépendante de $p21^{waf1/cip1}$.

En conclusion, le rôle de DDX3 dans le développement des cancers est complexe. Plusieurs facteurs peuvent expliquer la divergence des résultats décrits dans la littérature. La fonction de DDX3 dépend en effet des facteurs cellulaires qui lui sont associés, des modifications post-traductionnelles qu'elle subit, et par conséquent de sa localisation subcellulaire – autant de facteurs qui sont susceptibles de varier d'un type de cancer à un autre (Fuller-Pace, 2013).

iv. Rôle de DDX3 lors d'infections virales

Enfin, DDX3 est associée à plusieurs infections virales (Ariumi, 2014; Valiente-Echeverría et al., 2015). Outre son rôle dans l'infection par le VIH-1, qui sera détaillé plus loin, des rôles de DDX3 ont été décrits dans les infections par le virus de l'hépatite C (VHC) et B (VHB), le virus du Nile occidental (WNV), le virus de l'encéphalite japonaise (VEJ), le norovirus, le pestivirus, le virus de la Vaccine et le cytomégalovirus.

DDX3 participe à la réplication de plusieurs flaviviridae. Ainsi, elle interagit directement avec la protéine core du VHC, une protéine nécessaire à l'encapsidation du virus. Cette interaction se fait via le domaine C-terminal de DDX3. La réplication du VHC est réduite en absence de DDX3, mais cet effet semble restreint à un sous-ensemble. L'expression de la protéine core du VHC entraîne la relocalisation de DDX3 autour des gouttelettes lipidiques, qui sont des sites d'assemblage du virus (Ariumi, 2014). L'infection par le WNV et par le VEJ, deux autres flaviviridae, entraîne également la relocalisation de DDX3 au niveau des sites d'assemblage du virus. DDX3 interagit notamment avec les protéines virales NS3 et NS5 du VEJ ainsi qu'avec les régions 5' et 3' non traduites de l'ARN viral et stimule sa traduction. DDX3 est aussi nécessaire à la réplication du norovirus, qui appartient à la famille des caliciviridae. Les caliciviridae, comme les flaviviridae, sont des virus à ARN simple brin de polarité positive. DDX3 a été identifiée parmi les protéines cellulaires associées à l'ARN du norovirus, et là encore elle est retrouvée au niveau des sites de réplication au cours de l'infection (Valiente-Echeverría et al., 2015).

A l'inverse, DDX3 inhibe la réplication du VHB, un epadnavirus dont le génome est composé d'une molécule d'ADN circulaire double-brin. La réplication du VHB implique une

étape de transcription inverse. DDX3 interagit avec la polymérase virale et inhibe la transcription inverse de l'ARN viral. DDX3 inhibe également la transcription à partir du promoteur du VHB, indépendamment de son activité hélicase (Ariumi, 2014).

v. Conclusion

DDX3 intervient dans de nombreux processus moléculaires au sein de la cellule. En particulier, DDX3 est impliquée dans la signalisation antivirale et dans la régulation du cycle cellulaire. Il semble donc assez clair que les virus, comme les autres pathogènes intracellulaires, évoluent sous la forte pression sélective de DDX3. De fait, le cycle réplicatif du VIH comporte des mécanismes qui bloquent l'effet de DDX3, ou au contraire détournent ses fonctions cellulaires. A ce titre, DDX3 est à la fois un facteur antiviral et proviral.

4) Le VIH détourne les fonctions cellulaires de DDX3 lors *de l'infection*

i. Export de l'ARN viral.

Une fois intégré dans le génome de l'hôte, le génome du VIH est transcrit par la Pol II cellulaire comme n'importe quel ARNm (Purcell 1993). Parmi la centaine de transcrits différents retrouvé dans le cytoplasme, certains sont totalement épissés, d'autres non. Les transcrits totalement épissés suivent la voie d'export dépendante de TAP à l'instar des ARNm cellulaires. En revanche, la protéine TAP ne reconnaît pas les transcrits non épissés, qui ne présentent pas d'EJC, tels que l'ARN génomique. Comme présenté plus haut, ces derniers sont exportés suivant un mécanisme dépendant de la protéine virale Rev (Cullen, 2003). Rev reconnaît le motif RRE présent dans les ARN viraux partiellement ou non épissés, et interagit avec CRM1 via un motif riche en leucine. Des auteurs ont montré que la surexpression de DDX3 augmente l'accumulation cytoplasmique d'un transcrit issu d'un gène rapporteur contenant le motif RRE en présence de Rev uniquement (Yedavalli et al., 2004). Le virus détourne donc la machinerie d'export des ARNm dépendante de CRM1 et DDX3 via la protéine Rev pour exporter des ARNm non épissés, échappant ainsi au contrôle cellulaire.

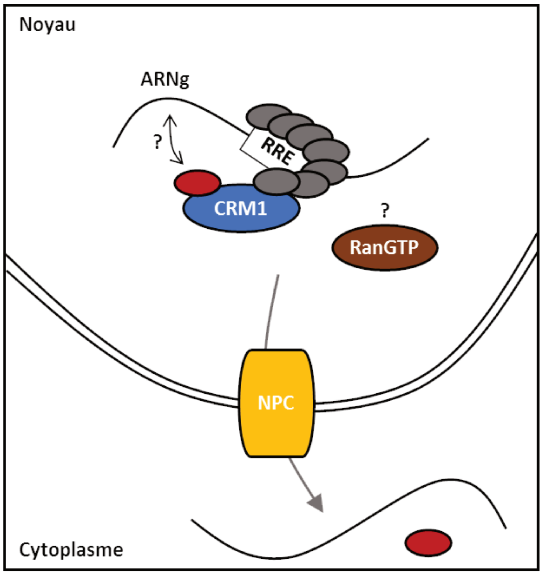


Figure 31 : Export des ARN du VIH partiellement ou pas épissés via CRM1 et DDX3 (en rouge). Seules les interactions CRM1 – Rev et DDX3 – CRM1 ont été mises en évidences. L'implication de RanGTP, associé au mécanisme canonique d'export par CRM1, n'est pas décrite dans le cas de l'export de l'ARNg du VIH. Dans ce complexe, le mode d'interaction de DDX3 avec l'ARN viral n'est pas connu.

ii. Traduction de l'ARN génomique.

Comme présenté plus haut, la 5'UTR des transcrits du VIH est longue et très structurée. Elle présente à son extrémité 5' une tige-boucle de 57 nucléotides remarquablement stable, TAR. Cette structure inhibe fortement la traduction in vitro, vraisemblablement en rendant la coiffe inaccessible à la protéine de liaison à la coiffe eIF4E ou au ribosome (Soto-Rifo et al., 2012c). L'inhibition de la traduction induite par TAR observée in vitro et atténuée dans des cellules, ce qui suggère qu'il existe des facteurs cellulaires capable de lever cette inhibition (Ohlmann et al., 2014).

Plusieurs études ont montré que DDX3 stimule la traduction des transcrits du VIH (Lai et al., 2013; Liu et al., 2011; Soto-Rifo et al., 2012b). La surexpression de DDX3 entraîne l'augmentation de la synthèse de Gag ; à l'inverse, dans des cellules où l'expression de DDX3 est réduite par interférence ARN, la production de Gag est diminuée. Dans ces expériences, l'ARN génomique est transfecté dans le cytoplasme des cellules. Les résultats montrent donc que la stimulation de la traduction par DDX3 est indépendante de son rôle dans l'export nucléaire de l'ARN viral non épissé. Le recours à un transcrit bicistronique révèle que DDX3 stimule principalement l'initiation de la traduction au niveau de l'IRES de la 5'UTR (Liu et al., 2011). Concernant l'initiation de la traduction de Gag dépendante de la coiffe, il est possible

que DDX3 soit impliquée dans la levée de l'inhibition due à TAR observée in vitro. En effet, dans des cellules n'exprimant pas DDX3, la traduction d'un gène rapporteur sous contrôle de la 5'UTR du VIH-1 est inhibée ; en revanche, l'ajout de 15 nucléotides en amont de TAR suffit à rendre la traduction du rapporteur indépendante de DDX3 (Soto-Rifo et al., 2012b). Par ailleurs, les auteurs montrent que DDX3 a de l'affinité pour la coiffe et interagit avec TAR. Ils proposent un modèle où DDX3 déstabilise TAR et recrute la machinerie d'initiation de la traduction en se substituant à eIF4E (Soto-Rifo et al., 2013). Enfin, une autre étude suggère que DDX3 stimule la traduction de Gag en interagissant avec Tat. En effet, Tat est associés aux complexes d'initiation et aux polysomes uniquement lorsque DDX3 est exprimée (Lai et al., 2013). Il est possible que DDX3 soit recrutée au niveau de TAR par Tat, mais le mécanisme exact de la stimulation de la traduction par DDX3 reste incertain. Enfin, le rôle de DDX3 dans la traduction de l'ARNg n'est probablement pas indépendant de son rôle dans l'export de l'ARNg. En effet, l'extension N-terminale de DDX3 inhibe la traduction de l'ARNg de façon dépendante de CRM1 (Fröhlich et al., 2016).

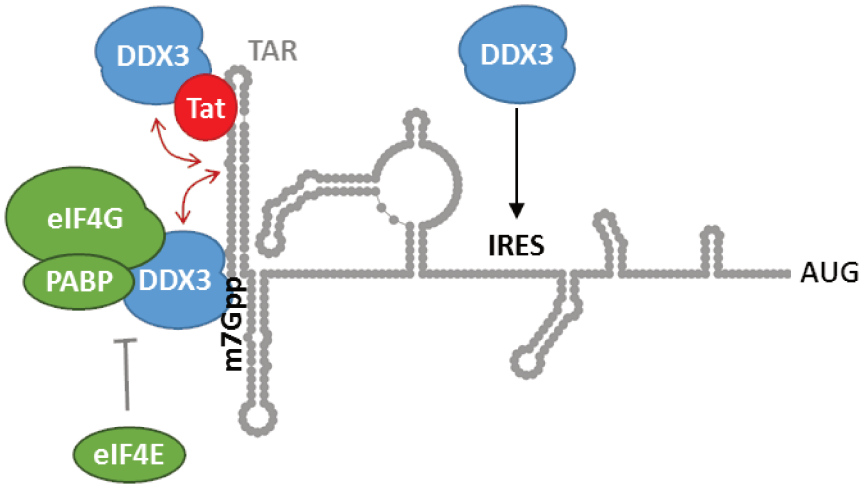


Figure 32 : *DDX3 stimule la traduction de l'ARN génomique du VIH-1*. *La 5'UTR est* représentée en gris. Les flèchent rouges indiquent *l'effet* possible de DDX3 sur la structure de TAR.

On peut finalement envisager que DDX3 intervienne dans le mécanisme d'inhibition de la réponse NMD (Nonsense Mediated Decay). L'ARN génomique du VIH-1 est en effet un substrat propice à la dégradation par cette réponse cellulaire qui assure à la cellule de ne pas traduire des ARNm présentant un défaut d'épissage ou des mutations non-sens. DDX3 est associée à un complexe Rev – UPF1, une hélicase nécessaire à la réponse NMD. Il semble que ce complexe inhibe le recrutement des autres facteurs requis pour la dégradation de l'ARN génomique (Ajamian et al., 2015).

iii. Inhibition de la signalisation antivirale

Lors de l'infection de cellules dendritiques par le VIH, la signalisation antivirale médiée par DDX3 est inhibée (Gringhuis et al., 2016). DDX3 interagit avec les transcrits abortifs du VIH et interagit avec MAVS, possiblement via son domaine N-terminal. En général, comme décrit plus haut, MAVS active la réponse IRF3 à l'origine d'une sécrétion d'IFN-1. Il semble cependant que le virus puisse contrôler la réponse immunitaire puisque l'activation de MAVS par DDX3 échoue à entraîner une telle réponse. En effet, la kinase PLK-1 devient activée suite à l'activation du récepteur membranaire DC-SIGN par la protéine virale gp120. PLK-1 activée se lie à MAVS et inhibe son interaction avec TRAF-3 (TNF Receptor Associated Factor 3) ainsi que l'activation consécutive de la réponse IRF3. De façon importante, la réponse DC-SIGN permet simultanément au virus de stimuler la transcription via l'activation du facteur d'élongation P-TEFb. Notons également que la transfection d'un transcrit abortif composé de la tige-boucle TAR seule suffit à rendre les cellules dendritiques incapables de sécréter de l'IFN-1.

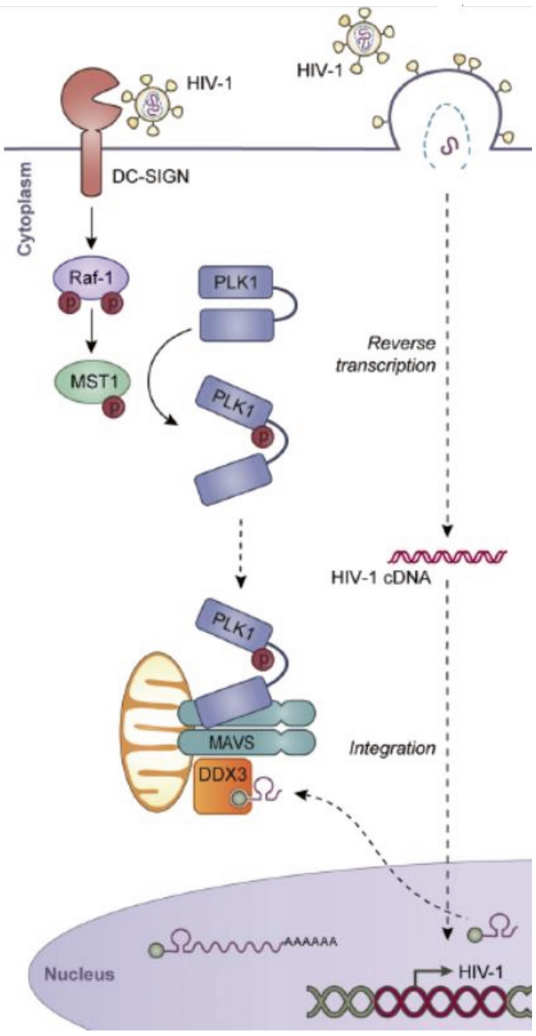


Figure 33 : Le VIH-1 inhibe la signalisation antivirale médiée par DDX3 dans les cellules dendritiques (Stunnenberg et al., 2018)

5) Conclusion

DDX3 joue de nombreux rôles dans le métabolisme des ARN, une pléiotropie qui se reflète certainement dans son implication dans plusieurs voies de signalisation cellulaire, dans la régulation du cycle cellulaire et dans des infections virales. La complexité du réseau d'interactions dans lequel elle s'inscrit explique que, suivant le modèle expérimental utilisé, on lui attribue aussi bien des rôles anti- et pro-tumoraux comme anti- et pro-viraux. DDX3 intervient au cours de différentes étapes du cycle de réplication du VIH-1 et tout particulièrement au niveau de l'expression de son ARN génomique. Comprendre les bases moléculaires de cette interaction revêt une importance décisive dans la perspective d'une utilisation comme cible thérapeutique contre le VIH, mais aussi de la compréhension de son rôle dans une grande variété de pathologies.

Objectifs

Dans la cellule, les hélicases à boîte DEAD jouent des rôles très divers et présentent une multitude d'ARN cibles et de protéines associées. Cependant, pour la plupart d'entre elles, on ignore ce qui détermine leur spécificité de substrat. Ainsi, DDX3 intervient dans l'épissage, l'export nucléaire d'ARN et la traduction ; elle participe à des mécanismes de signalisation cellulaire et elle est la cible de nombreux virus. En particulier, DDX3 est impliquée dans l'expression de l'ARN génomique du VIH-1 dont elle stimule l'export nucléaire et la traduction, mais les bases moléculaires de ces interactions ne sont pas connues.

L'objectif de ma thèse est de caractériser l'interaction de DDX3 avec l'ARN génomique du VIH-1 afin d'identifier des déterminants de spécificité. Au préalable, afin de décrire ce système de la façon la plus exhaustive, les aspects controversés ou mal maîtrisés concernant les propriétés biochimiques de DDX3 et la dimérisation de l'ARN génomique doivent être précisés. En effet, peu d'études ont été consacrées à la biochimie de DDX3 du fait de la difficulté d'obtenir une forme recombinante active. Il s'agit notamment de caractériser les activités hélicase et ATPase de DDX3 en utilisant des substrats modèles. D'autre part la dimérisation de l'ARN génomique du VIH-1 n'est encore que partiellement comprise. Une des raisons vient sans doute de la multiplicité des tailles et des séquences des constructions utilisées dans les différentes études qui lui sont consacrées. J'ai donc cherché à préciser le mécanisme de dimérisation au regard de ces deux aspects. Dans un second temps, l'interaction de DDX3 avec l'ARN génomique pourra être étudiée. En particulier, on pourra chercher des motifs de spécificité dans l'ARN génomique du VIH-1 en comparant le comportement biochimique de DDX3 en présence de cet ARN et en présence de substrats modèles. De plus, en utilisant la technique de SHAPE, on pourra chercher à localiser les sites d'interaction de DDX3 sur l'ARN génomique et à caractériser les changements structuraux qu'elle est susceptible d'entraîner. Enfin, l'effet de DDX3 sur la traduction in vitro peut être examiné au regard de cette description moléculaire.

Résultats

Etude structurale de l'ARN génomique du VIH-1

1) Les dimères formés peuvent être de différentes natures

i. Comparaison de différents protocoles de repliement in vitro de l'ARN

Un ARN peut adopter une multitude de structures différentes, mais généralement seules certaines d'entre elles sont fonctionnelles. Il est donc important de tenir compte de ce « RNA folding problem » (Herschlag, 1995) dans l'analyse in vitro. In vivo, le repliement des ARN est co-transcriptionel : les structures apparaissent au fur et à mesure de l'élongation de la chaîne nucléotidique, de l'extrémité 5' vers l'extrémité 3'. On peut supposer que le repliement se produise de façon similaire in vitro lorsque la synthèse d'ARN est effectuée au moyen d'une polymérase comme celle du phage T7, bien que la différence de vitesse de transcription puisse aussi avoir un effet sur le repliement. Une autre différence importante entre les repliements adoptés in vivo et in vitro réside dans la présence ou l'absence de protéines de liaison à l'ARN qui assistent le repliement (Schroeder et al., 2004). Le repliement d'un ARN transcrit in vitro peut donc être cinétiquement bloqué dans un état non fonctionnel du fait de l'absence de protéine chaperone.

Deux approches sont généralement utilisées pour permettre à l'ARN d'adopter un repliement natif in vitro. La première consiste à purifier les produits de transcription de façon non dénaturante ; la seconde inclut une étape de dénaturation puis une renaturation de l'ARN dans des conditions mimant les conditions physiologiques (Chillón et al., 2015). Autrement dit, la première respecte le repliement co-transcriptionnel tandis que la seconde entraîne la formation simultanée de toutes les structures. Aucun consensus ne se dégage quant aux protocoles de dénaturation – renaturation in vitro. Il semble au contraire que chaque ARN nécessite la mise au point d'un protocole qui permette d'obtenir le repliement le plus pertinent d'un point de vue biologique (Uhlenbeck, 1995).

J'ai donc comparé plusieurs protocoles de repliements inspirés de la littérature en vue d'étudier la dimérisation de l'ARN du VIH-1. Dans tous les cas, l'ARN une fois transcrit est précipité dans 2,5 M de LiCl, dessalé dans de l'éthanol à 70% puis resuspendu dans l'eau. Après une dernière étape de dessalage par filtration sur gel, l'échantillon est congelé à -20°C. Le premier protocole de repliement consiste à dénaturer l'ARN à 95°C pendant 3 minutes, puis à refroidir rapidement la suspension en plaçant les tubes dans un mélange de glace et d'eau. Ce

« chauffe – trempe » est supposé détruire tous les éléments de structure secondaire. On ajoute du magnésium avant d'incuber l'ARN 10 minutes à 37°C pour permettre la formation des structures secondaires et tertiaires (Dubois et al., 2018b). Le groupe de M. Summers rapporte que dans certains cas, l'absence de chauffage / refroidissement rapide donne des résultats similaires (Tran et al., 2015). J'ai donc également essayé cette approche, bien que le protocole utilisé pour la purification après transcription diffère. La troisième approche consiste à dénaturer l'ARN et à laisser la suspension refroidir lentement avant d'ajouter du magnésium et d'incuber l'ARN 10 minutes à 37°C (Deforges et al., 2012). Dans cette approche, les structures secondaires sont supposées se former pendant le refroidissement tandis que le repliement tertiaire est adopté au cours de l'incubation à 37°C. Les échantillons sont ensuite analysées par électrophorèse en conditions natives, à froid et à faible puissance afin de ne pas dénaturer les dimères formés. La méthode employée pour analyser la dimérisation par électrophorèse en conditions natives est présentée en annexe.

Après dénaturation, les refroidissements lent et rapide donnent des résultats assez différents (Figure 34). Si la constante d'affinité des dimères est semblable, la fraction maximale de dimère obtenue dans la gamme de concentration explorée est nettement inférieure lorsque le refroidissement est lent. Il est possible que le refroidissement lent permette la formation de structures thermodynamiquement plus stables mais moins favorables, voire complètement réfractaires à la dimérisation. Par ailleurs, l'absence de dénaturation donne des résultats assez semblables à ceux obtenus avec dénaturation puis refroidissement rapide.

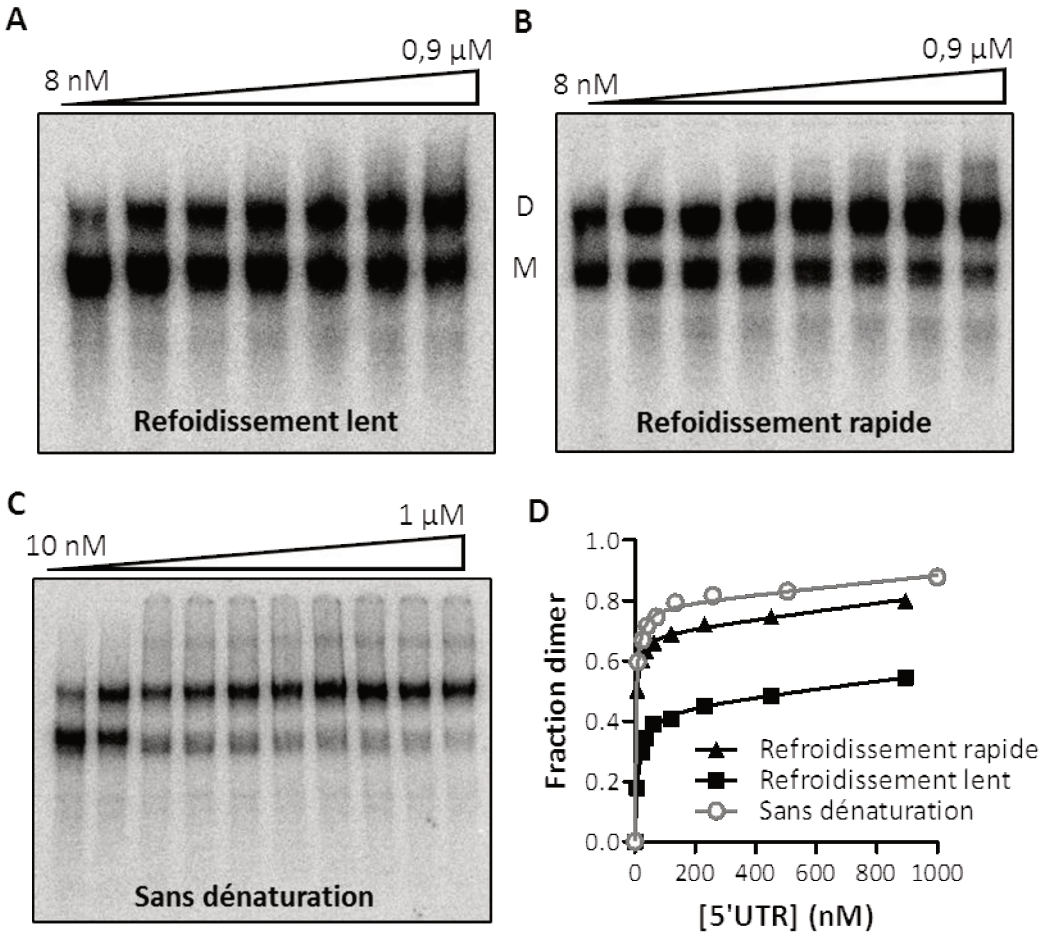


Figure 34 ***: Comparaison de l'effet de différents protocoles de repliements sur la dimérisation de la 5'UTR (1*-416) de NL4-3.** A/ Dénaturation puis refroidissement rapide (Kd = 3,1 nM), B/ Dénaturation puis refroidissement lent (Kd = 11,5 nM) et C/ Sans dénaturation (Kd = 3,5 nM). D/ Quantification de la fraction dimère. Le tampon de repliement est composé de 10 mM Tris pH7, 10 mM NaCl, 140 mM KCl et 1 mM $MgCl_2$ (Tran et al., 2015).

J'ai également comparé l'effet de la dénaturation sur le profil de réactivité SHAPE et sur l'allure de la courbe de fusion de l'ARN (Figure 35). Le principe de la technique de SHAPE et la méthode utilisée sont précisés en annexe. J'ai utilisé deux constructions, un ARN correspondant à la 5'UTR suivie des 81 premiers nucléotides de la région codante (1-416) et un ARN comprenant la 5'UTR et les 1301 premiers nucléotides de la région codante (1-1636). En effet, il est possible que le processus de repliement de deux ARN de taille significativement différente varie du fait de contraintes topologiques. Les courbes de fusion ont été obtenues en mesurant la fluorescence émise par du RiboGreen lors d'un gradient de température de 37°C à 95°C. Cette molécule fluorescente a de l'affinité pour l'ARN simple-brin, et son rendement quantique augmente lorsqu'elle est liée à l'ARN. Lors du gradient, la fusion des structures conduit à une augmentation du nombre de sites accessibles au RiboGreen, et donc à une

augmentation de l'intensité de fluorescence. Parallèlement, l'augmentation de la température déstabilise l'interaction entre le RiboGreen et l'ARN et contribue à diminuer la fluorescence. Si on calcule la variation de la fluorescence en fonction de la température, on obtient des courbes tout à fait similaires aux courbes de fusion traditionnelles obtenues en mesurant l'absorbance UV (Silvers et al., 2015). Pour chacune des constructions, on constate que les courbes de fusion sont très similaires en présence ou en absence de dénaturation ce qui suggère que la structure secondaire est globalement identique. A l'inverse, concernant les réactivités, l'effet semble différer selon la construction utilisée. En effet, la corrélation entre les réactivités obtenues avec ou sans dénaturation est significativement meilleure pour la construction qui inclut la région codante de Gag. Ces résultats suggèrent que la présence de la région codante de Gag, et pas seulement des premiers nucléotides impliqués dans l'interaction U5 – AUG, a un effet sur le repliement de la 5'UTR.

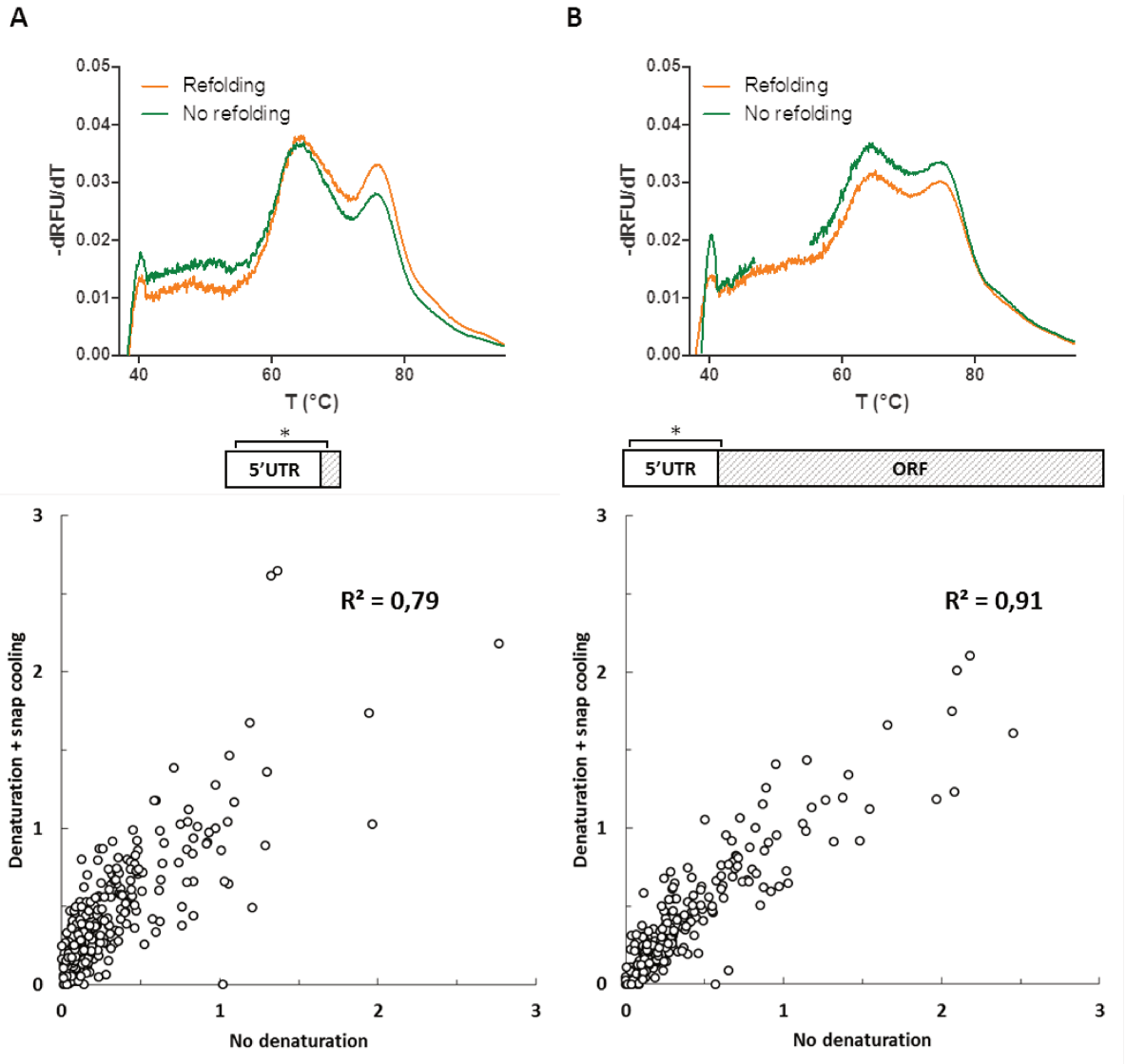


Figure 35 (page précédente) : Effet de la dénaturation sur la courbe de fusion (haut) et sur la réactivité au BzCN (bas) *pour la 5'UTR seule (1-416, A) et l'ARN '*de pleine longueur*'* (1-1636, B). Les courbes de fusion re*présentent l'opposé de la variation de la fluorescence* du RiboGreen en fonction de la température. La région examinée par SHAPE est indiquée par une étoile sur le schéma des deux constructions. Les réactivités obtenues sans dénaturation initiale sont indiquées en abscisse, et les réactivités obtenues après dénaturation et refroidissement rapide sont en ordonnée. Le tampon de repliement est composé de 40 mM Tris pH8, 50 mM NaCl, 2 mM DTT, 0,5 mM MgCl2, 0,01% NP-40.

ii. Comparaison des dimères formés par deux souches du VIH-1

J'ai cherché à comparer la dimérisation in vitro de deux souches du VIH-1 : le clone moléculaire NL4-3, dont la partie 5' correspond au sous-type B, et l'isolat MAL qui correspond au sous-type A. L'alignement des séquences des constructions utilisées révèle que les 5'UTR de MAL et NL4-3 sont très similaires, avec près de 92% d'identité. La différence majeure réside dans une insertion de 20 nucléotides dans la structure PBS de MAL (Figure 36). D'autres différences apparaissent dans des sites critiques pour la dimérisation. Ainsi, la boucle apicale de DIS correspond à 255-AAGCGCGCA-263 pour le clone NL4-3 et à 272-AGGUGCACA-280 pour l'isolat MAL. A l'inverse, la séquence de l'élément U5 à laquelle DIS s'hybride potentiellement dans la conformation monomère (105-UGUGC-109 dans NL4-3) est parfaitement conservée. Cet appariement est composé de 3 paires GC chez le sous-type B contre 2 chez le sous-type A, ce qui suggère qu'il est moins stable chez ce dernier. Toujours dans l'élément U5, on observe la mutation 112G>A (numérotation NL4-3). Les modèles de structure du dimère proposent un appariement entre la région U5 dans laquelle survient cette mutation et la région de l'AUG qui, elle, est parfaitement conservée. La mutation 112G>A transforme donc une paire GU en une paire AU. Ces deux liaisons étant de stabilité équivalente, il est donc peu probable que 112G>A affecte la stabilité de l'appariement U5-AUG. Les palindromes des boucles apicales de TAR (31-GGGAGCUCUC-40) et PolyA (77-AAGCUU-82), également suspectées d'intervenir dans la dimérisation, sont parfaitement conservés. Ainsi, le palindrome de TAR est susceptible de former un complexe kissing plus stable dans le cas du clone NL4-3 (28-CCUGGG-33) que dans celui de l'isolat MAL (27-CCCGGG-32). Du reste, quelques mutations ponctuelles affectent les séquences de TAR et PolyA. En particulier, la structure de TAR peut former 23 paires de base avec la séquence de NL4-3 contre 22 avec la séquence de MAL, avec chacune 12 paires GC. Enfin, les séquences des tige-boucles MSD et Psi sont conservées.

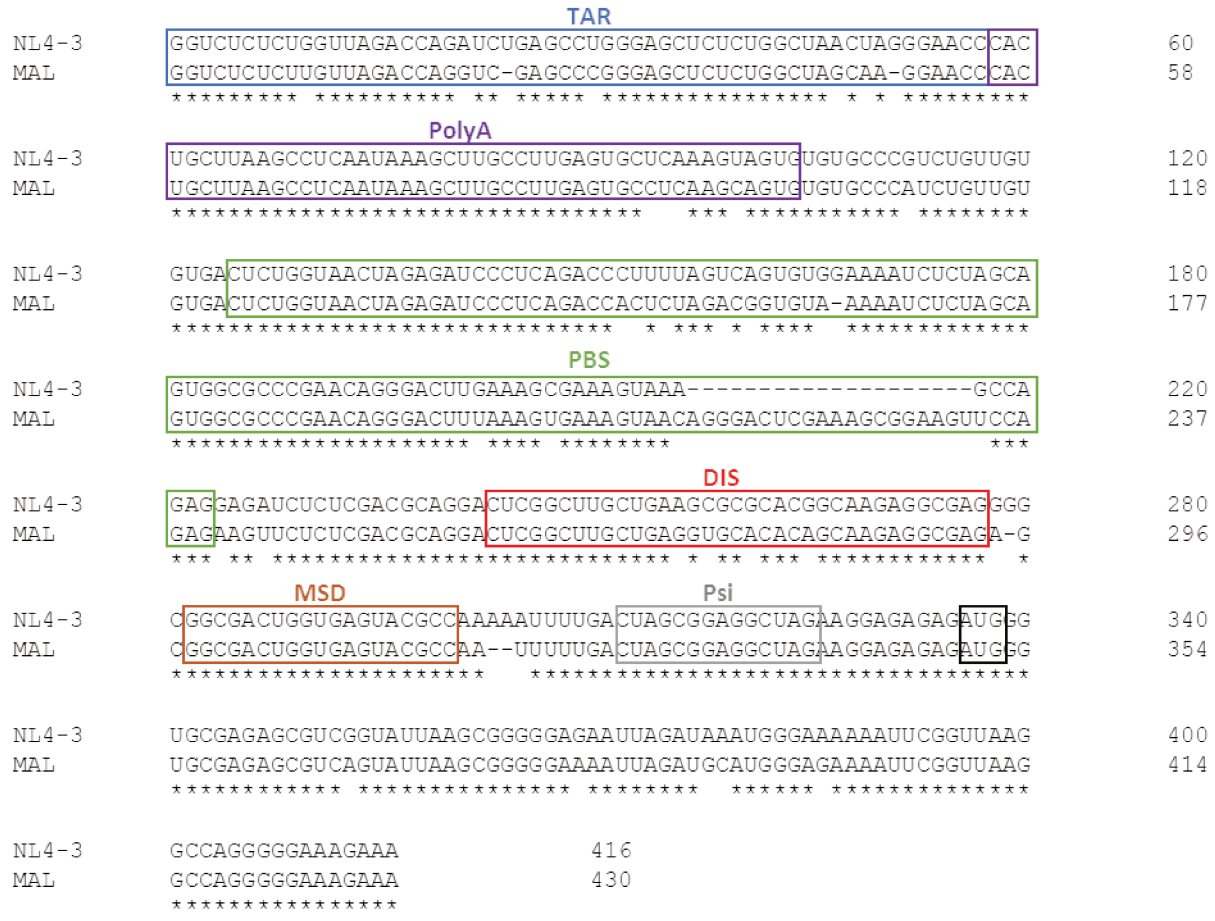


Figure 36 : Alignement des séquences de régions 5' de NL4-3 et MAL obtenu avec Clustal Omega. Les principaux éléments de structure secondaire sont indiqués par des cadres de couleur.

La capacité de ces deux isotypes à former des dimères in vitro est évaluée par électrophorèse en conditions natives. L'ARN est préalablement dénaturé puis rapidement refroidi à 4°C. Cette étape est supposée permettre la dénaturation totale de l'ARN. Après ajout de magnésium, l'ARN est incubé pendant 30 minutes à 37°C de façon à laisser les structures secondaire et tertiaire se former. Les échantillons sont analysés par électrophorèse en conditions natives comme précédemment. Pour chacun des isotypes, on observe deux bandes caractéristiques qui correspondent à la conformation monomère et dimère (Figure 37). Le dimère présente en effet un retard de migration dû à sa taille plus importante. La proportion de dimère formé augmente avec la concentration, ce qui permet de déduire une constante de dissociation du dimère. On constate que, dans ces conditions, les deux isotypes dimérisent avec une affinité similaire.

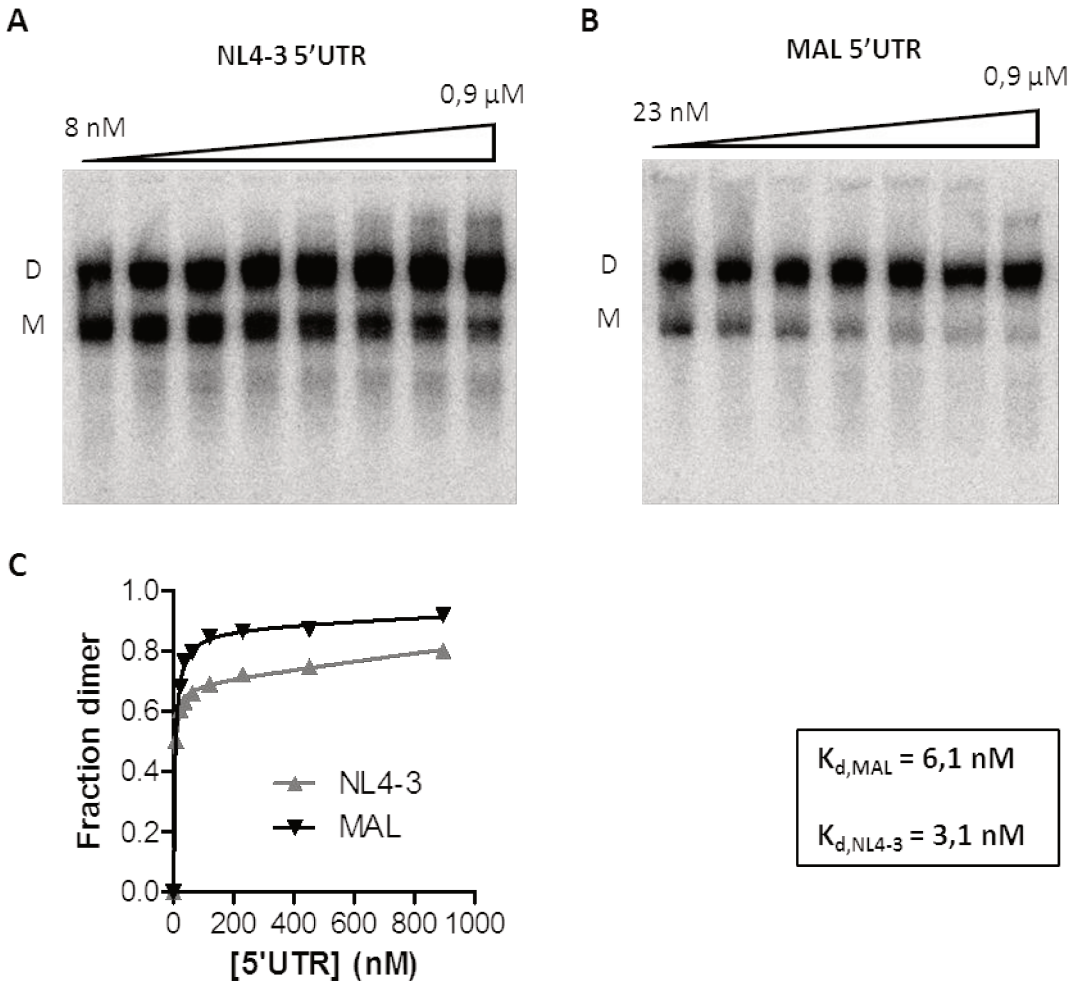


Figure 37 : Analyse des dimères de NL4-3 (A) et MAL (B) par électrophorèse en conditions natives. *Différentes concentrations d'ARN dénaturé puis renaturé sont déposé*es sur gel. La proportion de dimères formés est mesurée après électrophorèse en conditions natives. C/ La *constante d'affinité des dimères est calculée à partir de* la quantification des bandes correspondant au monomère et au dimère. Le tampon de repliement es composé de 10 mM Tris pH7, 10 mM NaCl, 140 mM KCl et 1 mM $MgCl_2$.

Nous avons vu que le processus de dimérisation in vitro est complexe et mal maîtrisé, et il n'est pas exclu que les dimères kissing puissent être convertis en dimères étendus. Cette première approche ne permet pas de distinguer ces deux types de dimères. Certains auteurs soutiennent qu'on peut distinguer deux types de dimères au moyen de gels natifs en présence ou en absence de magnésium dans le tampon de migration (Takahashi et al., 2000; Tran et al., 2015). En présence de magnésium, les dimères kissing sont présumés stables et présentent le retard sur gel attendu pour un dimère. Lorsque l'ARN est replié en présence de magnésium mais que l'électrophorèse est conduite avec un tampon de migration sans magnésium, le magnésium chélaté par les nucléotides se dissocie et migre vers la cathode, en sens contraire de

l'ARN et seul les dimères étendus sont supposés rester stables. La nature exacte de ces deux types de dimères reste cependant incertaine.

J'ai donc comparé les dimères formés par la 5'UTR de NL4-3 et celle de MAL en utilisant ces deux types de gels (Figure 38). Les constructions utilisées comprennent les 416 et 430 premiers nucléotides de l'ARN génomique, respectivement. En utilisant un tampon de repliement natif contenant entre autres du magnésium, on constate que la 5'UTR de NL4-3 forme des dimères stables sur gels sans magnésium. A l'inverse, si on utilise un tampon de repliement non natif, dépourvu de magnésium, l'ARN de NL4-3 ne forme pas de dimère stable sur des gels sans magnésium. On observe par ailleurs que la 5'UTR de MAL ne forme de dimère stable dans aucune de ces conditions. Cette dernière observation a par ailleurs déjà été effectuée à partir d'une construction plus grande (1-615) de l'isolat MAL (Marquet et al., 1994). Deux hypothèses peuvent être avancées pour expliquer ces résultats. Tout d'abord, on peut penser que la 5'UTR de NL4-3 forme spontanément des dimères étendus in vitro en présence de magnésium, alors que la 5'UTR de MAL en est incapable. Cette hypothèse est cohérente avec les résultats présentés en introduction. En effet le sous-type A (MAL) présente un site de liaison au magnésium supplémentaire au niveau de la boucle apicale de DIS (Ennifar et al., 2001; Jossinet et al., 1999); or le magnésium stabilise le dimère kissing, et la transition du dimère kissing au dimère étendu requiert la déstabilisation de DIS. Par conséquent NL4-3 pourrait être plus enclin à former des dimères étendus que MAL en absence de nucléocapside. Il est également envisageable que les dimères kissing formés par NL4-3 soient plus stables que ceux formés par MAL. En effet, 6 liaisons GC assurent l'interaction DIS – DIS chez NL4-3 contre 4 GC et deux AU chez MAL. Ces deux hypothèses seront discutées dans la dernière partie de ma thèse.

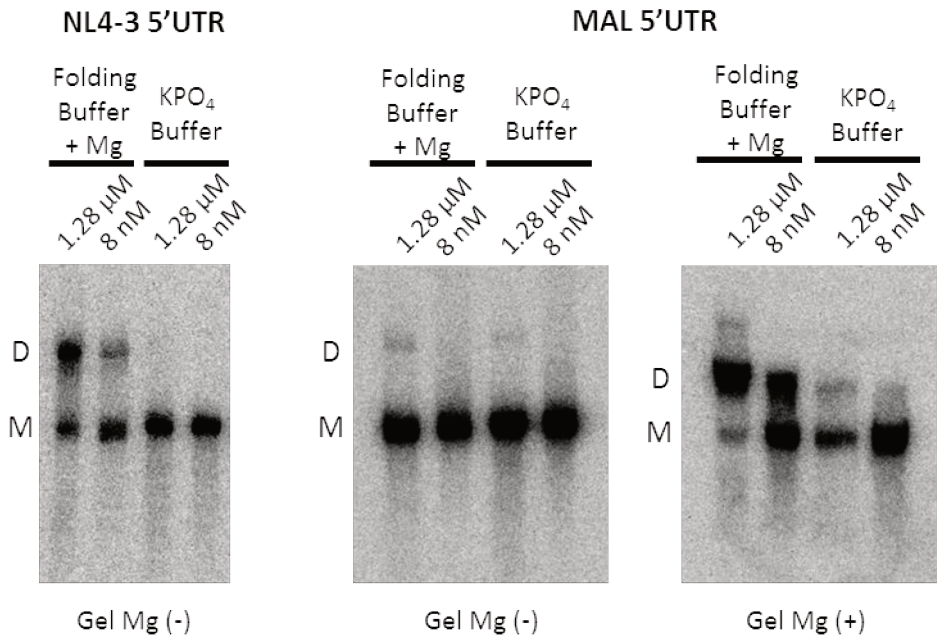
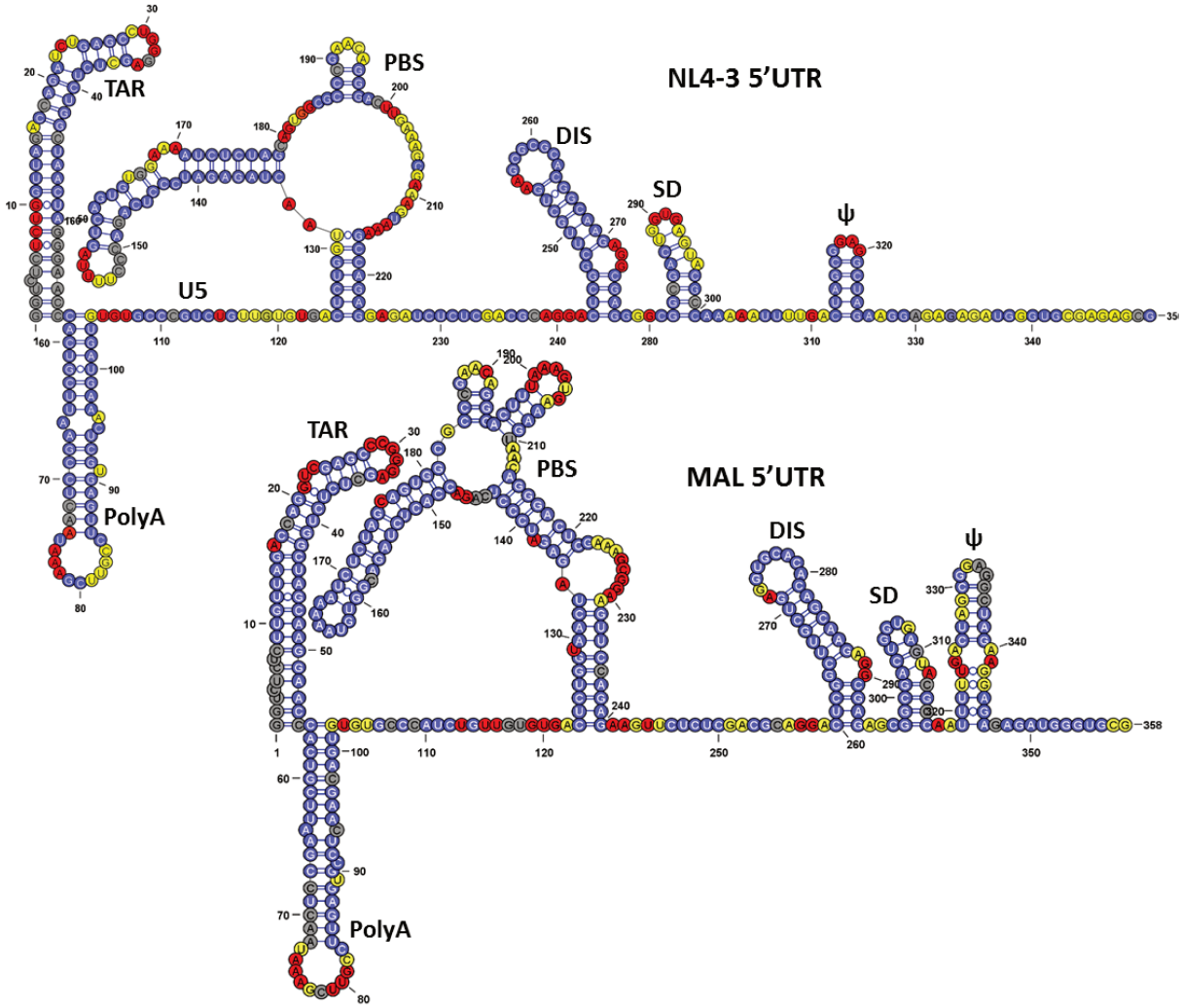


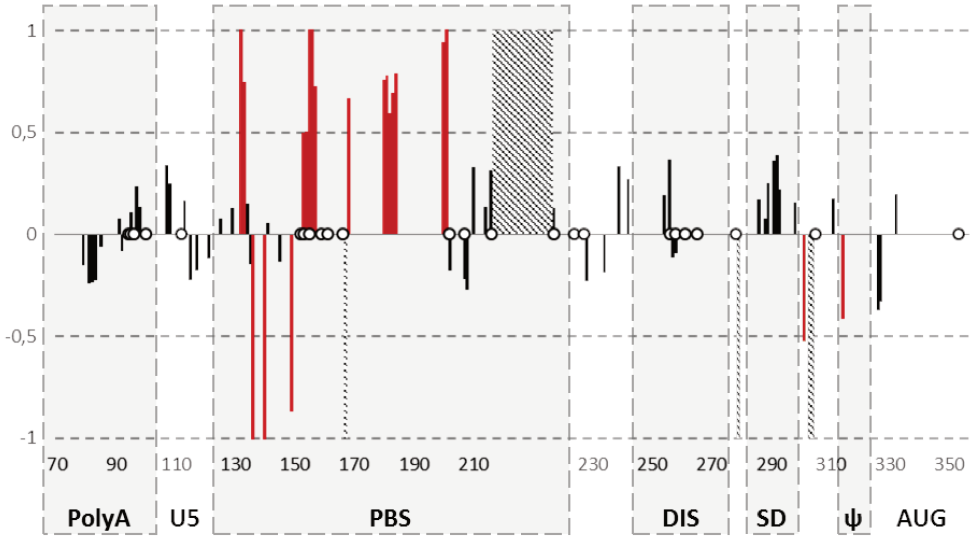
Figure 38 : Comparaison des dimères formés par NL4-3 et MAL. *L'ARN est dénaturé dans* un tampon de repliement *'natif'* puis renaturé en présence de magnésium, ou bien dénaturé et renaturé dans du tampon phosphate. Les dimères formés sont analysés par électrophorèse en conditions natives en présence ou en absence de magnésium. *Tampon de repliement 'natif'* : 10 mM Tris pH7, 10 mM NaCl, 140 mM KCl et 1 mM $MgCl_2$. Tampon phosphate : 10 mM KH_2PO_4 pH7,4.

Dans le but de proposer une explication moléculaire de ces différences, j'ai réalisé le profil de réactivité SHAPE de ces deux ARN (Figure 39). Les résultats révèlent des différences importantes tout au long de la structure. Les variations dans la séquence conduisent dans certains cas à des changements concomitants de réactivités, qui reflètent des modifications de la structure secondaire. Par exemple, C207U (numérotation NL4-3) située dans le PBS entraîne clairement la formation d'une tige-boucle formée d'une hélice de 5 paires de bases dans MAL. Cette même hélice ne comporterait que 4 paires de bases dans NL4-3, et les fortes réactivités observées indiquent qu'elle est très peu représentée in vitro. Cependant, 15 des 24 substitutions de la région analysée n'entraînent pas de changement de réactivité, et près d'une soixantaine de nucléotides conservés changent de réactivité. Cela signifie que certaines régions sont conservées, mais se replient différemment sous l'effet de mutations en trans. Par exemple, la séquence C125 – U145 (NL4-3) est parfaitement identique à la région C123 – U143 (MAL), mais s'apparie avec des séquences différentes du fait d'une insertion de 20 nucléotides entre les positions 216 et 217 (NL43). Ainsi, A132 et A133 sont très réactives dans NL4-3 alors que les bases correspondantes A130 et A131 ne sont pas réactives.

Si la structure secondaire du PBS diffère clairement entre l'ARN de MAL et celui de NL4-3, les variations observées dans les autres éléments de structure secondaire sont nettement moins importantes. Les profils de réactivité de l'élément U5 et de DIS sont globalement similaires. Les variations de réactivité observées dans la région de l'AUG ne sont pas significatives ($p_{Student}<0,05$). On remarque toutefois que des variations significatives autour de la boucle apicale de DIS, notamment en ce qui concerne les nucléotides G254 et A256 de NL4-3 qui sont plus exposés que leurs équivalents G271 et G273 de MAL ainsi que les nucléotides G257 et C258 de NL4-3 qui sont moins exposés que G274 et U275 de MAL. Ces observations sont en accord avec l'hypothèse selon laquelle l'interface du dimère est plus stable avec le palindrome GCGCGC qu'avec le palindrome GUGCAC. Elles sont également cohérentes avec la stabilisation du reste de la boucle par le magnésium. En particulier, le nucléotide G273 apparaît lier le magnésium dans les structures du dimère kissing et du dimère étendu de MAL (Ennifar et al., 1999, 2001).

En conclusion, l'isolat MAL et le clone NL4-3, représentant respectivement les sous-types A et B du VIH-1, présentent des propriétés de dimérisation différentes bien que leur structure secondaire soit globalement conservée. Dans la suite de ma thèse, j'ai utilisé exclusivement le clone NL4-3.





Légende page suivante.

Figure 39 : Comparaison des réactivités au 1M7 de la 5'UTR de NL4-3 (haut) et MAL (bas). Gris : réactivité indéterminée ; bleu : 0<R<0,4 ; jaune : 0,4<R<0,7 ; rouge : R>0,7. L'histogramme représente l'écart entre les réactivités des deux constructions : les barres positives indiquent des sites où le NL4-3 est significativement plus réactif, et inversement. Les ronds indiquent des mutations ponctuelles. Les parties hachurées correspondent aux insertions de MAL (positif) et NL4-3 (négatif). Les barres rouges correspondent aux différences de plus *de 0,4 unités de réactivité et d'un ratio supérieur à* 2. La numérotation employée est celle de NL4-3. Les différents éléments de structures secondaires sont indiqués.

2) Interdépendance du repliement de la région codante et *de la 5'UTR*

Les constructions utilisées pour étudier la dimérisation de l'ARN du VIH-1 in vitro varient d'un groupe à l'autre suivant les contraintes expérimentales. L'ARNg dans son intégralité a pu être analysé par microscopie électronique, mais à faible résolution (Höglund et al., 1997). Les approches par gels natifs et par sondage de la structure permettent d'étudier des fragments intermédiaires, mais ils restent toutefois souvent restreint à la 5'UTR (Huthoff and Berkhout, 2001; Kenyon et al., 2013; Seif et al., 2013). Les approches par RMN, cristallographie et FRET limite davantage la taille de l'ARN, et de ce fait les constructions utilisées se restreignent souvent à la tige-boucle DIS (Ennifar et al., 1999; Mujeeb et al., 1999; Mundigala et al., 2014) bien qu'un groupe ait mis au point une méthode pour analyser un fragment du site de dimérisation d'une centaine de nucléotides (Keane et al., 2015). Cependant des résultats montrent que la capacité de l'ARN à dimériser peut être modifiée par les séquences situées de part et d'autre du site de dimérisation, qu'il s'agisse de séquences natives (Skripkin et al., 1994), de gènes rapporteurs (Deforges et al., 2012) et même de quelques résidus inclus à des fins de clonage (Tran et al., 2015). Afin de proposer des éclaircissements supplémentaires sur l'influence de la construction utilisée sur la dimérisation, j'ai étudié l'interdépendance du repliement de la région codante et de la 5'UTR. Pour ce faire, j'ai utilisé trois constructions correspondant à la 5'UTR suivie des 81 premiers nucléotides de la région codante (1-416), aux 1301 premiers nucléotides du cadre de lecture de Gag (336-1636) et à l'ARN 'de pleine longueur' (1-1636) comprenant la 5'UTR suivie des 1301 premiers nucléotides du cadre de lecture de Gag (Figure 40).

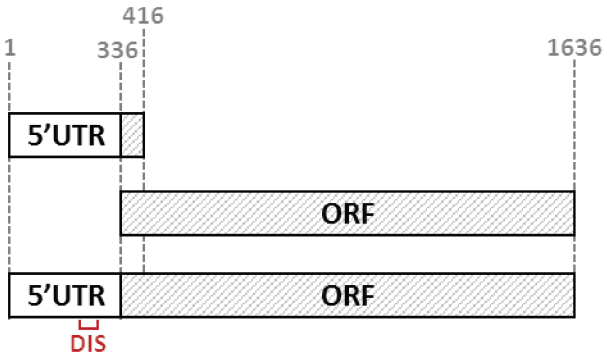


Figure 40 : Schéma des construction*s utilisées pour étudier l'interdépendance du repliement de la région codante et de la 5'UTR.* Le site de dimérisation (DIS) est indiqué en rouge.

On constate que les réactivités de la région codante présentent une bonne corrélation entre les deux constructions (Figure 41 A). A l'inverse, les réactivités de la 5'UTR sont moins bien corrélées (Figure 41 B). Ces résultats indiquent que la présence de la région codante affecte de façon importante le repliement de la 5'UTR, et que cet effet n'est pas réciproque. Une analyse plus détaillée des différences de réactivité qui surviennent dans la 5'UTR révèle que les structures TAR, PolyA et PBS sont peu affectées et que la plupart des changements significatifs de réactivité surviennent dans les structures directement impliquées dans la dimérisation (Figure 41 C). Ainsi, on observe trois baisses significatives de réactivité dans DIS, deux au niveau du palindrome de la boucle apicale et une au niveau de l'hélice, ainsi que dans la région de l'AUG au niveau de quatre nucléotides consécutifs. Ces observations sont en faveur d'une stabilisation de la structure du dimère en présence de l'ORF. A l'inverse, on observe une série de nucléotides en 3' et en 5' de la tige-boucle Psi pour lesquels la réactivité est plus forte, en position 308, 311, 326, 327 et 329. Tous les modèles de structure secondaire décrivent ces nucléotides en simple-brin. Notons toutefois que le groupe de M. Summers rapporte des signaux NOE (Nuclear Overhauser Effect) entre les nucléotides A326 et A327 d'une part, et A311 d'autre part (Keane et al., 2015). Ces signaux montrent que ces résidus sont maintenus à proximité les uns des autres, et par conséquent possiblement moins réactifs. Il est donc possible que la structure adoptée par l'ARN de pleine longueur soit plus éloignée du modèle de Summers que la structure de l'ARN ne comprenant que la 5'UTR.

Plusieurs hypothèses permettent d'expliquer ces résultats. On peut penser que, dans la construction '5'UTR', la troncation induise un repliement aberrant qui modifie le repliement du reste de l'ARN. En particulier, les résidus 363 – 367, 382 – 385, 394 – 397 et 399 – 407 sont appariés avec des résidus situés en dehors de la construction 5'UTR dans le modèle de structure

de l'ARN génomique présenté Figure 41. Par conséquent ces résidus, qui ont été intégrés à la construction pour permettre la mesure de la réactivité SHAPE sur la totalité de la 5'UTR, peuvent adopter un repliement non natif dans la construction 5'UTR. Cependant, le logiciel de prédiction de structure secondaire RNAStructure prédit pour ces nucléotides la formation d'un élément de structure secondaire indépendant du reste de la structure. De plus, le fragment 5'UTR (1-416) dimérise avec la même affinité que le fragment 1-343 (résultats non présentés), ce qui suggère que les résidus 344 à 416 n'interfèrent pas avec la structure du site de dimérisation. Une autre hypothèse repose sur l'existence d'autres sites de dimérisation possibles dans la région codante de Gag. La présence de ces sites pourrait induire une maturation du site de dimérisation de la 5'UTR. Cette hypothèse est discutée dans la dernière partie de ce manuscrit.

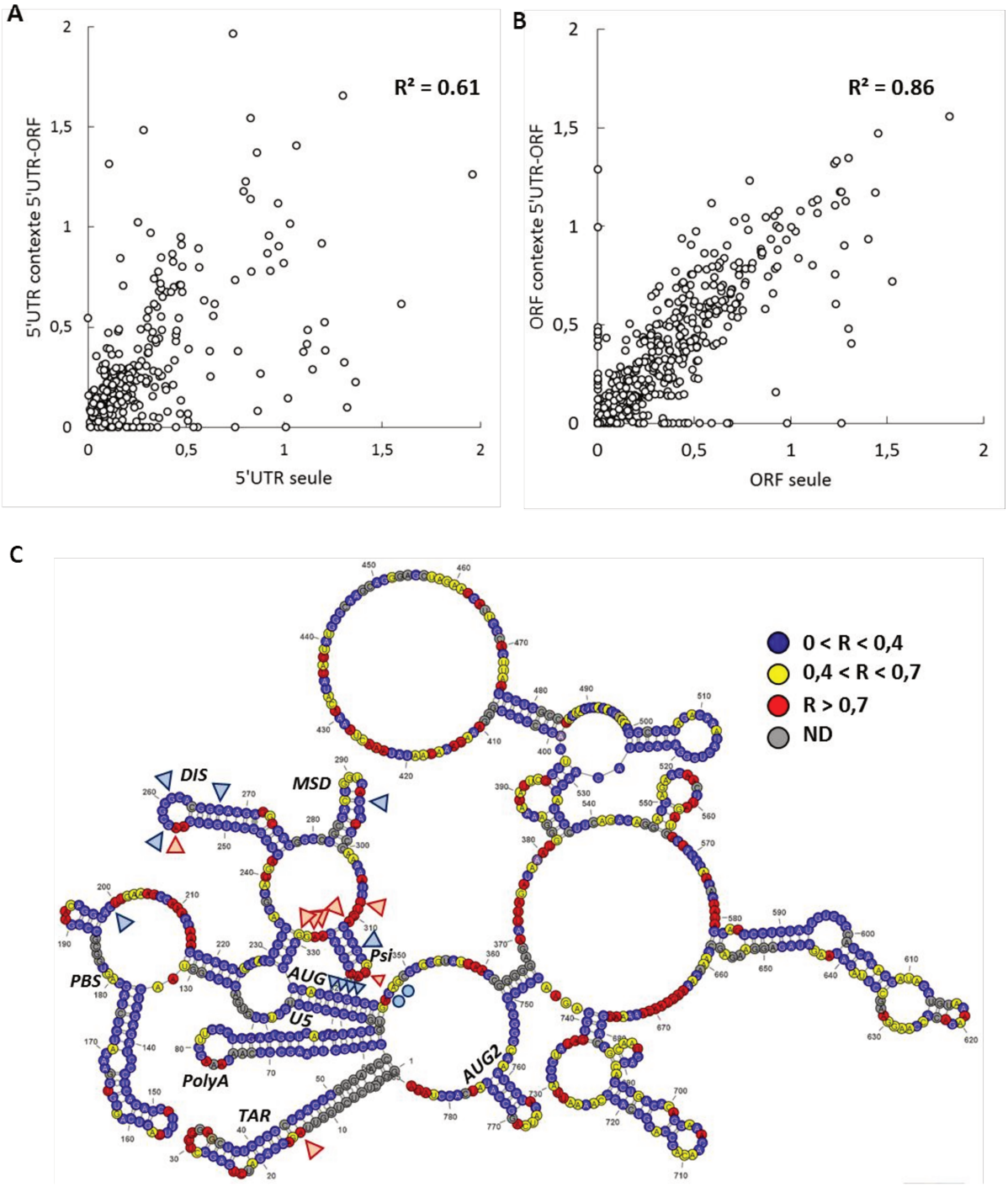


Figure 41: Comparaison des réactivités au BzCN de la 5'UTR (A) et de la région codante (B) suivant la construction utilisée. *Les réactivités de la 5'UTR et de la région codan*te sont en *abscisse, tandis que la réactivité de l'ARN de* pleine longueur est en ordonnée. Le tampon de repliement est composé de 40 mM Tris pH8, 50 mM NaCl, 2 mM DTT, 0,5 mM MgCl2, 0,01% NP-40. C/ Modèle de structure secondaire des nucléotides 1 à 785 représentant la réactivité au BzCN de chaque nucléotide *obtenue à partir de l'ARN de pleine longueur*. Les différences de réactivités les plus significatives ($p<0,05$, écart>0,4, ratio>3) *avec la construction '5'UTR' et la construction 'ORF'* sont indiquées par des flèches et des ronds, respectivement ; rouge *lorsque le nucléotide est plus réactif dans le contexte de l'ARN de pleine longueur et bleu dans* le cas contraire.

3) Effet de la méthylation des adénosines en N6 sur la dimérisation

L'ARN génomique du VIH-1 est méthylé in vivo au niveau des région 5'UTR, 3'UTR et du RRE (Riquelme-Barrios et al., 2018). L'effet de ces méthylations semble être ambivalent : les résidus m6A stimulent la dégradation de l'ARNg lors du stade précoce de l'infection, mais ils stimulent la synthèse de Gag lors des stades tardifs. Le groupe de R. Soto-Rifo a donc cherché à identifier l'effet de la méthylation des adénosines sur la traduction et l'encapsidation de l'ARNg. De façon étonnante, les méthylations sont détectées dans les cellules infectées mais pas dans les particules virales. En effet, la méthylation stimule la traduction de Gag et inhibe l'encapsidation ; en retour Gag interagit avec la déméthylase FTO et stimule la déméthylation de l'ARNg, ce qui a pour effet de favoriser l'encapsidation. Cet effet n'est pas sans rappeler l'hypothèse selon laquelle le devenir de l'ARNg est gouverné par un équilibre conformationnel, le monomère étant efficacement traduit et le dimère encapsidé. Dans le cadre de cette collaboration, nous avons donc cherché à évaluer l'effet de la méthylation des adénosines sur la structure de la 5'UTR. L'approche initiale consiste à substituer toutes les adénosines de la 5'UTR (1-416) par des adénosines méthylées et à comparer le profil SHAPE de l'ARN méthylé à celui de l'ARN non méthylé. Les résultats révèlent que la méthylation des adénosines modifie localement la réactivité, et pas uniquement au niveau de résidus A ou U, ce qui suggère que la méthylation affecte le repliement de structures particulières. Nous avons donc étudié l'effet de la méthylation sur la dimérisation in vitro, et nous montrons que la 5'UTR méthylée dimérise moins efficacement. Ces travaux font l'objet d'une publication en cours de révision, et dont la version preprint est présentée ici.

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

An epitranscriptomic switch at the 5´-UTR controls genome selection during HIV-1 genomic RNA packaging

Camila Pereira-Montecinos[1,2]; Daniela Toro-Ascuy[1,2#]; Cecilia Rojas-Fuentes[1,2]; Sebastián Riquelme-Barrios[1,2]; Bárbara Rojas-Araya[1,2]; Francisco García-de-Gracia[1,2]; Paulina Aguilera-Cortés[1,2]; Catarina Ananías-Sáez[1,2]; Grégoire de Bisschop[3]; Jonás Chaniderman[1]; Mónica L. Acevedo[1]; Bruno Sargueil[3]; Fernando Valiente-Echeverría[1,2] and Ricardo Soto-Rifo[1,2*]

[1] Laboratory of Molecular and Cellular Virology, Virology Program, Institute of Biomedical Sciences, Faculty of Medicine, Universidad de Chile, Santiago, Chile

[2] HIV/AIDS Workgroup (CHAIR), Faculty of Medicine, Universidad de Chile, Santiago, Chile

[3] CNRS UMR 8038, Laboratoire "Cibles Thérapeutiques et Conception de Médicaments" - CitCoM, Université Paris Descartes, 4 Avenue de l'Observatoire, 75270 Paris Cedex 06, France.

* To whom correspondence should be addressed:
Ricardo Soto-Rifo: Tel: (56) 2 2978 68 69; Fax: (56) 2 2978 61 24;
rsotorifo@uchile.cl

Current address: Instituto de Ciencias Biomédicas, Facultad de Ciencias de la Salud, Universidad Autónoma de Chile, Santiago, Chile.

Keywords: N^6-methyladenosine, HIV-1 genomic RNA, RNA Packaging, 5´-UTR, Gag

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

ABSTRACT

During retroviral replication, the full-length RNA serves both as mRNA and genomic RNA (gRNA). While the simple retrovirus MLV segregates its full-length RNA into two functional populations, the HIV-1 full-length RNA was proposed to exist as a single population used indistinctly for protein synthesis or packaging. However, the mechanisms by which the HIV-1 Gag protein selects the two RNA molecules that will be packaged into nascent virions remain poorly understood. Here, we demonstrate that HIV-1 full-length RNA packaging is regulated through an epitranscriptomic switch requiring demethylation of two conserved adenosine residues present within the 5′-UTR. As such, while m^6A deposition by METTL3/METTL14 onto the full-length RNA was associated with increased Gag synthesis and reduced packaging, FTO-mediated demethylation was required for the incorporation of the full-length RNA into viral particles. Interestingly, HIV-1 Gag associates with the RNA demethylase FTO in the nucleus and drives full-length RNA demethylation. Finally, the specific inhibition of the FTO RNA demethylase activity suppressed HIV-1 full-length RNA packaging. Together, our data propose a novel epitranscriptomic mechanism allowing the selection of the full-length RNA molecules that will be used as viral genomes.

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

INTRODUCTION

Retroviral full-length RNA plays two key functions in the cytoplasm of infected cells. First, it is used as the mRNA template for the synthesis of Gag and Gag-Pol precursors and, second, it serves as the genomic RNA (gRNA) packaged into newly produced viral particles [1-3]. In contrast to the simple retrovirus murine leukemia virus (MLV), which was shown to segregate its full-length RNA into two functionally different populations serving as template for translation (mRNA) or packaging (gRNA), the HIV-1 and HIV-2 full-length RNA were proposed to exist as a single population acting indistinctly as mRNA and gRNA [4-6]. However and despite several years of efforts, there is still an important gap in our knowledge regarding the molecular mechanisms behind the selection of the full-length RNA molecules that will be incorporated into assembling viral particles.

The 5′-untranslated region (5′-UTR) present within the HIV-1 full-length RNA is the most conserved region of the viral genome and contains several high order structural motifs involved in different steps of the viral replication cycle from transcription, reverse transcription, splicing, translation to dimerization and packaging [3, 7, 8]. Since the full-length RNA serves both as mRNA and gRNA, translation and packaging are expected to be mutually exclusive events [2]. The Gag protein recognizes *cis*-acting RNA elements present at the 5′-UTR and the beginning of the Gag coding region and drives the selective incorporation of two copies of the gRNA into assembling viral particles. Indeed, there is accumulating evidence showing that dimerization and packaging of the HIV-1 full-length RNA are two tightly interconnected processes dependent on the Gag protein [9-11]. Structural and mutational analyses proposed that a conformational switch within the 5′-UTR regulates the transition from translation to dimerization and packaging *in vitro* [12-15]. In such models, the 5′-UTR alternates in conformations that occlude the dimerization initiation signal (DIS) or the Gag start codon thus, favoring translation or dimerization and packaging, respectively [3, 8]. However, chemical probing performed in cells and purified viral particles showed that a single structure, in which DIS is accessible for dimerization and packaging, predominates in these biological states [16, 17]. Moreover, the packaging prone structure does not interfere with full-length RNA translation suggesting that other factors rather than structural rearrangements are involved in the regulation of the cytoplasmic sorting of the HIV-1 full-length RNA [18, 19].

It was recently reported that the HIV-1 full-length RNA contains N^6-methyladenosine (m^6A) residues located at the 5′- and the 3′-UTR as well as at internal positions such as the Rev response element (RRE) [20-22]. Methylation of adenosines at the RRE and the 3′-UTR was shown to promote Gag synthesis by favoring nuclear export and/or the intracellular

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

accumulation of viral transcripts at late stages of the replication cycle [20-23]. However, it was also reported that the presence of m^6A could also induce the degradation of the incoming gRNA early during infection [22, 23]. These controversial data prompted us to study whether m^6A could serve as a mark that defines the functions of the HIV-1 full-length RNA as template for translation or packaging during viral replication.

Here, we show that methylation of two adenosine residues within the 5′-UTR by the METTL3/METLL14 complex inhibits full-length RNA packaging. m^6A-seq analysis revealed that the full-length RNA present in purified viral particles lacks m^6A at the 5′-UTR suggesting the existence of two populations that differ in their m^6A patterns. Further bioinformatic analyses identified two highly conserved nucleotides A_{198} and A_{242} within the 5′-UTR as the key residues involved in the m^6A-mediated regulation of gRNA packaging. We also observed that the full-length RNA is a substrate for the RNA demethylase FTO, which together with Gag drives RNA demethylation to promote packaging. Finally, the pharmacological targeting of FTO activity resulted in impaired full-length RNA metabolism and a strong inhibition of packaging. Together, our data reveal a novel mechanism by which Gag selects the molecules of gRNA that will be used for packaging, which is regulated by an epitranscriptomic switch within the 5′-UTR.

RESULTS

The presence of m^6A alters the *in vitro* folding and dimerization of the HIV-1 full-length RNA 5′-UTR

Since previous studies suggesting that a conformational switch of the 5′-UTR could regulate the ability of the HIV-1 full-length RNA to function as mRNA or gRNA have not considered the presence of RNA modifications such as m^6A, we first sought to determine the impact of adenosine methylation on the folding of the 5′-UTR. As a first approach to study the impact of m^6A on RNA structure, we generated an *in vitro*-transcribed 5'-UTR containing unmodified adenosines or a 5′-UTR in which all adenosines were replaced by m^6A. Both m^6A-5'-UTR and A-5'-UTR were submitted to 1M7 SHAPE analysis in parallel as described in Methods. Reactivity towards SHAPE reagents reveals the ribose flexibility, and as a consequence, the pairing status of each nucleotide. As such, higher reactivity of a given nucleotide means a higher probability to be in a single strand conformation. Comparative analysis of the SHAPE reactivity profiles indicates that the presence of m^6A significantly alters the folding of the 5′-UTR (Fig. 1a and Supplementary Fig. 1a). The first interesting

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

observation from our SHAPE data is that we do not only observe a reactivity modification for As or Us.

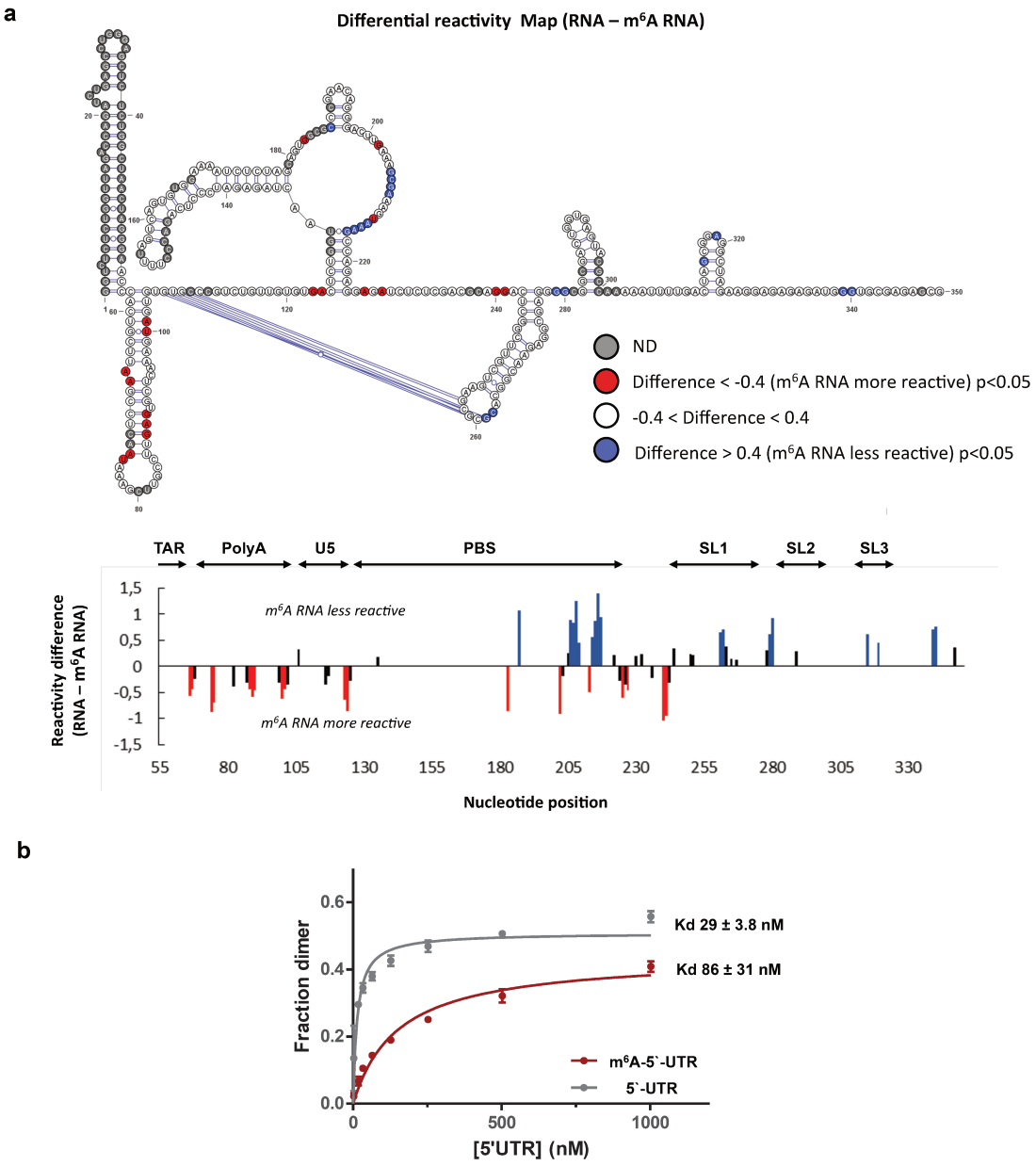


Figure 1: The presence of m^{6}A alters the *in vitro* folding and dimerization of the HIV-1 full-length RNA 5′-UTR. a, SHAPE reactivity differences between *in vitro* transcribed 5′-UTR RNA containing 0% m^{6}A or 100% m^{6}A plotted on the secondary structure model of the HIV-1 full-length RNA 5'-UTR (upper panel). Nucleotides in red are significantly more reactive with m^{6}A whereas those in blues are significantly less reactive (p-value < 0.05 and reactivity difference > 0.4). Nucleotides in white are of equivalent reactivity. Sites where the reactivity could not be determined are depicted in grey. The histogram shows the reactivity difference between 0% m^{6}A 5'-UTR and 100% m^{6}A 5'-UTR (lower). Blue and red bars highlight nucleotides with a significantly lower or higher reactivity in the 100% m6A RNA than in the 0% m6A RNA (p-value < 0.05 and reactivity difference > 0.4), respectively. Significative differences that are below the threshold of 0.4 are indicated in black. **b,** Fraction of dimers determined by electromobility shift assay for different concentrations of *in* vitro transcribed 5'-UTR (1 µM, 0.5 µM, 0.25 µM, 127 nM, 65 nM, 33 nM, 18 nM, 2 nM) harboring either 0% of m^{6}A (grey) or 100% of m^{6}A (red) with the dissociation constants obtained from the data points (see Methods for details).

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Moreover, although all adenosines were methylated in the m^{6}A-5'-UTR RNA only local reactivity changes, often clustered, were observed suggesting that the presence of m^{6}A influences the folding of specific domains or motifs within the 5´-UTR. The presence of m^{6}A is predicted and has been observed to destabilize A-U pairings embedded in helical regions. This is mostly in agreement with the destabilizations (increase in reactivity) that are often observed for As, Us or in nucleotides in close proximity with proposed A-U base pairs. In particular, this could explain the reactivity enhancement within the poly-A stem, which is likely to be globally destabilized by m^{6}A. However, none of the other nucleotides for which we observe a reactivity increase is involved or are close to A-U pairings and most of them are not even predicted to be base paired in most of the published models. In contrast, m^{6}A has been shown to stabilize A-U pairings when preceded in 5' by a bulged nucleotide. This does not offer a rationale for any of the reactivity drop we observe which are mostly predicted to be in single strand regions thus, suggesting that m^{6}A modulates a higher order structure and/or tertiary pairings yet to be identified. This prompted us to monitor dimerization of the 5'-UTR and m^{6}A-5´-UTR *in vitro*. We observed that the presence of m^{6}A reduces but not abolish the efficiency of dimerization (Fig. 1b). The structure of the HIV-1 full-length RNA dimer is still poorly defined and a matter of debate, thus the dimerization deficiency observed does not provide a straightforward explanation for all the alterations of the SHAPE profile but could reveal unknown rearrangements.

Together, these data indicate that the presence of m^{6}A may play an important role in the folding and dimerization of the 5´-UTR and prompted us to study this feature in a cellular context.

The presence of m^{6}A within the full-length RNA favors Gag synthesis but interferes with packaging

In order to study the role of m^{6}A on the cytoplasmic fate of the full-length RNA during viral replication, we first determined the effects of METTL3/14 overexpression on Gag and full-length RNA levels obtained from cell extracts and purified viral particles (see scheme in Supplementary Fig. 2a). m^{6}A-RIP analysis from METTL3/14 overexpressing cells showed an increase in the m^{6}A/A ratio of the full-length RNA compared to the control indicating that the viral transcript is hypermethylated under these experimental conditions (Supplementary Fig. 2b). Consistent with the positive role of m^{6}A on Gag synthesis previously described [20-22], we observed that increased methylation of the full-length RNA by METTL3/14 overexpression results in increased levels of Gag and its processing products with minor effects on the

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

intracellular levels of the full-length RNA (Fig. 2a). Quantification of viral particles produced from the same cells revealed a slight increase in Gag levels (as judged by anti-CAp24 ELISA) from METTL3/14 overexpressing cells, which could be attributed to the increased Gag synthesis observed (Fig. 2b).

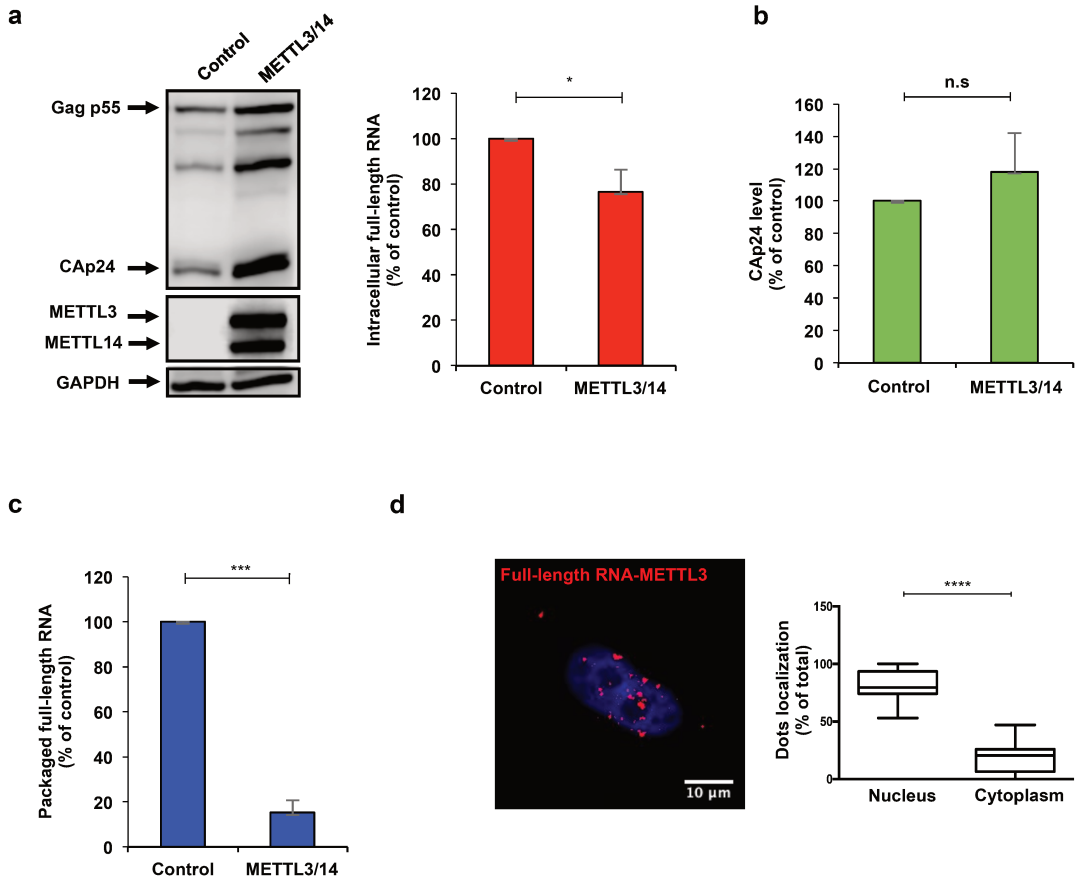


Figure 2: The presence of m^{6}A within the full-length RNA favors Gag synthesis but interferes with packaging. HEK293T cells were transfected with pNL4.3 and pCMV-VSVg together with pCDNA-Flag-METTL3 and pCDNA-Flag-METTL14 or pCDNA-d2EGFP as a control. **a,** At 24 hpt cells extracts were used to detect Gag, Flag-METTL3 and Flag-METTL14 by Western blot. GAPDH was used as a loading control (left panel). In parallel, cells extracts were used to perform RNA extraction and the full-length RNA was quantified by RT-qPCR (right panel). Intracellular full-length RNA was normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (*$P<0.05$, *t*-test). **b,** Supernatants from cell cultures in (a) were filtered and viral particles were purified by ultracentrifugation. The level of CAp24 was quantified by an anti-CAp24 ELISA. The level of CAp24 was normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (n.s; non-significant, *t*-test). **c,** Viral particles purified from (b) were used to perform RNA extraction and the packaged full-length RNA from CAp24 equivalents was quantified by RT-qPCR. Packaged full-length RNA was normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (***$P<0.001$, *t*-test). **d,** HeLa cells were transfected with pNL4.3, pCMV-VSVg and pCDNA-Flag-METTL3. At 24 hpt, the interaction between the full-length RNA and the Flag-tagged METTL3 was analyzed by ISH-PLA as described in the Methods section. Red dots indicate the interactions between the full-length RNA and the METTL3. Scale bar 10 mm. A quantification of the red dots in the nucleus

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

(co-localizing with the DAPI staining) and the cytoplasm of 14 cells is presented on the right (****$P<0.0001$, Mann-Whitney test).

Then, we quantified the full-length RNA from equal amounts of viral particles and observed that viral particles produced under METTL3/14 overexpression contain around 3-fold less packaged gRNA indicating that hypermethylation of the full-length RNA impedes its packaging into nascent particles (Fig. 2c). We were not able to observe a similar effect of murine METTL3/14 overexpression on the simple retrovirus MLV, which was shown two segregate their full-length RNA into two specialized populations for translation and packaging further suggesting that MLV and HIV-1 might evolved diverse mechanisms for the cytoplasmic sorting of their full-length RNA (Supplementary Fig. 2c).

Since results presented above indicate that m^{6}A deposition by METTL3/14 affects the cytoplasmic fate of the full-length RNA, we wanted to investigate where within the cell the m^{6}A writer complex modifies the viral RNA. For this, we analyzed the interaction between the full-length RNA and METTL3 by *in situ* hybridization coupled to the proximity ligation assay (ISH-PLA)[24]. Confocal microscopy analyses revealed a predominant interaction within the nucleus, which suggests that the full-length RNA must be methylated in the nucleus and reach the cytoplasm in a methylated form (Fig. 2d and Supplementary Fig. 2d).

Together, these data suggest that nuclear methylation of the HIV-1 full-length RNA by METTL3/14 favors its use as mRNA for Gag synthesis but interferes with its incorporation into viral particles.

Methylation of A_{198} and A_{242} within the 5′-UTR interferes with HIV-1 full-length RNA packaging

From data presented above, it seems that the presence of m^{6}A interferes with the function of the HIV-1 full-length RNA as gRNA. Thus, to gain further insights into this regulation, we employed the m^{6}A-seq strategy to determine the m^{6}A patterns of the intracellular and viral particle-associated HIV-1 full-length RNA. In agreement with previous data reported for the NL4.3 and LAI.2 strains in T-lymphocytes and HEK293T cells [20-22], we identified m^{6}A peaks mainly at the 5′-UTR and a cluster of peaks at the 3′ end of the intracellular full-length RNA (Fig. 3a, see intracellular full-length RNA). Interestingly, we observed that the full-length RNA from purified viral particles maintains the m^{6}A peak at the 3′-UTR but lacks the m^{6}A peak at the 5′-UTR. This observation together with data from Fig. 2c suggests that the presence of m^{6}A at the 5′-UTR interferes with the incorporation of the full-length RNA into

bioRxiv preprint first posted online Jun. 19, 2019; doi: http://dx.doi.org/10.1101/676031. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

viral particles (Fig. 3a, see packaged full-length RNA). Of note, this difference in the methylation patterns between intracellular and packaged RNA was not observed in the host 7SL RNA, which is also packaged at high levels into HIV-1 particles [25], indicating a very specific effect of m^6A on full-length RNA packaging (Supplementary Fig. 3a). These data strongly suggest that full-length RNA molecules lacking m^6A at the 5´-UTR are primarily selected by Gag as gRNA to be incorporated into viral particles.

A bioinformatic prediction of the potentially methylated residues within the 5-UTR of the NL4.3 strain identified A_{198} and A_{242} in a very favorable methylation context (Supplementary Fig. 3b). Both residues are contained within the m^6A peak we have identified within the 5´-UTR of the intracellular full-length RNA (Supplementary Fig. 3c) and were also identified in previous m^6A-seq data obtained from T-lymphocytes and HEK293T cells [20, 22]. Thus, we deleted A_{198}, A_{242} or both from the NL4.3 provirus in order to determine the role of these adenosine residues on the m^6A-mediated regulation of full-length RNA packaging. We observed that ΔA_{198} and ΔA_{242} single mutant proviruses were slightly resistant to the effects of METTL3/14 overexpression on full-length RNA packaging (Supplementary Fig. 3d). However, the $\Delta A_{198}/\Delta A_{242}$ double mutant provirus was refractory to the positive effects of METTL3/14 overexpression on Gag synthesis and the negative effects on full-length RNA packaging observed with the wild type provirus indicating that A_{198} and A_{242} are key residues involved in this m^6A-mediated regulation (Fig. 3b and Supplementary Fig. 3e). Next, we sought to investigate the role of A_{198} and A_{242} on HIV-1 full-length RNA metabolism. A comparison between wild type and the $\Delta A_{198}/A_{242}$ provirus showed that the double mutant virus accumulates more intracellular Gag but releases significantly less viral particles (Fig. 3c and Supplementary Fig. 3f). Moreover, quantification of the full-length RNA from equal amounts of viral particles revealed a defect in packaging in the $\Delta A_{198}/A_{242}$ double mutant provirus thus, confirming the critical relevance of these two adenosine residues for the incorporation of the HIV-1 full-length RNA into viral particles (Fig. 3c). An analysis of 890 sequences from the HIV database (www.hiv.lanl.gov) indicate that A_{198} and A_{242} are highly conserved within the 5´-UTR of isolates suggesting that this epitranscriptomic regulation must be a common feature of different HIV-1 subtypes including the highly prevalent subtypes C and B as well as circulating recombinant forms (Fig. 3d and Supplementary Fig. 3g). Taking together, these results indicate that the HIV-1 full-length RNA may exist as two different populations that differ at least in the m^6A residues present within the 5´-UTR. Only

the full-length RNA molecules lacking m^{6}A at positions 198 and 242 might be recognized by Gag for packaging.

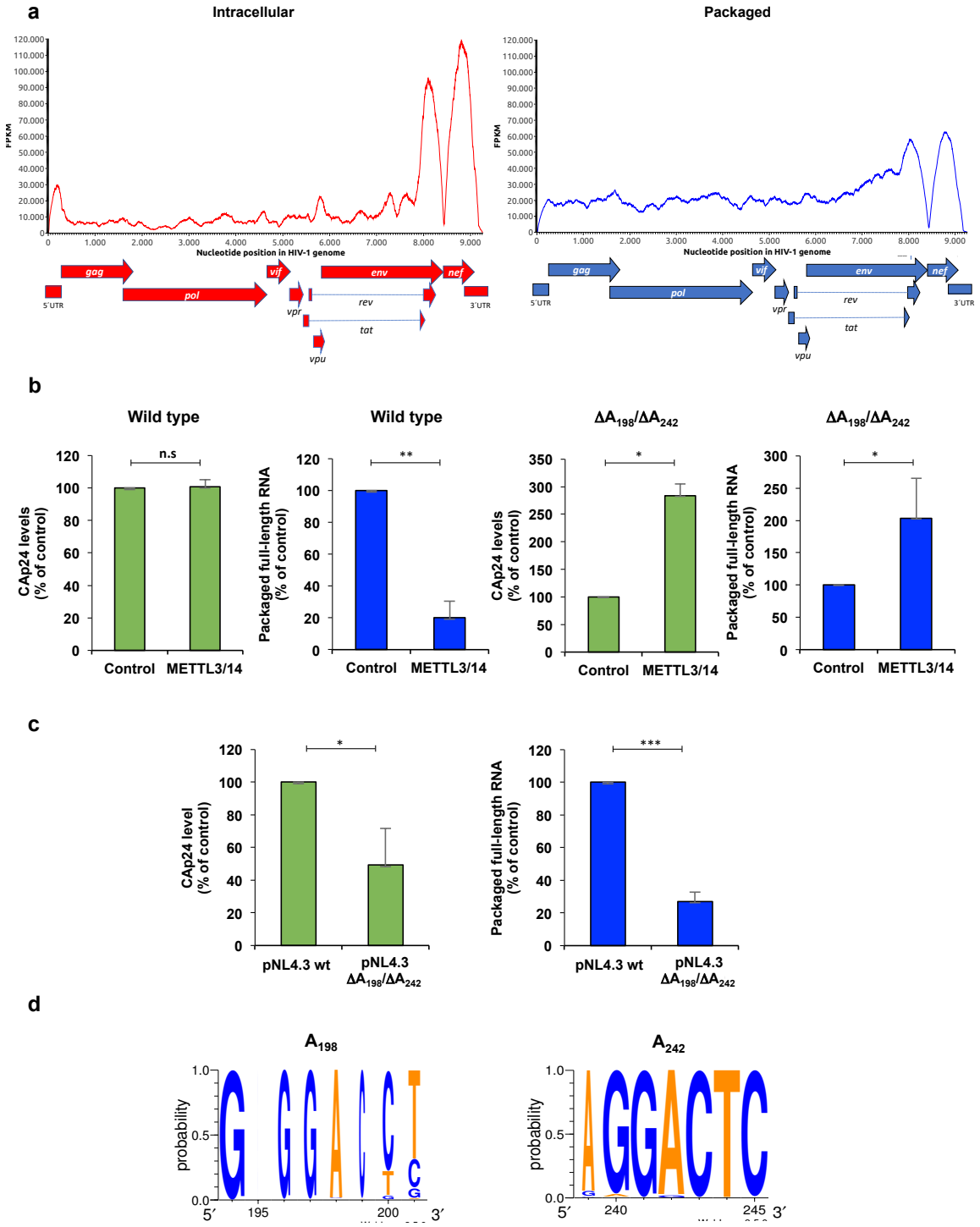


Figure 3: Methylation of A_{198} and A_{242} within the 5′-UTR interferes with HIV-1 full-length RNA packaging. a, HEK293T cells were transfected with pNL4.3 and pCMV-VSVg. Intracellular polyA RNA or viral particle-associated RNA was extracted at 24 hpt, fragmented and used for m^{6}A-seq as described in Methods. Peak calling results for the intracellular (left) and packaged (right) full-length RNA is shown. **b,** HEK293T cells were transfected with pNL4.3 wild type or pNL4.3

DA_{198}/DA_{242} together with pCMV-VSVg, pCDNA-Flag-METTL3 and pCDNA-Flag-METTL14 or pCDNA-d2EGFP as a control. At 24 hpt supernatants were filtered and viral particles were purified by ultracentrifugation. Purified viral particles were used to perform an anti-CAp24 ELISA and RNA extraction and RT-qPCR analysis as described above. The levels of CAp24 and the packaged full-length RNA (per CAp24 equivalents) were normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (**P<0.05*; ***P<0,01*; n.s, non-significant, *t*-test). **c,** HEK293T cells were transfected with pNL4.3 wild type or pNL4.3 DA_{198}/DA_{242} together with pCMV-VSVg and the supernatant was filtered and ultracentrifuged at 24 hpt. Purified viral particles were used to perform an anti-CAp24 ELISA and for RNA extraction and RT-qPCR analysis as described above. The levels of CAp24 and the packaged full-length RNA (per CAp24 equivalents) were normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (**P<0.05*; ****P<0,001*, *t*-test). **d,** Conservation analyses of adenosines 198 and 242 from 879 sequences from the HIV-1 database.

Demethylation by a Gag-FTO complex favors HIV-1 full-length RNA packaging

Then, we sought to determine whether this m^6A-mediated regulation of full-length RNA packaging was a dynamic process. This was important considering that the reversible nature of adenosine methylation in cellular mRNA has been challenged [26]. For this, we overexpressed the RNA demethylase FTO and the analysis of the m^6A/A ratio of the full-length RNA in control and FTO overexpressing cells revealed that the viral RNA is indeed a substrate for this m^6A eraser (Supplementary Fig. 4a). Consistent with a positive role of m^6A on Gag synthesis, we observed that FTO-induced demethylation of the full-length RNA results in a reduction of Gag levels despite a slight increase in intracellular full-length RNA levels (Fig. 4a). We also observed minimal changes in the CAp24 levels from purified viral particles produced under RNA demethylation conditions (Fig. 4b). However and in agreement with a negative role of m^6A on full-length RNA packaging, we observed that viral particles produced from FTO overexpressing cells contain around 3-fold more packaged gRNA compared to the control (Fig. 4c). It should be mentioned that we were not able to observe similar results with the RNA demethylase ALKBH5 suggesting that full-length RNA demethylation by FTO is important for packaging (Supplementary Fig. 4b).

Considering that m^6A demethylation favors full-length RNA packaging, we wanted to know where within the cell the viral RNA became demethylated by FTO. For this, we analyzed the interaction of the between the full-length RNA and FTO in cells by ISH-PLA but despite several attempts, we were not able to detect a direct interaction regardless all the components were correctly expressed within the cells (Supplementary Fig. 4c). This observation suggests that either there is no a massive interaction between the full-length RNA and FTO or that such interactions occurs very transiently (or at very low rates) being below the detection limit of our ISH-PLA strategy.

The lack of a detectable interaction between the full-length RNA and FTO prompted us to investigate whether Gag could interact with FTO and drive full-length RNA demethylation. In order to test this possibility, we employed the proximity ligation assay (PLA) and observed that Gag and FTO indeed form complexes in cells (Fig. 4d). Interestingly, quantification of the dots per cell localizing with the nuclear staining as well as 3D reconstitutions of representative images indicate that Gag and FTO mostly associates within the nucleus (Figs. 4e and 4f). Indeed, we observed that FTO overexpression increases the nuclear localization of Gag (Supplementary Fig. 4d).

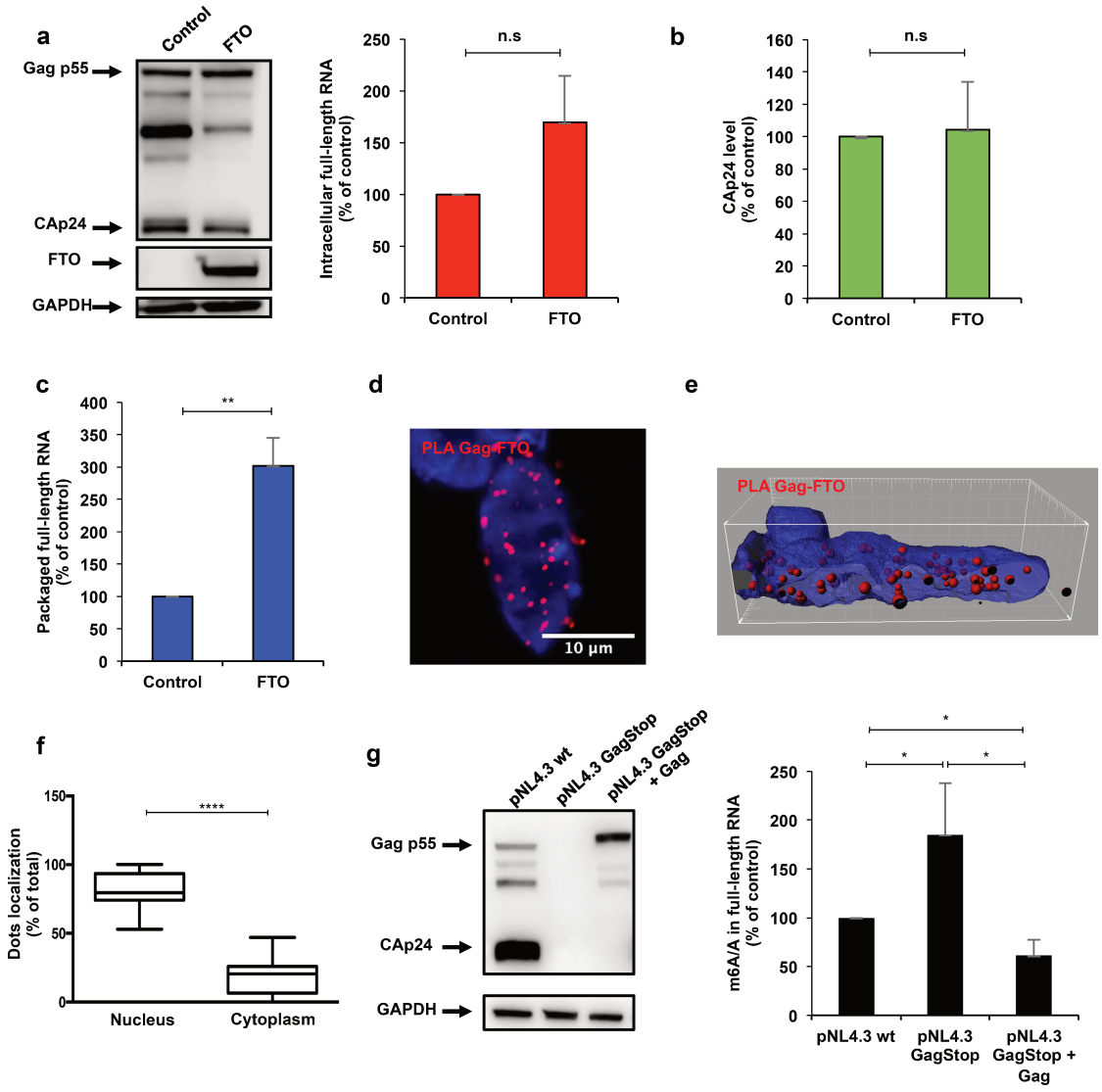


Figure 4: Demethylation by a Gag-FTO complex favors HIV-1 full-length RNA packaging. HEK293T cells were transfected with pNL4.3 and pCMV-VSVg together with pCDNA-3XFlag-FTO or pCDNA-3XFlag-d2EGFP as a control. **a,** At 24 hpt cells extracts were used to detect Gag and 3XFlag-FTO by Western blot. GAPDH was used as a loading control (left panel). In parallel, cells extracts were used to perform RNA extraction and the full-length RNA was quantified by RT-qPCR (right panel). Intracellular full-length RNA was normalized to the control (arbitrary set to 100%) and presented as the

mean +/- SD of three independent experiments (*$P<0.05$, *t*-test). **b,** Supernatants from cell cultures in (a) were filtered and viral particles were purified by ultracentrifugation. The level of CAp24 was quantified by an anti-CAp24 ELISA, normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (n.s; non-significant, *t*-test). **c,** Viral particles purified from (b) were used to perform RNA extraction and the packaged full-length RNA from CAp24 equivalents was quantified by RT-qPCR. Packaged full-length RNA was normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (**$P<0.01$, *t*-test). **d,** HeLa cells were co-transfected with pNL4.3, pCMV-VSVg and pCDNA-3XFlag-FTO. At 24 hpt, the interaction between Gag and Flag-tagged FTO was analyzed by PLA as described in Methods. Red dots indicate the interactions between Gag and FTO (left panel). Scale bar 10 mm. **e,** Three-dimensional reconstitution of the PLA results shown in (d) was performed to determine the subcellular localization of the interaction between Gag and FTO. **f,** Quantification of the red dots in the nucleus (co-localizing with the DAPI staining) and the cytoplasm of 15 cells is presented on the right (****$P<0,0001$, Mann-Whitney test). **g,** HEK293T cells were transfected with the pNL4.3 wild type, pNL4.3-GagStop or pNL4.3-GagStop together with pCDNA-Gag. At 24 hpt cells extracts were used to detect Gag and GAPDH was used as a loading control. In parallel, cells extracts were used to perform RNA extraction followed by an immunoprecipitation using an anti-m^6A antibody (m^6A-RIP as described in Methods). The full-length RNA from the input ("A" fraction) and from the immunoprecipitated material ("m^6A" fraction) was quantified by RT-qPCR. The m^6A/A ratio was normalized to pNL4.3 wild type (arbitrary set to 100%) and presented as the mean +/-SD of three independent experiments (*$P<0.05$, *t*-test).

From these results, it was tempting to speculate that the full-length RNA is methylated within the nucleus by METLL3/14 and Gag interacts with FTO in the nucleus to drive demethylation of the full-length RNA molecules that will be incorporated into assembling viral particles. Thus, we analyzed the m^6A content of the full-length RNA in the presence or absence of Gag by using the wild type NL4.3 provirus and a mutant provirus containing premature stops codons that abolish Gag synthesis (GagStop provirus). Compared to the wild type full-length RNA, the level of m^6A increases when Gag is absent and is restored or even decreased when the Gag protein was expressed *in trans* indicating that Gag regulates the methylation status of the full-length RNA (Fig. 4g).

Together, these results strongly indicate that the FTO-mediated demethylation is required for full-length RNA packaging in a process supported by the Gag precursor.

Inhibition of FTO demethylase activity impacts full-length RNA metabolism and blocks packaging

We finally sought to determine whether this epitranscriptomic regulation of the HIV-1 full-length RNA packaging was a potential therapeutic target for pharmacological intervention. For this, we took advantage of the ester form of meclofenamic acid (MA2), which was shown to specifically interfere with FTO-mediated m^6A demethylation [27]. Therefore, we analyzed Gag and the full-length RNA in cells treated with DMSO (as a control) or MA2 and observed a reduction in Gag synthesis and the intracellular levels of the full-length RNA indicating that

FTO-mediated demethylation is required for proper metabolism of the full-length RNA within the cell (Fig. 5a). Consistent with a perturbed intracellular full-length RNA metabolism, we observed a decrease in the viral particles released from MA2-treated cells (Fig. 5b). Strikingly, quantification of the packaged full-length RNA from equal amounts of viral particles indicates that inhibition of FTO activity by MA2 almost abolished packaging (Fig. 5c).

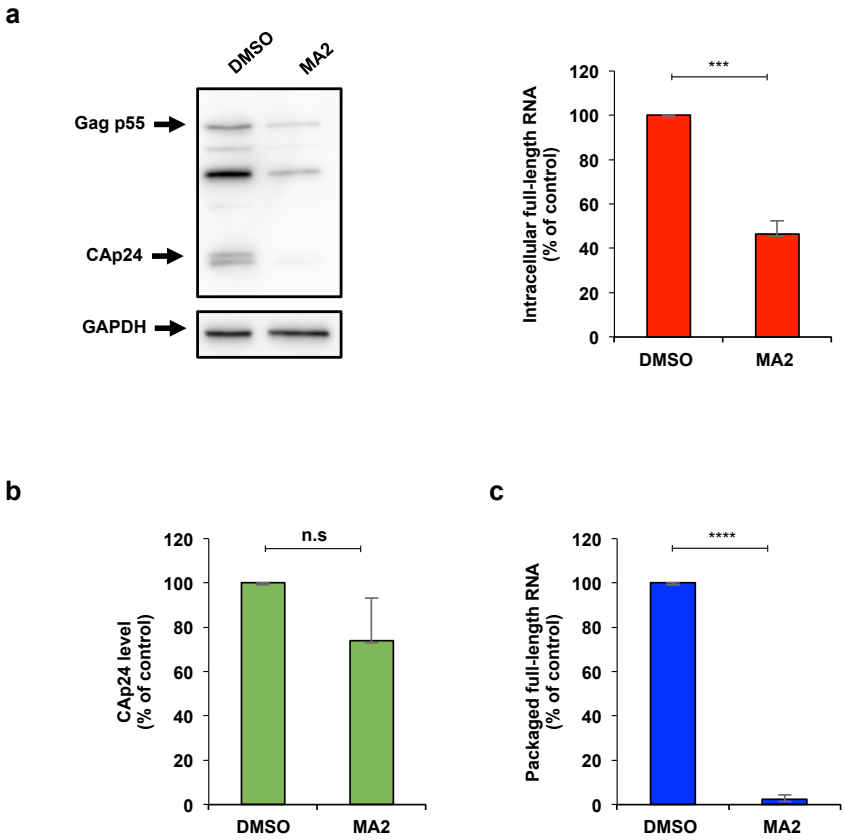


Figure 5: Inhibition of FTO demethylase activity impacts full-length RNA metabolism and blocks packaging. HEK293T cells were transfected with pNL4.3 and pCMV-VSVg and were treated with MA2 or DMSO as a control. **a,** At 24 hpt cells extracts were used to detect Gag and GAPDH as a loading control (left panel). In parallel, cells extracts were used to perform RNA extraction and the full-length RNA was quantified by RT-qPCR (right panel). The intracellular full-length RNA was normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (****P<0,001,* *t*-test). **b,** At 24 hpt the supernatant was filtered and viral particles were purified by ultracentrifugation. The level of CAp24 was quantified by an anti-CAp24 ELISA, normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (n.s; non-significant, *t*-test). **c,** Purified viral particles from (b) were used to perform an RNA extraction and the packaged full-length RNA from CAp24 equivalents was quantified by RT-qPCR. Packaged full-length RNA was normalized to the control (arbitrary set to 100%) and presented as the mean +/- SD of three independent experiments (*****P<0.0001,* *t*-test).

Taking together, these results confirm that FTO-mediated demethylation is critical for HIV-1 full-length RNA packaging and this process is a potential target for the design of novel antiretroviral drugs.

DISCUSSION

Assembly of human immunodeficiency virus type-1 particles is a highly regulated process in which the major structural polyprotein Gag together with other viral and cellular components are recruited to the plasma membrane for the release of the viral progeny. The assembly process occurs in multiple steps driven by the different functional domains that compose the Gag precursor. As such, while the nucleopcapsid (NC) domain specifically recruits two copies of the full-length RNA, the matrix (MA) domain allows targeting of the complex to specific plasma membrane micro-domains and the capsid (CA) domain drives Gag multimerization at such sites. Packaging of two copies of full-length RNA by the NC domain of Gag is highly specific and occurs selectively over thousands of cellular and viral RNA species. This selectivity was proposed to be possible by the presence of *cis*-acting RNA signatures spanning the 5′-UTR and the beginning of the Gag coding region. However, the full-length RNA also serves as mRNA for the synthesis of Gag and Gag-Pol precursors and thus, translation and packaging are expected to be two mutually exclusive processes. Although the adoption of a branched multiple hairpin (BMH) conformation of the 5′-UTR was initially proposed to favor dimerization and packaging over translation [12-14], it was later demonstrated that translation of the full-length RNA is under positive selection and thus, not regulated by a conformational switch of the 5′-UTR [18, 19]. Additional structural studies carried out in cells and virions also argued against structural rearrangements as drivers of the transition between translation and packaging of the HIV-1 full-length RNA [16, 17, 28]. Therefore, the mechanism by which Gag selects “packageable” from “translatable” full-length RNA molecules still remains as one of the long-lasting questions in Retrovirology. In this work, we showed that demethylation of two highly conserved adenosine residues within the 5′-UTR is critical for packaging of the HIV-1 full-length RNA. Interestingly, we observed that Gag associates with the RNA demethylase FTO in the nucleus and promotes demethylation of the full-length RNA, suggesting that Gag may drive FTO-mediated demethylation of those RNA molecules that will be used as gRNA to be incorporated into assembling viral particles (Fig. 6). This differential epitranscriptomic regulation exerted on the full-length RNA depending on its functions (mRNA or gRNA) may also help to explain the controversies reported in the literature [29]. As such, while the presence of m^6A favors Gag synthesis through YTHDF

proteins acting on the full-length RNA molecules destined to serve as mRNA [21], the same cytoplasmic m^{6}A readers may recognize specific features and drive degradation of the incoming viral RNA early upon infection (i.e., when the full-length RNA acts as gRNA) [22].

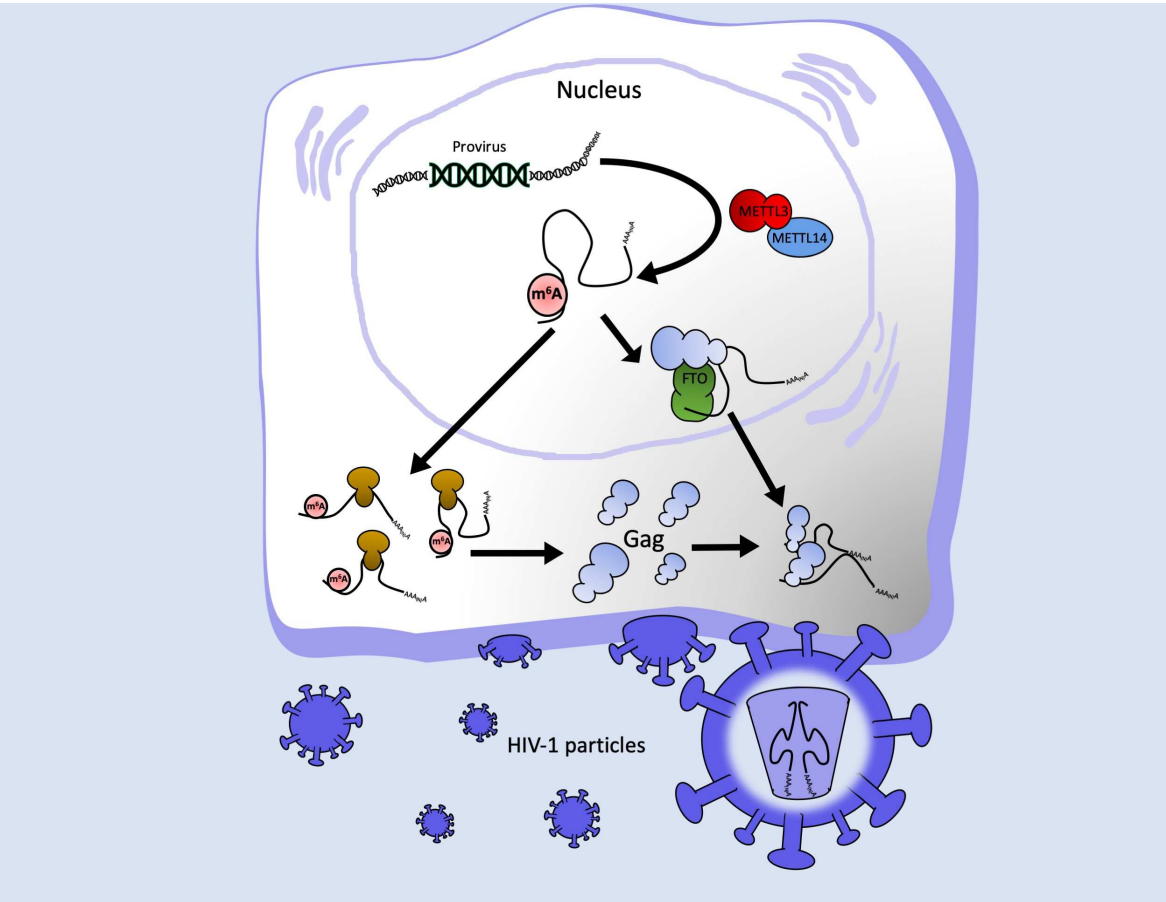


Figure 6 Working model for the epitranscriptomic regulation of HIV-1 full-length RNA packaging. The HIV-1 full-length RNA is methylated by METTL3/14 complex in the nucleus (for simplicity, only the presence of m^{6}A on the 5′-UTR is shown). However, the structural protein Gag interacts with the m^{6}A eraser FTO and drives demethylation of adenosines residues present at the 5′-UTR in a process required for full-length RNA packaging.

Further studies are required to elucidate the mechanism by which the binding of YTHDF proteins to m^{6}A residues negatively impact full-length RNA metabolism in the absence of translation and whether methylation of the 5′-UTR is involved. In addition, the molecular mechanism by which the presence of m^{6}A at the 5′-UTR interferes with full-length RNA packaging also deserves further investigation. One of the most plausible explanations is that recognition of m^{6}A residues at the 5′-UTR by reader proteins interferes with Gag binding and/or full-length RNA dimerization. However, it is also possible that the presence of m^{6}A itself directly repeals the recruitment of Gag or alters the optimal RNA conformation of the dimerization and/or packaging signal. Indeed, our *in vitro* structural and dimerization analyses

support this last idea as they suggest that the presence of m^6A affects the folding and dimerization of the 5′-UTR. Of note, although neither A_{198} nor A_{242} showed a significant reactivity alteration this does not exclude that their modification may have influenced the reactivity of close nucleotides such as G_{240} and G_{241}. To date no studies have monitored the effect of m^6A on such a complex structure and here we clearly show that this modification can have not only local but more global effects on RNA folding. Indeed, we observed that the introduction of m^6A alters dimerization of the 5′-UTR, which could at least partly explain why methylated full-length RNA are not recovered within viral particles. Nevertheless, our experiments using the $\Delta A_{198}/\Delta A_{242}$ provirus strongly indicate that these two adenosine residues play a major role in packaging. While A_{198} is located within the region complementary to the $tRNA^{Lys3}$ at the primer binding site (PBS), A_{242} is located in the AGGA bulge at the base of the SL1 region and corresponds to a Gag-binding domain previously shown by chemical probing *in vitro* and *in virio* as well as by CLIP-seq studies [17, 30, 31]. Further studies are required to fully understand the role of these residues in full-length RNA packaging.

In contrast to the simple retrovirus MLV, which segregates its full-length RNA into two separate populations for translation and packaging, the full-length RNA of the human lentivirus HIV-1 was proposed to exist as a single population that can indistinctly serve as mRNA or gRNA. While the presence of two specialized full-length RNA populations supports the lack of an epitranscriptomic regulation of MLV packaging, our data obtained with HIV-1 suggest that its full-length RNA may also exist as two populations with different m^6A patterns. From these two populations, those molecules lacking m^6A at the 5′-UTR will be preferentially selected by Gag for packaging (Fig. 6). In this regard, the HCV core protein was also shown to bind preferentially to RNA molecules lacking m^6A most probably to avoid YTHDF proteins-mediated degradation upon viral entry [32]. Whether the packaging of hypomethylated RNA genomes is a conserved mechanism evolved by RNA viruses to avoid early degradation by cytoplasmic m^6A readers must be determined.

Although the reversibility of adenosine methylation as well as the main target of the RNA demethylase FTO have been recently challenged [26, 33], we showed that the HIV-1 full-length RNA is a substrate for FTO and this RNA demethylase regulates the incorporation of the viral genome into released viral particles. Interestingly, treatment of HIV-1 producer cells with the specific FTO inhibitor meclofenamic acid resulted in an impairment of full-length RNA metabolism with a potent effect on packaging, confirming the critical role of the demethylase activity of FTO during viral replication. In addition to meclofenamic acid, two small

molecules developed by structure-based rational design were recently described as specific inhibitors of FTO m^6A demethylase activity with the potential to be used as a treatment for adult myeloid leukemia [34]. Therefore, this novel epitranscriptomic mechanism regulating packaging of the HIV-1 full-length RNA could also be exploited as a target for pharmacological intervention.

METHODS

Selective 2´-hydroxyl acylation analyzed by primer extension (SHAPE): The HIV-1 full-length RNA 5`-UTR was *in vitro* transcribed using the T7 RNA polymerase as described [35]. For m^6A RNAs, ATP was substituted by N^6-methyl-ATP (Jena Bioscience). RNA was quantified by measurement of the OD_{256} using a BioSpec-nano (Shimadzu). RNA integrity was assessed by agarose gel electrophoresis. SHAPE probing was conducted essentially as in [17] with minor modifications. Briefly, 6 pmol of *in vitro* transcribed RNA containing 100% m^6A or 0% m^6A were diluted in 24 μl of water, denatured at 80ºC for 2 minutes and ice cooled. After addition of 3 μl of 10X Folding Buffer (HEPES Ph7.5 400 nM, KCl 1M, $MgCl_2$ 50 mM), samples were incubated for 10 minutes at room temperature and then 10 minutes at 37ºC. Then, the RNA solution was added to 3 μl of 20 mM 1M7 (2mM final) (AEchem Scientific Corporation) or to 3 μl of DMSO (control) and incubated for 6 minutes at 37ºC. RNA was subsequently precipitated in presence of 1 μl of 20 mg/ml of glycogen, 3 μl of 5M sodium acetate, 100 μl of ETOH for 1 hour at -20ºC and then washed with ETOH 70% and resuspended in water. For primer extension, RNA was treated with 1 μl of DMSO and denatured for 3 minutes at 95ºC and cooled at 4ºC. Samples were mixed with 3 μl of 2μM fluorescent primer (D4-5'-TTTCTTTCCCCCTGGCCTT for the probed/control samples and D2-5'-TTTCTTTCCCCCTGGCCTT for the sequencing reaction, Sigma) and incubated for 5 minutes at 65ºC, 10 minutes at 35ºC and then 1 minute on ice. 5 μl of Reverse Transcription Mix (10 μl of 10 mM dNTPs and 40 μl of 5X MMLV RT Buffer Promega), 1 μl of 10 mM ddTTP for the sequencing reaction and 1 μl of MMLV Reverse Transcriptase RNAse H minus (Promega) were finally added and the reverse transcription was performed at 35ºC for 2 minutes, 42ºC for 30 minutes and 55ºC for 5 minutes. Samples were then ice cooled and precipitated with ETOH for 2 hours. Pellets were resuspended in 40 μl of Sample Loading Solution (Beckman Coulter). Reverse transcription products were resolved on a CEQ-8000 sequencer (Beckman Coulter). Electropherograms were analyzed using QuSHAPE [36]. Raw data were processed by excluding the 2% of the highest values and normalizing the remaining

values by the mean of the next 8% highest values [37]. The experiments were performed three times and reproducibility was assessed by calculating the standard error of the mean. Secondary structure was drawn using VaRNA [38].

Dimerization assay: *In vitro* transcribed 5`-UTR was serially diluted in 10 mM Tris pH 7, 10 mM NaCl, 140 mM KCl to obtain a final concentration ranging from 0 to 1 μM. 20 fmol (2 nM final) of radiolabeled RNA was added. Samples were denatured at 95ºC for 3 minutes and then ice cooled. After addition of 1 mM $MgCl_2$, RNA was allowed to fold for 30 minutes at 37°C. Samples were subsequently chilled on ice, mixed with a 5X native loading buffer (glycerol 20%, xylene cyanol 0.1%, bromophenol blue 0.1%) and loaded on a native 4% acrylamide gel. Samples were run for 1 hour at 100V on 4% native acrylamide mini-gels containing 34 mM Tris, 54 mM HEPES, 0.1 mM EDTA and 2.5 mM $MgCl_2$. Dried gels were quantified using a BAS-5000 phosphorimager and MultiGauge 3.0 (Fujifilm). The fraction of dimer was calculated as the ratio "dimer – background" over "dimer + monomer – 2*background". Data were fitted to Fraction dimer = Bmax*[RNA]/(Kd+[RNA]) using Prism 5.02 (GraphPad Software).

Cell culture, DNA transfection and viral particle purification: HEK293T and HeLa cells were maintained in DMEM (Life technologies) supplemented with 10% FBS (Hyclone) and antibiotics (Hyclone) at 37ºC and 5% CO_2 atmosphere. Cells growing in 6-well plates ($2.5x10^5$ cells/well) were transfected using linear PEI ~25000 Da (Polyscience) as described previously [24]. Cells were transfected using a ratio μg DNA/μl PEI of 1/15 and the DNA/PEI mix was incubated for 20 min at room temperature before adding to the cells. For experiments involving the FTO inhibitor, the culture medium was replaced by medium containing dimethyl sulfoxide (DMSO) as a control or 80 μM of an ethyl ester form of meclofenamic acid diluted in DMSO (MA2) prior DNA transfection. For viral particle purification, the supernatant was collected and filtered by passing through a 0.22 μm filter and then ultracentrifugated at 25.000 rpm for 2 hours at 4ºC in a 20% sucrose cushion (prepared previously and stored at 4 ºC). Purified viral particles were resuspended in 100 μl of PBS and stored in aliquots at -80 ºC to then perform anti-CAp24 ELISA (HIV-1) or Western blot (MLV) or RNA extraction. Cells were also collected to perform Western blot and RNA extraction as described in Supplementary Methods.

m^{6}A-seq: Poly(A) RNA was purified from 100 μg of total RNA extracted from HEK293T cells previously transfected with pNL4.3 and pCMV-VSVg. Briefly, total RNA in 500 μl of water was incubated at 65ºC for 10 minutes and incubated with 3 μl of oligo dT-Biotin (dT-B; IDT Technologies) (50 pmol/μl) and 13 μl of SSC Buffer 20X (Santa Cruz Biotechnology) and allowed to cool at room temperature. 600 μg of Dynabeads™ Streptavidin (60 μl; Thermo Fisher) were washed three times with Buffer SSC 0.5X and resuspended in 100 μl of Buffer SSC 0.5X. Then, the RNA/oligo dT-B mix was incubated with the streptavidin beads at room temperature for 10 minutes in head-over-tail rotation. RNA-beads were washed four times with 300 μl Buffer SSC 0.1X and bound RNA was eluted twice with 100 μl of water. The RNA was precipitated with 10 mM $MgCl_2$, 20 μg glycogen (Thermo Fisher) and 2.5 volume of ETOH 100% overnight at -20 ºC and then washed with ETOH 70%. Poly(A) RNA as well as RNA obtained from purified viral particle were fragmented using Fragmentation Reagent (Thermo Fisher). For this, 2 μg of RNA in 9 μl of water was incubated with 1 μl of Fragmentation Buffer 10X for 15 minutes at 70 ºC, then 1 μl of Stop solution was added and incubated on ice. The RNA fragmented was precipitated overnight as described above. Fragmented RNA diluted in 380 μl was heated at 70 ºC for 5 minutes, placed on ice for 3 minutes. The denatured RNA was mixed with 1 μl of rRNasin® (Promega), 5 μl VRC, 100 μl of IP Buffer 5X (50 mM Tris-Hcl pH7.4, 750 mM NaCl and 0.5% NP-40) and 5 μl of an anti-m^{6}A antibody (0,5 mg/mL; Synaptic System #202003) and incubated for 2 hours at 4 ºC with head-over-tail rotation. At the same time, 600 μg of Dynabeads™ Protein A magnetic beads (20 μl; Thermo Fisher) were washed in 1 ml of IP Buffer 1X with 1 μl of VRC and were incubated with 500 μl of Buffer IP 1X with 0.5 mg/ml of BSA for 2 hours at 4 ºC with head-over-tail rotation. Then, beads were washed with 500 μl of IP Buffer 1X and added to the RNA/anti-m^{6}A antibody mix. The RNA-beads mix was incubated for 2 hours at 4ºC in head-over-tail rotation. After incubation, the RNA-beads mix was washed twice with 500 μl IP Buffer 1X. Bound RNA was eluted with 100 μl of Elution Buffer (5mM Tris-HCl, 1mM EDTA and 0.05% SDS) and 1 μl of Proteinase K (New England BioLabs) and incubated for 1.5 hour at 50ºC. RNA was extracted from supernatant using TRIzol™ (ThermoFisher). The RNA recovered was precipitated with 10 mM $MgCl_2$, 20 μg glycogen (Thermo Fisher) and 2.5 volumes of ETOH 100% overnight at -20 ºC and then washed with ETOH 70%. Equal amounts of RNA from input and immunoprecipitation were used for RT-qPCR. cDNA libraries preparations and RNAseq was performed at Genoma Mayor. All the samples were

sequenced in an Illumina HiSeq2000 platform with paired-end 100 bp read length. Read quality was evaluated with *Fastqc* and the *Burrows-Wheeler Alignment Tool (BWA -Mem)* was used for mapping reads to the HIV-1 genome with default parameters.

The alignment data were analyzed with *MACS2* to call peaks with *-f BAMPE --nomodel --SPMR* options for generated viral peaks data and to generate FPKM (Fragments Per Kilobase per Million mapped reads). Predicted peaks were sorted by average coverage.

DATA AVAILABILITY

m^6A-seq data were deposited at the GEO upon accession number GSE130687.

REFERENCES

1. Karn, J. & Stoltzfus, C.M. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harbor Perspectives in Medicine* **2**, a006916 (2012).
2. Butsch, M. & Boris-Lawrie, K. Destiny of unspliced retroviral RNA: ribosome and/or virion? *J Virol* **76**, 3089-3094 (2002).
3. Mailler, E. *et al.* The Life-Cycle of the HIV-1 Gag-RNA Complex. *Viruses* **8** (2016).
4. Levin, J.G., Grimley, P.M., Ramseur, J.M. & Berezesky, I.K. Deficiency of 60 to 70S RNA in murine leukemia virus particles assembled in cells treated with actinomycin D. *J Virol* **14**, 152-161 (1974).
5. Levin, J.G. & Rosenak, M.J. Synthesis of murine leukemia virus proteins associated with virions assembled in actinomycin D-treated cells: evidence for persistence of viral messenger RNA. *Proc Natl Acad Sci U S A* **73**, 1154-1158 (1976).
6. Dorman, N. & Lever, A. Comparison of viral genomic RNA sorting mechanisms in human immunodeficiency virus type 1 (HIV-1), HIV-2, and Moloney murine leukemia virus. *J Virol* **74**, 11413-11417. (2000).
7. Berkhout, B. Structure and function of the human immunodeficiency virus leader RNA. *Prog Nucleic Acid Res Mol Biol* **54**, 1-34 (1996).
8. Kuzembayeva, M., Dilley, K., Sardo, L. & Hu, W.S. Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology* **454-455**, 362-370 (2014).
9. Laughrea, M. *et al.* Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. *J Virol* **71**, 3397-3406 (1997).
10. Chen, J. *et al.* HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. *Proc Natl Acad Sci U S A* **113**, E201-208 (2016).
11. Clever, J.L. & Parslow, T.G. Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. *J Virol* **71**, 3407-3414 (1997).
12. Huthoff, H. & Berkhout, B. Two alternating structures of the HIV-1 leader RNA. *RNA* **7**, 143-157 (2001).
13. Berkhout, B. *et al.* In vitro evidence that the untranslated leader of the HIV-1 genome is an RNA checkpoint that regulates multiple functions through conformational changes. *J Biol Chem* **277**, 19967-19975 (2002).

14. Abbink, T.E. & Berkhout, B. A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. *J Biol Chem* **278**, 11601-11611 (2003).
15. Keane, S.C. *et al.* RNA structure. Structure of the HIV-1 RNA packaging signal. *Science* **348**, 917-921 (2015).
16. Paillart, J.C. *et al.* First Snapshots of the HIV-1 RNA Structure in Infected Cells and in Virions. *J Biol Chem* **279**, 48397-48403 (2004).
17. Wilkinson, K.A. *et al.* High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. *PLoS Biol* **6**, e96 (2008).
18. Abbink, T.E., Ooms, M., Haasnoot, P.C. & Berkhout, B. The HIV-1 leader RNA conformational switch regulates RNA dimerization but does not regulate mRNA translation. *Biochemistry* **44**, 9058-9066 (2005).
19. Boeras, I. *et al.* The basal translation rate of authentic HIV-1 RNA is regulated by 5'UTR nt-pairings at junction of R and U5. *Sci Rep* **7**, 6902 (2017).
20. Lichinchi, G. *et al.* Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. *Nat Microbiol* **1**, 16011 (2016).
21. Kennedy, E.M. *et al.* Posttranscriptional m(6)A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. *Cell Host Microbe* **19**, 675-685 (2016).
22. Tirumuru, N. *et al.* N(6)-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *eLife* **5** (2016).
23. Lu, W. *et al.* N(6)-methyladenosine-binding proteins suppress HIV-1 infectivity and viral production. *J Biol Chem* (2018).
24. Toro-Ascuy, D. *et al.* A Rev-CBP80-eIF4AI complex drives Gag synthesis from the HIV-1 unspliced mRNA. *Nucleic Acids Res* **46**, 11539-11552 (2018).
25. Onafuwa-Nuga, A.A., Telesnitsky, A. & King, S.R. 7SL RNA, but not the 54-kd signal recognition particle protein, is an abundant component of both infectious HIV-1 and minimal virus-like particles. *RNA* **12**, 542-546 (2006).
26. Darnell, R.B., Ke, S. & Darnell, J.E., Jr. pre-mRNA processing includes N6 methylation of adenosine residues that are retained in mRNA exons and the fallacy of "RNA epigenetics". *RNA* (2017).
27. Huang, Y. *et al.* Meclofenamic acid selectively inhibits FTO demethylation of m6A over ALKBH5. *Nucleic Acids Res* **43**, 373-384 (2015).
28. Watts, J.M. *et al.* Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**, 711-716 (2009).
29. Riquelme-Barrios, S., Pereira-Montecinos, C., Valiente-Echeverria, F. & Soto-Rifo, R. Emerging Roles of N(6)-Methyladenosine on HIV-1 RNA Metabolism and Viral Replication. *Frontiers in Microbiology* **9**, 576 (2018).
30. Abd El-Wahab, E.W. *et al.* Specific recognition of the HIV-1 genomic RNA by the Gag precursor. *Nat Commun* **5**, 4304 (2014).
31. Kutluay, S.B. *et al.* Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. *Cell* **159**, 1096-1109 (2014).
32. Gokhale, N.S. *et al.* N6-Methyladenosine in Flaviviridae Viral RNA Genomes Regulates Infection. *Cell Host Microbe* **20**, 654-665 (2016).
33. Mauer, J. & Jaffrey, S.R. FTO, m(6) Am , and the hypothesis of reversible epitranscriptomic mRNA modifications. *FEBS Lett* (2018).
34. Huang, Y. *et al.* Small-Molecule Targeting of Oncogenic FTO Demethylase in Acute Myeloid Leukemia. *Cancer Cell* **35**, 677-691 e610 (2019).

35. de Bisschop, G. *et al.* HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity. *Biochimie* (2019).
36. Karabiber, F., McGinnis, J.L., Favorov, O.V. & Weeks, K.M. QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA* **19**, 63-73 (2013).
37. Low, J.T. & Weeks, K.M. SHAPE-directed RNA secondary structure prediction. *Methods* **52**, 150-158 (2010).
38. Darty, K., Denise, A. & Ponty, Y. VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics* **25**, 1974-1975 (2009).

AKNOWLEDGEMENTS

Authors wish to thank Professor Cai-Guang Yang (CAS Key Laboratory of Receptor Research, Shangai Institute of Materia Medica, Chinese Academy of Sciences) for providing MA2. Authors will also thank to Dr. Chuan He (University of Chicago), Dr. Yun-Gui Yang (Beijing Institute of Genomics, Chinese Academy of Sciences), Dr. Jin Crystal Zhao and Dr. Monica Roth (University of Wisconsin) for providing expression vectors as well as Dr. Gloria Arriagada (Universidad Andrés Bello) for providing the anti-MLV CAp30 antibody.

FUNDING

This work was funded through the FONDECYT Program (Nº 1160176 and 1190156 to RSR and 1180798 to FVE). CPM, FGdeG, SRB, BRA and PAC are recipients of a National Doctoral fellowship from CONICYT. Research at BS laboratory was funded by CNRS, Université Paris Descartes and a grant from Fondation pour la Recherche Médicale (FRM #DBI20141231337). Exchanges between RSR and BS laboratories were funded by a LIA grant from CNRS. GdB is a recipient of a fellowship from the French Ministry for "Enseignement Superieur et Recherche".

AUTHOR CONTRIBUITIONS

CP-M designed and performed experiments, analyzed data and wrote the paper. DT-A, SR-B, BR-A, FGdeG, PA-C, CA-S, MLA and JCh performed experiments and analyzed data. CR-F performed bioinformatic analyses. GdeB performed *in vitro* experiments and analyzed data. FV-E and BS contributed to data analysis and manuscript writing. RS-R contributed to the concept and design, data analysis and manuscript writing. All the authors read and approved the final version of the manuscript.

COMPETING FINANCIAL INTERESTS

Authors declare no competing financial interests.

Etude de l'interaction de DDX3 avec l'ARN génomique du VIH

DDX3 est une hélicase cellulaire qui contribue à la réplication du VIH-1. Bien que plusieurs études aient mis en évidence son intervention dans l'export de l'ARNg et dans sa traduction, très peu de résultats caractérisent à ce jour l'interaction de cette hélicase avec l'ARNg. L'obstacle majeur à cet objectif est la difficulté d'obtenir une forme purifiée de la protéine in vitro, et la difficulté technique d'étudier la structure de l'interaction entre une protéine et un long ARN.

1) Caractérisation biochimique de DDX3

Une première partie de mon projet a consisté à effectuer une caractérisation biochimique de DDX3 à partir d'une forme recombinante de la protéine. Des études antérieures qui ont rapporté des résultats parfois contradictoires sur les propriétés biochimiques de DDX3 ont toutes été conduites sur des substrats modèles comprenant quelques dizaines de nucléotides (Franca et al., 2007; Garbelli et al., 2011; Sharma et al., 2017; Yedavalli et al., 2004). Nous confirmons que l'activité hélicase de DDX3 est spécifiquement stimulée par l'ARN, et que l'activité hélicase de DDX3 requiert une extension simple brin en 3' ou en 5'. Le rôle clef de DDX3 dans le métabolisme de l'ARNg du VIH-1 nous a cependant amené à faire l'hypothèse que cet ARN présente des déterminants qui spécifient l'interaction et l'activation de DDX3. Nous montrons ainsi que la stimulation de l'activité ATPase par des substrats modèles n'est que marginale par rapport à la stimulation induite par l'ARNg. En effectuant des délétions successives, nous identifions un fragment minimal requis pour cette stimulation. En absence des extensions N- et C-terminales, l'activité ATPase du cœur catalytique n'est pas stimulée par l'ARNg. Enfin nos résultats révèlent que la liaison de DDX3 à l'ARN n'est pas directement corrélée à la stimulation de l'activité biochimique de DDX3. D'une part, DDX3 dissocie plus efficacement des duplex avec une extrémité 3' libre mais elle a moins d'affinité pour ces duplex que pour un duplex avec une extrémité 5' libre ; d'autre part, la stimulation de l'activité ATPase de DDX3 par des fragments de l'ARN du VIH n'est pas corrélée à l'affinité de la liaison.

Biochimie xxx (xxxx) xxx

Contents lists available at ScienceDirect

Biochimie

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

Research paper

HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity

Grégoire de Bisschop, Mélissa Ameur, Nathalie Ulryck, Fatima Benattia, Luc Ponchon, Bruno Sargueil*, Nathalie Chamond*

CiTCOM, Université Paris Descartes, CNRS UMR 8038, Paris, France

ARTICLE INFO

Article history:
Received 18 December 2018
Accepted 12 March 2019
Available online xxx

Keywords:
DDX3
HIV-1
ATPase
Biological substrate
RNA specificity

ABSTRACT

DEAD-box helicases play central roles in the metabolism of many RNAs and ribonucleoproteins by assisting their synthesis, folding, function and even their degradation or disassembly. They have been implicated in various phenomena, and it is often difficult to rationalize their molecular roles from *in vivo* studies. Once purified *in vitro,* most of them only exhibit a marginal activity and poor specificity. The current model is that they gain specificity and activity through interaction of their intrinsically disordered domains with specific RNA or proteins. DDX3 is a DEAD-box cellular helicase that has been involved in several steps of the HIV viral cycle, including transcription, RNA export to the cytoplasm and translation. In this study, we investigated DDX3 biochemical properties in the context of a biological substrate. DDX3 was overexpressed, purified and its enzymatic activities as well as its RNA binding properties were characterized using both model substrates and a biological substrate, HIV-1 gRNA. Biochemical characterization of DDX3 in the context of a biological substrate identifies HIV-1 gRNA as a rare example of specific substrate and unravels the extent of DDX3 ATPase activity. Analysis of DDX3 binding capacity indicates an unexpected dissociation between its binding capacity and its biochemical activity. We further demonstrate that interaction of DDX3 with HIV-1 gRNA relies both on specific RNA determinants and on the disordered N- and C-terminal regions of the protein. These findings shed a new light regarding the potentiality of DDX3 biochemical activity supporting its multiple cellular functions.

© 2019 Elsevier B.V. and Société Française de Biochimie et Biologie Moléculaire (SFBBM). All rights reserved.

1. Introduction

RNA helicases are involved in all major steps of RNA metabolism from transcription, pre-mRNA processing to RNA transport and decay. Helicases are classified into super-families and families according to conserved sequence motifs, domain composition and oligomeric state [1]. The DEAD-box protein family, which belongs to superfamily 2, forms the largest group of helicases [2]. DEAD-box proteins are characterized by the signature sequence D-E-A-D (Asp-Glu-Ala-Asp) in their conserved Walker B motif [3] and are present in numerous prokaryotes and in all eukaryotes [4,5]. DEAD-box proteins are ATP-dependent RNA binding proteins that remodel RNA structures and RNA-protein complexes [6,7]. DDX3 (encoded by the *ddx3x* gene) and Ded1p, its yeast orthologue, define the Ded1/DDX3 subfamily with additional conserved N− and C-terminal regions involved in RNA and protein interactions [8].

DDX3 is involved in numerous physiological cellular processes such as cell cycle progression, apoptosis, RNA metabolism, from transcription to translation, splicing, nuclear export and mRNA degradation [8−10]. In addition, DDX3 has been identified as a key molecule in pathological situations such as in various cancer types [11] as well as during diverse viral infections [12,13], including HIV-1 infection [14]. DDX3 is essential to HIV-1 replication through its role both in the CRM1-mediated nuclear export of the unspliced HIV-1 gRNA [14,15] and in the viral genome translation [16]. Indeed, DDX3 shuttles from the nucleus to the cytoplasm where it allows the ATP-dependent formation of cytoplasmic granules that are required for HIV-1 translation [17].

HIV-1 full length RNA, known as the genomic RNA (gRNA), is multifunctional. It acts both as an mRNA that is translated to yield the Gag and Gag-Pol polyproteins and as a genome that is encapsidated within the virions. Translation of eukaryotic mRNA relies on

* Corresponding authors.
E-mail addresses: bruno.sargueil@parisdescartes.fr (B. Sargueil), nathalie.chamond@parisdescartes.fr (N. Chamond).

https://doi.org/10.1016/j.biochi.2019.03.008
0300-9084/© 2019 Elsevier B.V. and Société Française de Biochimie et Biologie Moléculaire (SFBBM). All rights reserved.

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

the recognition of the 5′-cap structure to recruit the 43S pre-initiation complex (composed of the 40S subunit, eIF1, eIF1A, eIF3 and eIF2-Met-tRNAi^{Met}). This is mediated by the eIF4F complex that allows cap recognition through eIF4E and unwinding of mRNA secondary structures by the eIF4A helicase. The eIF4G platform coordinates 40S and mRNA association through its connection with eIF4E, eIF3 and the Poly(A) Binding Protein (PABP). The 5′-UTR of HIV-1 gRNA is long, highly structured and thus predicted to impede the scanning step of the canonical cap-dependent initiation pathway. More specifically, the presence of the 5′ proximal extra-stable stem loop TAR strongly inhibits translation of HIV-1 gRNA [18]. The exact molecular mechanism responsible for this inhibition remains undetermined; it could hinder the cap accessibility to eIF4E, and/or prevent ribosome binding or progression on the mRNA. Nonetheless, HIV-1 can initiate translation either by internal recruitment of the ribosome through RNA domains called IRESs (internal ribosome entry sites) or by the classical cap-dependent mechanism [19]. DDX3 has been shown to be required for HIV-1 gRNA efficient translation *in cellulo*. More precisely, the helicase is required for 43S initiation complexes recruitment to the 5′ cap structure of HIV-1 gRNA [16]. However, its exact role in HIV-1 translation and in the initiation mechanisms involved are yet to be defined.

DEAD-box helicases share 12 conserved motifs that are involved in ATP binding and hydrolysis, in RNA binding and in the coupling of ATP- and RNA-binding; which are biochemical properties supporting their biological functions. Ded1p displays ATP-dependent helicase and RNA-dependent ATPase activities [20] as well as additional properties including RNP remodeling [21,22], stable RNA binding [23], AMP sensing [24] and strand annealing [25] activities. The accepted model of RNA duplex unwinding is by “local strand separation” [26] where DEAD-box helicases load directly onto the duplex and separate a limited number of base pairs in an ATP-dependent fashion (ATP binding). ATP hydrolysis is necessary for RNA release and enzyme recycling. More recently, it has been shown that Ded1p activity relies on the assembly of three Ded1p “protomers” [27]. Two “protomers” loaded onto single-stranded RNA stimulate the binding of a third one onto duplex RNA, allowing for strand separation.

The demonstration of DDX3 basic properties (RNA-dependent ATPase and RNA unwinding activities) was first attained by Yedavalli et al., in 2004 [14] although the protein had been identified in 1998 by Park et al. [28]. Nonetheless, through three seminal papers approaching the biochemical characterization of DDX3 [14,29,30], the recombinant protein was described as displaying distinct features compared to its yeast orthologue. For instance, DDX3 was reported to present a relatively high level of RNA-independent ATPase activity and an ATPase activity that could be stimulated by RNA and DNA (DNA-stimulated ATPase activity is unusual for DEAD-box helicases). In addition, a DDX3 fragment lacking critical RNA binding motifs and motif VI (which contains the catalytically critical arginine finger) maintained intrinsic ATPase activity. Finally, DDX3 was reported to hydrolyze all NTPs and not to be restricted to ATP. More recently, Sharma et al. performed a thorough characterization of DDX3 helicase activity using short model substrates, thereby identifying DDX3 as a “true” DEAD-box helicase. In addition, the enzyme displays some specific features as compared to its yeast orthologue Ded1p such as reduced strand annealing activity and preference for 3′ single-strand RNA [31].

To appreciate the link between the biochemical properties of DDX3 and its role in HIV-1 gRNA translation, we explored the relevance of DDX3 biochemical parameters in the context of a biological substrate. We produced the recombinant protein and performed a thorough characterization of its ATPase activity when stimulated by HIV-1 gRNA; activity of which cannot be fully attained with short model substrates. The production of the DQAD mutant allowed us to analyze the binding properties of DDX3 to RNA and to unravel an unexpected dissociation between its binding capacity and its biochemical activities. We further demonstrate that interaction of DDX3 with HIV-1 RNA relies both on specific RNA determinants and on the disordered N− and C-terminal regions of the protein. Altogether, our results specify the biochemical properties of DDX3 and the RNA requirements for a biological target.

2. Material and methods

2.1. Cloning, expression and purification

The pET21 construct containing the complete open reading frame of DDX3 (*ddx3x*) was a kind gift of T. Olhmann and R. Soto-Rifo. The coding sequence was sub-cloned into pETM41 expression vector leading to an N-terminally hexahistidine-tagged fusion protein of MBP (Maltose Binding Protein) and DDX3, with a TEV protease recognition site in between. QuickChange site-directed mutagenesis (Stratagene) was used to generate the ATPase-deficient Motif II mutant (DQAD). Wild-type and mutant proteins were produced following the same protocol. Proteins were expressed in BL21 (DE3) cells (Invitrogen), grown in LB medium and protein expression was induced at $OD_{600} = 0.6$ upon addition of 0.1 mM IPTG. Bacteria were grown for 2 h at 37 °C. Cells were harvested by centrifugation, cell pellets were resuspended in Buffer A containing 50 mM Tris-Cl pH 8/300 mM KCl/10 mM Imidazole in the presence of proteases inhibitor (Roche) and cell disruption was obtained after 3 passages in the Eaton-press. The lysate was cleared at 47000 g at 10 °C for 20 min and the supernatant was loaded on a Ni-NTA His Trap column (GE Healthcare), pre-equilibrated in Buffer A. The column was washed with 10 Column Volumes (CV) in Buffer A and 10 CV in Buffer A containing 1 M KCl to avoid nucleic acids contaminations. Recombinant protein was eluted in a gradient of Buffer A supplemented with 500 mM Imidazole. Fractions containing MBP-DDX3 were pooled, dialyzed overnight against 20 mM Tris-Cl pH 8/200 mM KCl/1 mM DTT/5% glycerol (Buffer B) and loaded on a MonoQ 10/100 GL column (GE Healthcare) pre-equilibrated in Buffer B. The protein was eluted in a linear gradient to MonoQ Buffer B but supplemented with 1 M KCl. Whenever indicated, the protein was cleaved with the TEV protease to remove the MBP fusion protein. The protein was further purified by the use of an Heparin column pre-equilibrated in Buffer C (20 mM Tris-Cl pH 8, 200 mM KCl, 5% glycerol and 1 mM DTT) and a linear gradient to Buffer C supplemented with 2 M KCl. Fractions containing pure protein were pooled, dialyzed against 20 mM Tris-Cl pH8, 250 mM KCl, 1 mM DTT and 30% glycerol, aliquoted and conserved at −60 °C. All purifications steps were monitored by 8% SDS-PAGE and Coomassie-blue staining. The purified protein was quantified by performing a 230−320 nm OD spectrum and examination of the OD 260/280. Purified Ded1p is a kind gift of K. Tanner and J. Banroques.

2.2. In vitro transcription

RNAs were directly transcribed using the T7 RNA polymerase from polymerase chain reaction (PCR) products containing the T7 promoter sequence (5′-TAATACGACTCACTATA-3′). *In vitro* transcription was conducted in a final volume of 100 μl, using T7 RNA polymerase, 5 mM DTT, 5 mM rNTPs, 1x transcription buffer (40 mM Tris-Cl pH 8.0, 25 mM $MgCl_2$, 1 mM spermidine) and 20 U RNAsin (Promega) and incubated 2 h at 37 °C. Upon synthesis RNAs were treated with 2 U DNAse RQ1 (Promega) for 30 min at 37 °C. RNA was precipitated in the presence of 2.5 M LiCl, centrifuged at 16000 g for 30 min at 4 °C, washed with 70% ethanol, resuspended

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

in 50 μl of nuclease-free water and purified through Sephadex G50 columns. RNAs concentrations were determined spectrophotometrically (Biospec NanoDrop Technology), and RNA integrity was monitored by electrophoresis on agarose gels. Radiolabeled RNAs were transcribed and purified as above but in the presence of 1 mM UTP and 3000 mCi/mmol of α^{32}-UTP.

2.3. ATP hydrolysis assay

ATPase assays were performed in 20 μl reactions containing 100–500 nM of purified protein in 20 mM Tris-Cl pH 8, 80 mM KCl, 2.5 mM $MgCl_2$ and 1 mM DTT in the presence of 1–1000 nM RNA. The reaction was initiated by the addition of an ATP solution (1–3200 μM of cold ATP/$MgCl_2$ and 3 nM of 800 Ci/mmol α^{32}P-ATP as the final concentrations in the reaction) and allowed to proceed for 10–20 min at 37 °C. Aliquots (2 μl) were removed and spotted on polyethylenimine cellulose plates at various time points, hydrolysis products were separated by thin layer chromatography (TLC) in 0.5 M KH_2PO_4 pH 3.4, exposed and analyzed using a BAS-5000 scanner (Fujifilm). For competition assays, the ATP solution (10 μM) was supplemented by 100 μM of the different NTPs. The different nucleic acids used for ATPase activity stimulation are indicated in the Supplemental Table 1. The apparent affinity for the ATP substrate and the apparent rate of ATP hydrolysis were calculated by measuring the variation of initial velocities of the ATPase reaction, as a function of ATP concentration. Data were analyzed according to the Michaelis-Menten equation: $V = (k.E_0.S^n)/(Km + S^n)$, where k is the catalytic rate, E_0 is the input enzyme concentration, S is the ATP concentration and n is the cooperativity index. Experimental data and statistical analyses were performed using GraphPad Prism.

2.4. Duplex unwinding assay

The RNA:RNA and RNA:DNA substrates for the unwinding assays were formed by annealing a 5′-end radiolabeled short RNA or DNA oligonucleotide to either the 5′ or 3′ end of a longer RNA oligonucleotide (Supplemental Table 1). Oligonucleotides were purchased from Eurogentec. Unwinding assays were performed in unwinding buffer (40 mM Tris-Cl pH 8, 50 mM NaCl, 0.5 mM $MgCl_2$, 2 mM DTT and 0.01% NP40) in 20 μl reactions containing 100 nM of purified protein, 2 mM ATP/$MgCl_2$ and 5 nM of the different substrates (as indicated in the Figures). Reactions were initiated by the addition of ATP, incubated for 5 min at 37 °C and stopped with a solution containing 1% SDS, 50 mM EDTA, 0.1% xylene cyanol, 0.1% Bromophenol Blue and 20% glycerol. Samples were loaded on 15% native acrylamide gels, run at 4 °C for 1 h, the gels were dried, and the radiolabeled RNAs were visualized and analyzed using a BAS-5000 scanner and the software MutliGauge (Fujifilm). The primer release (R) is calculated from endpoint data as the following: $(P_f\text{-}P_i)/(1\text{-}P_i)$ where P is the initial (i) or final (f) fraction of free primer.

2.5. Elecrophoretic mobility shift assay

Various concentrations of purified protein were mixed with 2 mM ATP, AMP-PNP or ATPγS and 7 nM radiolabeled RNA:RNA or RNA:DNA substrate and incubated in unwinding buffer for 5 min at 37 °C. Samples were then chilled on ice, stopped with a 5X solution containing 50 mM EDTA, 0.1% Xylene Cyanol, 0.1% Bromophenol Blue and 20% glycerol. Samples were loaded on 4% native polyacrylamide gels, which were run in cold for 1 h at 100 V, dried, exposed and analyzed using a BAS-5000 scanner and the software MultiGauge (Fujifilm).

2.6. Filter binding assay

Radio-labeled RNA (2.5–10 nM) was denatured by heating to 80 °C for 2 min and then cooled to room temperature in FB buffer (20 mM Tris-Cl pH 8, 100 mM KOAc, 100 mM KCl, 2.5 mM $MgCl_2$, 2 mM DTT). Serial dilutions of recombinant protein were prepared extemporaneously, added to a 10 μl reaction and incubated at 37 °C for 20 min. The reactions were then used for filter binding assays. Filter binding was accomplished essentially as previously described [32] using two filters: 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 BAS-5000 scanner and the software MultiGauge (Fujifilm). To determine the apparent *K*d, the data were fit to the Langmuir isotherm described by the equation: $\theta = P/[P + Kd]$, where θ is the fraction of RNA bound and P is the protein concentration.

3. Results

3.1. Production of soluble and active DDX3

To facilitate the biochemical characterization of DDX3, we overexpressed the protein in *E. coli*. In first attempts, the full-length protein was cloned in pET vectors fused to a His_6-tag in C– or N-terminus. Using standard protocols for protein expression and extraction, we collected large amounts of DDX3 but entirely in the insoluble fraction. All our efforts aiming at either reducing protein expression or synthesis rate, or facilitating protein folding by chaperonin (GroEL/ES) overexpression failed to yield the protein in the soluble fraction. Similarly, varying the lysis conditions or solubilizing the protein with denaturants such as urea or guanidium chloride did not result in the production of sufficient amounts of soluble active recombinant protein. We then utilized the His_6-Maltose Binding Protein (MBP) fused with the N-terminus of DDX3 (Fig. 1A). Most of the protein was in the soluble fraction of bacterial extracts. The recombinant protein was purified on a nickel-agarose column, its apparent molecular mass (about 120 kD) was in agreement with its estimated weight and the identity of the protein was confirmed by western blot analysis using anti-DDX3 antibody (data not shown). Fractions enriched in MBP-DDX3 were further pooled and applied onto a MonoQ column. The resulting recombinant protein was about 90% pure as judged by Coomassie staining (Supplemental Fig. S1). The absence of nucleic acids contamination was obtained by performing a wash step with 1 M salt during the nickel column purification followed by the MonoQ column which is known to trap nucleic acids; and ascertained by a 230–320 nm OD spectrum (see Methods). To exclude a drastic influence of the MBP-tag on DDX3 properties, we performed a TEV-cleavage step followed by a third purification step on a heparin column to separate the cleaved MBP-tag and the TEV protease from DDX3. In addition, to corroborate the importance of the DEAD motif on DDX3 properties, we constructed a plasmid harboring the DQAD motif. Molecular replacement of the glutamate moiety by a glutamine in Motif II results in the absence of ATP hydrolysis but does not excessively interfere with ATP binding or with nucleic acid interaction [33]. The DQAD mutant was purified in the same conditions as the wild-type protein. The yield and purity were similar to those obtained with the wild-type (Supplemental Fig. S2).

3.2. DDX3 is an active RNA-dependent ATPase

The ATPase activity was measured by using a TLC assay, in which ATP hydrolysis is directly assessed by Pi release quantification.

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

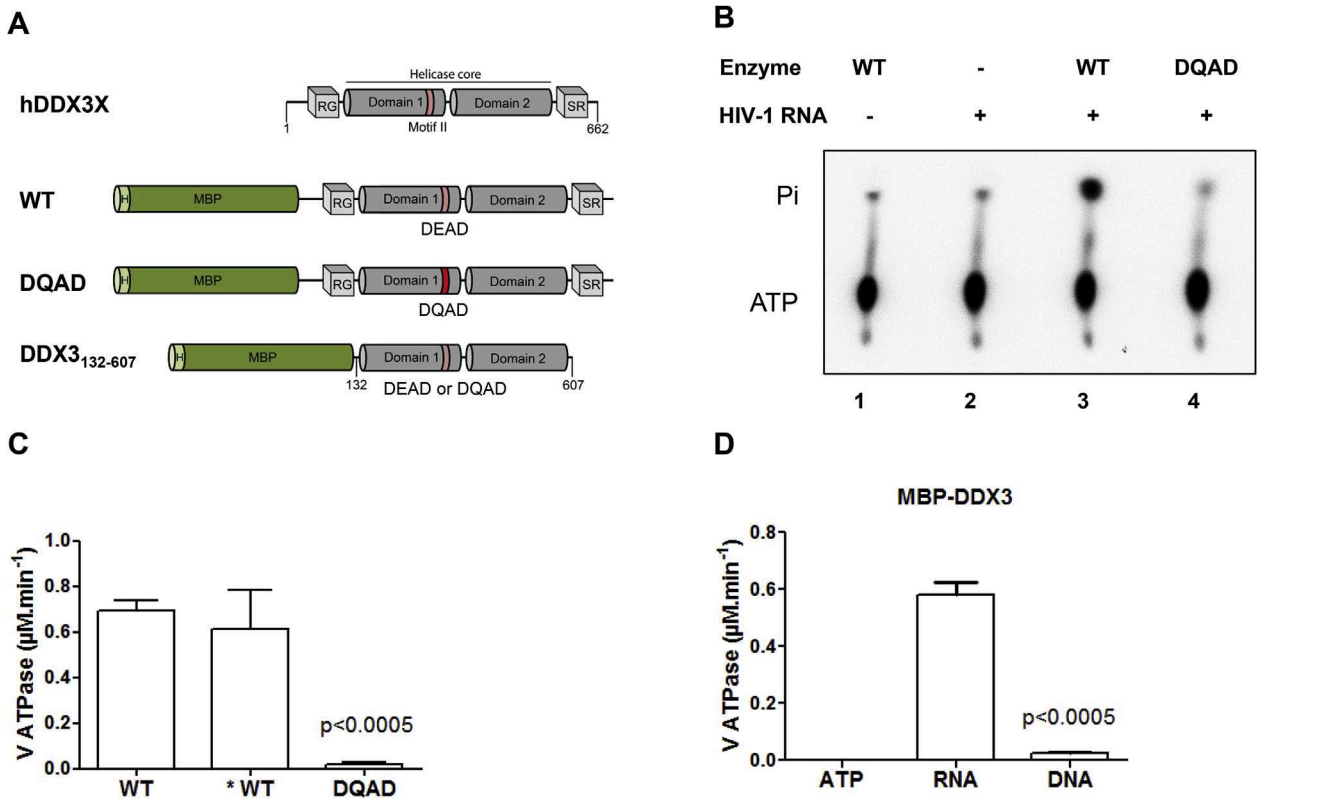


Figure 1. Purification of an active recombinant DDX3 protein. (A) Schematic representation of DDX3 and of the derivative constructs used in this study. H, His6x tag; RG, arginine and glycine rich N-terminal domain; SR, serine rich C-terminal domain. (B) The ATPase activity of wild-type (WT) or mutant (DQAD) MBP-DDX3 (100 nM) was assessed by thin layer chromatography measurement of the amount of Pi released after 10 min in the absence (lane 1) or presence (lanes 2–4) of 1 μM HIV-1 RNA and 10 μM ATP. (C) The initial rate of Pi release was measured in the presence of 100 nM of the indicated protein, 100 μM ATP and 100 nM HIV-1 RNA. (D) Same as C but in the absence of nucleic acid (ATP) or in the presence of 100 nM HIV-1 RNA (RNA) or its corresponding dsDNA (DNA) and 100 nM of MBP-DDX3. HIV-1 RNA corresponds to the 5′UTR of HIV-1 followed by Gag ORF (nt 1–1636). The values represent the average and standard deviation of at least three independent experiments. p indicates the P value issued from a paired *t*-test.

Fig. 1B represents typical end-point data generated by the ATPase assay. In absence of RNA or of the enzyme, only traces of Pi were detected contrary to what is observed in the presence of DDX3 and RNA (lane 3); therefore proving that DDX3 has no intrinsic ATPase activity which is consistent with the absence of contaminating nucleic acids in our preparation. This activity was not observed with the DDX3-DQAD mutant (Fig. 1B, lane 4) thereby establishing that measured ATPase activity is that of DDX3. Importantly, when the MBP-tag was removed from DDX3 (*WT), similar yields of Pi release were obtained indicating that the presence of the MBP at the N-terminus of DDX3 does not alter its activity (Fig. 1C). Since the cleaved protein was less stable in our hands, for practical reason, but also to ensure that the protein does not precipitates during the assays, all subsequent experiments were performed with MBP-DDX3 (further referred to as DDX3). In addition, when dsDNA was used to stimulate the ATPase activity, no significant increase in Pi release was observed (Fig. 1D) therefore demonstrating that DDX3 ATPase activity is strictly RNA-dependent.

To test for a possible specificity of the ATPase reaction toward the nucleic acid substrate, we compared a variety of nucleic acids (Fig. 2A, sequences are indicated in the Supplemental Table 1) that differ by their nature (DNA or RNA), length or structure (ss, single-stranded or ds, double-stranded). As shown in Fig. 2B and C, while DNAs are unable to stimulate DDX3 ATPase activity, all the RNAs tested activated the ATPase activity, although with different levels (ranging from 15 to 90% of Pi release). The highest activity was observed with RNA HIV-1, which could result from the size difference of the RNAs used or from their structural status, or from both. In order to further evaluate if one of these parameters influences the ATPase activity of DDX3, we expressed the amount of RNA as the nucleotide concentration and compared the RNA stimulation obtained with RNA HIV-1 and RNA TAR (57 nt long). As can be observed in Fig. 2D, for equivalent nucleotides concentrations, RNA HIV-1 induces the release of twice the amount of Pi as compared to RNA TAR (25% and 13%, respectively). Importantly, RNA HIV-1 contains the extra-stable stem loop TAR that has been shown to be a critical determinant for HIV-1 gRNA interaction with DDX3 [17]. However, in our ATPase assay, the isolated TAR is not sufficient to recapitulate HIV-1 gRNA stimulation of DDX3 activity. In addition, when we compared two RNAs that are similar in size (RNA TAR and RNA 64) but that differ from their structural status (extra-stable TAR stem loop vs essentially single-stranded RNA of random sequence), both species stimulated the ATPase activity of DDX3 to a similar extent. Taken together, our results suggest that neither single- nor double-stranded RNA are an optimal substrate to stimulate DDX3 ATPase activity but rather, more complex RNAs alternating single- and double-stranded regions.

Next, we performed Michaelis-Menten kinetics and characterized the ATP hydrolysis performed by DDX3 (100 nM) in the presence of RNA HIV-1 (10–1000 nM) and ATP (1–1600 μM, Supplemental Fig. S3). The linear regression of Pi release versus time allows for the initial reaction rate determination V (Fig. 3A). First, kinetics were assayed in the presence of various RNA concentrations (1–1000 nM) and saturating ATP (2 mM ATP) conditions. As can be observed in Fig. 3B, saturation of the apparent Vmax is attained in the presence of RNA HIV-1 with a $k_{1/2\ [RNA]}$ of

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

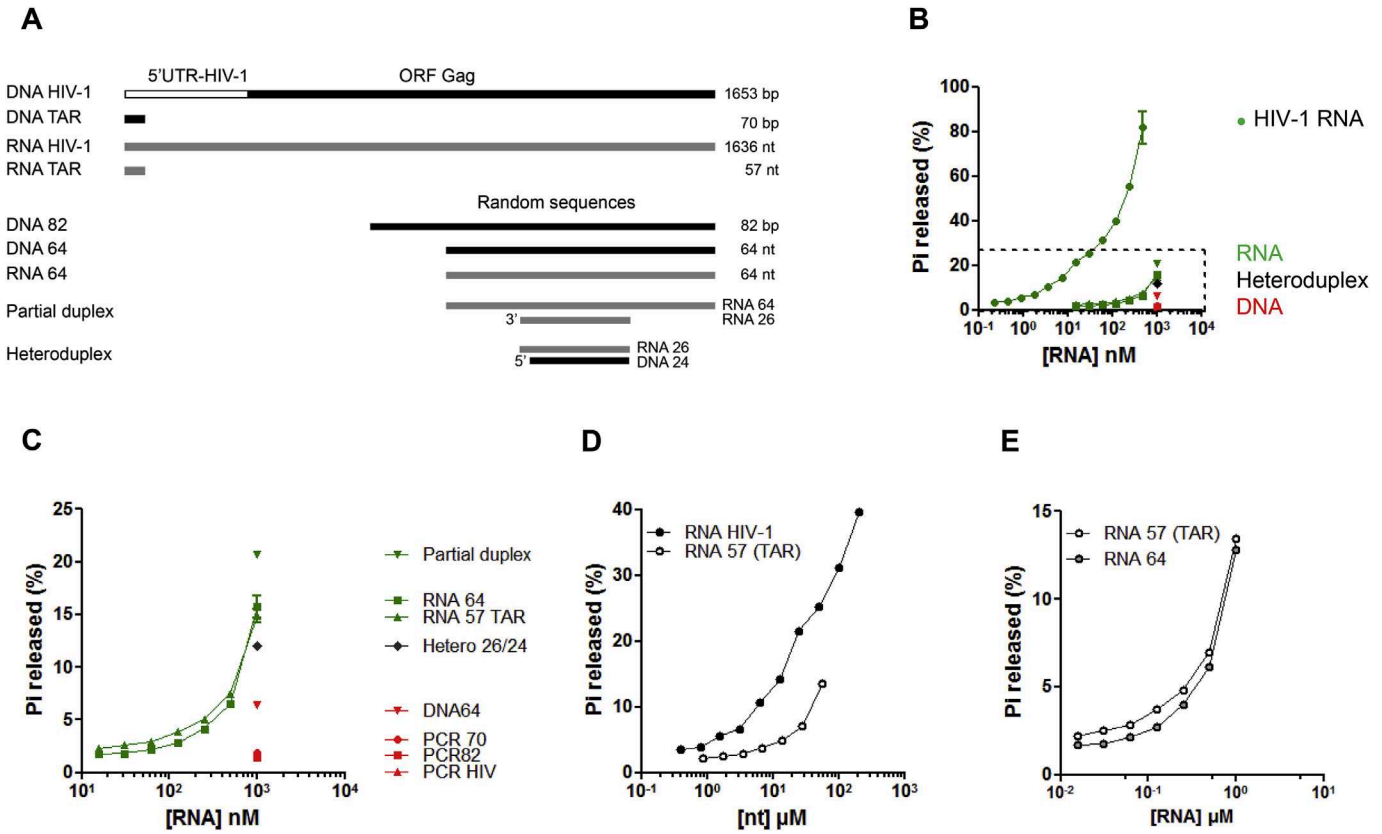


Fig. 2. Stimulation of DDX3 ATPase activity with various RNA or DNA substrates. (A) Schematic representation of the different substrates used. (B) ATP hydrolysis was measured in the presence of RNA HIV-1 (green circle), less than 100 nt RNA (green), DNA (red) or heteroduplex NA substrates (black) and 200 nM ATP and 600 nM enzyme. For clarity, short substrates (dotted square) are enlarged in (C). (D) and (E) ATP hydrolysis was measured in the presence of the indicated substrates at different concentrations expressed as nucleotide (nt) or molar (RNA) concentrations to account for the size difference whenever relevant.

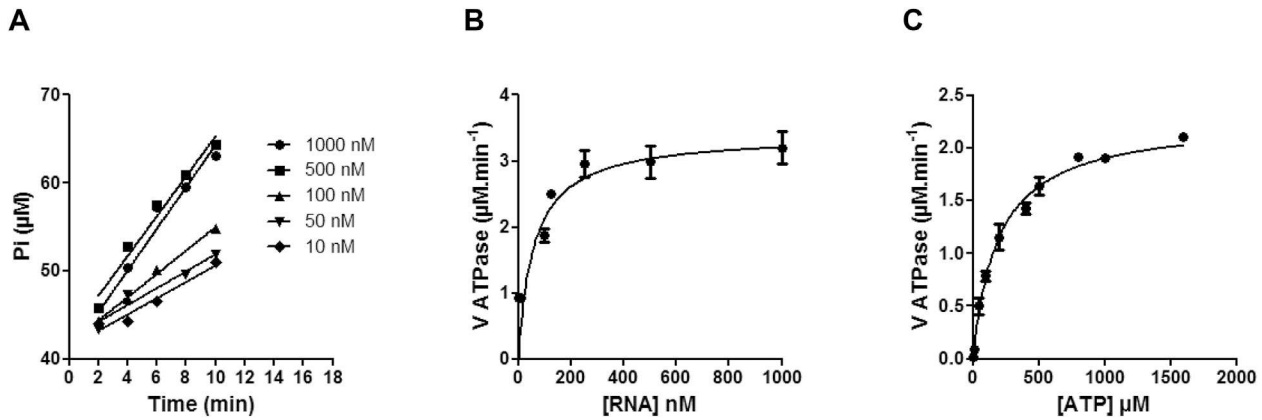


Fig. 3. Determination of the kinetic parameters of DDX3 ATPase activity. (A, B) ATPase reactions were performed in the presence of 100 nM DDX3, 10–1000 nM HIV-1 RNA and 2 mM ATP. The linear regression of Pi release versus time allows for the initial rate (V) determination at different RNA concentrations. (C) In the presence of 100 nM DDX3, 100 nM HIV-1 RNA and 1–1600 µM ATP, DDX3 exhibits Michaelis-Menten kinetics with an apparent KmATP = 196 ± 26 µM and an apparent Vmax of 2.3 ± 0.1 µM.min-1. The values represent the average and standard deviation of at least three independent experiments.

59.6 ± 21.0 nM. This result differ by several orders of magnitude from what was recently reported using short model substrates ($K_{1/2\ [RNA]}$ of 2.3–3.4 µM, [31]). It is important to note here that we and others [30,31] could not attain RNA saturating conditions with short synthetic substrates, even at high concentrations (up to 10 µM). We next tested ATP hydrolysis at unsaturated to saturated ATP concentrations to determine the kinetic parameters of DDX3 towards ATP. As can be observed in Fig. 3C, DDX3 exhibits Michaelis-Menten kinetics with an apparent K_{mATP} of 193 ± 26 µM and an apparent K_{cat} of 23 ± 10 min^{-1}.

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

3.3. DDX3 is an active ATP-dependent RNA helicase

To validate the helicase activity of our recombinant protein, duplexes previously used (Fig. 4A [27,34]), were incubated in the presence of saturating ATP (2 mM) and 140 nM enzyme (Fig. 4 and Supplemental Fig. S4). As can be observed in Fig. 4B, DDX3 efficiently unwinds RNA/RNA duplexes in the presence of ATP, similarly to Ded1p. In addition, our results are consistent with a previous report of DDX3 helicase activity in the presence of the D3 duplex [34]. The DDX3-DQAD mutant which cannot hydrolyze ATP did not display any helicase activity and this was observed independently of the substrate used. Nonetheless, we observed with the mutant a higher band shift that could reflect its interaction with the duplexes in the presence of ATP (see further). Importantly, when duplexes (RNA/RNA or RNA/DNA) with either 5′ or 3′ single stranded overhangs were tested, DDX3 displays a marked preference for 3′-tails, with an apparent two-fold decrease in DDX3 helicase activity whenever 5′-tailed substrates were used. Finally, blunt duplexes, independently of their nature, are not substrates for the helicase activity even if incubated for longer periods of time (data not shown). We did not evaluate further the kinetic parameters of DDX3 helicase activity. Nonetheless, taken together, these results confirm that DDX3 is an active ATP-dependent RNA helicase although it can also be active on RNA/DNA heteroduplexes, and it displays a preference for substrates presenting 3′ single-stranded regions.

3.4. Characterization of DDX3 RNA binding activity on synthetic substrates

Our results on the helicase activity of DDX3 indicate that the functional binding for RNA depends on the identity of the molecule used. However, apparent affinities deduced from the helicase activity potentially reflect the actual binding step as well as steps of the ATP-hydrolysis reaction, or of the helicase reaction or of the products release. Therefore, to investigate the preference of DDX3 for 3′-single stranded substrates, we examined its binding capacity through mobility shift assays (Fig. 5). In the presence of ATP, it is not possible to observe DDX3 bound to its RNA substrate, which could be a consequence of the helicase rapid turnover. More surprisingly, when the biochemical activities of DDX3 were compromised by the use of non- or poorly-hydrolysable ATP-analogs such as AMP-PNP or ATPγS, we could not observe stable complexes even in the presence of increased DDX3 concentrations (up to 1.7 μM, data not shown). Nevertheless, whenever the binding capacities of the DQAD mutant were assessed in the same conditions (5 nM substrate, 50 nM protein), a clear band shift was only observed in the presence of ATP and to a lesser extent in the presence of ATPγS, but not in the presence of AMP-PNP. These results suggest that the DQAD-mutant, and most probably the wild-type DDX3, forms complexes with its RNA substrate in the presence of ATP but not in the presence of non-hydrolysable analogs, potentially because of an incorrect geometry at the γ−phosphate position. Therefore,

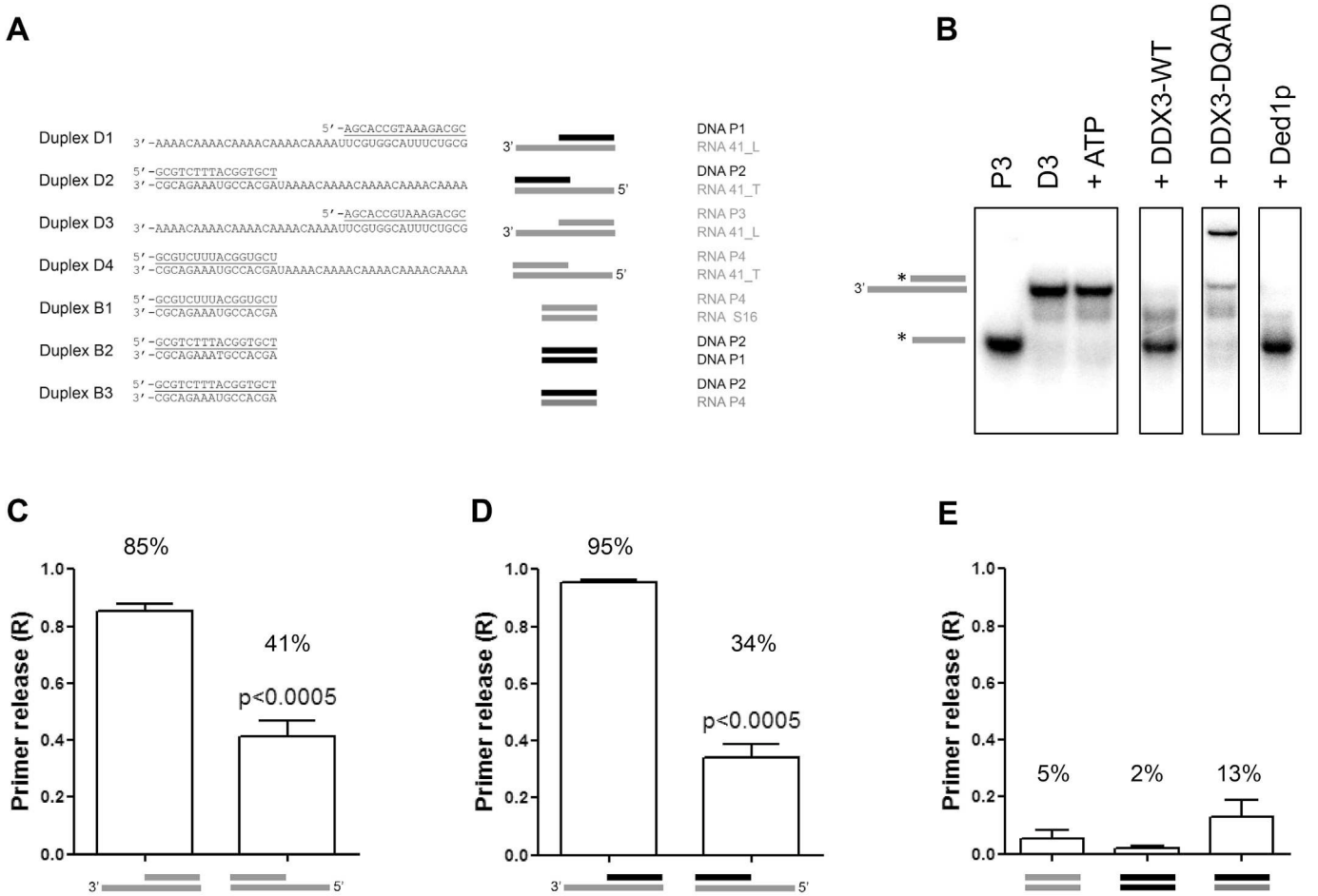


Fig. 4. DDX3 is an active helicase. (A) Sequences and schematic representation of the substrates used, from [27]. (B) Representative PAGE (complete gels in Supplemental Figure S4) for unwinding reactions performed in the absence (+ATP) or in the presence of DDX3 (+DDX3-WT), of the DQAD mutant (+DDX3-DQAD) or of Ded1p (+Ded1p). (C) (D) (E) Quantification of duplex unwound by 140 nM of DDX3 in the presence of 5 nM substrate and saturating 2 mM ATP. The cartoons represent the helicase substrates described in (A). The values represent the average and standard deviation of at least three independent experiments. p indicates the P value issued from an unpaired *t*-test.

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

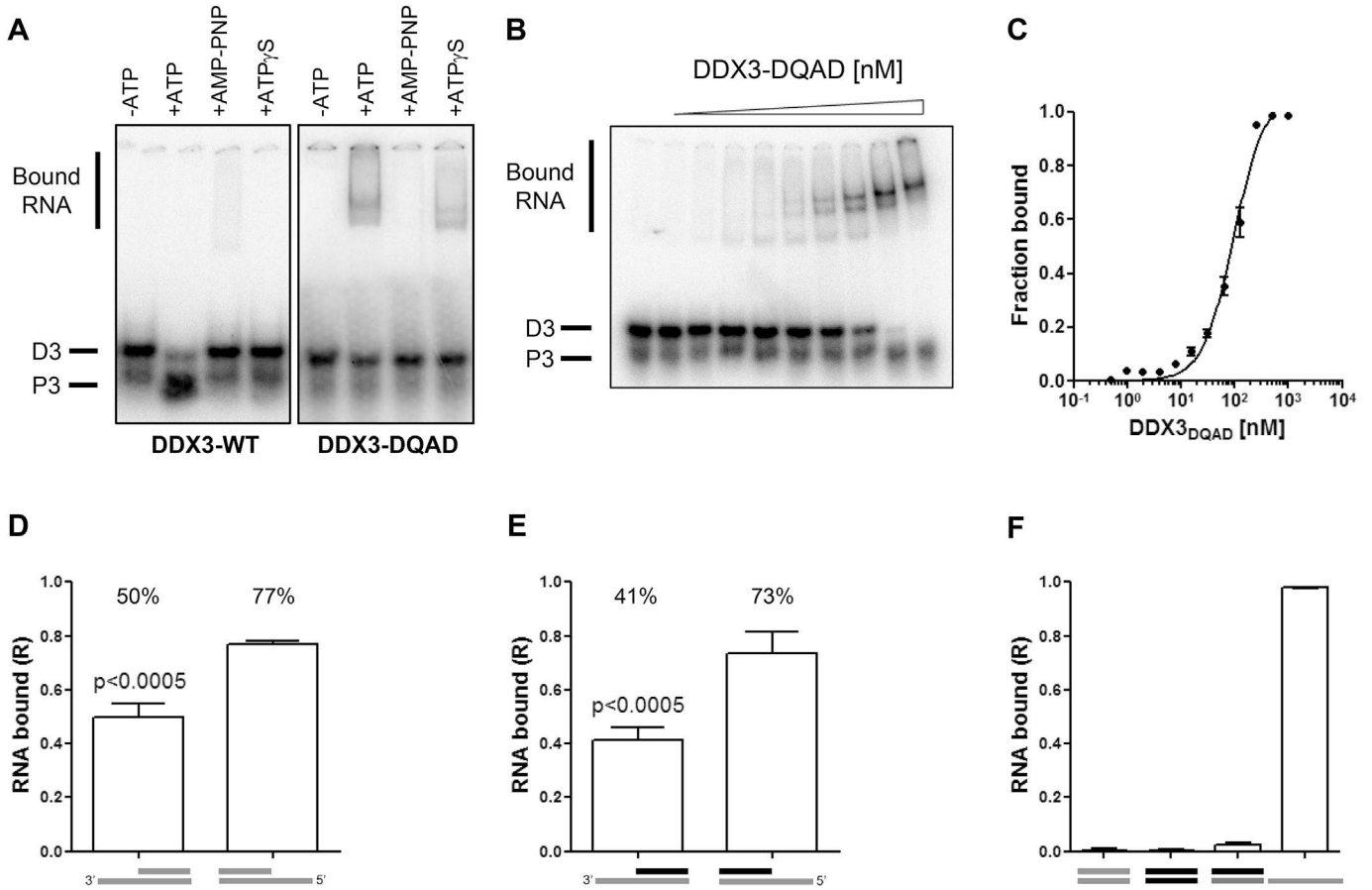


Fig. 5. DDX3-DQAD interaction with model substrates. 16 bp RNA/RNA duplexes (D3, 5 nM) were incubated in the presence of (A) 50 nM protein or of (B) increasing concentrations of DDX3-DQAD (0–500 nM) and 2 mM ATP. (C) The quantitation of bound duplexes allows for the determination of the following binding parameters: Bmax = 1.00 ± 0.02, h = 1.7 ± 0.2 and Kd = 80.6 ± 5.1 nM. (D) (E) and (F) The different duplexes were tested for their capacity to form stable binary complexes. 5 nM of the indicated duplexes were incubated with 50 nM of DDX3- DQAD for 5 min and the fraction of shifted RNA was quantified. The values represent the average and standard deviation of at least three independent experiments. p indicates the P value issued from an unpaired *t*-test.

substrate binding was characterized by gel-shift assays performed with the DDX3-DQAD mutant (Fig. 5B). Worth noting, at least two shifted species can be observed on Fig. 5A and B, suggesting that at least two molecules of DDX3 can bind on these substrates, although we cannot rule out an additional effect of the RNA structure. We used the standard model substrates, either the 16 bp duplexes with a 25 nt 5′-or 3′-tail, the blunt 16 bp duplexes or a 41 nt ss-RNA. Quantitative analysis of the fraction of bound duplexes (R) upon DDX3 titration in the presence of the D3 substrate (RNA/RNA substrate with a 3′-overhang) reveals an apparent equilibrium constant (K_d) of 80.6 ± 5.1 nM (Fig. 5C). Surprisingly, when the different helicase substrates were tested, a 1.5-fold increase in binding is observed with substrates presenting a 5′-tail as compared to those presenting a 3′-tail, independently of the nature of the top strand (RNA or DNA) (Fig. 5D–E). This result suggests that the RNA selectivity exhibited by DDX3 helicase activity is not directly proportional to the level of RNA binding. Complexes can be observed on single-strand RNA, while they fail to form with blunt-end RNA or DNA/RNA duplexes (Fig. 5F). Altogether, our data show that the presence of ssRNA is sufficient to promote RNA binding and that the presence of a ss-ds junction can modulate its efficiency. This observation is consistent with the RNA requirements described for Ded1p [20].

3.5. Determination of the shortest HIV-1 substrate necessary to activate DDX3 activity

As mentioned previously, the ATPase activity of DDX3 is only moderate on short synthetic substrates. Therefore, to evaluate the link between the RNA binding capacity and the ATPase activity of DDX3, we produced several RNA substrates derived from HIV-1 gRNA or from a control RNA (Fig. 6A, Supplemental Table 1) and measured their apparent Kd by filter binding assay (Table 1). Of note, this technique is performed in presence of high monovalent salt concentration (300 mM) to prevent high background. Consequently, reported Kds are more elevated than those observed in the low salt concentration (80 mM) used in gel-shift assay, which is less applicable to long RNA substrates. The obtained Kds are not proportional to the size of the RNA molecules (Fig. 6B and Table 1) but indicate that a minimal size of approximatively 200 nt is required for RNA to efficiently bind DDX3 in this assay. In addition, RNAs encoding the renilla luciferase (control RNAs) coding region display Kds that are similar to those obtained with HIV-1-derived RNAs. We then compared the capacity of the different fragments to stimulate the ATPase activity of DDX3 (Fig. 6C, Table 2). As described above for short synthetic substrates, HIV-1 or control RNAs of less than 200 nt fail to efficiently stimulate DDX3 ATPase activity. Over a threshold

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

Table 1
DDX3-DQAD binding affinities to HIV-1 gRNA fragments. Filter binding assays were performed in the presence of 5 nM RNAs and increasing concentrations of DDX3-DQAD. The results are the mean and standard deviation of two to three independent experiments. The normalization per nucleotide is shown to assess the size effect.

RNA	Nucleotides	Kd (nM) Mean ± s.e.m.	Length (nt)	Kd per nt ($pM.nt^{-1}$)
FL	G_1-U_{1636}	30.5 ± 4.8	1636	18.6
UTR	G_1-A_{416}	89.4 ± 14.7	416	215.0
Gag ORF	A_{336}-U_{1636}	9.6 ± 3.6	1301	7.4
	G_{690}-U_{1636}	37.1	947	39.2
Gag IRES	A_{336}-U_{851}	32.2 ± 5.8	516	62.4
	A_{336}-A_{715}	37.1 ± 5.0	380	97.6
	A_{336}-G_{656}	60.7 ± 8.5	321	189.1
	A_{336}-C_{540}	54.3 ± 17.6	205	264.9
	G_{406}-C_{540}	>500	135	–
	A_{336}-C_{477}	59.2 ± 12.7	152	389.5
Δ Site1 (Δ1)	A_{336}-U_{881}Δ363-485	39.0 ± 11.4	1178	33.1
Δ Site2 (Δ2)	A_{336}-U_{881}Δ632-756	43.2 ± 8.2	1195	36.2
CTRL 1		63.0 ± 29.6	1031	61.1
CTRL 2		31.8 ± 5.2	505	63.0
CTRL 3		60.8 ± 31.6	243	250.2
CTRL 4		106.8 ± 34.8	183	583.6
Duplex D3		>500	57	–
RNA L		>500	41	–

around 200 nt we do not observe any correlation between the ATPase activity and the length of the molecule used to stimulate. Interestingly, control RNAs do not stimulate the ATPase activity to the same extent as HIV-1 derived RNAs of similar size; therefore suggesting the presence of some specific elements in HIV-1 RNA. In order to identify a shorter specific substrate, HIV-1 G_1-U_{1636} was subjected to progressive deletion. Removal of the 5′-UTR in the G_1-U_{1636} (leading to A_{336}-U_{1636}) is beneficial to the ATPase activity stimulation. Worth noting, the 5′-UTR by itself is a rather poor substrate (G_1-A_{416}). Additional 3′ deletions lead to an optimal fragment (A_{336}-U_{851}) corresponding to Gag-IRES [35] whereas the remaining 3′ region (G_{690}-U_{1636}) fails to stimulate DDX3 ATPase activity efficiently. Further 5′ or 3′ deletions of this fragment (A_{336}-U_{851}) drastically reduce the ATPase activity of DDX3 defining an optimal fragment of 516 nt (A_{336}-U_{851}) and a minimal fragment of 205 nt (A_{336}-C_{540}). Altogether, these results clearly demonstrate that HIV-1 RNA holds specific determinants for DDX3.

3.6. Evaluation of the requirement for DDX3X terminal extensions

A previous study performed on a truncated version of DDX3 (aa 132–607) that removes the N– and C-terminal intrinsically disordered domains of DDX3 indicated that this mutant only retains partial helicase activity [34]. To evaluate the influence of these regions located outside the helicase core on the biochemical properties of DDX3, we produced and purified the MBP-$DDX3_{132\text{-}607}$ version of the protein as well as the corresponding DQAD mutant (Fig. 1A). In stark contrast with the full length DDX3 ($DDX3_{1\text{-}662}$), the ATPase activity of the truncated $DDX3_{132\text{-}607}$ (in the presence or absence of the MBP-tag) is not significantly stimulated in the presence of 100 nM HIV-1 RNA (Fig. 7A). HIV-1 RNA concentration has to be increased up to 1 μM to observe a significant ATPase activity with $DDX3_{132\text{-}607}$ (Fig. 7B). In the presence of short synthetic substrates (D3 duplex, RNA/RNA duplex with a 3′-tail) $DDX3_{132\text{-}607}$ shows a weak activity comparable to what is observed with the full length protein (Fig. 7C–D). In addition and consistent with previous results [34], $DDX3_{132\text{-}607}$ does not display a significant helicase activity on any of the short substrates (not shown). Taken together, these results show that the helicase core of DDX3 is not sufficient to sustain robust ATPase and helicase activities of DDX3. Finally, in order to evaluate the binding capacities of the truncated version of the protein, we performed gel shift assays. Similarly to what was observed with the full-length protein, we were not able to detect a significant amount of high molecular weight complex with the wild-type version of $DDX3_{132\text{-}607}$. In stark contrast, $DDX3_{132\text{-}607}$-DQAD in presence of ATP forms stable complexes with the model substrates, although with a significantly reduced affinity ($K_d = 138$ nM). With this protein, three to four high molecular weight complexes are observed, suggesting a stoichiometry of 3–4 per substrate (Fig. 7E). The reduced affinity holds true for HIV-1 derived RNAs (Fig. 7F, Table 3) with a consistent 10-fold increase in the observed Kds. These results indicate that the flanking N and C-terminal regions of DDX3 are also required for DDX3 interaction with its RNA substrate.

4. Discussion

DDX3 is multi-functional both in the cell and during HIV-1 replication, as a consequence the understanding of its molecular role(s) necessitates the development of an *in vitro* system to independently study separate functions. In an attempt to elucidate further the role of DDX3 in HIV-1 gRNA translation initiation and to initiate the molecular dissection of this mechanism, we first overexpressed DDX3 in *E. coli* and performed the characterization of its biochemical properties. It is important to note that in the absence of the MBP-module, the isolated full-length DDX3 is poorly soluble and hardly amenable to extensive biochemical assays. In spite of these difficulties, our purification scheme allows us to produce mg amounts of an active protein. For instance, we performed ATPase assays [20,30] and confirm that DDX3 is active and that this activity relies on the integrity of its DEAD domain. However, we and others [31] could not reproduce a basal ATPase activity of DDX3 in absence of nucleic acids, nor its stimulation by dsDNA; which are controversial properties for DEAD-box proteins. Short nucleic acids substrates only induce a weak stimulation of the ATPase activity and saturation could not be obtained even in presence of up to 10 μM of such substrates. In contrast, the stimulation of the ATPase activity saturates in presence of 200 nM of the first 1636 nucleotides of HIV-1 gRNA. We were then able to determine the kinetic parameters of DDX3 ATPase activity, and obtained a significantly increased K_{cat}/K_M value of 1.12 μM min^{-1} as compared to the values reported by Garbelli et al. or by Sharma et al. (0.07 and 0.08 μM min^{-1} respectively). The K_{cat}/K_M ratio is a measure of the catalytic efficiency; nonetheless, it can mask differences in individual parameters,

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

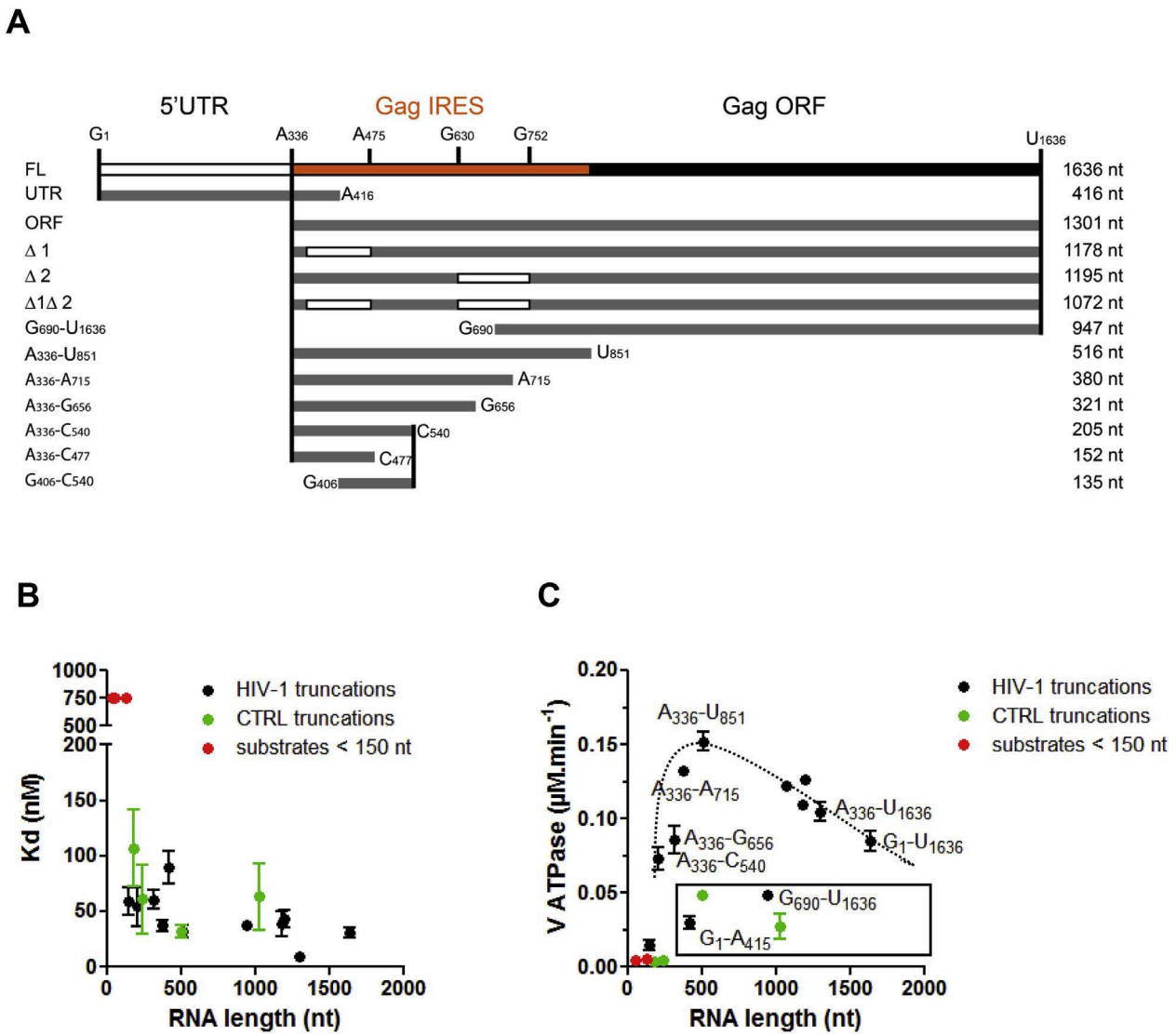


Fig. 6. HIV-1 RNA is a preferential substrate. (A) Schematic representation of the RNA HIV-1 fragments used (on scale). (B) Representation of the results obtained in Table 1. RNAs derived from HIV-1 (black), from the luciferase control (green) or substrates < 150 nt (red) are indicated in the Figure. Whenever the Kd cannot be determined, it was artificially set at 750 nM. The results are the mean and standard deviation of at least three independent experiments. (C) Kinetic ATPase reactions were performed in the presence of 100 nM DDX3-WT, 100 nM RNA and 10 μM ATP for 10 min. Initial rates are indicated in Table 2 as the average and standard deviation of independent experiments.

Table 2
DDX3-WT ATPase activity after stimulation by HIV-1 derived RNAs. Kinetic ATPase assays were performed with 100 nM DDX3, 100 nM RNA and 10 μM ATP. The results are the mean and standard deviation of at least three independent experiments. The normalization per nucleotide is shown to assess the size effect.

RNA	Nucleotides	Vi ($nM.min^{-1}$)	Length (nt)	Vi per nt ($pM.min^{-1}.nt^{-1}$)
FL	G_1-U_{1636}	85 ± 7	1636	52.0
UTR	G_1-A_{416}	30 ± 4	416	72.1
Gag ORF	A_{336}-U_{1636}	105 ± 6	1301	80.7
	G_{690}-U_{1636}	48.6	947	51.3
Gag IRES	A_{336}-U_{851}	152 ± 6	516	294.6
	A_{336}-A_{715}	132 ± 3	380	347.4
	A_{336}-G_{656}	86 ± 9	321	267.9
	A_{336}-C_{540}	73 ± 8	205	356.1
	G_{406}-C_{540}	5 ± 1	135	37.0
	A_{336}-C_{477}	14 ± 4	152	92.1
Δ Site1 (Δ1)	A_{336}-U_{881}Δ363-485	109	1178	92.5
Δ Site2 (Δ2)	A_{336}-U_{881}Δ632-756	127	1195	106.3
Δ1 Δ2	A_{336}-U_{881}Δ363-485Δ632-756	122	1072	113.8
CTRL 1		27.5	1031	26.7
CTRL 2		48.4	505	95.8
CTRL 3		4	243	16.5
CTRL 4		4	183	21.9
Duplex D3		5 ± 1	57	87.7

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

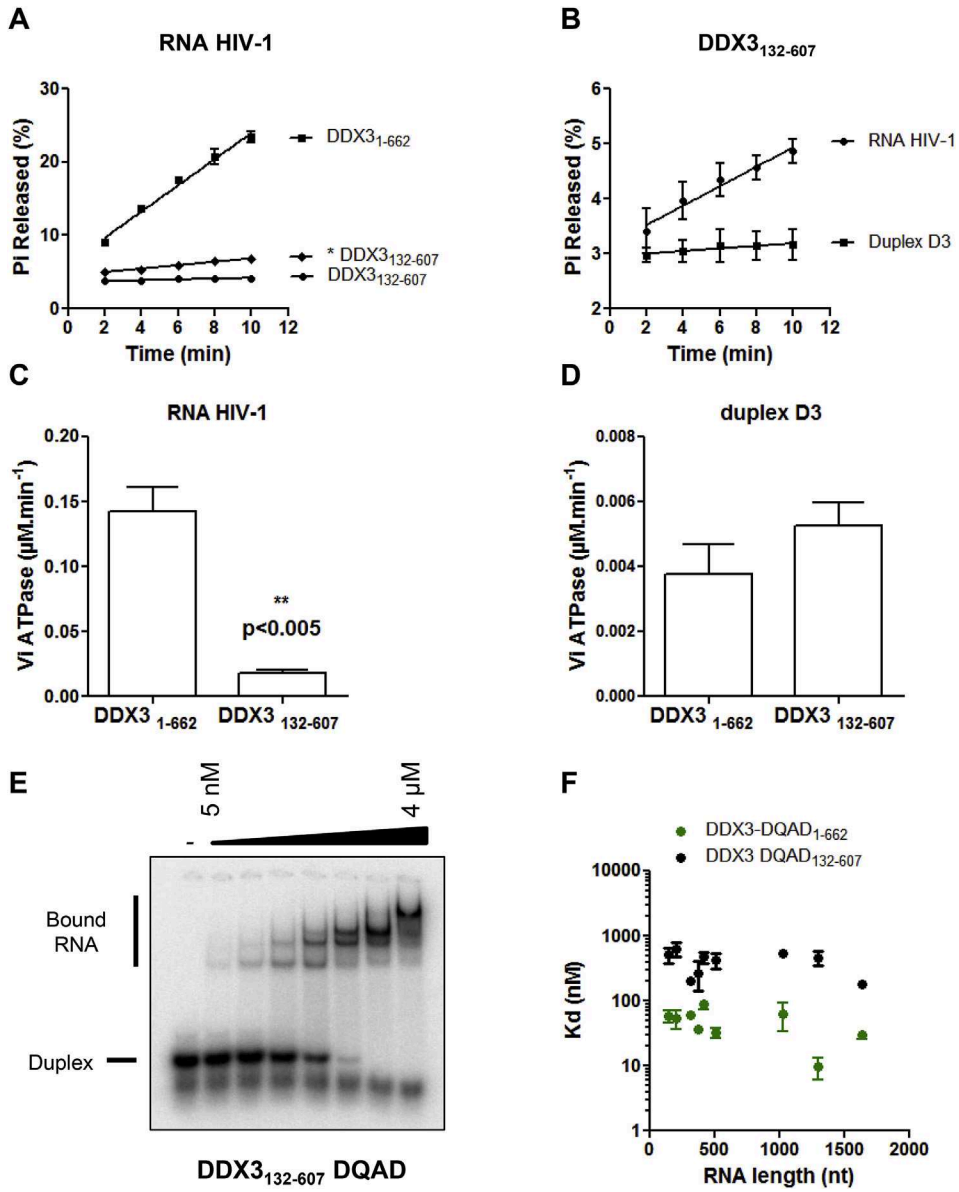


Fig. 7. The N- and C-terminal domains of DDX3X are required for its activity. Kinetic ATPase reactions were performed in the presence of 100 nM $DDX3_{1-662}$, $DDX3_{132-607}$ or $DDX3_{132-607}$ after TEV cleavage (*$DDX3_{132-607}$) and 100 nM (A) or 1 µM (B–D) substrate RNA. (E) 16 bp RNA/RNA duplexes (D3, 5 nM) were incubated in the presence of increasing concentrations of $DDX3_{132-607}$ -DQAD (0–4000 nM) and 2 mM ATP. The quantitation of bound duplexes allows for the determination of the following binding parameters: Bmax = 1.00 ± 0.04, h = 1.4 ± 0.1 and Kd = 138 ± 12 nM. (F) Representation of the results obtained in Table 3.5 nM of RNA were incubated in the presence of 0–4000 nM recombinant protein and 2 mM (see Supplemental Material and Methods section). The values represent the average and standard deviation of at least three independent experiments. p indicates the P value issued from an unpaired *t*-test whenever appropriate.

which could result from the quality/stability of the produced protein, or probably more relevant from the nature of the nucleic acid used to stimulate DDX3 ATPase activity.

To establish that our recombinant protein is an active helicase, we used short synthetic duplexes and confirm the requirement for a single-stranded region for the enzyme to be active. This is in agreement with the current model of "local strand separation" and enzyme loading on single-stranded nucleic acids [27]. Nonetheless, DDX3 displayed a preference for substrates presenting a 3′ tail as previously reported [31]. Increase in the helicase activity on substrates presenting a 3′ overhang could result from a differential substrate binding or from any of the steps involved in ATP hydrolysis, in duplex unwinding or in the coupling of both activities. Although we could not reliably measure the ATPase activity triggered by synthetic helicase substrates, in one case we determined that an increase in helicase activity is associated with a 1.5 fold decrease in binding to nucleic acids. This could result from an increased stability of the enzyme/substrate complex and render the duplex harder to destabilize or else, the reaction product resulting from unwinding of a 3′ overhang substrate may have more affinity for the enzyme than the product of a 5' overhang substrate. Nonetheless, at this point, it seems premature to make further considerations.

In contrast to what is observed with Ded1p, we could not detect any RNA/DDX3 complexes when using non-hydrolysable ATP analogs. We had to resort to the DQAD mutant impaired for ATP hydrolysis to observe RNA-protein complexes in presence of ATP. Complex formation with the DDX3 DQAD is strongly reduced in

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

Table 3
$DDX3_{132-607}$ DQAD binding affinities to HIV-1 gRNA fragments. Filter binding assays were performed in the presence of 5 nM RNAs and increasing concentrations of DDX3-DQAD. The results are the mean and standard deviation of two to three independent experiments. The normalization per nucleotide is shown to assess the size effect.

RNA	Nucleotides	Kd (nM) Mean ± s.e.m.	Length (nt)	Kd per nt (pM/nt)
FL	G_1-U_{1636}	183 ± 24	1636	112
UTR	G_1-A_{416}	473 ± 90	416	1137
Gag ORF	A_{336}-U_{1636}	462 ± 113	1301	355
Gag IRES	A_{336}-U_{851}	423 ± 116	516	820
	A_{336}-A_{715}	270 ± 130	380	711
	A_{336}-C_{540}	633 ± 157	205	3088
	G_{406}-C_{540}	870 ± 187	135	6444
	A_{336}-C_{477}	510 ± 133	152	3355

presence of AMP-PNP and ATP-γS indicating that a stable nucleotide-DDX3-RNA ternary complex strictly requires ATP. Modification of ATP by a non-hydrolysable analog is not trivial as shown by structural analysis of myosin complexed with AMP-PNP which revealed that the presence of the bridging nitrogen in AMP-PNP change the hydrogen bonding pattern and could explain the lowered affinity [36].

When testing synthetic substrates of different sequences, we could not observe any difference in the stimulation of ATP hydrolysis. Since we tested only a very small set of RNAs, it is possible that we did not find the correct sequence to specifically activate DDX3. One can also envisage that DDX3, as most DEAD-box helicases, displays a low sequence specificity towards its RNA substrate and has to be promiscuous to fulfill its multiple cellular functions. To our knowledge, the bacterial DbpA is the only DEAD-box helicase which activity strictly relies on the anchoring on a specific sequence which in the present case consists in the hairpin 92 of the 23S RNA [37]. Instead, our results with a biological substrate rather suggest the presence of determinants in the A_{336}-U_{851} region of Gag ORF. Counter-intuitively, adding extra sequences to this domain seem to decrease DDX3 stimulation, which could reflect the formation of alternative structures, or result from a non-productive titration of DDX3. Deleting additional sequences from this domain allows us to define a minimal A_{336}-C_{540} fragment that retains significant capacity to stimulate DDX3 ATPase activity. Further 5′ or 3′ deletion of this domain (leading to G_{406}-C_{540} or A_{336}-C_{477}) results in low ATPase levels that are similar to those obtained with control RNAs. This could result from the size of the fragments (135 and 152 nt, respectively) or else, such a bipartite signal could reflect the presence of a structural element involving sequences from each of the above-mentioned regions. Such structural element would match the P2 domain of HIV-1 RNA (G_{363}-A_{368} and U_{534}–U_{539}) that is involved in a recently described three-way junction [38]. Taken together, our results strongly suggest that structural determinants are involved in DDX3X interaction with HIV-1 RNA and their identification will necessitate further examination.

In this work, we have identified HIV-1 gRNA as a preferred substrate which strongly stimulates DDX3 ATPase activity as compared to the short model substrates used so far. We were able to narrow down to the A_{336}-C_{851} region of the molecule that contains the Gag-IRES. This region could conceal a specific motif yet to be characterized or else, the specificity could be defined by the presence of alternating structures. The rationale for this peculiar specificity could reside in DDX3 function in the cell such as selectively facilitating the translation of structured mRNAs [16] or selectively interacting with G-quadruplex containing RNAs [39]. In most DEAD-box proteins, the core is flanked by additional N and C-terminal extensions, which contribute to the functional diversity of this protein family. Many of these extensions direct individual DEAD-box proteins to their functional targets by interaction with protein or RNA components of the target and some extensions modulate the activity of the helicase core [40,41]. Our results suggest that HIV-1 RNA is one of these targets, able to optimally stimulate DDX3 enzymatic activity provided the presence of the N– and C-terminal extensions. DDX3 has been described in complex with ubiquitous translation initiation factors but does stimulate the translation of only a subset of cellular or viral mRNAs. This helicase may therefore combine intrinsic specificity for some structural motif(s) and interaction with other proteins to act on different but yet specific substrates.

Conflict of interest

The authors report no conflict of interest.

Author agreement

G. de B., B.S. and N.C. designed and planned experiments; G. de B., M.A., N.U., F.B., L.P. and N.C. carried out experiments; G. de B., B.S. and N. C. analyzed the results obtained in this study and wrote the paper. All authors have made significant contributions to the study, have read, and approved the manuscript. The material submitted is original, all authors are in agreement to have the article published.

Acknowledgements

The authors would like to thank Josette Banroques and Kyle Tanner for the kind gift of purified Ded1p and fruitful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biochi.2019.03.008.

Funding

Research in B.S. laboratory was financed by the CNRS, the University Paris Descartes and a grant from la Fondation pour la Recherche Médicale (FRM #DBI20141231337). G. B. and M.A are recipient of a PhD fellowship from the French ministry for research and education (Allocation MRE).

References

[1] B. Gilman, P. Tijerina, R. Russell, Distinct RNA-unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs, Biochem. Soc. Trans. 45 (2017) 1313–1321, https://doi.org/10.1042/BST20170095.
[2] A.E. Gorbalenya, E.V. Koonin, Helicases: amino acid sequence comparisons and structure-function relationships, Curr. Opin. Struct. Biol. 3 (1993) 419–429, https://doi.org/10.1016/S0959-440X(05)80116-2.
[3] P. Linder, P.F. Lasko, M. Ashburner, P. Leroy, P.J. Nielsen, K. Nishi, J. Schnier, P.P. Slonimski, Birth of the D-E-A-D box, Nature 337 (1989) 121–122, https://doi.org/10.1038/337121a0.
[4] M.E. Fairman-Williams, U.-P. Guenther, E. Jankowsky, SF1 and SF2 helicases: family matters, Curr. Opin. Struct. Biol. 20 (2010) 313–324, https://doi.org/10.1016/j.sbi.2010.03.011.
[5] S. Rocak, P. Linder, DEAD-box proteins: the driving forces behind RNA metabolism, Nat. Rev. Mol. Cell Biol. 5 (2004) 232–241, https://doi.org/10.1038/nrm1335.
[6] P. Linder, Dead-box proteins: a family affair—active and passive players in RNP-remodeling, Nucleic Acids Res. 34 (2006) 4168–4180, https://doi.org/10.1093/nar/gkl468.
[7] P. Linder, E. Jankowsky, From unwinding to clamping — the DEAD box RNA helicase family, Nat. Rev. Mol. Cell Biol. 12 (2011) 505–516, https://doi.org/10.1038/nrm3154.
[8] D. Sharma, E. Jankowsky, The Ded1/DDX3 subfamily of DEAD-box RNA helicases, Crit. Rev. Biochem. Mol. Biol. 49 (2014) 343–360, https://doi.org/

ARTICLE IN PRESS

10.3109/10409238.2014.931339.
[9] Y. Ariumi, Multiple functions of DDX3 RNA helicase in gene regulation, tumorigenesis, and viral infection, Front. Genet. 5 (2014), https://doi.org/10.3389/fgene.2014.00423.
[10] R. Soto-Rifo, T. Ohlmann, The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism, Wiley Interdiscip. Rev. RNA 4 (2013) 369–385, https://doi.org/10.1002/wrna.1165.
[11] L. Zhao, Y. Mao, J. Zhou, Y. Zhao, Y. Cao, X. Chen, Multifunctional DDX3: dual roles in various cancer development and its related signaling pathways, Am. J. Cancer Res. 6 (2016) 387–402.
[12] M. Schröder, Viruses and the human DEAD-box helicase DDX3: inhibition or exploitation? Biochem. Soc. Trans. 39 (2011) 679–683, https://doi.org/10.1042/BST0390679.
[13] F. Valiente-Echeverría, M.A. Hermoso, R. Soto-Rifo, RNA helicase DDX3: at the crossroad of viral replication and antiviral immunity, Rev. Med. Virol. 25 (2015) 286–299, https://doi.org/10.1002/rmv.1845.
[14] V.S.R.K. Yedavalli, C. Neuveut, Y. Chi, L. Kleiman, K.-T. Jeang, Requirement of DDX3 DEAD box RNA helicase for HIV-1 rev-RRE export function, Cell 119 (2004) 381–392, https://doi.org/10.1016/j.cell.2004.09.029.
[15] A. Fröhlich, B. Rojas-Araya, C. Pereira-Montecinos, A. Dellarossa, D. Toro-Ascuy, Y. Prades-Pérez, F. García-de-Gracia, A. Garcés-Alday, P.S. Rubilar, F. Valiente-Echeverría, T. Ohlmann, R. Soto-Rifo, 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 BBA - gene Regul, Mech 1859 (2016) 719–730, https://doi.org/10.1016/j.bbagrm.2016.03.009.
[16] R. Soto-Rifo, P.S. Rubilar, T. Limousin, S. de Breyne, D. Décimo, T. Ohlmann, DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs: translation initiation mediated by DDX3, EMBO J. 31 (2012) 3745–3756, https://doi.org/10.1038/emboj.2012.220.
[17] R. Soto-Rifo, P.S. Rubilar, T. Ohlmann, 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 (2013) 6286–6299, https://doi.org/10.1093/nar/gkt306.
[18] N.T. Parkin, E.A. Cohen, A. Darveau, C. Rosen, W. Haseltine, N. Sonenberg, Mutational analysis of the 5' non-coding region of human immunodeficiency virus type 1: effects of secondary structure on translation, EMBO J. 7 (1988) 2831–2837.
[19] N. Chamond, N. Locker, B. Sargueil, The different pathways of HIV genomic RNA translation, Biochem. Soc. Trans. 38 (2010) 1548–1552, https://doi.org/10.1042/BST0381548.
[20] I. Iost, M. Dreyfus, P. Linder, Ded1p, a DEAD-box protein required for translation initiation in *Saccharomyces cerevisiae*, is an RNA helicase, J. Biol. Chem. 274 (1999) 17677–17683, https://doi.org/10.1074/jbc.274.25.17677.
[21] H.A. Bowers, P.A. Maroney, M.E. Fairman, B. Kastner, R. Lührmann, T.W. Nilsen, E. Jankowsky, Discriminatory RNP remodeling by the DEAD-box protein DED1, RNA 12 (2006) 903–912, https://doi.org/10.1261/rna.2323406.
[22] M.E. Fairman, P.A. Maroney, W. Wang, H.A. Bowers, P. Gollnick, T.W. Nilsen, E. Jankowsky, Protein displacement by DExH/D "RNA helicases" without duplex unwinding, Science 304 (2004) 730–734, https://doi.org/10.1126/science.1095596.
[23] F. Liu, A.A. Putnam, E. Jankowsky, DEAD-box helicases form nucleotide-dependent, long-lived complexes with RNA, Biochemistry 53 (2014) 423–433, https://doi.org/10.1021/bi401540q.
[24] A.A. Putnam, E. Jankowsky, AMP sensing by DEAD-box RNA helicases, J. Mol. Biol. 425 (2013) 3839–3845, https://doi.org/10.1016/j.jmb.2013.05.006.
[25] Q. Yang, E. Jankowsky, ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1 †, Biochemistry 44 (2005) 13591–13601, https://doi.org/10.1021/bi0508946.
[26] Q. Yang, M.D. Campo, A.M. Lambowitz, E. Jankowsky, DEAD-box proteins unwind duplexes by local strand separation, Mol. Cell. 28 (2007) 253–263, https://doi.org/10.1016/j.molcel.2007.08.016.
[27] A.A. Putnam, Z. Gao, F. Liu, H. Jia, Q. Yang, E. Jankowsky, Division of labor in an oligomer of the DEAD-box RNA helicase Ded1p, Mol. Cell. 59 (2015) 541–552, https://doi.org/10.1016/j.molcel.2015.06.030.
[28] S.H. Park, S.G. Lee, Y. Kim, K. Song, Assignment of a human putative RNA helicase gene, DDX3, to human X chromosome bands p11.3->p11.23, Cytogenet. Cell Genet. 81 (1998) 178–179, https://doi.org/10.1159/000015022.
[29] R. Franca, A. Belfiore, S. Spadari, G. Maga, Human DEAD-box ATPase DDX3 shows a relaxed nucleoside substrate specificity, Proteins Struct. Funct. Bioinforma 67 (2007) 1128–1137, https://doi.org/10.1002/prot.21433.
[30] A. Garbelli, S. Beermann, G. Di Cicco, U. Dietrich, G. Maga, 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 (2011) e19810, https://doi.org/10.1371/journal.pone.0019810.
[31] D. Sharma, A.A. Putnam, E. Jankowsky, Biochemical differences and similarities between the DEAD-box helicase orthologs DDX3X and Ded1p, J. Mol. Biol. 429 (2017) 3730–3742, https://doi.org/10.1016/j.jmb.2017.10.008.
[32] J.S. Kieft, A. Grech, P. Adams, J.A. Doudna, Mechanisms of internal ribosome entry in translation initiation, Cold Spring Harbor Symp. Quant. Biol. 66 (2001) 277–283.
[33] A. Pause, N. Sonenberg, Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A, EMBO J. 11 (1992) 2643–2654.
[34] S.N. Floor, K.J. Condon, D. Sharma, E. Jankowsky, J.A. Doudna, Autoinhibitory interdomain interactions and subfamily-specific extensions redefine the catalytic core of the human DEAD-box protein DDX3, J. Biol. Chem. (2015), https://doi.org/10.1074/jbc.M115.700625 jbc.M115.700625.
[35] N. Locker, N. Chamond, B. Sargueil, 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 (2011) 2367–2377, https://doi.org/10.1093/nar/gkq1118.
[36] A.M. Gulick, C.B. Bauer, J.B. Thoden, I. Rayment, X-ray structures of the MgADP, MgATPγS, and MgAMPPNP complexes of the *Dictyostelium discoideum* myosin motor domain †, ‡, Biochemistry 36 (1997) 11619–11628, https://doi.org/10.1021/bi9712596.
[37] C.M. Diges, O.C. Uhlenbeck, Escherichia coli DbpA is an RNA helicase that requires hairpin 92 of 23S rRNA, EMBO J. 20 (2001) 5503–5512, https://doi.org/10.1093/emboj/20.19.5503.
[38] J. Deforges, S. de Breyne, M. Ameur, N. Ulryck, N. Chamond, A. Saaidi, Y. Ponty, T. Ohlmann, B. Sargueil, Two ribosome recruitment sites direct multiple translation events within HIV1 Gag open reading frame, Nucleic Acids Res. 45 (2017) 7382–7400, https://doi.org/10.1093/nar/gkx303.
[39] B. Herdy, C. Mayer, D. Varshney, G. Marsico, P. Murat, C. Taylor, C. D'Santos, D. Tannahill, S. Balasubramanian, Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts, Nucleic Acids Res. (2018), https://doi.org/10.1093/nar/gky861.
[40] O. Cordin, J. Banroques, N.K. Tanner, P. Linder, The DEAD-box protein family of RNA helicases, Gene 367 (2006) 17–37, https://doi.org/10.1016/j.gene.2005.10.019.
[41] M. Hilbert, A.R. Karow, D. Klostermeier, The mechanism of ATP-dependent RNA unwinding by DEAD box proteins,, Biol. Chem. 390 (2009), https://doi.org/10.1515/BC.2009.135.

Please cite this article as: G. de Bisschop et al., HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity, Biochimie, https://doi.org/10.1016/j.biochi.2019.03.008

2) Caractérisation de l'*interaction de* DD*X3 à l'ARN* génomique

Afin d'étudier plus en détail l'interaction de DDX3 avec l'ARNg, j'ai utilisé la technique de footprint. L'ARN du VIH-1 comprenant la 5'UTR et la région codante de Gag (nucléotides 1 à 1636) est incubé en présence ou en absence du mutant DQAD, puis mis à réagir avec une sonde SHAPE. En effet, nous avons montré que le mutant DQAD forme des complexes stables avec l'ARN, probablement parce que l'absence d'hydrolyse de l'ATP empêche le relargage de l'ARN.

Les différences de réactivité entre l'ARN seul et en présence de $DDX3_{DQAD}$ sont représentées sur le modèle de structure secondaire de l'ARN (Figure 42 A et B). Les différences sont considérées comme significatives lorsque l'écart entre les réactivités avec et sans DDX3 est supérieur à 0,2, le ratio supérieur à 1,5 et que $p_{Student} < 0,05$. De plus les nucléotides dont la réactivité est significativement différente en présence du tampon de DDX3 seul ($p_{Student} > 0,05$) ne sont pas considérés (Figure 42 C). Les nucléotides pour lesquels la réactivité diminue en présence du mutant DQAD peuvent correspondre à des sites de liaison de DDX3, ou à des sites qui se structurent suite à l'interaction de DDX3. Les nucléotides dont la réactivité augmente peuvent être interprétés comme des régions déstabilisées par DDX3. De façon intéressante, les nucléotides dont la réactivité diminue significativement sont toujours de sites modélisés en simple-brin. En effet, il est plus délicat de détecter une diminution de réactivité au niveau de nucléotides de faible réactivité, généralement modélisés en double brin. On ne peut donc pas exclure que DDX3 lie des régions d'ARN double-brin, non détectées en footprint. Au demeurant, les résultats obtenus sont cohérents avec ceux obtenus sur des substrats modèles, pour lesquels nous avons montré que la présence d'ARN simple brin est indispensable à la liaison de DDX3. De plus, les sites d'empreinte ne dépassent pas 3 nucléotides consécutifs ce qui semble conforme aux structures cristallographiques d'hélicases à boîte DEAD en complexe avec de l'ARN où 4 nucléotides consécutifs sont liés par leur résidu 2'OH. Cette observation n'exclut pas qu'il puisse y avoir d'autres sites de liaison à l'ARN que celui du cœur catalytique. A l'inverse, les nucléotides dont la réactivité augmente significativement sont presque tous modélisés en double-brin, ce qui est cohérent avec l'hypothèse d'une déstabilisation de la structure secondaire sous l'effet de la liaison du mutant DQAD. En effet, le modèle de séparation locale des brins propose que l'hydrolyse de l'ATP soit requise uniquement pour le relargage de l'ARN et pas pour la séparation des brins. Autrement dit, le mutant DQAD est

capable de lier et de dissocier des brins d'ARN, mais incapable de libérer l'ARN (absence de turn-over) (Liu et al., 2008, 2014).

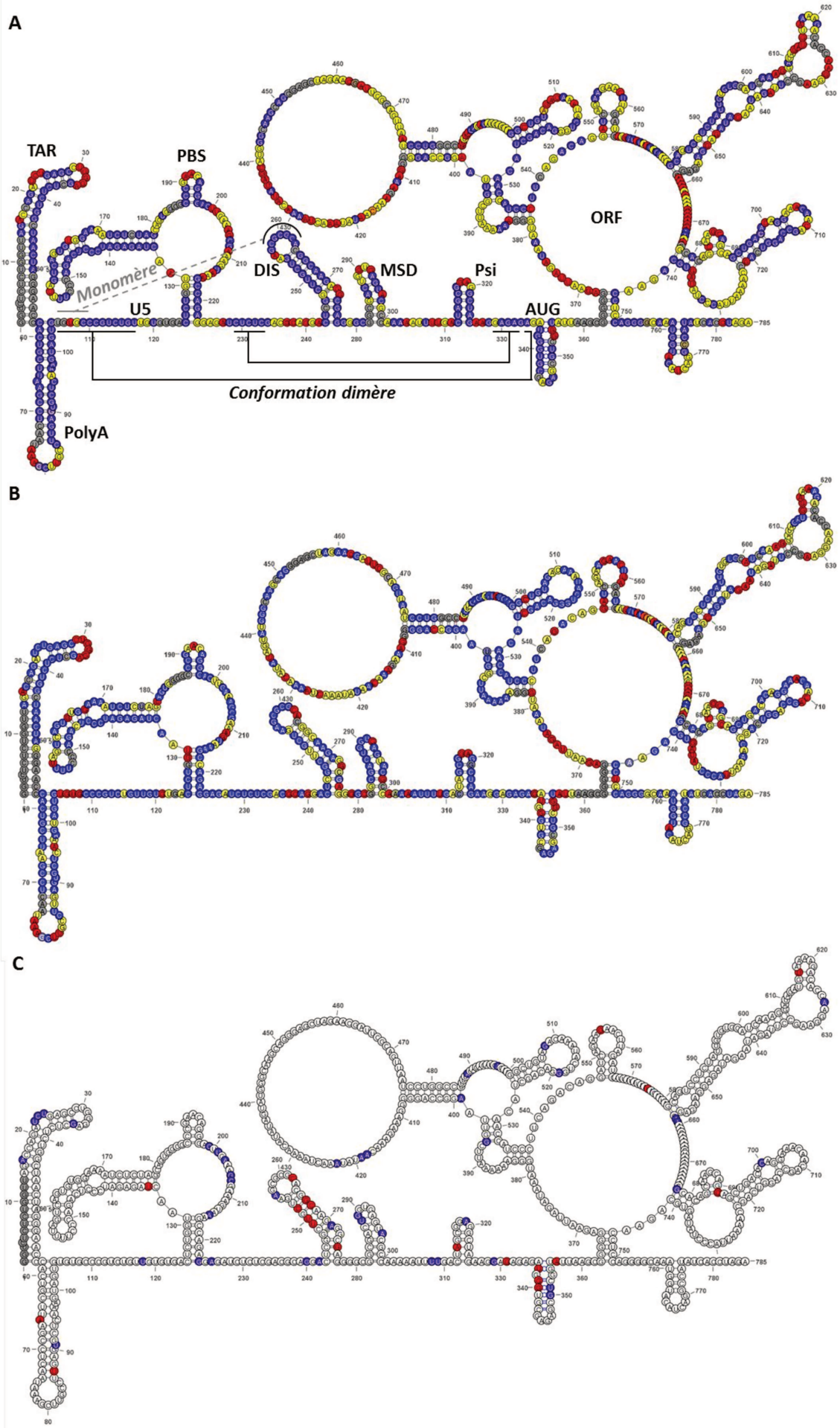
Parmi les sites d'empreinte majeurs détectés, deux rejoignent des résultats publiés par l'équipe de T. Ohlmann (Soto-Rifo et al., 2012b) : la boucle interne de TAR, positions 22 à 24, et la boucle du PBS en positions 199 à 207. En effet, Soto-Rifo et al. ont mis en évidence la liaison de DDX3 à TAR par des expériences de retard sur gels. Ces résultats précisent donc le site de liaison de DDX3. Ils ont également identifié un signal de toeprint en positions 202 – 214. Le toeprint permet de localiser les sites d'arrêts de la polymérase sous la contrainte d'un complexe ARN – protéine d'intérêt. Par conséquent le signal apparaît légèrement en 3' du site de liaison effectif. L'empreinte observée en footprint semble donc concorder parfaitement avec le signal de toeprint décrit par Soto-Rifo et al. D'autres baisses significatives de réactivité se répartissent le long de la 5'UTR, sans qu'on puisse clairement les interpréter comme des sites de liaison ou de remodelage. Le début de la région codante (nucléotides 339 à 785) présente nettement moins de changements significatifs de réactivités. En comparant les profils de réactivité, on constate toutefois des baisses de réactivités importantes mais non significatives au niveau de trois sites, 387-392, 420-440 et 485-500. La variation de réactivité de ces deux derniers sites n'est pas significative elle survient lors de l'ajout du tampon de DDX3 seul. La variation observée en 387-392 n'est quant à elle pas significative du fait d'une trop grande variabilité dans les réactivités mesurées.

De façon particulièrement remarquable, cinq nucléotides de l'hélice apicale de DIS présentent une forte augmentation de réactivité. Ces résultats suggèrent que DIS est au moins partiellement déstructurée lors de l'interaction avec le mutant DQAD. Il est possible que cette hélice passe du type A, le type plus fréquemment adopté par les hélices d'ARN, à une conformation non-A suite à l'interaction. De telles hélices peuvent en effet avoir une réactivité plus forte vis-à-vis des sondes qui réagissent au niveau du ribose (Bindewald et al., 2011). On observe aussi plusieurs augmentations significatives de la réactivité autour du codon AUG (nucléotides 337 à 340). Dans le modèle structural de transition entre dimère et monomère de B. Berkhout (Abbink and Berkhout, 2003) ainsi que dans celui de M. Summers (Keane et al., 2016), ces nucléotides passent d'une structure locale en tige-boucle à un appariement longue distance avec la région U5. Il est possible que cette augmentation de réactivité traduise un état intermédiaire entre ces deux appariements, mais il est difficile d'expliquer la baisse de réactivité observée plus loin, positions 350 – 351. Enfin, il est possible que les autres modifications

observées, plus ponctuelles, correspondent à de légères modifications de la structure secondaire et/ou tertiaire.

En conclusion, le footprint de l'ARNg par DDX3 confirme l'existence de sites de liaison dans la tige-boucle TAR et dans le PBS. L'interaction de DDX3 avec TAR a été reliée à l'initiation de la traduction de l'ARNg, où DDX3 déstabilise TAR pour promouvoir l'initiation dépendante de la coiffe. A l'inverse, la liaison de DDX3 au PBS semble indépendante de son rôle dans l'initiation de la traduction (Soto-Rifo et al., 2012b). Mes résultats révèlent un effet inattendu de DDX3 sur la tige-boucle DIS et la région de l'AUG qui sont toutes deux des structures impliquées dans la dimérisation. Ces sites d'interaction suggèrent des rôles possibles de DDX3 dans la régulation de la dimérisation. J'ai donc étudié l'effet de DDX3 sur la dimérisation de l'ARN du VIH-1.

Figure 42 (page suivante) : Empreinte de l'ARNg (1-1636) par DDX3$_{DQAD}$. Réactivités au BzCN en absence (A) et en présence (B) de 1 µM de DDX3$_{DQAD}$. La concentration en ARN est de 50 nM. Le tampon de repliement est composé de 40 mM Tris pH8, 50 mM NaCl, 2 mM DTT, 0,5 mM MgCl2, 0,01% NP-40. Code couleur : bleu < 0,4 ; 0,4 < jaune < 0,7 ; rouge > 0,7. C/ Les nucléotides en bleu indiquent de sites où la réactivité diminue de façon significative en présence de DDX3, en rouge les sites où la réactivité augmente significativement. Les éléments de structures secondaires sont annotés en A.


A
TAR
PBS
DIS
MSD
Psi
ORF
Monomère
U5
AUG
Conformation dimère
PolyA
B
C

DDX3 déstabilise des dimères d'ARN du VIH-1

1) *DDX3 déstabilise des dimères d'ARN du VIH*-1

Afin d'observer l'effet de DDX3 sur la dimérisation, des dimères de 5'UTR formés in vitro sont incubés en présence de différentes concentrations de DDX3 et d'ATP (Figure 43 A). On constate que DDX3 dissocie les dimères en monomères. En absence de extensions N et C-terminales, il faut employer une concentration cent fois plus élevée de protéine tronquée pour observer un effet similaire (Figure 43 B). Par conséquent les extensions terminales contribuent de façon importante à l'effet de DDX3 sur la dissociation des dimères de 5'UTR, de même que nous avons montré qu'elles sont requises pour la stimulation optimale de l'activité ATPase par des fragments de l'ARN du VIH. Ded1 dissocie des dimères de 5'UTR avec une efficacité similaire (Figure 43 C). Cette observation suggère que le mode d'interaction de DDX3 avec la 5'UTR du VIH repose sur des déterminants de structure et/ou de séquence qui ne sont pas uniquement propres au VIH. Il est intéressant de constater que l'ajout de la région codante de Gag après la 5'UTR (5'UTR-ORF) rend les dimères beaucoup plus résistants à l'action de DDX3 (Figure 43 D). Ceci permet de formuler deux hypothèses quant à la structure du dimère de 5'UTR-ORF. D'une part, il est possible que ce dimère soit plus stable que le dimère de 5'UTR seule du fait de l'existence de sites de dimérisation additionnels dans la région codante. D'autre part, nous avons vu que la présence de la région codante influe sur le repliement de la 5'UTR ; on peut penser qu'en présence de la région codante, la structure du site de dimérisation est plus stables. En conséquence, il est possible que les dimères soient dissociés plus lentement ou reformés plus rapidement après dissociation. Notons que ces deux hypothèses ne sont pas mutuellement exclusives. On retrouve un effet plus proche de celui observé avec la 5'UTR seule en ajoutant un excès d'oligonucléotide ADN complémentaire de la boucle apicale de DIS (5'-TGCGCGCTT) (Figure 43 D). Il est possible que la présence d'anti-DIS augmente l'effet observé de DDX3 en empêchant la reformation des complexes kissing après la dissociation des dimères – c'est-à-dire en stabilisant le produit de la réaction. Il est aussi envisageable que l'oligonucléotide anti-DIS augmente l'efficacité de DDX3 stabilisant des conformations partiellement dissociés. En d'autres termes, l'anti-DIS diminuerait la proportion de cycles futiles.

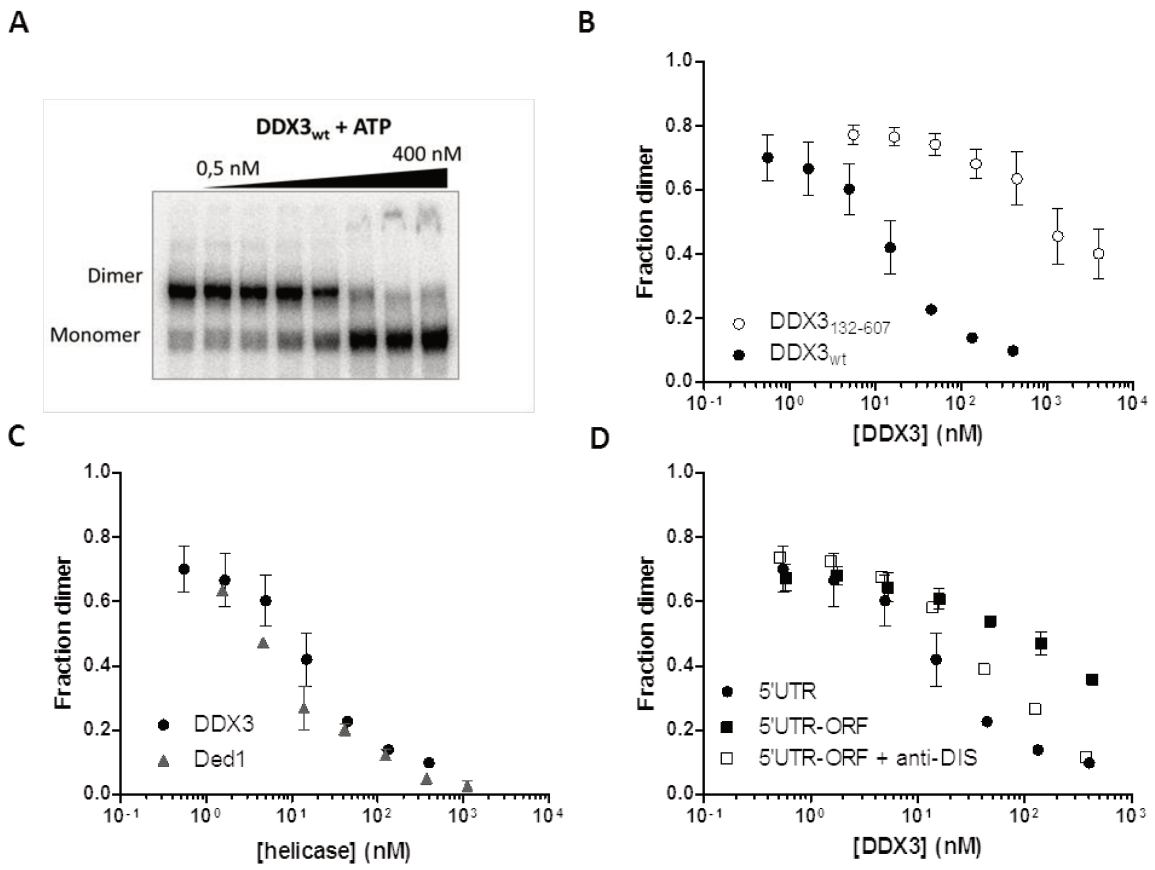


Figure 43 : DDX3 déstabilise des dimères d'ARN du VIH-1. A/ Analyse sur gel natif des dimères formés (50 nM) *en présence de concentrations croissante de DDX3 et d'ATP. Les quantifications permettent de comparer l'effet de la protéine tronquée (B) et de Ded1 (C), et d'évaluer la stabilité relative conférée par la région codante (D).* Le milieu réactionnel est composé de 40 mM Tris pH8, 50 mM NaCl, 2 mM DTT, 0,5 mM MgCl2, 0,01% NP-40

2) Analyse cinétique des paramètres de dissociation des dimères

Les résultats présentés montrent que DDX3 dissocie les dimères y compris lorsque l'ARN est en excès. Je me suis donc intéressé aux paramètres cinétiques en mesurant la dissociation des dimères de 5'UTR au cours du temps en présence de différentes concentrations de DDX3 (Figure 44). La proportion de dimères est analysée par électrophorèse native en utilisant un tampon de migration avec ou sans magnésium, afin d'obtenir des informations sur la nature des dimères dissociés.

On constate que, pour des concentrations en DDX3 de 1, 5 et 10 nM, la concentration de dimères dissociés au terme de la réaction est supérieure à la concentration en DDX3. Par conséquent, DDX3 dissocie des dimères de 5'UTR dans des conditions de turn-over multiple (Figure 44 C). La dissociation des dimères de 5'UTR requiert la présence d'ATP, puisqu'on

n'observe pas de variation de la proportion de dimère en absence d'ATP (Figure 44 A et B). En présence d'AMPPNP, un analogue stable de l'ATP, une faible fraction de dimère est dissociée au cours du temps. Il est possible que DDX3 conserve un reliquat d'activité hélicase en présence d'AMPPNP, comme c'est le cas pour l'hélicase DExD/H YxiN (Kossen et al., 2002).

On observe que la proportion maximale de dimères dissociés varie en fonction de la concentration en DDX3 (Figure 44 C). Il se peut que la protéine, qui a une forte tendance à l'agrégation, se désactive en précipitant de façon significative au-delà de 10 minutes d'incubation à 37°C. On peut également envisager que l'ARN puisse reformer spontanément des dimères au cours du temps, de sorte qu'un équilibre s'établisse progressivement entre l'activité de dissociation des dimères de DDX3 et la réassociation des molécules d'ARN. Cette dernière interprétation pose la question de la nature des dimères dissociés par DDX3. Enfin, il se peut que la quantité d'ATP disponible entre en jeu. Aux faibles concentrations de DDX3, on peut supposer que les cycles futiles soient plus abondants, ce qui aurait pour conséquence un épuisement plus rapide de la quantité d'ATP initialement disponible.

On constate que les dimères dissociés par DDX3 sont stables sur des gels sans magnésium. En effet, la proportion de dimères à l'instant initial est similaire sur les deux types de gels, 69±5% en absence de magnésium contre 79±3% en présence de magnésium. On remarque également que la concentration maximale d'ARN dissociés varie en fonction du type de gel utilisé pour l'analyse des conformations (Figure 44 C). Elle est supérieure lorsqu'on utilise un tampon de migration dépourvu de magnésium. Cela suggère que des dimères se reforment lors de la migration en présence de magnésium, ou bien que des dimères se dissocient lors de la migration en absence de magnésium. Il est peu vraisemblable que des dimères se reforment lors de la migration en présence de magnésium. Par conséquent, il se peut que des dimères restants se dissocient lors de la migration dans un tampon sans magnésium. Ces dimères instables sont probablement des dimères kissing. Par ailleurs, l'écart entre les fractions maximales de dimères dissociés mesurées sur les deux types de gels est plus important aux fortes concentrations de DDX3. En conséquence ces dimères kissing résultent possiblement de la réassociation des dimères stables initialement présents.

Ces résultats permettent de mesurer une vitesse initiale pour chaque concentration en DDX3. La constante d'affinité apparente pour l'ARN $K_{1/2}$ mesurée en activité ATPase au moyen de l'ARNg (1-1636) est d'environ 50 nM, concentration à laquelle ont été menées les expériences. On ne peut donc pas considérer ces conditions comme saturantes puisque [ARN] ≈ $K_{1/2}$. On peut toutefois extraire la constante catalytique apparente $k_{cat, app}$ à partir de la

pente de la droite $v_i = f([DDX3])$. Une régression linéaire permet de calculer $k_{cat,\ app}$ = 2,0 monomère min^{-1} (R^2 = 98%) soit $k_{cat,\ app}$ = 1,0 dimère min^{-1}. Cela implique que la constante catalytique réelle soit supérieure à 1,0 dimère min^{-1}. Cette valeur peut être comparée à la constante catalytique de l'activité ATPase $k_{cat} \approx 20\ min^{-1}$. En ordre de grandeur, on pourrait donc considérer qu'au maximum 20 molécules d'ATP qui sont requises pour la dissociation d'un dimère. Néanmoins, une étude plus précise et dans les conditions adéquates serait nécessaire pour affiner ces résultats encore très préliminaires.

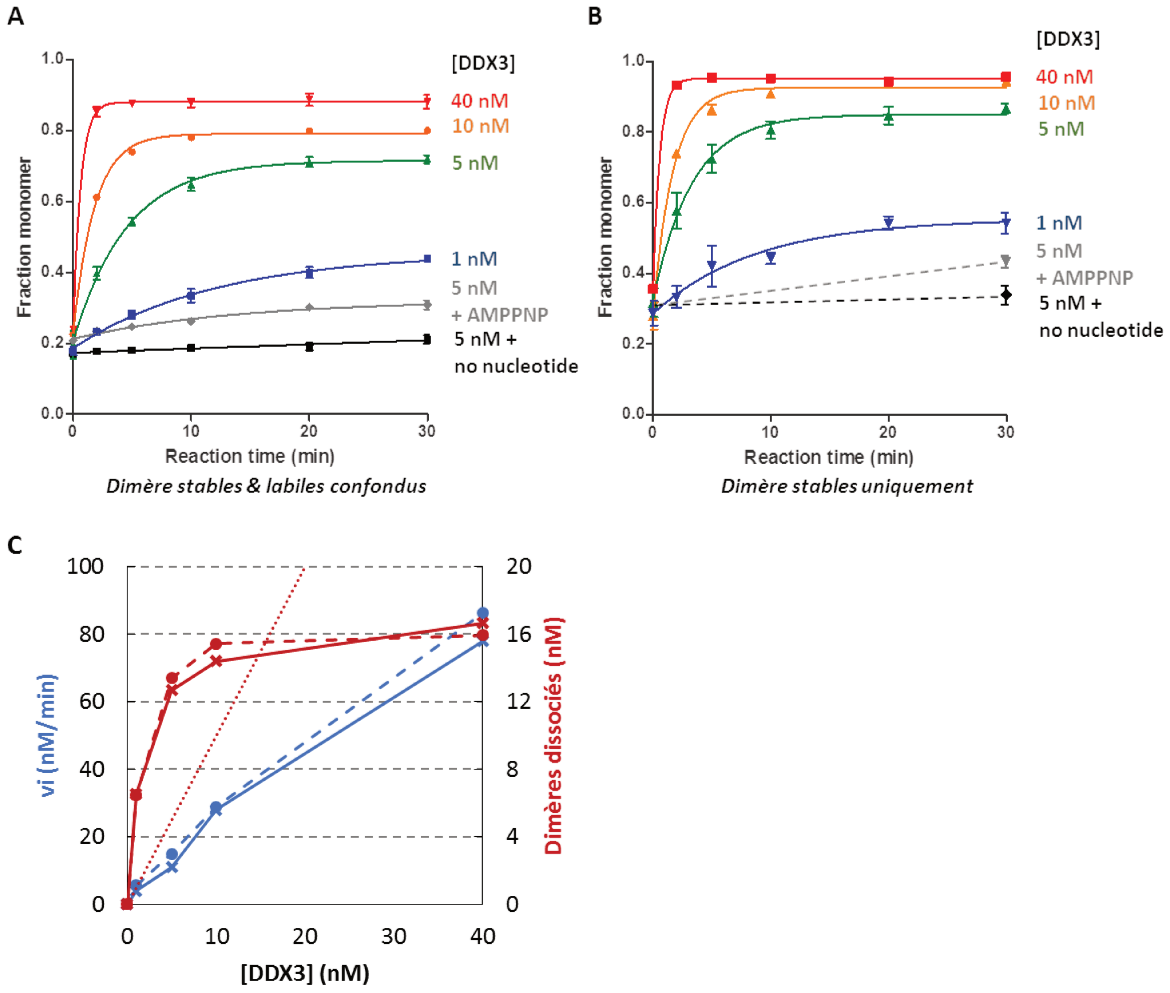


Figure 44 : DDX3 dissocie des dimères stables sur des gels sans magnésium dans des conditions de turn-over multiple. A et B/ Quantification des monomères formés au cours du *temps lors de l'incubation avec DDX3, après migration sur gel* natif avec (A) ou sans magnésium (B). La concentration en ARN est de 50 nM. C/ Analyse de la vitesse initiale de la réaction (bleu) et de la concentration de dimères dissociés (rouge) en fonction de la concentration en DDX3. La concentration de dimères dissociés est calculée comme le rapport entre la fraction de dimère à t=0 et la fraction de dimères *à* $t \rightarrow \infty$ multiplié par la concentration initiale de dimères. Les traits continus et hachurés correspondent aux valeurs calculées à partir des gels en présence et en absence de magnésium, respectivement. La droite [DDX3] = [dimères dissociés] est indiquée en pointillés rouges.

Conformément aux résultats présentés Figure 43, DDX3 dissocie moins efficacement des dimères d'ARN « de pleine longueur » (1-1636) que des dimères de 5'UTR (1-416). La vitesse initiale de la réaction ainsi que la fraction maximale de dimères dissociés sont moins importantes. En présence de 10 µM d'oligonucléotide anti-DIS, la fraction de dimères dissociés passe à 97%. Ces résultats confirment que les dimères d'ARN de pleine longueur sont formés par une interaction DIS – DIS. Il est possible que l'oligonucléotide anti-DIS permette d'observer strictement l'activité hélicase en inhibant la reformation de dimères, ou qu'il assiste la déstabilisation des dimères par DDX3. Par ailleurs, DDX3 dissocie aussi des dimères formés par le Core Encapsidation Signal (CES), une région qui s'étend de la région U5 à l'AUG (105 – 343) capable de dimériser et d'être encapsidée (Keane et al., 2015) (Figure 45 B). L'efficacité de DDX3 est sensiblement supérieure à celle qu'on observe pour la construction comprenant toute la 5'UTR, ce qui suggère que DDX3 dissocie des dimères de 5'UTR en interagissant principalement avec la région 105 – 343. De plus, la mutation du palindrome de DIS GCGCGC en GAGA inhibe la formation de dimères, ce qui confirme des résultats déjà décrits (Keane et al., 2015) et qui confirme que DDX3 dissocie des dimères formés par DIS. Ensemble, ces résultats indiquent que les régions situées en 3' et en 5' du site de dimérisation (105-343) contribuent à stabiliser les dimères vis-à-vis de DDX3. Cette stabilisation reflète soit que les molécules dissociées se réassocient plus vite, soit que les dimères sont thermodynamiquement plus stables.

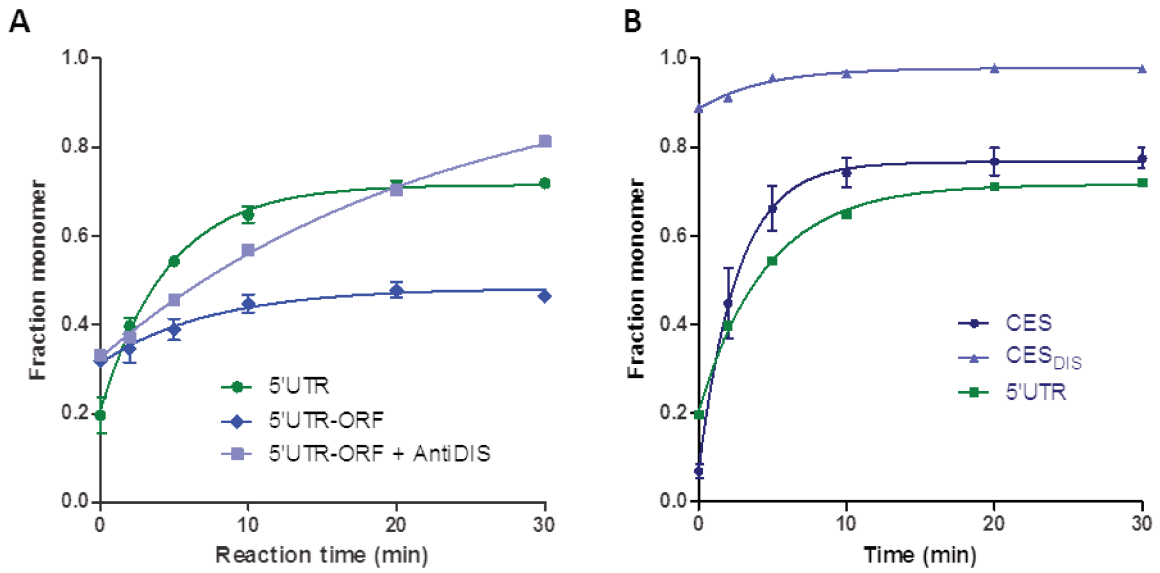


Figure 45 *: Cinétique de dissociation des dimères d'ARN de pleine longueur* avec ou sans anti-DIS (A) et de CES (105-343) ou de CES$_{DIS}$ muté au niveau du palindrome de DIS (B). 50 *nM d'ARN sont incubés en p*résence de 5 nM de DDX3 et de 2 *mM d'ATP* et des échantillons sont prélevés à différents instants puis déposés sur gel en présence de magnésium. La cinétique *de dissociation des dimères de 5'UTR est indiquée en vert pour comparaison.*

3) Caractérisation du rema*niement de l'ARN par SHAPE résolu en* temps

Ces conditions de turn-over multiple permettent de suivre le remaniement structural de l'ARN par SHAPE sans observer l'empreinte de DDX3. En effet, en se plaçant à 50 nM de 5'UTR (1-416) et 5 nM de DDX3, on s'attend à ce que près de 50% des molécules d'ARN passent de l'état dimère à l'état monomère (Figure 44). De plus, avec une concentration en ARN dix fois plus élevée que celle de protéine, on peut négliger le signal lié à la liaison de DDX3 sur l'ARN. Enfin l'utilisation d'une sonde de SHAPE à demie vie courte telle que le BzCN ($t_{1/2}$ = 0,25 s) permet d'atteindre une résolution temporelle suffisante pour établir les profils de réactivité à 0, 2, 5, 10 et 30 minutes après ajout de DDX3.

Dans le but de regrouper les nucléotides en fonction de leur variation de réactivité, j'ai utilisé le logiciel Cluster 3.0 (Eisen et al., 1998) essentiellement comme décrit dans une autre étude (Grohman et al., 2013). Ce logiciel, initialement conçu pour analyser des profils d'expression génétique, permet de classer des données selon la méthode des k-moyennes. Les nucléotides analysés sont d'abord répartis aléatoirement dans un nombre donné de groupes (ou clusters). L'algorithme procède en deux temps : 1/ le vecteur moyen de chaque groupe est

calculé (ici, la réactivité moyenne à chaque instant), et 2/ chaque nucléotide est réaffecté au groupe dont il est le plus proche. La mesure de distance utilisée est le coefficient de corrélation de Pearson. Les nucléotides dont la réactivité présente un écart-moyen supérieur à 50% sont omis, ainsi que ceux dont la réactivité est indéterminée. Le nombre de clusters est fixé arbitrairement à 3. Un nombre inférieur conduit à des clusters trop hétéroclites, et un nombre supérieur à des clusters insuffisamment représentatifs. On distingue ainsi des nucléotides dont la réactivité croît (respectivement, decroît) puis décroît (respectivement, croît), et des nucléotides dont la réactivité croît jusqu'à l'état final. Les nucléotides sont ensuite filtrés pour ne garder que les variations d'amplitude supérieure à 0,2 unités de réactivité et de ratio supérieur à 1,5 (Figure 46).

On constate que plusieurs nucléotides autour de l'hélice H3 sont exposés, en particulier les nucléotides 230 et 333. Cette exposition est compatible avec le modèle de passage du dimère au monomère proposé par M. Summers dans lequel cette hélice disparaît (Figure 14). De plus, Kenyon et al., en effectuant le probing du monomère et du dimère séparément après purification sur gel en conditions natives (Kenyon et al., 2013), suggèrent que l'exposition du nucléotide 230 est caractéristique de la structure du monomère. L'exposition des nucléotides 327 et 235 pourrait également être corrélée aux données de RMN qui mettent en évidence un couplage NOE avec les nucléotides 311-312 et 284, respectivement, dans la structure du dimère (Keane et al., 2015). Il se pourrait en effet que ces nucléotides soient spatialement contraints dans la structure du dimère, mais pas dans le monomère formé par DDX3, ce qui pourrait expliquer l'augmentation de leur réactivité. Enfin le nucléotide 291 est également exposé sous l'effet de DDX3. Cette exposition correspond à une autre signature de la conformation monomère précédemment mise en évidence (Deforges et al., 2012). Ces résultats sont donc cohérents avec l'hypothèse selon laquelle DDX3 stimule le changement de conformation du dimère au monomère.

Curieusement, les nucléotides 127 et 220 sont exposés, ce qui signale un changement de conformation de l'hélice située à la base du PBS. A l'inverse, Kenyon et al. rapportent que le nucléotide 219 est moins exposé dans la conformation monomère que dans la conformation dimère. Ces derniers emploient un protocole de dénaturation puis refroidissement lent de manière à obtenir une proportion égale de monomère et de dimère Cette contradiction apparente pourrait donc suggérer que les monomères produits par DDX3 diffèrent des monomères obtenus par dénaturation – renaturation lente.

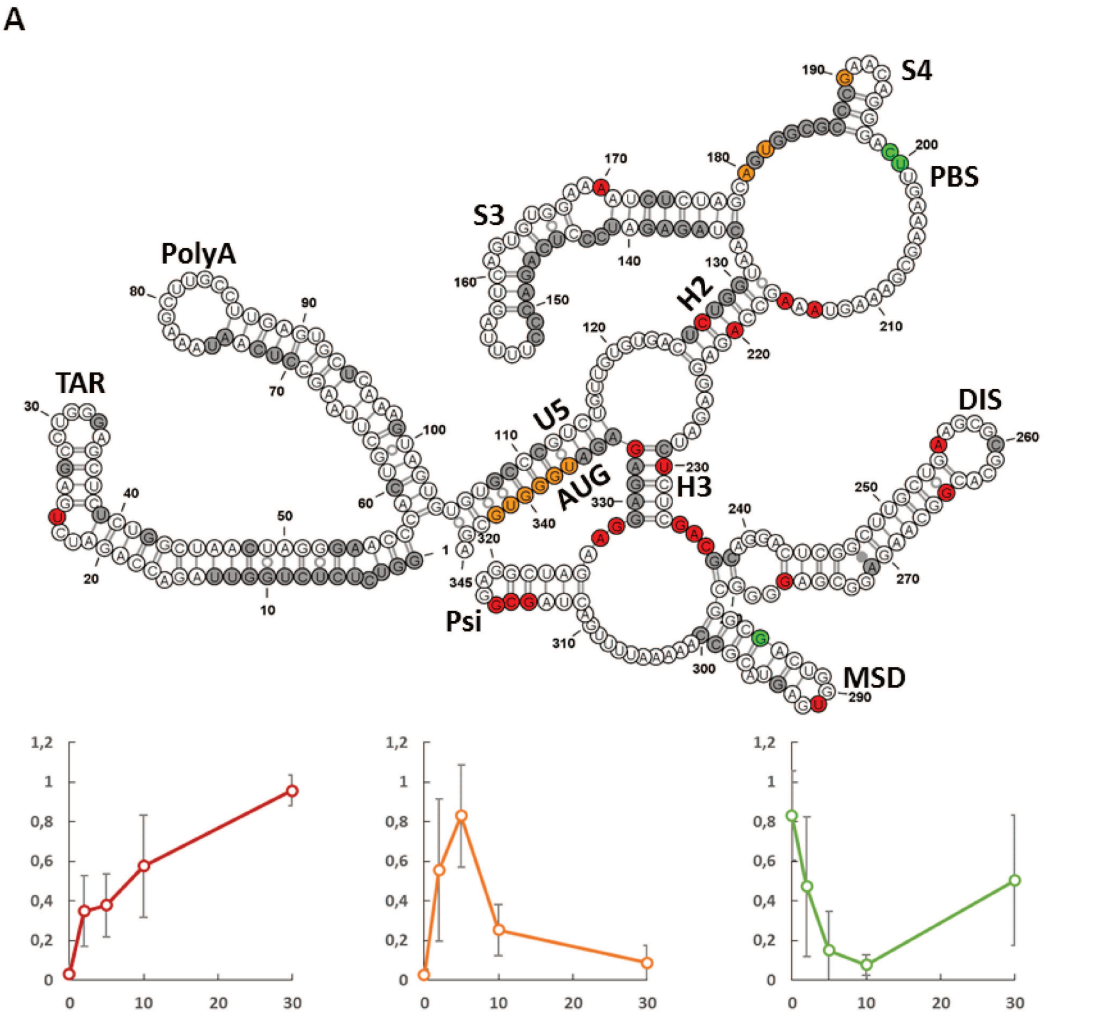
Les nucléotides 337 à 342, modélisés dans l'appariement U5:AUG, sont exposés au cours des 2 à 5 minutes après ajout de DDX3 puis reviennent à une réactivité plus basse. Dans le modèle de Summers, l'appariement AUG:U5 est déstabilisé au profit de la formation d'une tige-boucle comprenant les nucléotides 336 à 355. D'après les résultats obtenus, il est possible qu'il existe un intermédiaire relativement stable entre ces deux états dans lequel il semble qu'aucun de ces deux appariements ne soit présent. D'autre part, les réactivités mesurées résultent d'un mélange de conformations. Le dimère étant majoritaire à l'instant initial, puis le monomère devient majoritaire. Il se peut donc aussi que l'effet de DDX3 soit d'exposer la région de l'AUG dans la conformation dimère, et de structurer cette région dans la conformation monomère. Un comportement similaire est observé pour les nucléotides 180, 182 et 190 du PBS.

En parallèle, j'ai effectué analysé l'empreinte de $DDX3_{DQAD}$ sur le même ARN par SHAPE au moyen du réactif 1M7 (Figure 46 B). On observe deux empreintes principales, une dans la boucle du PBS (positions 206 – 208) et une dans une boucle interne de DIS (239 – 242). Ces sites pourraient correspondre à des sites de liaison de DDX3. On remarque qu'ils ne semblent pas significativement remaniés lors de la dissociation du dimère (Figure 46 A). Ils pourraient donc correspondre à des sites où l'hélicase est recrutée pour permettre la déstabilisation de structures en *trans*. Cependant, comme expliqué précédemment, il est délicat d'observer des empreintes au niveau de régions de faible réactivité, modélisées en double-brin. A l'inverse, les expositions apparaissent dans un contexte où la réactivité est initialement basse. Par conséquent il est difficile d'observer des sites liés par DDX3 et déstabilisés en *cis* – c'est-à-dire une empreinte en Figure 46 B qui corresponde à une empreinte en Figure 46 A. On peut toutefois relever les positions 218 dans l'hélice H2, 230 et 236 autour de l'hélice H3, 257 et 265 dans DIS et 293-294 dans MSD qui pourrait correspondre à cette situation. On remarque par ailleurs que les nucléotides 340, 342 et 343 dans la région de l'AUG sont exposés, ce qui suggère que la seule liaison de $DDX3_{DQAD}$ suffit à déstabiliser cette région. Cette région présente une augmentation de réactivité transitoire en présence de $DDX3_{wt}$. Par conséquent, l'intégrité du site catalytique actif est requise pour permettre un remaniement en plusieurs étapes.

Les résultats diffèrent de ceux présentés en Figure 42, à l'exception de l'exposition de la région de l'AUG et de l'empreinte dans la boucle du PBS. En particulier, on n'observe pas d'empreinte dans la boucle interne de TAR ni d'exposition dans l'hélice de DIS dans les résultats présentés ici, Figure 46 B. Plusieurs facteurs peuvent expliquer ce constat et limiter la

portée des interprétations. Tout d'abord, la sonde utilisée diffère entre le BzCN (Figure 42) et le 1M7 (Figure 46). Le 1M7 présente en effet une demi-vie plus longue dans l'eau, et semble plus sensible à la structure tertiaire (Bindewald et al., 2011). De plus le footprint de l'ARN de pleine longueur a été effectué en présence de 1 µM de $DDX3_{DQAD}$ (Figure 42) contre 150 nM pour le footprint de la 5'UTR (Figure 46).

Figure 46 (page suivante) ***: A/ Analyse par SHAPE du remodelage de la 5'UTR par DDX3*** **au cours du temps.** Les nucléotides dont la réactivité varie significativement (ratio supérieur à 1,5 et écart supérieur à 0,2) sont représentés en couleur. Les nucléotides de réactivité indéterminée ou trop variable sont représentés en gris. Les graphiques indiquent, pour chaque *cluster, l'évolution de la réactivité normalisée entre 0 et 1 au* cours du temps. B/ Empreinte de $DDX3_{DQAD}$ *sur la 5'UTR analysée par SHAPE avec le réactif 1M7.*


A
S4
PBS
S3
PolyA
H2
TAR
U5
AUG
H3
DIS
Psi
MSD
B
S4
PBS
S3
PolyA
H2
TAR
U5
AUG
H3
DIS
Psi
MSD

Rôle de DDX3 dans la traduction de l'ARN génomique du VIH-1

Plusieurs rapports, présentés en introduction, mettent en évidence un rôle de DDX3 dans la stimulation de la traduction de l'ARNg du VIH-1 (Fröhlich et al., 2016; Lai et al., 2013; Liu et al., 2011; Soto-Rifo et al., 2012b). D'autre part, la possibilité d'une régulation de la traduction de l'ARNg par la dimérisation a été proposée (Baudin et al., 1993; Huthoff and Berkhout, 2001; Keane et al., 2015). Etant donné la capacité de DDX3 de dissocier des dimères d'ARNg, DDX3 pourrait stimuler la traduction en déstabilisant les dimères. J'ai donc cherché à observer l'effet de DDX3 sur la traduction de l'ARNg in vitro afin d'évaluer cette hypothèse.

1) *Recherche des interactions de DDX3 avec des facteurs de l'initiation* de la traduction

Dans un premier temps, j'ai cherché à valider des interactions déjà décrites dans la littérature (eIF3, eIF4E, eIF4A, eIF4G, PABP1 et eIF2α) en effectuant des expériences de pull-downs à partir de protéines recombinantes disponibles au laboratoire (Figure 47). DDX3 est immobilisée sur une résine d'amylose grâce à son étiquette MBP, puis différentes protéines sont passées sur la résine. L'élution est effectuée par ajout de maltose, qui entre en compétition avec la résine pour la liaison à la MBP. Les protéines co-éluées avec DDX3 lors de l'élution interagissent directement avec elle. Le facteur recombinant eIF4G n'étant pas produit efficacement chez E. coli, j'ai utilisé le fragment p100 correspondant aux résidus 621 à 1560 et résultant du clivage d'eIF4G par la protéase du poliovirus (résidus 612 à 1560) (Ali et al., 2001). p100 ne contient plus le site de liaison à eIF4E mais peut encore interagir avec eIF3. eIF4E est produite sous forme recombinante chez E. coli, tandis que eIF3 est purifiée à partir d'extraits cytoplasmiques de cellules HeLa.

Après avoir contrôlé la liaison de MBP-DDX3 à la résine d'amylose ainsi que l'absence d'interaction aspécifique entre p100 et la MBP (Figure 47 A, gel de gauche), on constate que p100 est co-éluée avec MBP-DDX3 (Figure 47 A, gel de droite) ce qui traduit une interaction directe entre ces deux protéines. Cette observation confirme les résultats obtenus dans des extraits de cellules HeLa et dans des lysats de réticulocytes de lapin, où DDX3 interagit avec les résidus 682-1086 de eIF4G (Soto-Rifo et al., 2012c). De plus, ils permettent de préciser que cette interaction est directe. L'interaction entre DDX3 et eIF4E est difficilement observable

(Figure 47 B), bien qu'eIF4E soit légèrement enrichi dans l'élution en présence de MBP-DDX3. Sachant que DDX3 s'associe au complexe eIF4F comprenant eIF4E et eIF4G, il est possible que eIF4E soit plus efficacement retenue par le complexe DDX3 – p100. J'ai donc immobilisé MBP-DDX3 puis passé successivement p100 et eIF4E sur la résine (Figure 47 C). Il apparaît cependant qu'eIF4E n'est pas plus efficacement retenue, ce qui suggère que l'interaction entre DDX3 est eIF4E est de faible affinité voire même que ces deux protéines n'interagissent pas dans ces conditions. De même, je n'ai pas observé d'interaction nette entre eIF3 et DDX3, eIF3 étant largement élué lors des lavages (Figure 47 D).

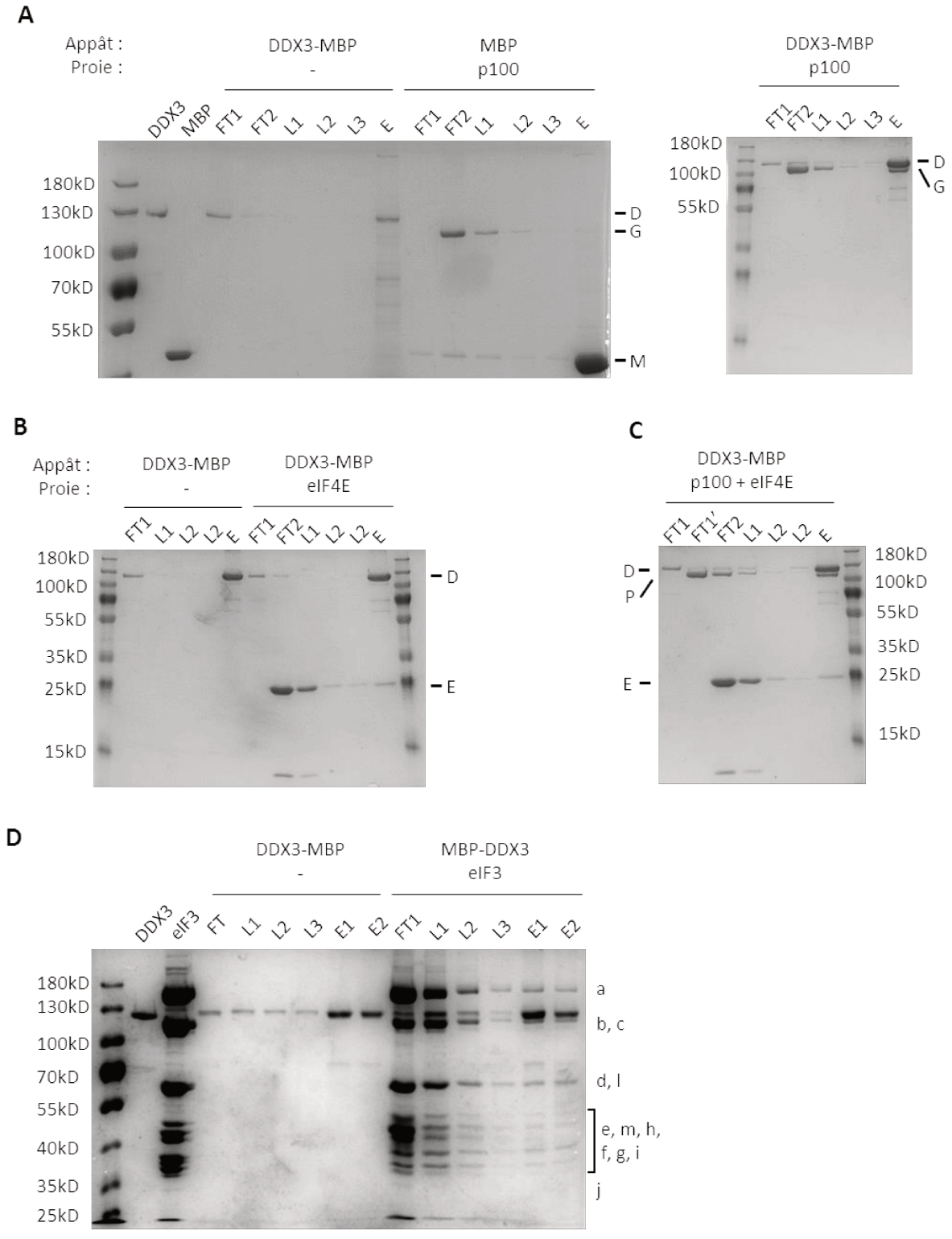


Figure 47 : ***Interaction de DDX3 avec des facteurs d'initiation de la traduction***. MBP-DDX3 est immobilisée sur une *résine d'amylose* équilibrée avec 40 mM Tris pH8, 50 mM NaCl, 2 mM DTT, 0,5 mM MgCl2, 0,01% NP-40. A, B :Pull-downs des protéines recombinantes p100 (A) et eIF4E (B) séparément, ou simultanément (C). La position des protéines sont indiquées sur les gels (D=DDX3, G=p100, E=eIF4E, M=MBP). D/ Pull-down du complexe eIF3. Les lettres en minuscule désignent les sous-unités du complexe eIF3. FT : fraction non retenue ; L :lavage ;E :élution.

Dans un second temps, j'ai cherché à contrôler la validité de ces résultats en effectuant des expériences de pull-down à partir d'extraits cytoplasmiques de cellules HeLa. (Figure 48). Comme précédemment, DDX3 est immobilisée sur une résine d'amylose grâce à son étiquette MBP. Les protéines qui semblent spécifiquement co-éluées avec DDX3 après une succession de lavages ont été identifiées par spectrométrie de masse.

Parmi les protéines identifiées, il ressort plusieurs sous-unités d'eIF3 ainsi que l'hélicase DHX9. Il est intéressant de constater que l'interaction DDX3 – eIF3 a déjà décrite dans des extraits cytoplasmiques de cellules HeLa traitées par la RNase A (Lee et al., 2008), ce qui révèle une interaction protéine – protéine mais pas nécessairement directe. Mes résultats précisent donc que l'interaction DDX3 – eIF3 n'est pas directe, puisqu'elle n'est pas observée à partir de protéines purifiées (Figure 47 D). L'interaction avec DHX9 n'a en revanche jamais été observée. DHX9 est une hélicase à boîte DExD/H qui intervient dans de nombreuses étapes du métabolisme de l'ARN et en particulier dans le balayage de la 5'UTR de certains ARNm lors de l'initiation de la traduction (Lee et al., 2016). DDX3 est identifiée dans toutes les bandes analysées, ce qui résulte probablement de l'abondance relativement élevée de DDX3 recombinante par rapport aux protéines cellulaires, et peut être également de sa protéolyse. De même, l'analyse par spectrométrie de masse relève plusieurs protéines « de ménage » telles que la glycogène phosphorylase et la tubuline. Ces protéines sont également présentes dans les extraits en quantité relativement abondante et contaminent vraisemblablement l'élution de manière non spécifique. En revanche, on ne retrouve pas l'interaction avec eIF4G. Il est possible que, dans des extraits, eIF4G soit trop peu abondante pour que son interaction avec DDX3 soit détectable. Il se peut aussi qu'eIF4G soit rendue inaccessible à DDX3 du fait d'interactions avec d'autres protéines telles que eIF4A, PABP, eIF3 ou encore eIF4E.

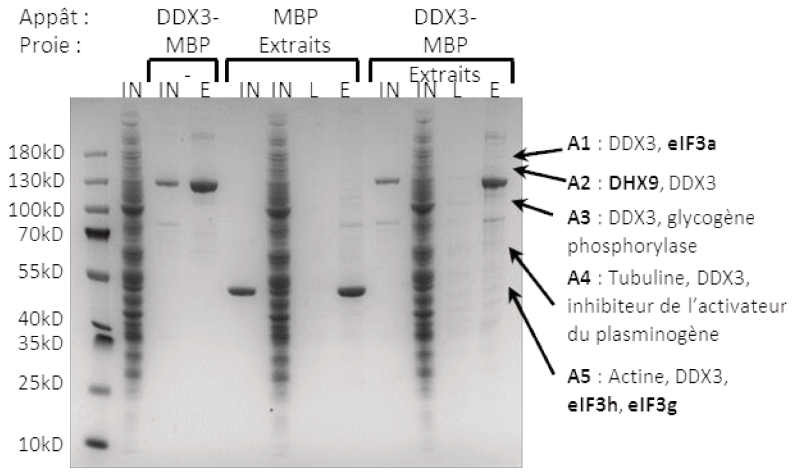


Figure 48 : Recherche des partenaires protéiques de DDX3. A/ Pull-*down d'extraits* cytoplasmiques de cellules HeLa au moyen de MBP-DDX3 immobili*sée sur résine d'amylose.* Les bandes analysées par spectrométrie de masse et les protéines identifiées sont indiquées par des flèches.

Enfin, J'ai contrôlé la capacité de DDX3 à interagir avec la coiffe, décrite précédemment (Soto-Rifo et al., 2012b, 2013). J'ai utilisé une résine de chromatographie d'affinité comportant des résidus de guanosine 5' triphosphate méthylés en N7. Les protéines retenues sont éluée en ajoutant un excès de m7GTP libre. Comme attendu, le facteur eIF4E est efficacement retenu. A l'inverse, DDX3 seule ou en présence d'eIF4E n'est pas retenue (Figure 49). Ces résultats suggèrent que DDX3 n'interagit avec la coiffe ni directement, ni via eIF4E in vitro. Ils contredisent l'hypothèse selon laquelle DDX3 interagirait directement avec la coiffe (Soto-Rifo et al., 2013). Les raisons de cette divergence ne sont pas claires, et il est possible que des différences expérimentales non explicitées dans l'article de Soto-Rifo et al. telles que la nature du tampon, la concentration en DDX3 ou la qualité de la préparation soient en cause.

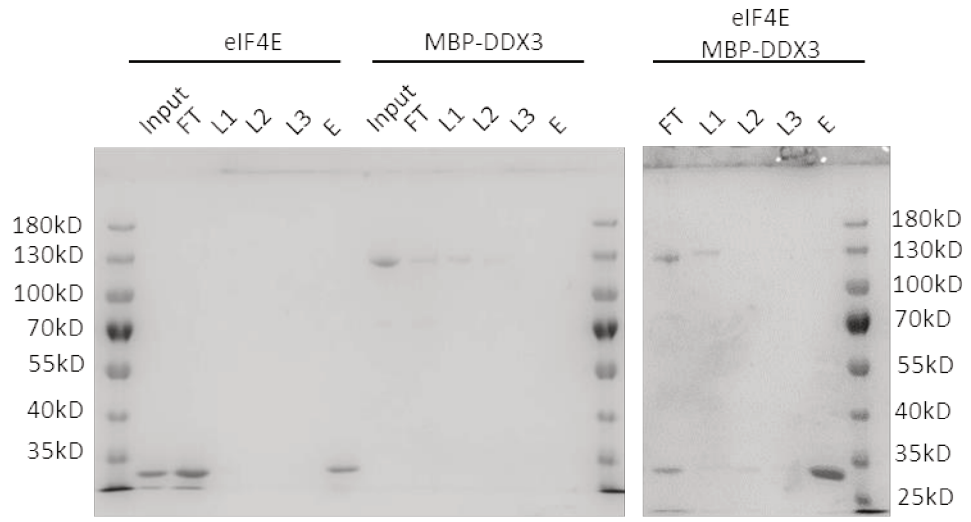


Figure 49 : DDX3 n'interagit pas avec la coiffe in vitro. Pull-downs sur résine de m7GTP *d'*eIF4E et DDX3 (à gauche). A droite : eIF4E est incubé en présence de DDX3 et passé sur résine de m7GTP.

2) *Evaluation de l'effet de DDX3 sur la traduction in vitro*

Afin d'évaluer l'effet de DDX3 sur la traduction de l'ARN génomique in vitro, j'ai effectué des tests de traductions sur des lysats de réticulocytes de lapin (RRL) en présence ou en absence de DDX3 recombinante. Notons qu'on ne détecte pas de DDX3 endogène par western-blot dans les RRL utilisés, ce qui suggère que DDX3 n'est pas présente dans les réticulocytes (Nathalie Chamond, communication personnelle). J'ai utilisé comme témoin un ARN constitué de la 5'UTR de la globine suivie de la région codante de la luciférase, un rapporteur fréquemment utilisé dans les expériences de traduction. La quantité de protéine produite est estimée au moyen de l'activité spécifique du ^{35}S utilisé pour le marquage.

Comme attendu, la traduction de la luciférase produit principalement une protéine de 39 kD, mais aussi deux protéines plus petites de 37 et 35 kD qui résultent de l'initiation au niveau de codons initiateurs situés légèrement un aval du premier codon initiateur. Quant à l'ARNg du VIH-1, sa traduction donne lieu à deux protéines de 40 et 55 kD. La première est l'isoforme de Gag traduite à partir du deuxième codon initiateur (759-AUG-761), GagPr40 (Figure 50). L'initiation au niveau de ce codon résulte de l'entrée interne du ribosome entre les deux codons initiateurs sur l'IRES Gag. La deuxième isoforme, GagPr55, est la protéine de pleine longueur traduite à partir du premier codon initiateur (336-AUG-338), soit par entrée interne du ribosome, soit par l'initiation dépendante de la coiffe.

La traduction dépendante de la coiffe est stimulée par la présence d'une coiffe et d'une queue polyA. On s'attend donc à une augmentation de la production de GagPr55 et de la luciférase en présence d'un ARN coiffé et polyadénylé. Curieusement, mes résultats indiquent que la coiffe et la queue polyA ne semblent pas avoir d'effet notable sur l'efficacité de la traduction. Il est possible que seule une faible proportion d'ARN aient été effectivement coiffé et polyadénylé in vitro.

On constate que GagPr40 est traduit plus efficacement que GagPr55. Cette observation est cohérente avec l'inhibition de la traduction dépendante de la coiffe in vitro. L'ajout de DDX3 ne semble pas avoir d'effet significatif sur l'efficacité de traduction des deux isoformes, pas plus que sur la traduction de la luciférase. Cependant, ces résultats préliminaires pourraient être approfondis. Il est possible que certains paramètres expérimentaux nécessitent un ajustement pour permettre d'observer un effet de DDX3 sur la traduction de l'ARNg du VIH-1.

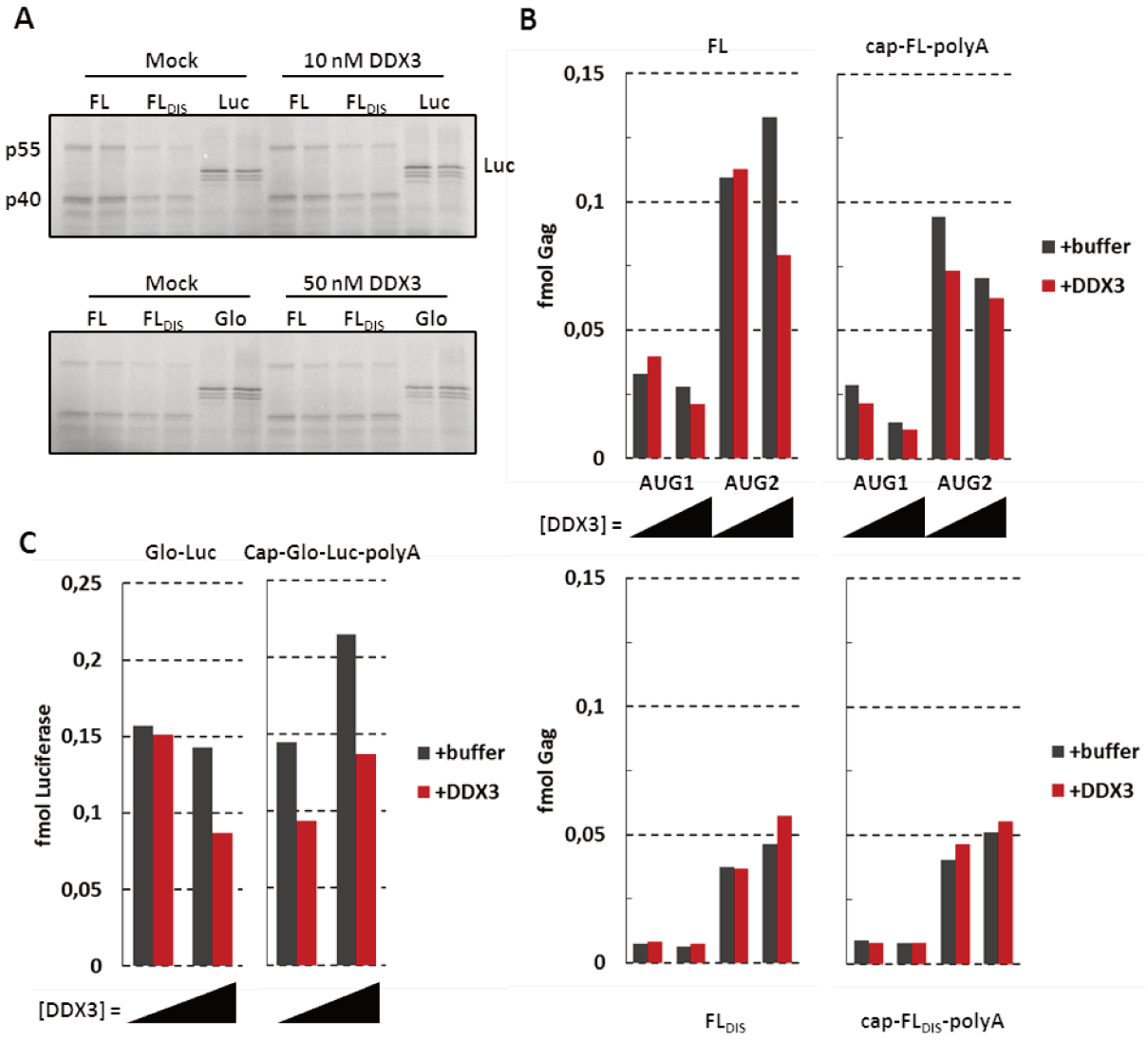


Figure 50 : Effet de DDX3 sur la traduction dans des lysats de réticulocytes. A/ Les protéines synthétisées sont analysées par SDS-PAGE. Dans chacune des conditions testées, chaque ARN est utilisé coiffé et polyadénylé (à droite) ou non (à gauche). La quantification ne révèle pas *d'effet significatif de DDX3 sur la traduction de l'ARNg ('FL')* (B, haut), ni sur la traduction de *l'ARNg muté dans DIS (B, bas)* ni sur la traduction de la luciférase sous contrôle de la *5'UTR de la globine (C).*

Discussion

Relations structure – fonction de l'ARNg du VIH-1

1) Structures secondaire et tertiaire du PBS

La comparaison des structures secondaires des isotypes MAL et NL4-3 révèle des différences importantes au niveau du PBS qui résultent d'une insertion d'une vingtaine de nucléotides dans la séquence de MAL. Lors de l'encapsidation, le PBS permet de recruter l'ARN_t^{Lys3} grâce à une séquence complémentaire. Cet ARN_t interagit dans la cellule avec l'ARN_t-lysyl synthétase. Une région du PBS, le tRNA Like Element (TLE) interagit avec l'ARN_t-lysyl synthétase pour libérer l'ARN_t^{Lys3} et faciliter son recrutement (Jones et al., 2013). De façon intéressante, si la séquence de la boucle de l'anticodon de l'ARN_t^{Lys3} ressemble à celle de la boucle du TLE de NL4-3, elle est nettement différente de celle du TLE de MAL (Figure 51). La séquence complémentaire de l'ARN_t^{Lys3} est quant à elle parfaitement conservée. Cependant les parties 5' et 3' de cette séquence est impliquée dans un appariement pour l'isolat MAL mais pas pour le clone NL4-3, ce qui suggère que l'hybridation de l'ARN_t^{Lys3} entraîne pour l'isolat MAL un remaniement différent de la structure.

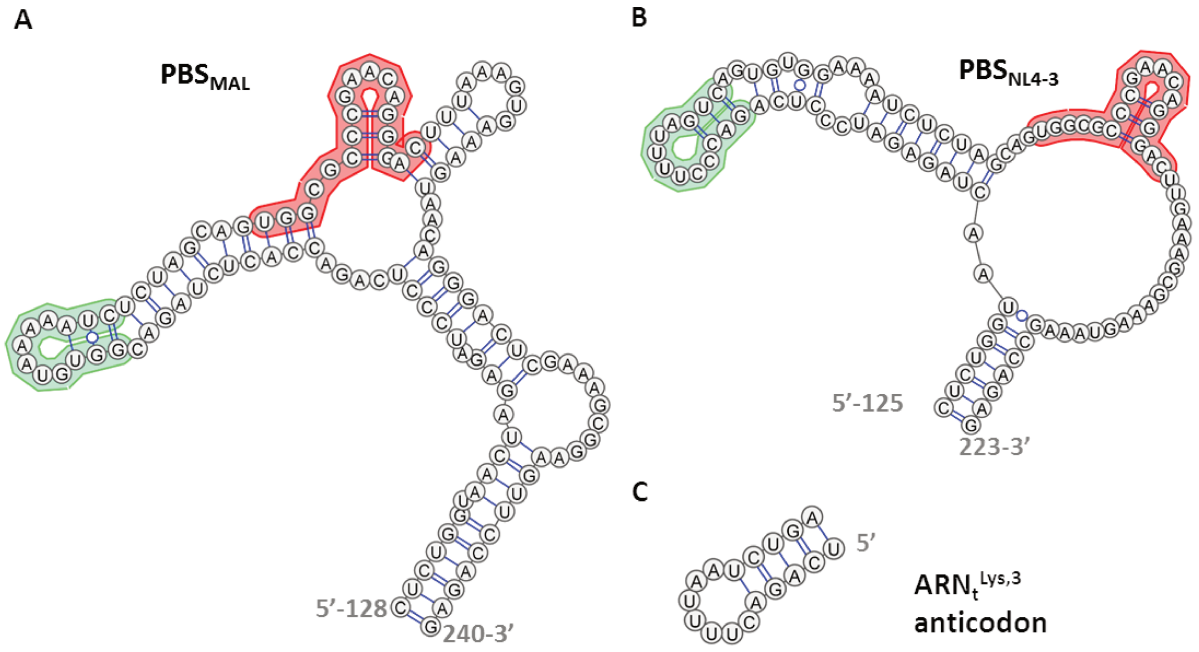


Figure 51 : Structure secondaire du TLE et du PBS de MAL (A) et NL4-3 (B). Le TLE en *entouré en vert tandis que le PBS (au sens strict) est entouré en rouge. C/ Boucle de l'anticodon de l'ARN_t^{Lys3}.*

Des études de SAXS ont récemment montré que les structures globales du PBS des souches NL4-3 (125-223) et MAL (123-218) sont proches, qu'elles ont une forme similaire à celle d'un ARN_t, et que cette structure est critique dans le recrutement de l'ARN_t^{Lys3} (Figure 52) (Comandur et al., 2017). Ces résultats indiquent donc que les structures secondaires du PBS du sous-type A, représenté par l'isolat MAL, et celle du sous-type B, représenté par le clone NL4-3, ont évolué sous la contrainte d'une structure tertiaire fonctionnelle.

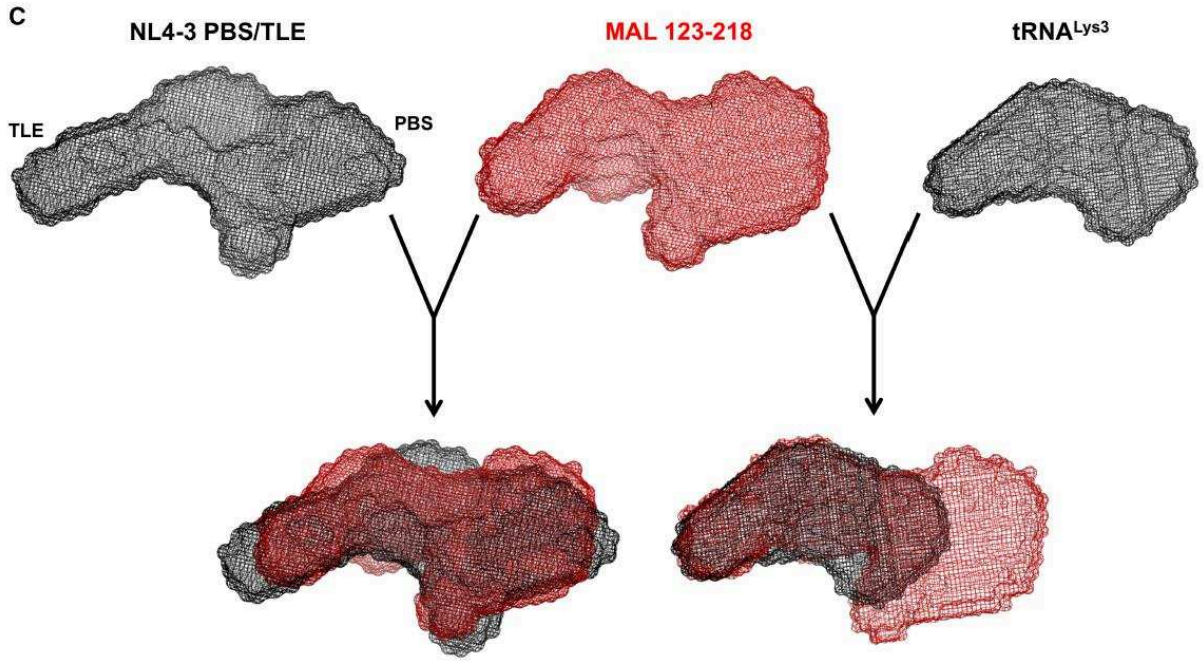


Figure 52 : Enveloppes des PBS de NL4-3 (125-223) et MAL (123-218) et de la boucle de *l'anticodon de l'*ARN_t^{Lys3}. Les enveloppes des PBS ont été obtnues par SAXS. Celle de la boucle *de l'a*nticodon ARN_t^{Lys3} a été obtenue par cristallographie (Comandur et al., 2017).

2) Différents sous-types, différents mécanismes de dimérisation?

A la différence du clone NL4-3, les dimères formés par l'isolat MAL se dissocient lors d'une électrophorèse native en absence de magnésium. Ces observations rejoignent celles d'autres auteurs qui ont L'examen de la séquence de la région 5' de ces deux variants du VIH-1 révèle, outre une insertion de 20 nucléotides dans le PBS de l'isolat MAL, une séquence différente au niveau de la boucle apicale de DIS. Mes résultats aboutissent à deux hypothèses : ou bien le complexe kissing de NL4-3 est stable lors d'une électrophorèse en absence de magnésium, à la différence du complexe kissing de MAL, en raison d'une plus forte proportion de paires GC intermoléculaires ; ou bien le nombre inférieur de sites de liaison au magnésium de la tige-boucle DIS du clone NL4-3 lui confère une instabilité favorable à la formation spontanée des dimères étendus.

Ces hypothèses permettent d'expliquer certaines divergences décrites dans la littérature. En effet, une étude montre que les dimères étendus ne se forment pas spontanément à 37°C en présence de magnésium, en recourant à une construction de 615 nucléotides issue de l'isolat MAL (Paillart et al., 1996). Toutes les études qui témoignent de la formation de dimères étendus à partir de l'isolat MAL utilisent une construction réduite à la seule tige-boucle DIS (Aduri et al., 2013; Baba et al., 2005; Ennifar et al., 1999; Windbichler et al., 2003). A l'inverse, d'autres auteurs ont montré que le dimère étendu peut se former spontanément à 37°C en utilisant le clone NL4-3. Ces derniers proposent que seuls les ARN dont le palindrome est GCGCGC peuvent former des dimères stables sur des gels sans magnésium, indépendamment de la présence de nucléocapside (Tran et al., 2015). En contradiction apparente, d'autres rapports soutiennent que la formation de dimères étendus à partir de l'isolat Lai, également rattaché au sous-type B, n'est pas spontanée à 37°C et requiert le domaine nucléocapside (Muriaux et al., 1996a) ou une incubation à 55°C (Muriaux et al., 1996b). Outre la différence de constructions utilisée dans ces deux études (HIV-1 NL4-3 1-356 contre HIV-1 Lai 77-402, respectivement), dans un cas la formation spontanée de dimères étendus à 37°C est observée en présence de magnésium (Tran et al., 2015), tandis que dans l'autre la nécessité d'une incubation à 55°C ou de la présence de nucléocapside pour étendre des dimères kissing est observée en absence de magnésium (Muriaux et al., 1996b, 1996a). Les résultats de Muriaux et al. n'excluent donc pas la formation spontanée de dimères étendus du sous-type B en présence de magnésium. Ensemble, ces résultats indiquent que le sous-type A ne peut pas former de dimères étendus en présence de magnésium, au contraire du sous-type B.

Une étude par RMN a montré que la boucle interne de DIS déstabilise la partie supérieure de l'hélice. La liaison du magnésium stabilise la structure et rend la transition du dimère kissing au dimère étendu possiblement plus dépendante de la nucléocapside (Sun et al., 2007). La séquence de la boucle interne de DIS étant conservée chez les sous-types A et B, on peut penser que ces conclusions s'appliquent aux deux sous-types. En revanche, la seule liaison du magnésium au niveau de la boucle interne n'inhibe pas la conversion spontanée du dimère kissing. En effet, en utilisant une tige-boucle correspondant au sous-type B, des auteurs ont montré que la conversion spontanée du dimère kissing en dimère étendu est même stimulée par le magnésium (Windbichler et al., 2003). De façon intéressante, la tige-boucle DIS du sous-type A présente un site de liaison au magnésium supplémentaire modélisé au niveau des purines en 5' du palindrome (Figure 53) (Ennifar et al., 2001; Jossinet et al., 1999). On peut penser que ce site de liaison supplémentaire augmente encore le niveau de stabilisation induit par le

magnésium et inhibe la conversion spontanée en dimères étendus. En conclusion, la propension du sous-type B à former spontanément des dimères étendus peut s'expliquer par le comportement plus instable de la structure du dimère kissing. A l'inverse, la structure du dimère kissing du sous-type A est plus stable et la conversion vers un dimère étendu plus dépendante de la nucléocapside.

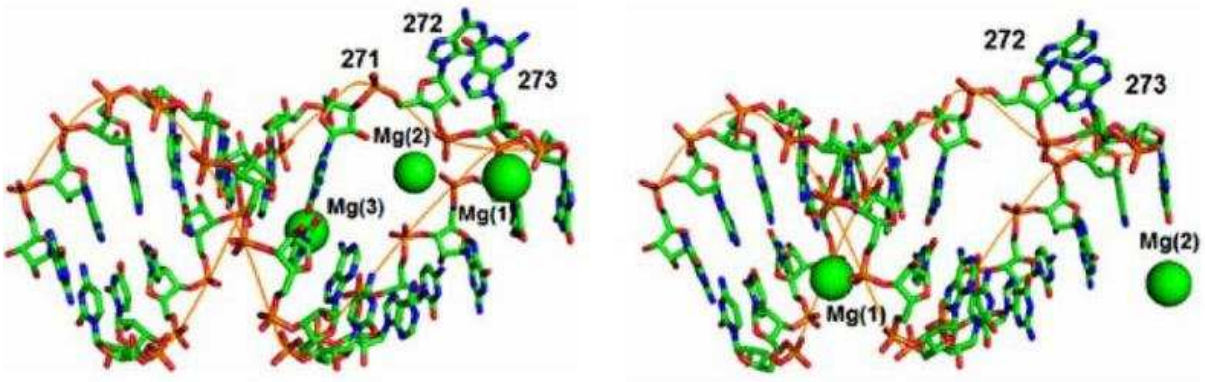


Figure 53 : Liaison du magnésium par la boucle apicale de DIS du sous-type A (gauche) et B (droite) dans la conformation monomère. Tiré de (Kim and Shapiro, 2013).

D'autre part, des expériences menées au moyen de pinces optiques ont montré que les dimères kissing des tige-boucles DIS du sous-type B sont plus résistants que ceux du sous-type A. Ces deux dimères kissing sont stabilisés par le magnésium (Li and Tinoco, 2009). Ces résultats sont tout à fait cohérents avec le plus grand nombre de paires GC présentes dans l'interface du dimère du sous-type B (Figure 54). Il est donc également envisageable que les dimères kissing formés par le clone NL4-3 soient plus stables que ceux formés par l'isolat MAL sur des gels en absence de magnésium. Notons que cette hypothèse et l'hypothèse précédente ne sont pas mutuellement exclusives et peuvent simultanément expliquer les résultats obtenus.

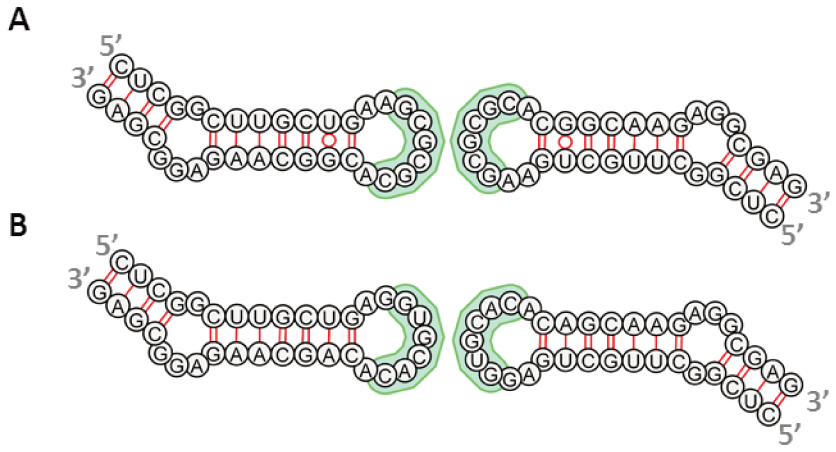


Figure 54 : Complexes kissing des tige-boucles DIS du sous-type A (Mal) et B (NL4-3). Les nucléotides du palindrome sont entourés en vert.

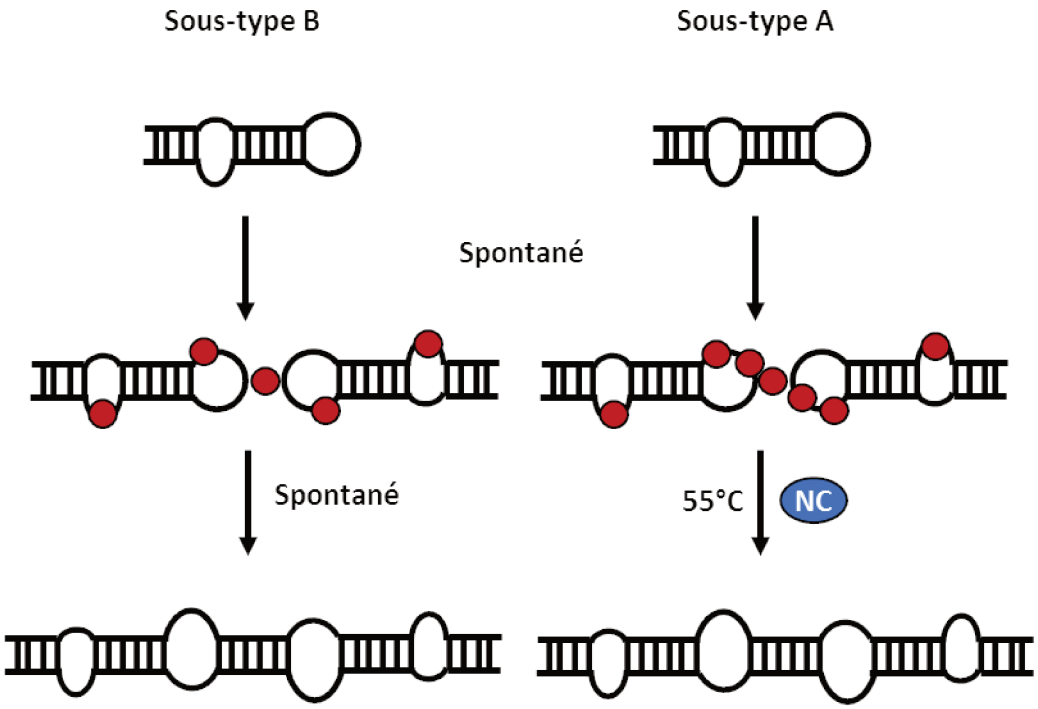


Figure 55 : Modèle de dimérisation des sous-types A et B. Seule la tige-boucle DIS est schématisée. Les ions magnésium sont indiqués en rouge.

A l'instar de l'ARN du VIH-1, d'autres ARN forment des dimères. En particulier, l'ARN génomique du VHC (Virus de l'Hepatite C) dimérise in vitro via son domaine 3'X situé dans la 3'UTR, selon un mécanisme étonnamment similaire. En effet, le VHC n'appartient pas à la famille des retroviridae. Les particules virales sont supposées ne contenir qu'un brin d'ARN à la différence des retroviridae, encore que cette affirmation soit discutée. La dimérisation de l'ARN du VHC est également supposée procéder en deux temps, avec la formation d'un dimère kissing susceptible d'être converti en un dimère étendu. La dimérisation du VHC pourrait

réguler l'encapsidation, la traduction ou encore la réplication du virus. Cependant, à ce jour, il n'est pas certain que la dimérisation de l'ARN du VHC se produise effectivement dans les cellules (Cantero-Camacho et al., 2017). In vitro, suite à une incubation à 22°C en présence de magnésium, le domaine 3'X forme des dimères instables sur gel sans magnésium. A l'inverse, une incubation en présence de magnésium à 55°C ou avec la protéine core du VHC conduit à la formation de dimères stables sur des gels sans magnésium (Shetty et al., 2010). D'autres auteurs rapportent que le domaine 3'X forme des dimères étendus en absence de magnésium, et des dimères kissing en présence de magnésium (Cantero-Camacho and Gallego, 2015; Cantero-Camacho et al., 2017). Ces résultats suggèrent que le magnésium stabilise les dimères kissing et inhibe la formation spontanée de dimères étendus, de façon assez semblable à ce qui est observé pour le sous-type A du VIH.

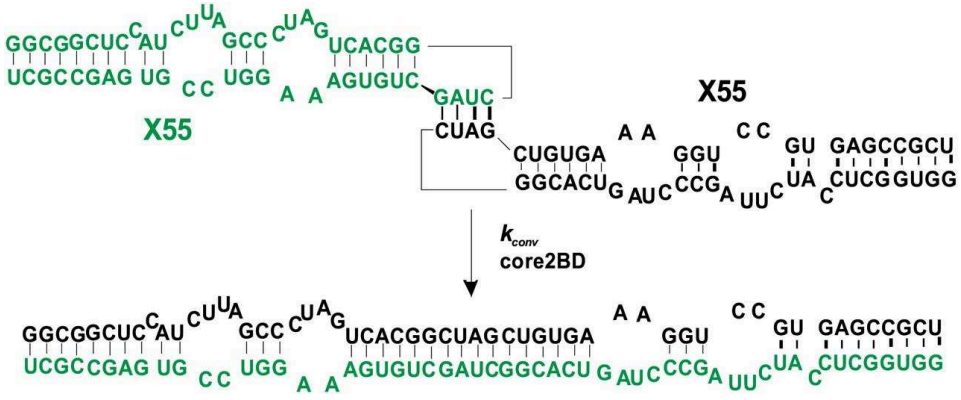


Figure 56 ***: Modèle de conversion du dimère kissing d'une construction de 55 nucléotides du domaine 3'X du VHC (haut) en dimère étendu (bas).*** Les deux molécules sont colorées en vert et noir, respectivement. Tiré de (Shetty et al., 2010)

Enfin, il est possible que les différences observées pour la dimérisation de l'isolat MAL et du clone NL4-3 s'expliquent également par la différence de structure secondaire adoptée par le domaine PBS. En effet, l'hybridation de l'ARN_t^{Lys3} modifie la structure de la boucle du PBS (Jones et al., 2014) et favorise la dimérisation du NL4-3 (Seif et al., 2013). D'autre part, DDX3 modifie la réactivité SHAPE de nucléotides situés dans le PBS, notamment autour du site de liaison à l'ARN_t^{Lys3} (Figure 46). Il est donc possible que la structure du PBS régule la dimérisation in vitro. In vivo, le recrutement de l'ARN_t^{Lys} requiert la liaison de la lysyl-ARN_t synthétase au niveau du « tRNA Like Element » (TLE) situé en 5' du site de liaison de l'ARN_t de la lysine. La liaison de la lysyl-ARN_t synthétase au TLE est indépendante de l'état monomère ou dimère de l'ARN (Comandur et al., 2017). On peut donc imaginer que la dimérisation in vivo requiert le recrutement de l'ARN_t^{Lys3}. Cependant, un ARN dépourvu de la majeure partie

du domaine PBS est encapsidé aussi efficacement que l'ARN sauvage (Heng et al., 2012), ce qui questionne cette hypothèse.

D'un point de vue biologique, ces différences suggèrent que ces sous-types ont mis au point différents mécanismes de contrôle de la dimérisation. C'est une étape critique du cycle réplicatif, puisque la dimérisation semble elle-même réguler l'encapsidation, et peut-être aussi la traduction des protéines Gag et Gag-Pol. Une meilleure compréhension des différents mécanismes de dimérisation est donc un prérequis nécessaire pour envisager la dimérisation comme une éventuelle cible thérapeutique.

3) Régulation de la dimérisation par la méthylation

J'ai montré que la méthylation de toutes les adénosines au niveau de l'azote 6 diminue l'efficacité de la dimérisation de NL4-3. D'autres études rapportent l'effet de modifications des bases sur la dimérisation. Ainsi, chez l'isolat MAL, la méthylation des guanosines du palindrome en position N7 inhibe également la dimérisation, mais pas la méthylation des adénosines flanquant le palindrome en N1 (Jossinet et al., 1999). L'approche utilisée ne permet pas d'identifier précisément les méthylations dont l'effet sur la dimérisation est critique. In vivo, il semble que les adénosines méthylées dans la 5'UTR sont principalement A198 et A242. La réactivité de A198 est modifiée de façon transitoire au cours de l'incubation en présence de DDX3 (Figure 46), ce qui suggère que ce nucléotide est impliqué dans la dimérisation. Par conséquent, la méthylation de ce résidu pourrait avoir un effet sur la dimérisation, mais cette hypothèse reste très incertaine. Afin de clarifier l'effet de la méthylation sur la dimérisation, il serait possible de méthyler plus spécifiquement l'ARN in vitro en purifiant le complexe METTL3/14 responsable de la méthylation de la 5'UTR in vivo. Les différences de structures pourraient être examinées en effectuant une expérience de SHAPE sur l'ARNg purifié à partir de cellules d'un côté, et transcrit in vitro de l'autre.

4) *Interdépendance du repliement de la 5'UTR et de la régio*n codante de Gag

Deux observations mettent en évidence une interdépendance entre le repliement de la 5'UTR et de la région codante. D'une part, les profils de réactivités de ces deux fragments diffèrent par rapport au profil de réactivité de l'ARNg. Notons que la région dont la réactivité

varie le plus suivant la construction est le site de dimérisation. D'autre part, les dimères d'ARN comprenant la 5'UTR et la région codante de Gag sont dissociés moins efficacement par DDX3 que les dimères de 5'UTR seule. Ces résultats sont cohérents avec des études antérieures qui ont mis en évidence des différences de stabilités suivant la taille de la construction utilisée. Ainsi, les dimères formés par le fragment 1-311 sont moins stables que les dimères formés par les fragments 1-615 et 1-707 (Marquet et al., 1994). Ensemble, ces données suggèrent que la région codante de Gag est nécessaire au bon repliement du site de dimérisation de la 5'UTR, ou bien qu'il existe des sites de dimérisation additionnels dans la région codante.

Comme présenté en introduction, il est possible qu'un quadruplex de G puisse se former entre deux brins et contribuer à la dimérisation (Sundquist and Heaphy, 1993). Cette hypothèse pourrait permettre d'expliquer la stabilité accrue de l'ARN de pleine longueur vis-à-vis de DDX3, bien que sa pertinence biologique a été remise en question dans un autre rapport (Fu et al., 1994). La séquence de G supposée intervenir dans cet appariement se trouve dans le fragment 732-825 qui seul est capable de dimériser in vitro en présence de potassium, et dans une moindre mesure en présence de sodium. Dans mes conditions expérimentales qui incluent également du sodium, le fragment 336 – 1636 forme effectivement des dimères, mais de façon quasi-indétectable (Figure 57). Il est possible que les plus faibles concentrations d'ARN (50 nM contre 1,5 µM) et de sodium (50 mM contre 100 mM) utilisées expliquent cette différence. Par ailleurs, lors de l'incubation en présence de DDX3, la bande qui correspondrait au dimère formé par un quadruplex de G disparaît. Ces résultats, quoique préliminaires, semblent en faveur de l'existence d'un quadruplex de G intermoléculaire dans la région codante de Gag, et de la capacité de DDX3 d'interagir et de déstabiliser des quadruplex de G.

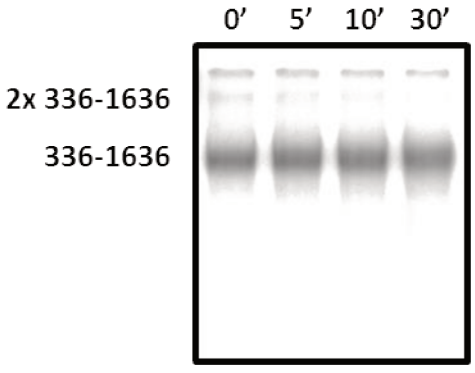


Figure 57 : Analyse sur gel natif du fragment 336-1636 incubé en présence de DDX3. 50 nM *d'ARN sont incubés en présence d*e 5 nM de DDX3, puis un aliquot est prélevé aux temps indiqués et déposé sur gel.

La boucle apicale du polyA forme une interaction de longue distance avec des nucléotides situés dans la région codante (79 – 85 avec 442 – 448 dans NL4-3) (Paillart et al., 2002). De façon cohérente, j'ai observé que les nucléotides 81, 82 et 83 sont significativement plus exposés dans la 5'UTR seule, et que le nucléotide 446 est plus exposé dans la région codante seule. Ces différences sont toutefois d'amplitude plus modérée que les différences rapportées en Figure 41. Les auteurs ont démontré que cette interaction est intramoléculaire (Paillart et al., 2002), ce qui exclut l'hypothèse d'une stabilisation des dimères d'ARNg par cette interaction. En revanche, il est possible que cette interaction contribue indirectement à stabiliser les dimères d'ARNg en modifiant la structure tertiaire de l'ARN ou en limitant le degré de liberté de la structure.

Les résultats des footprints diffèrent suivant la construction. En particulier, avec l'ARN de pleine longueur, l'hélice de DIS est exposée vis-à-vis du 1M7 (résultats non présentés) comme vis-à-vis du BzCN (Figure 42) en présence de $DDX3_{DQAD}$ alors qu'elle n'est pas exposée avec la 5'UTR seule (Figure 46 B). Ces observations paraissent contradictoires avec le fait que DDX3 dissocie plus efficacement des dimères de 5'UTR que des dimères d'ARNg. Il est possible que la stabilité de DIS diffère suivant la construction utilisée ; il se peut également que le mode de liaisons de $DDX3_{DQAD}$ diffère.

La dimérisation de l'ARNg est proposée comme un mécanisme contribuant à la reconnaissance spécifique de l'ARN non épissé par Gag (Moore and Hu, 2009). Par ailleurs, une étude a montré que les transcrits épissés dimérisent également in vitro (Sinck et al., 2007). Pourtant, les transcrits épissés sont très faiblement encapsidés comparé à l'ARNg (Houzet et al., 2007). Outre les déterminants structuraux propres à l'ARN non épissé, il se pourrait que la réplication du VIH-1 s'appuie sur des facteurs cellulaires pour garantir la spécificité de l'encapsidation. Dans cette perspective, il est intéressant de constater que DDX3 dissocie moins efficacement des dimères comprenant la région codante de Gag. DDX3, comme d'autres hélicases cellulaires, pourrait inhiber l'encapsidation des transcrits épissés en inhibant leur dimérisation. L'existence de déterminant de dimérisation dans la région 3' constituerait donc un mécanisme de contrôle qualité.

DDX3 régule l'expression de l'ARNg

1) Effet de DDX3 sur la *traduction de l'ARN génomique*

i. DDX3 pourrait réguler la traduction et l'encapsidation de l'ARNg

Je n'ai pas observé d'effet de DDX3 sur la traduction de l'ARN génomique dans des lysats de réticulocytes. Ces résultats contrastent avec ceux obtenus in cellulo par Soto-Rifo et al, qui ont observé un effet de l'expression de DDX3 sur la traduction de l'ARN du VIH. Il est possible que des facteurs cellulaires absents de réticulocytes soient nécessaires à la stimulation de la traduction par DDX3.

Dans le cytoplasme, DDX3 pourrait réguler les niveaux relatifs de traduction et d'encapsidation de l'ARNg. En effet, certaines études suggèrent que les dimères d'ARNg sont moins efficacement traduits que les monomères in vivo. Cette inhibition ne résulte pas directement de la conformation monomère ou dimère de l'ARN (Abbink et al., 2005; Boeras et al., 2017), mais plutôt de l'interaction avec Gag (Anderson and Lever, 2006). Par conséquent, si DDX3 dissocie les dimères d'ARNg, elle pourrait diminuer l'interaction de Gag avec l'ARNg et stimuler ainsi la traduction de façon indirecte. Cette hypothèse est cohérente avec une récente étude qui a montré que la dimérisation se produit au niveau de la membrane (Chen et al. 2016). DDX3 quant à elle est présente dans le noyau et dans le cytoplasme, principalement dans des granules de stress. DDX3 pourrait donc inhiber la dimérisation dans le cytoplasme, mais pas au niveau de la membrane. Notons que les résultats de mes tests de traduction en extraits de réticulocytes ne permettent pas d'alimenter cette hypothèse. GagPr55 est effectivement produite dans les extraits de réticulocytes, quoique faiblement. Cependant, dans mes conditions expérimentales, la concentration maximale de DDX3 est de 50 nM pour 50 nM d'ARN. A ces concentrations, DDX3 n'a qu'un faible effet sur la dissociation des dimères d'ARN génomique (Figure 43).

ii. Interactions de DDX3 avec des facteurs de traduction

La recherche de partenaires protéiques de DDX3 s'est avérée délicate. La technique de pull-down permet de mettre en évidence des interactions de forte affinité. Or, comme présenté en introduction, DDX3 interagit avec de très nombreuses protéines de l'immunité innée, de la traduction, de l'export nucléaire ou encore de l'épissage. Afin de garantir la multiplicité de ses

rôles, on peut penser que les interactions de DDX3 sont de faible affinité. Autrement dit, ces interactions présentent un k_{off}/k_{on} élevé, ce qui faciliterait l'échange de partenaires protéiques et le passage d'un état à un autre dans un processus moléculaire donné. L'absence d'identification flagrante de partenaires protéiques par pull-down conforte cette hypothèse. In vivo, il est possible que ces interactions soient régulées par des modifications post-traductionnelles.

Les résultats obtenus confirment toutefois l'interaction entre DDX3 et eIF4G. Cette interaction est également décrite chez la levure (Hilliker et al., 2011). Par ailleurs, la co-purification de Ded1, eIF4G (humaine) et eIF4A (humaine) à partir de bactéries sur-exprimant chacune de ces protéines conduit à la formation d'un complexe stable sur une colonne d'exclusion stérique (Luc Ponchon, communication personnelle). Ces résultats suggèrent que l'interaction avec eIF4G est conservée chez Ded1 et DDX3

2) *Autres effets possibles de DDX3 sur l'expression de l'ARN génomique*

i. DDX3 pourrait réguler l'épissage de l'ARNg

Le compartiment cellulaire où se produit la dimérisation n'est cependant pas certain, et il n'est pas exclu qu'elle se produise dans le noyau (Sinck et al., 2007). On peut alors supposer un rôle de DDX3 dans l'épissage. En effet, indépendamment de son effet sur la dimérisation, DDX3 remodèle des structures aux environs du MSD (Figure 46). Par ailleurs l'épissage peut être régulé par la structure de l'ARN (Jablonski et al., 2008; Mueller et al., 2014). DDX3 pourrait stimuler l'activité du site majeur donneur d'épissage en augmentant son accessibilité à machinerie d'épissage. Elle favoriserait ainsi la production des protéines virales régulatrices telles que Vpr, qui induit un arrêt du cycle cellulaire en phase G2/M. Au cours de cette phase, DDX3 est phosphorylée et relocalisée dans le cytoplasme (Choi and Lee, 2012). Par conséquent, en réponse, DDX3 ne stimulerait plus l'épissage de l'ARN de pleine longueur et participerait à la progression du cycle réplicatif du virus.

ii. DDX3 pourrait stimuler l'expression de Gag en inhibant la réponse NMD

Une autre hypothèse pourrait permettre d'expliquer qu'on n'observe pas d'effet de DDX3 sur la traduction in vitro. La réplication du VIH, comme celle de tous les rétrovirus,

nécessite de contourner la réponse NMD (Non-sens Mediated Decay). Cette réponse assure à la cellule de dégrader les ARNm qui présentent un codon stop prématuré, issu d'une mutation ou d'un défaut d'épissage. Les ARNm substrats de la réponse NMD présentent donc une région 3'UTR très longue, des introns dans la 3'UTR ou encore de multiples cadres de lecture. Pourtant, l'ARNg du VIH échappe à la dégradation par la réponse NMD. Il semble que le virus inhibe la réponse NMD en détournant UPF1, une hélicase impliquée dans ce mécanisme. Une RNP contenant l'ARNg, Rev, UPF1 et DDX3 a été identifiée (Ajamian et al., 2015). Ces résultats indiquent un rôle possible de DDX3 dans la réponse NMD.

Perspectives

Il serait intéressant d'étudier l'effet des modifications post-traductionnelles de DDX3 sur son activité biochimique. Des travaux en cours au laboratoire visent à purifier DDX3 à partir d'extraits cellulaires, pour caractériser les modifications post-traductionnelles qu'elle présente et la comparer à la protéine recombinante. En particulier, il est possible que les modifications post-traductionnelles modulent son affinité pour l'ARN ou pour des facteurs cellulaires tels que ceux que j'ai étudié dans ma thèse. Cependant les modifications post-traductionnelles sont susceptibles de varier au cours du cycle et suivant les compartiments cellulaires, ce qui rend cette étude complexe.

De plus, concernant l'effet de DDX3 sur la traduction de l'ARN génomique, les expériences présentées de traduction in vitro pourraient être poursuivies. Il serait notamment intéressant de reproduire ces expériences à plus forte concentration en DDX3, et de contrôler en parallèle l'état monomère ou dimère de l'ARN et l'interaction de Gag avec l'ARN. Si la traduction n'est pas stimulée alors que les dimères sont efficacement dissociés, on pourra conclure que l'état monomère ou dimère n'influe pas sur la traduction in vitro.

Enfin, l'étude mécanistique des hélicases à boîte DEAD est limitée, entre autres, par leur faible activité in vitro. Par exemple, peu d'études de FRET en molécule unique ont pu être menées sur des hélicases à boîte DEAD. Récemment, des auteurs ont proposé un système expérimental reposant sur des substrats modèles permettant d'étudier Ded1 au moyen de cet outil (Raj et al., 2019). Ma thèse met en évidence l'existence de substrats biologiques préférentiels qui stimulent fortement l'activité de DDX3. En particulier, j'ai observé une activité hélicase sans précédent en présence de dimères de 5'UTR. Dans ces conditions de turn-over multiple, il serait envisageable de concevoir une expérience de FRET pour détailler le mécanisme sous-jacent.

Annexes

1) La technique de SHAPE

La technique de SHAPE (Selective Hydroxyl-acylation Analyzed by Primer Extension) permet de mesurer la réactivité de chaque nucléotide vis-à-vis d'une molécule susceptible d'acyler le résidu 2'OH du ribose (Figure 58). La conformation adoptée par le ribose des nucléotides non appariés le rend plus réactif. Une fois modifié, l'ARN est rétrotranscrit. Les modifications stoppent l'élongation d'une transciptase inverse et peuvent donc être quantifiées au nucléotide près par extension d'amorce. En parallèle, un ARN non traité est rétrotranscrit pour être utilisé comme témoin et indiquer les arrêts spontanés de la transcriptase inverse (Merino et al., 2005).

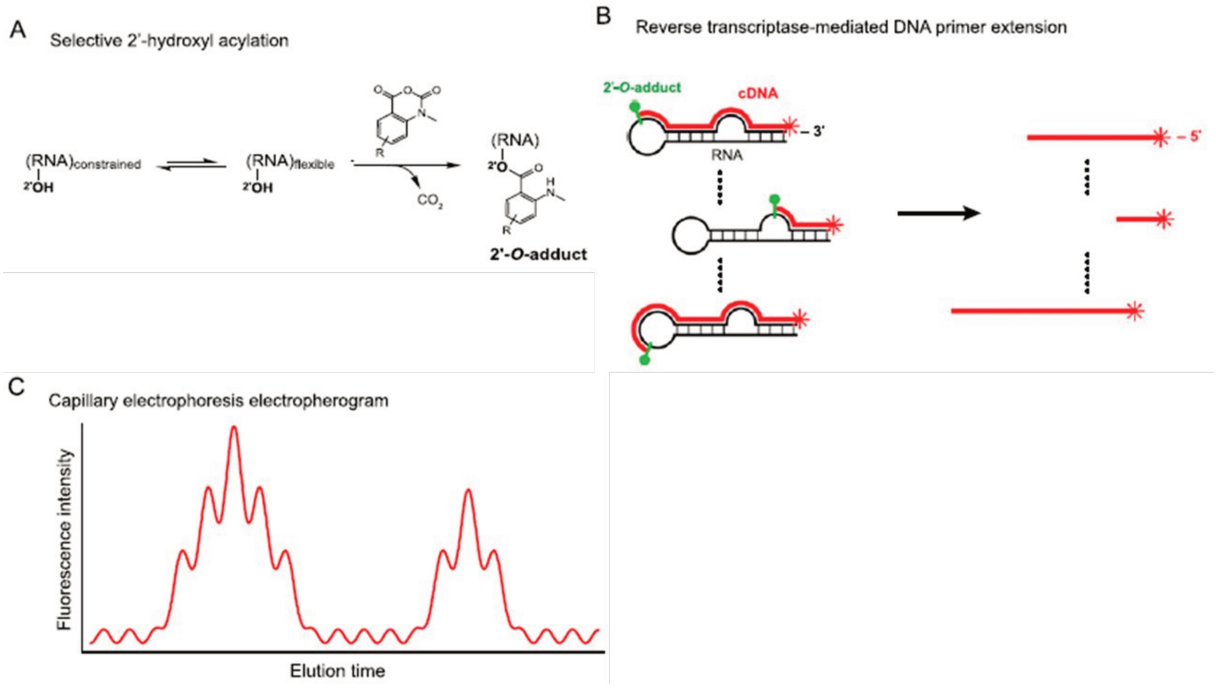


Figure 58 : ***Principe d'une expérience de SHAPE*** (Karabiber et al., 2013). *A/ L'ARN est modifié au moyen d'une molécule susceptible d'acyler le résidu 2'OH des nucléotides* non *appariés. B/ L'extension d'amorce fluorescentes génère des fragments d'ADNc de taille variable. C/ Les fragments d'ADNc sont quantifiés par électrophorèse capillaire.*

J'ai suivi essentiellement le protocole mis au point par K. Weeks (Wilkinson et al., 2006). J'ai utilisé deux sondes SHAPE différentes, le 1M7 (anhydride 1-méthyl-7-nitroisatoïque) et le BzCN (cyanure de benzoyle) qui réagissent indistinctement avec les quatre ribonucléotides. Ces deux réactifs diffèrent par leur demi-vie, celle du BzCN (0,25 s) étant nettement plus courte que celle du 1M7 (14 s) (Weeks and Mauger, 2011). La quantité

nécessaire d'ARN est de 3 pmol, pour une concentration d'ARN de 50 nM ou 200 nM dans mes expériences. Le 1M7 est utilisé à une concentration finale de 2 mM et le BzCN à une concentration de 40 mM (Mortimer and Weeks, 2009). L'ARN témoin est traité par un volume équivalent de DMSO. Une fois modifié, l'ARN est précipité puis resuspendu dans l'eau, et rétrotranscrit au moyen de la transcriptase inverse du virus de la leucémie murine de Moloney à partir d'amorces fluorescentes. Les fragments d'ADNc sont analysés par électrophorèse capillaire. Les résultats sont traités au moyen du logiciel QuShape (Karabiber et al., 2013) afin de mesurer l'intensité de fluorescence associée à chaque position de la séquence. Pour chaque nucléotide, la réactivité non normalisée est la différence entre les intensités de fluorescence obtenues avec et sans traitement par un réactif SHAPE. 10% des nucléotides correspondant aux fragments les plus abondants dans la condition témoin sont considérés comme non analysables. Les réactivités les plus élevées situées entre le 90ème et le 98ème centile sont moyennées pour fournir le terme de normalisation, les 2% des réactivités les plus élevées étant considérées trop variables. Les expériences de SHAPE sont généralement répliquées trois fois de façon indépendante.

2) Analyse de la dimérisation par électrophorèse en conditions natives

L'électrophorèse en conditions natives permet d'identifier différentes conformations adoptées par un ARN ainsi que des complexes ARN – ARN et ARN – protéine. Les complexes migrent plus ou moins vite sous l'effet du champ électrique suivant leur taille et la taille des pores du gel (Petrov et al., 2013).

Pour analyser la formation des dimères d'ARN du VIH-1, j'ai utilisé des mini-gels d'acrylamide – bisacrylamide (ratio 19:1) à une concentration volumique de 4%. Le tampon de migration 10X est composé de 340 mM de Tris, 570 mM de HEPES, 1 mM d'EDTA et 25 mM de $MgCl_2$ (tampon THEM). Le montage de la cuve d'électrophorèse est effectué à l'avance et placé à 4°C. Avant le dépôt, les échantillons sont mélangés à une solution de dépôt 5X (20% de glycérol, 0,1% de bleu de bromophénol et 0,1% de xylène cyanole pour suivre la migration) et placés sur de la glace. La migration est effectuée à 100V, soit 2,5W, pendant une heure. Les gels sont ensuite fixés dans une solution contenant 10% d'acide acétique et 30% d'éthanol pendant 15 minutes, puis séchés sous vide à 70°C. Les ARN marqués radioactivement à l'aide d'UTP-α^{32}P sont révélés en exposant les gels sur un écran radioluminescent à mémoire. La quantification des bandes correspondant au monomère et au dimère et la soustraction du bruit de fond sont effectuée grâce au logiciel MultiGauge v3.0 (Fujifilm).

Références bibliographiques

Abaeva, I.S., Marintchev, A., Pisareva, V.P., Hellen, C.U.T., and Pestova, T.V. (2011). Bypassing of stems versus linear base-by-base inspection of mammalian mRNAs during ribosomal scanning. EMBO J. 30, 115–129.

Abaeva, I.S., Pestova, T.V., and Hellen, C.U.T. (2016). Attachment of ribosomal complexes and retrograde scanning during initiation on the Halastavi árva virus IRES. Nucleic Acids Res. 44, 2362–2377.

Abbink, T.E.M., and Berkhout, B. (2003). A Novel Long Distance Base-pairing Interaction in Human Immunodeficiency Virus Type 1 RNA Occludes the Gag Start Codon. J. Biol. Chem. 278, 11601–11611.

Abbink, T.E.M., Ooms, M., Haasnoot, P.C.J., and Berkhout, B. (2005). The HIV-1 Leader RNA Conformational Switch Regulates RNA Dimerization but Does Not Regulate mRNA Translation. Biochemistry (Mosc.) 44, 9058–9066.

Abd El-Wahab, E.W., Smyth, R.P., Mailler, E., Bernacchi, S., Vivet-Boudou, V., Hijnen, M., Jossinet, F., Mak, J., Paillart, J.-C., and Marquet, R. (2014). Specific recognition of the HIV-1 genomic RNA by the Gag precursor. Nat. Commun. 5, 4304.

Adachi, A., Gendelman, H.E., Koenig, S., Folks, T., Willey, R., Rabson, A., and Martin, M.A. (1986). Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J. Virol. 59, 284–291.

Aduri, R., Briggs, K.T., Gorelick, R.J., and Marino, J.P. (2013). Molecular determinants of HIV-1 NCp7 chaperone activity in maturation of the HIV-1 dimerization initiation site. Nucleic Acids Res. 41, 2565–2580.

Ajamian, L., Abel, K., Rao, S., Vyboh, K., García-de-Gracia, F., Soto-Rifo, R., Kulozik, A.E., Gehring, N.H., and Mouland, A.J. (2015). HIV-1 Recruits UPF1 but Excludes UPF2 to Promote Nucleocytoplasmic Export of the Genomic RNA. Biomolecules 5, 2808–2839.

Ali, I.K., McKendrick, L., Morley, S.J., and Jackson, R.J. (2001). Truncated initiation factor eIF4G lacking an eIF4E binding site can support capped mRNA translation. EMBO J. 20, 4233–4242.

Andersen, E.S., Contera, S.A., Knudsen, B., Damgaard, C.K., Besenbacher, F., and Kjems, J. (2004). Role of the Trans-activation Response Element in Dimerization of HIV-1 RNA. J. Biol. Chem. 279, 22243–22249.

Anderson, E.C., and Lever, A.M.L. (2006). Human Immunodeficiency Virus Type 1 Gag Polyprotein Modulates Its Own Translation. J. Virol. 80, 10478–10486.

Ariumi, Y. (2014). Multiple functions of DDX3 RNA helicase in gene regulation, tumorigenesis, and viral infection. Front. Genet. 5.

Awang, G., and Sen, D. (1993). Mode of dimerization of HIV-1 genomic RNA. Biochemistry (Mosc.) 32, 11453–11457.

Baba, S., Takahashi, K., Noguchi, S., Takaku, H., Koyanagi, Y., Yamamoto, N., and Kawai, G. (2005). Solution RNA Structures of the HIV-1 Dimerization Initiation Site in the Kissing-Loop and Extended-Duplex Dimers. J. Biochem. (Tokyo) 138, 583–592.

Basu, V.P., Song, M., Gao, L., Rigby, S.T., Hanson, M.N., and Bambara, R.A. (2008). Strand transfer events during HIV-1 reverse transcription. Virus Res. 134, 19–38.

Batisse, J., Guerrero, S., Bernacchi, S., Sleiman, D., Gabus, C., Darlix, J.-L., Marquet, R., Tisné, C., and Paillart, J.-C. (2012). The role of Vif oligomerization and RNA chaperone activity in HIV-1 replication. Virus Res. 169, 361–376.

Baudin, F., Marquet, R., Isel, C., Darlix, J.-L., Ehresmann, B., and 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.

Berkhout, B., and Jeang, K.T. (1989). trans activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the trans-acting-responsive hairpin: a quantitative analysis. J. Virol. 63, 5501–5504.

Berkowitz, R.D., Luban, J., and Goff, S.P. (1993). Specific binding of human immunodeficiency virus type 1 gag polyprotein and nucleocapsid protein to viral RNAs detected by RNA mobility shift assays. J. Virol. 67, 7190–7200.

Bindewald, E., Wendeler, M., Legiewicz, M., Bona, M.K., Wang, Y., Pritt, M.J., Grice, S.F.J.L., and Shapiro, B.A. (2011). Correlating SHAPE signatures with three-dimensional RNA structures. RNA 17, 1688–1696.

Boeras, I., Seufzer, B., Brady, S., Rendahl, A., Heng, X., and Boris-Lawrie, K. (2017). The basal translation rate of authentic HIV-1 RNA is regulated by 5'UTR nt-pairings at junction of R and U5. Sci. Rep. 7.

Bol, G.M., Raman, V., Groep, P. van der, Vermeulen, J.F., Patel, A.H., Wall, E. van der, and Diest, P.J. van (2013). Expression of the RNA Helicase DDX3 and the Hypoxia Response in Breast Cancer. PLOS ONE 8, e63548.

Bol, G.M., Xie, M., and Raman, V. (2015a). DDX3, a potential target for cancer treatment. Mol. Cancer 14.

Bol, G.M., Vesuna, F., Xie, M., Zeng, J., Aziz, K., Gandhi, N., Levine, A., Irving, A., Korz, D., Tantravedi, S., et al. (2015b). Targeting DDX3 with a small molecule inhibitor for lung cancer therapy. EMBO Mol. Med. 7, 648–669.

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

Bono, F., Ebert, J., Lorentzen, E., and Conti, E. (2006). The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on mRNA. Cell 126, 713–725.

Botlagunta, M., Vesuna, F., Mironchik, Y., Raman, A., Lisok, A., Winnard Jr, P., Mukadam, S., Van Diest, P., Chen, J.H., Farabaugh, P., et al. (2008). Oncogenic role of DDX3 in breast cancer biogenesis. Oncogene 27, 3912–3922.

Botlagunta, M., Krishnamachary, B., Vesuna, F., Jr, P.T.W., Bol, G.M., Patel, A.H., and Raman, V. (2011). Expression of DDX3 Is Directly Modulated by Hypoxia Inducible Factor-1 Alpha in Breast Epithelial Cells. PLOS ONE 6, e17563.

Bowers, H.A., Maroney, P.A., Fairman, M.E., Kastner, B., Lührmann, R., Nilsen, T.W., and Jankowsky, E. (2006). Discriminatory RNP remodeling by the DEAD-box protein DED1. RNA 12, 903–912.

Brasey, A., Lopez-Lastra, M., Ohlmann, T., Beerens, N., Berkhout, B., Darlix, J.-L., and 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.

Brennan, R., Haap-Hoff, A., Gu, L., Gautier, V., Long, A., and Schröder, M. (2018). Investigating nucleo-cytoplasmic shuttling of the human DEAD-box helicase DDX3. Eur. J. Cell Biol.

de Breyne, S., and Ohlmann, T. (2019). Focus on Translation Initiation of the HIV-1 mRNAs. Int. J. Mol. Sci. 20, 101.

Briggs, J.A.G., Grünewald, K., Glass, B., Förster, F., Kräusslich, H.-G., and Fuller, S.D. (2006). The Mechanism of HIV-1 Core Assembly: Insights from Three-Dimensional Reconstructions of Authentic Virions. Structure 14, 15–20.

Buck, C.B., Shen, X., Egan, M.A., Pierson, T.C., Walker, C.M., and Siliciano, R.F. (2001). The Human Immunodeficiency Virus Type 1gag Gene Encodes an Internal Ribosome Entry Site. J. Virol. 75, 181–191.

Burd, C.G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science 265, 615–621.

Busa, V.F., Rector, M.J., and Russell, R. (2017). The DEAD-Box Protein CYT-19 Uses Arginine Residues in Its C-Tail To Tether RNA Substrates. Biochemistry (Mosc.) 56, 3571–3578.

Cantero-Camacho, Á., and Gallego, J. (2015). The conserved 3′X terminal domain of hepatitis C virus genomic RNA forms a two-stem structure that promotes viral RNA dimerization. Nucleic Acids Res. 43, 8529–8539.

Cantero-Camacho, Á., Fan, L., Wang, Y.-X., and Gallego, J. (2017). Three-dimensional structure of the 3′X-tail of hepatitis C virus RNA in monomeric and dimeric states. RNA 23, 1465–1476.

Castelló, A., Franco, D., Moral-López, P., Berlanga, J.J., Álvarez, E., Wimmer, E., and Carrasco, L. (2009). HIV- 1 Protease Inhibits Cap- and Poly(A)-Dependent Translation upon eIF4GI and PABP Cleavage. PLOS ONE 4, e7997.

Chang, Y.N., Kenan, D.J., Keene, J.D., Gatignol, A., and Jeang, K.T. (1994). Direct interactions between autoantigen La and human immunodeficiency virus leader RNA. J. Virol. 68, 7008–7020.

Chao, C.-H., Chen, C.-M., Cheng, P.-L., Shih, J.-W., Tsou, A.-P., and Lee, Y.-H.W. (2006). DDX3, a DEAD Box RNA Helicase with Tumor Growth–Suppressive Property and Transcriptional Regulation Activity of the p21waf1/cip1 Promoter, Is a Candidate Tumor Suppressor. Cancer Res. 66, 6579–6588.

Chen, H.-H., Yu, H.-I., Cho, W.-C., and Tarn, W.-Y. (2015). DDX3 modulates cell adhesion and motility and cancer cell metastasis via Rac1-mediated signaling pathway. Oncogene 34, 2790–2800.

Chen, H.-H., Yu, H.-I., Yang, M.-H., and Tarn, W.-Y. (2018). DDX3 activates CBC-eIF3-mediated translation of uORF-containing oncogenic mRNAs to promote metastasis in HNSCC. Cancer Res. canres.0282.2018.

Chen, J., Rahman, S.A., Nikolaitchik, O.A., Grunwald, D., Sardo, L., Burdick, R.C., Plisov, S., Liang, E., Tai, S., Pathak, V.K., et al. (2016). HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. Proc. Natl. Acad. Sci. 113, E201–E208.

Chen, W.-J., Wang, W.-T., Tsai, T.-Y., Li, H.-K., and Lee, Y.-H.W. (2017). DDX3 localizes to the centrosome and prevents multipolar mitosis by epigenetically and translationally modulating p53 expression. Sci. Rep. 7, 9411.

Chillón, I., Marcia, M., Legiewicz, M., Liu, F., Somarowthu, S., and Pyle, A.M. (2015). Chapter One - Native Purification and Analysis of Long RNAs. In Methods in Enzymology, S.A. Woodson, and F.H.T. Allain, eds. (Academic Press), pp. 3–37.

Chin, M.P.S., Rhodes, T.D., Chen, J., Fu, W., and Hu, W.-S. (2005). Identification of a major restriction in HIV-1 intersubtype recombination. Proc. Natl. Acad. Sci. 102, 9002–9007.

Chiu, Y.-L., Ho, C.K., Saha, N., Schwer, B., Shuman, S., and Rana, T.M. (2002). Tat Stimulates Cotranscriptional Capping of HIV mRNA. Mol. Cell 10, 585–597.

Choi, Y.-J., and Lee, S.-G. (2012). The DEAD-box RNA helicase DDX3 interacts with DDX5, co-localizes with it in the cytoplasm during the G2/M phase of the cycle, and affects its shuttling during mRNP export. J. Cell. Biochem. 113, 985–996.

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

Clever, J.L., Wong, M.L., and Parslow, T.G. (1996). Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA. J. Virol. 70, 5902–5908.

Collins, R., Karlberg, T., Lehtiö, L., Schütz, P., Berg, S. van den, Dahlgren, L.-G., Hammarström, M., Weigelt, J., and Schüler, H. (2009). The DEXD/H-box RNA Helicase DDX19 Is Regulated by an α-Helical Switch. J. Biol. Chem. 284, 10296–10300.

Comandur, R., Olson, E.D., and Musier-Forsyth, K. (2017). Conservation of tRNA mimicry in the 5′-untranslated region of distinct HIV-1 subtypes. RNA 23, 1850–1859.

Comas-Garcia, M., Davis, S., Rein, A., Comas-Garcia, M., Davis, S.R., and Rein, A. (2016). On the Selective Packaging of Genomic RNA by HIV-1. Viruses 8, 246.

Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006). The DEAD-box protein family of RNA helicases. Gene 367, 17–37.

Craig, A.W., Svitkin, Y.V., Lee, H.S., Belsham, G.J., and Sonenberg, N. (1997). The La autoantigen contains a dimerization domain that is essential for enhancing translation. Mol. Cell. Biol. 17, 163–169.

Cruciat, C.-M., Dolde, C., Groot, R.E.A. de, Ohkawara, B., Reinhard, C., Korswagen, H.C., and Niehrs, C. (2013). RNA Helicase DDX3 Is a Regulatory Subunit of Casein Kinase 1 in Wnt–β-Catenin Signaling. Science 339, 1436–1441.

Cullen, B.R. (2003). Nuclear mRNA export: insights from virology. Trends Biochem. Sci. 28, 419–424.

Darlix, J.-L., Gabus, C., Nugeyre, M.-T., Clavel, F., and Barré-Sinoussi, F. (1990). Cis elements and Trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. J. Mol. Biol. 216, 689–699.

Datta, S.A.K., Clark, P.K., Fan, L., Ma, B., Harvin, D.P., Sowder, R.C., Nussinov, R., Wang, Y.-X., and Rein, A. (2016). Dimerization of the SP1 Region of HIV-1 Gag Induces a Helical Conformation and Association into Helical Bundles: Implications for Particle Assembly. J. Virol. 90, 1773–1787.

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

Deforges, J., de Breyne, S., Ameur, M., Ulryck, N., Chamond, N., Saaidi, A., Ponty, Y., Ohlmann, T., and Sargueil, B. (2017). Two ribosome recruitment sites direct multiple translation events within HIV1 Gag open reading frame. Nucleic Acids Res. 45, 7382–7400.

Del Campo, M., and Lambowitz, A.M. (2009). Structure of the Yeast DEAD Box Protein Mss116p Reveals Two Wedges that Crimp RNA. Mol. Cell 35, 598–609.

Dhote, V., Sweeney, T.R., Kim, N., Hellen, C.U.T., and Pestova, T.V. (2012). Roles of individual domains in the function of DHX29, an essential factor required for translation of structured mammalian mRNAs. Proc. Natl. Acad. Sci. 109, E3150–E3159.

Dorin, D., Bonnet, M.C., Bannwarth, S., Gatignol, A., Meurs, E.F., and Vaquero, C. (2003). The TAR RNA-binding Protein, TRBP, Stimulates the Expression of TAR-containing RNAs in Vitro and in VivoIndependently of Its Ability to Inhibit the dsRNA-dependent Kinase PKR. J. Biol. Chem. 278, 4440–4448.

Dubois, N., Marquet, R., Paillart, J.-C., and Bernacchi, S. (2018a). Retroviral RNA Dimerization: From Structure to Functions. Front. Microbiol. 9.

Dubois, N., Khoo, K.K., Ghossein, S., Seissler, T., Wolff, P., McKinstry, W.J., Mak, J., Paillart, J.-C., Marquet, R., and Bernacchi, S. (2018b). The C-terminal p6 domain of the HIV-1 Pr55Gag precursor is required for specific binding to the genomic RNA. RNA Biol. 0, 1–14.

Duesberg, P.H. (1968). Physical properties of Rous Sarcoma Virus RNA. Proc. Natl. Acad. Sci. U. S. A. 60, 1511–1518.

Dugré-Brisson, S., Elvira, G., Boulay, K., Chatel-Chaix, L., Mouland, A.J., and DesGroseillers, L. (2005). Interaction of Staufen1 with the 5′ end of mRNA facilitates translation of these RNAs. Nucleic Acids Res. 33, 4797–4812.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. 95, 14863–14868.

Emery, A., Zhou, S., Pollom, E., and Swanstrom, R. (2017). Characterizing HIV-1 Splicing by Using Next-Generation Sequencing. J. Virol. 91.

Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C., and 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., and Dumas, P. (2001). Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. Nat. Struct. Mol. Biol. 8, 1064–1068.

Epling, L.B., Grace, C.R., Lowe, B.R., Partridge, J.F., and Enemark, E.J. (2015). Cancer-Associated Mutants of RNA Helicase DDX3X Are Defective in RNA-Stimulated ATP Hydrolysis. J. Mol. Biol. 427, 1779–1796.

Fairman, M.E., Maroney, P.A., Wang, W., Bowers, H.A., Gollnick, P., Nilsen, T.W., and Jankowsky, E. (2004). Protein Displacement by DExH/D “RNA Helicases” Without Duplex Unwinding. Science 304, 730–734.

Fairman-Williams, M.E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. Curr. Opin. Struct. Biol. 20, 313–324.

Feng, S., and Holland, E.C. (1988). HIV-1 tat trans- activation requires the loop sequence within tar. Nature 334, 165.

Fischer, U., Huber, J., Boelens, W.C., Mattajt, L.W., and Lührmann, R. (1995). The HIV-1 Rev Activation Domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. Cell 82, 475–483.

Floor, S.N., Condon, K.J., Sharma, D., Jankowsky, E., and Doudna, J.A. (2015). Autoinhibitory Interdomain Interactions and Subfamily-Specific Extensions Redefine the Catalytic Core of the Human DEAD-box Protein DDX3. J. Biol. Chem. jbc.M115.700625.

Franca, R., Belfiore, A., Spadari, S., and Maga, G. (2007). Human DEAD-box ATPase DDX3 shows a relaxed nucleoside substrate specificity. Proteins Struct. Funct. Bioinforma. 67, 1128–1137.

Frankel, A.D., and Young, J.A. (1998). HIV-1: fifteen proteins and an RNA. Annu. Rev. Biochem. 67, 1–25.

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., Valiente-Echeverría, F., et al. (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 BBA - Gene Regul. Mech. 1859, 719–730.

Fu, W., Gorelick, R.J., and Rein, A. (1994). Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. J. Virol. 68, 5013–5018.

Fuller-Pace, F.V. (2013). DEAD box RNA helicase functions in cancer. RNA Biol. 10, 121–132.

Gao, J., Byrd, A.K., Zybailov, B.L., Marecki, J.C., Guderyon, M.J., Edwards, A.D., Chib, S., West, K.L., Waldrip, Z.J., Mackintosh, S.G., et al. (2019). DEAD-box RNA helicases Dbp2, Ded1 and Mss116 bind to G-quadruplex nucleic acids and destabilize G-quadruplex RNA. Chem. Commun. 55, 4467–4470.

Garbelli, A., Beermann, S., Cicco, G.D., Dietrich, U., and Maga, G. (2011). 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.

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

Gendron, K., Ferbeyre, G., Heveker, N., and 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, W.V. (2010). Alternative Ways to Think about Cellular Internal Ribosome Entry. J. Biol. Chem. 285, 29033–29038.

Gringhuis, S.I., Hertoghs, N., Kaptein, T.M., Zijlstra-Willems, E.M., Sarrami-Fooroshani, R., Sprokholt, J.K., van Teijlingen, N.H., Kootstra, N.A., Booiman, T., van Dort, K.A., et al. (2016). HIV-1 blocks the signaling adaptor MAVS to evade antiviral host defense after sensing of abortive HIV-1 RNA by the host helicase DDX3. Nat. Immunol.

Grohman, J.K., Gorelick, R.J., Lickwar, C.R., Lieb, J.D., Bower, B.D., Znosko, B.M., and Weeks, K.M. (2013). A Guanosine-Centric Mechanism for RNA Chaperone Function. Science 340, 190–195.

Gu, L., Fullam, A., Brennan, R., and Schröder, M. (2013). Human DEAD Box Helicase 3 Couples IκB Kinase ε to Interferon Regulatory Factor 3 Activation. Mol. Cell. Biol. 33, 2004–2015.

Guenther, U.-P., Weinberg, D.E., Zubradt, M.M., Tedeschi, F.A., Stawicki, B.N., Zagore, L.L., Brar, G.A., Licatalosi, D.D., Bartel, D.P., Weissman, J.S., et al. (2018). The helicase Ded1p controls use of near-cognate translation initiation codons in 5′ UTRs. Nature.

Guo, J.U., and Bartel, D.P. (2016). RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. Science 353, aaf5371.

Halls, C., Mohr, S., Del Campo, M., Yang, Q., Jankowsky, E., and Lambowitz, A.M. (2007). Involvement of DEAD-box Proteins in Group I and Group II Intron Splicing. Biochemical Characterization of Mss116p, ATP Hydrolysis-dependent and -independent Mechanisms, and General RNA Chaperone Activity. J. Mol. Biol. 365, 835–855.

Hauk, G., and Bowman, G.D. (2015). Formation of a Trimeric Xpo1-Ran[GTP]-Ded1 Exportin Complex Modulates ATPase and Helicase Activities of Ded1. PLOS ONE 10, e0131690.

Heng, X., Kharytonchyk, S., Garcia, E.L., Lu, K., Divakaruni, S.S., LaCotti, C., Edme, K., Telesnitsky, A., and Summers, M.F. (2012). Identification of a Minimal Region of the HIV-1 5′-Leader Required for RNA Dimerization, NC Binding, and Packaging. J. Mol. Biol. 417, 224–239.

Henn, A., Bradley, M.J., and De La Cruz, E.M. (2012). ATP Utilization and RNA Conformational Rearrangement by DEAD-Box Proteins. Annu. Rev. Biophys. 41, 247–267.

Henriet, S., Sinck, L., Bec, G., Gorelick, R.J., Marquet, R., and Paillart, J.-C. (2007). Vif is a RNA chaperone that could temporally regulate RNA dimerization and the early steps of HIV-1 reverse transcription. Nucleic Acids Res. 35, 5141–5153.

Herdy, B., Mayer, C., Varshney, D., Marsico, G., Murat, P., Taylor, C., D'Santos, C., Tannahill, D., and Balasubramanian, S. (2018). Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts. Nucleic Acids Res.

Herschlag, D. (1995). RNA Chaperones and the RNA Folding Problem. J. Biol. Chem. 270, 20871–20874.

Hilliker, A., Gao, Z., Jankowsky, E., and 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.

Högbom, M., Collins, R., van den Berg, S., Jenvert, R.-M., Karlberg, T., Kotenyova, T., Flores, A., Hedestam, G.B.K., and 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. Biol. 372, 150–159.

Höglund, S., öhagen, Å., Goncalves, J., Panganiban, A.T., and Gabuzda, D. (1997). Ultrastructure of HIV-1 Genomic RNA. Virology 233, 271–279.

Houzet, L., Paillart, J.C., Smagulova, F., Maurel, S., Morichaud, Z., Marquet, R., and Mougel, M. (2007). HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. Nucleic Acids Res. 35, 2695–2704.

Huthoff, H., and Berkhout, B. (2001). Two alternating structures of the HIV-1 leader RNA. RNA 7, 143–157.

Huthoff, H., and Berkhout, B. (2002). Multiple Secondary Structure Rearrangements during HIV-1 RNA Dimerization. Biochemistry (Mosc.) 41, 10439–10445.

Jablonski, J.A., Buratti, E., Stuani, C., and Caputi, M. (2008). The Secondary Structure of the Human Immunodeficiency Virus Type 1 Transcript Modulates Viral Splicing and Infectivity. J. Virol. 82, 8038–8050.

Jablonski, J.A., Amelio, A.L., Giacca, M., and Caputi, M. (2010). The transcriptional transactivator Tat selectively regulates viral splicing. Nucleic Acids Res. 38, 1249–1260.

Jackson, R.J., Hellen, C.U.T., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. 11, 113–127.

Jalalirad, M., Saadatmand, J., and Laughrea, M. (2012). Dominant Role of the 5′ TAR Bulge in Dimerization of HIV-1 Genomic RNA, but No Evidence of TAR–TAR Kissing during in Vivo Virus Assembly. Biochemistry (Mosc.) 51, 3744–3758.

Jankowsky, E., and Fairman, M.E. (2007). RNA helicases — one fold for many functions. Curr. Opin. Struct. Biol. 17, 316–324.

Jankowsky, A., Guenther, U.-P., and Jankowsky, E. (2011). The RNA helicase database. Nucleic Acids Res. 39, D338–D341.

Jeong, W.-J., Ro, E.J., and Choi, K.-Y. (2018). Interaction between Wnt/β-catenin and RAS-ERK pathways and an anti-cancer strategy via degradations of β-catenin and RAS by targeting the Wnt/β-catenin pathway. Npj Precis. Oncol. 2, 5.

Jones, C.P., Saadatmand, J., Kleiman, L., and Musier-Forsyth, K. (2013). Molecular mimicry of human tRNALys anti-codon domain by HIV-1 RNA genome facilitates tRNA primer annealing. RNA 19, 219–229.

Jones, C.P., Cantara, W.A., Olson, E.D., and Musier-Forsyth, K. (2014). Small-angle X-ray scattering-derived structure of the HIV-1 5′ UTR reveals 3D tRNA mimicry. Proc. Natl. Acad. Sci. 111, 3395–3400.

Jossinet, F., Paillart, J.C., Westhof, E., Hermann, T., Skripkin, E., Lodmell, J.S., Ehresmann, C., Ehresmann, B., and Marquet, R. (1999). Dimerization of HIV-1 genomic RNA of subtypes A and B: RNA loop structure and magnesium binding. RNA 5, 1222–1234.

Kadaré, G., and Haenni, A.L. (1997). Virus-encoded RNA helicases. J. Virol. 71, 2583–2590.

Kao, S.-Y., Calman, A.F., Luciw, P.A., and Peterlin, B.M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature 330, 489.

Karabiber, F., McGinnis, J.L., Favorov, O.V., and Weeks, K.M. (2013). QuShape: Rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. RNA 19, 63–73.

Karlsson Hedestam, G.B., Fouchier, R.A.M., Phogat, S., Burton, D.R., Sodroski, J., and Wyatt, R.T. (2008). The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. Nat. Rev. Microbiol. 6, 143–155.

Karn, J., and Stoltzfus, C.M. (2012). Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. Cold Spring Harb. Perspect. Med. 2, a006916.

Kasim, V., Wu, S., Taira, K., and Miyagishi, M. (2013). Determination of the Role of DDX3 a Factor Involved in Mammalian RNAi Pathway Using an shRNA-Expression Library. PLOS ONE 8, e59445.

Keane, S.C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., Barton, S., Hosic, A., Florwick, A., Santos, J., et al. (2015). Structure of the HIV-1 RNA packaging signal. Science 348, 917–921.

Keane, S.C., Van, V., Frank, H.M., Sciandra, C.A., McCowin, S., Santos, J., Heng, X., and Summers, M.F. (2016). NMR detection of intermolecular interaction sites in the dimeric 5′-leader of the HIV-1 genome. Proc. Natl. Acad. Sci. 113, 13033–13038.

Kellaris, G., Khan, K., Baig, S.M., Tsai, I.-C., Zamora, F.M., Ruggieri, P., Natowicz, M.R., and Katsanis, N. (2018). A hypomorphic inherited pathogenic variant in DDX3X causes male intellectual disability with additional neurodevelopmental and neurodegenerative features. Hum. Genomics 12, 11.

Kennedy, E.M., Bogerd, H.P., Kornepati, A.V.R., Kang, D., Ghoshal, D., Marshall, J.B., Poling, B.C., Tsai, K., Gokhale, N.S., Horner, S.M., et al. (2016). Posttranscriptional m6A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. Cell Host Microbe 19, 675–685.

Kenyon, J.C., Prestwood, L.J., Grice, S.F.J.L., and 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–e174.

Kharytonchyk, S., Monti, S., Smaldino, P.J., Van, V., Bolden, N.C., Brown, J.D., Russo, E., Swanson, C., Shuey, A., Telesnitsky, A., et al. (2016). Transcriptional start site heterogeneity modulates the structure and function of the HIV-1 genome. Proc. Natl. Acad. Sci. 113, 13378–13383.

Kieken, F., Paquet, F., Brulé, F., Paoletti, J., and Lancelot, G. (2006). A new NMR solution structure of the SL1 HIV-1 Lai loop–loop dimer. Nucleic Acids Res. 34, 343–352.

Kim, T., and Shapiro, B.A. (2013). The role of salt concentration and magnesium binding in HIV-1 subtype-A and subtype-B kissing loop monomer structures. J. Biomol. Struct. Dyn. 31, 495–510.

Kubo, Y., Hayashi, H., Matsuyama, T., Sato, H., and Yamamoto, N. (2012). Retrovirus Entry by Endocytosis and Cathepsin Proteases.

Lai, M.-C., Lee, Y.-H.W., and Tarn, W.-Y. (2008). The DEAD-Box RNA Helicase DDX3 Associates with Export Messenger Ribonucleoproteins as well asTip-associated Protein and Participates in Translational Control. Mol. Biol. Cell 19, 3847–3858.

Lai, M.-C., Chang, W.-C., Shieh, S.-Y., and 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., and Sun, H.S. (2013). Human DDX3 Interacts with the HIV-1 Tat Protein to Facilitate Viral mRNA Translation. PLoS ONE 8, e68665.

Laughrea, M., and Jetté, L. (1994). A 19-nucleotide sequence upstream of the 5' major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. Biochemistry (Mosc.) 33, 13464–13474.

Laughrea, M., and Jetté, L. (1996). Kissing-Loop Model of HIV-1 Genome Dimerization: HIV-1 RNAs Can Assume Alternative Dimeric Forms, and All Sequences Upstream or Downstream of Hairpin 248−271 Are Dispensable for Dimer Formation. Biochemistry (Mosc.) 35, 1589–1598.

LeBlanc, J., Weil, J., and Beemon, K. (2013). Posttranscriptional regulation of retroviral gene expression: primary RNA transcripts play three roles as pre-mRNA, mRNA, and genomic RNA. Wiley Interdiscip. Rev. RNA 4, 567–580.

Lee, C.-S., Dias, A.P., Jedrychowski, M., Patel, A.H., Hsu, J.L., and Reed, R. (2008). Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. Nucleic Acids Res. 36, 4708–4718.

Lee, T., Pelletier, J., Lee, T., and Pelletier, J. (2016). The biology of DHX9 and its potential as a therapeutic target. Oncotarget 7, 42716–42739.

Lever, A., Gottlinger, H., Haseltine, W., and Sodroski, J. (1989). Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J. Virol. 63, 4085–4087.

Li, P.T.X., and Tinoco, I. (2009). Mechanical Unfolding of Two DIS RNA Kissing Complexes from HIV-1. J. Mol. Biol. 386, 1343–1356.

Lichinchi, G., Gao, S., Saletore, Y., Gonzalez, G.M., Bansal, V., Wang, Y., Mason, C.E., and Rana, T.M. (2016). Dynamics of the human and viral m^6A RNA methylomes during HIV-1 infection of T cells. Nat. Microbiol. 1, 16011.

Linder, P., and Fuller-Pace, F.V. (2013). Looking back on the birth of DEAD-box RNA helicases. Biochim. Biophys. Acta BBA - Gene Regul. Mech. 1829, 750–755.

Linder, P., and Jankowsky, E. (2011). From unwinding to clamping — the DEAD box RNA helicase family. Nat. Rev. Mol. Cell Biol. 12, 505–516.

Linder, P., Lasko, P.F., Ashburner, M., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J., and Slonimski, P.P. (1989). Birth of the D-E-A-D box. Nature 337, 121–122.

Liu, F., Putnam, A., and Jankowsky, E. (2008). ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. Proc. Natl. Acad. Sci. 105, 20209–20214.

Liu, F., Putnam, A.A., and Jankowsky, E. (2014). DEAD-Box Helicases Form Nucleotide-Dependent, Long-Lived Complexes with RNA. Biochemistry (Mosc.) 53, 423–433.

Liu, J., Henao-Mejia, J., Liu, H., Zhao, Y., and He, J.J. (2011). Translational Regulation of HIV-1 Replication by HIV-1 Rev Cellular Cofactors Sam68, eIF5A, hRIP, and DDX3. J. Neuroimmune Pharmacol. 6, 308–321.

Locker, N., Chamond, N., and 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.

Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E.L., Kharytonchyk, S., Dorjsuren, B., Kulandaivel, G., Jones, S., Hiremath, A., et al. (2011a). NMR detection of structures in the HIV-1 5′-leader RNA that regulate genome packaging. Science 334, 242–245.

Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E.L., Kharytonchyk, S., Dorjsuren, B., Kulandaivel, G., Jones, S., Hiremath, A., et al. (2011b). NMR Detection of Structures in the HIV-1 5′-Leader RNA That Regulate Genome Packaging. Science 334, 242–245.

Lu, K., Heng, X., and Summers, M.F. (2011c). Structural Determinants and Mechanism of HIV-1 Genome Packaging. J. Mol. Biol. 410, 609–633.

Lu, W., Tirumuru, N., Gelais, C.S., Koneru, P.C., Liu, C., Kvaratskhelia, M., He, C., and Wu, L. (2018). N6-Methyladenosine–binding proteins suppress HIV-1 infectivity and viral production. J. Biol. Chem. 293, 12992–13005.

Makutonina, A., Voss, T.G., Plymale, D.R., Fermin, C.D., Norris, C.H., Vigh, S., and Garry, R.F. (1996). Human immunodeficiency virus infection of T-lymphoblastoid cells reduces intracellular pH. J. Virol. 70, 7049–7055.

Malim, M.H., Hauber, J., Le, S.-Y., Maizel, J.V., and Cullen, B.R. (1989). The HIV-1 rev trans -activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature 338, 254.

Mallam, A.L., Sidote, D.J., and Lambowitz, A.M. (2014). Molecular insights into RNA and DNA helicase evolution from the determinants of specificity for a DEAD-box RNA helicase. ELife 3, e04630.

Marquet, R., Baudin, F., Gabus, C., Darlix, J.L., Mougel, M., Ehresmann, C., and Ehresmann, B. (1991). Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. Nucleic Acids Res. 19, 2349–2357.

Marquet, R., Paillart, J.C., Skripkin, E., Ehresmann, C., and 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.

Merino, E.J., Wilkinson, K.A., Coughlan, J.L., and Weeks, K.M. (2005). RNA Structure Analysis at Single Nucleotide Resolution by Selective 2‘-Hydroxyl Acylation and Primer Extension (SHAPE). J. Am. Chem. Soc. 127, 4223–4231.

Merz, C., Urlaub, H., Will, C.L., and Lührmann, R. (2007). Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. RNA 13, 116–128.

Mestre-Fos, S., Penev, P.I., Suttapitugsakul, S., Hu, M., Ito, C., Petrov, A.S., Wartell, R.M., Wu, R., and Williams, L.D. (2019). G-Quadruplexes in Human Ribosomal RNA. J. Mol. Biol. 431, 1940–1955.

Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.-B., and Jaffrey, S.R. (2015). 5′ UTR m6A Promotes Cap-Independent Translation. Cell 163, 999–1010.

Mihailescu, M.-R., and Marino, J.P. (2004). A proton-coupled dynamic conformational switch in the HIV-1 dimerization initiation site kissing complex. Proc. Natl. Acad. Sci. 101, 1189–1194.

von Moeller, H., Basquin, C., and Conti, E. (2009). The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. Nat. Struct. Mol. Biol. 16, 247–254.

Mohr, S., Stryker, J.M., and Lambowitz, A.M. (2002). A DEAD-Box Protein Functions as an ATP-Dependent RNA Chaperone in Group I Intron Splicing. Cell 109, 769–779.

Monette, A., Valiente-Echeverría, F., Rivero, M., Cohen, É.A., Lopez-Lastra, M., and Mouland, A.J. (2013). Dual Mechanisms of Translation Initiation of the Full-Length HIV-1 mRNA Contribute to Gag Synthesis. PLOS ONE 8, e68108.

Montpetit, B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M., and Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and $InsP_6$ in mRNA export. Nature 472, 238–242.

Moore, M.D., and Hu, W.-S. (2009). HIV-1 RNA Dimerization: It Takes Two to Tango. AIDS Rev. 11, 91–102.

Moore, A.F.T., Gentry, R.C., and Koculi, E. (2017). DbpA is a region-specific RNA helicase. Biopolymers 107.

Mortimer, S.A., and Weeks, K.M. (2009). Time-resolved RNA SHAPE chemistry: quantitative RNA structure analysis in one-second snapshots and at single-nucleotide resolution. Nat. Protoc. 4, 1413–1421.

Mueller, N., van Bel, N., Berkhout, B., and Das, A.T. (2014). HIV-1 splicing at the major splice donor site is restricted by RNA structure. Virology 468–470, 609–620.

Muesing, M.A., Smith, D.H., and Capon, D.J. (1987). Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. Cell 48, 691–701.

Mujeeb, A., Clever, J.L., Billeci, T.M., James, T.L., and Parslow, T.G. (1998). Structure of the dimer a initiation complex of HIV-1 genomic RNA. Nat. Struct. Biol. 5, 432.

Mujeeb, A., Parslow, T.G., Zarrinpar, A., Das, C., and James, T.L. (1999). NMR structure of the mature dimer initiation complex of HIV-1 genomic RNA. FEBS Lett. 458, 387–392.

Mundigala, H., Michaux, J.B., Feig, A.L., Ennifar, E., and Rueda, D. (2014). HIV-1 DIS stem loop forms an obligatory bent kissing intermediate in the dimerization pathway. Nucleic Acids Res. 42, 7281–7289.

Muriaux, D., Girard, P.-M., Bonnet-Mathonière, B., and Paoletti, J. (1995). Dimerization of HIV-1Lai RNA at Low Ionic Strength AN AUTOCOMPLEMENTARY SEQUENCE IN THE 5′ LEADER REGION IS EVIDENCED BY AN ANTISENSE OLIGONUCLEOTIDE. J. Biol. Chem. 270, 8209–8216.

Muriaux, D., Rocquigny, H.D., Roques, B.-P., and Paoletti, J. (1996a). NCp7 Activates HIV-1Lai RNA Dimerization by Converting a Transient Loop-Loop Complex into a Stable Dimer. J. Biol. Chem. 271, 33686–33692.

Muriaux, D., Fossé, P., and Paoletti, J. (1996b). A Kissing Complex Together with a Stable Dimer Is Involved in the HIV-1Lai RNA Dimerization Process in Vitro. Biochemistry (Mosc.) 35, 5075–5082.

Neville, M., Stutz, F., Lee, L., Davis, L.I., and Rosbash, M. (1997). The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. Curr. Biol. 7, 767–775.

Nielsen, K.H., Chamieh, H., Andersen, C.B.F., Fredslund, F., Hamborg, K., Hir, H.L., and Andersen, G.R. (2009). Mechanism of ATP turnover inhibition in the EJC. RNA 15, 67–75.

Ohlmann, T., Mengardi, C., and López-Lastra, M. (2014). Translation initiation of the HIV-1 mRNA. Translation 2, e960242.

Okamura, M., Inose, H., and Masuda, S. (2015). RNA Export through the NPC in Eukaryotes. Genes 6, 124–149.

Oshiumi, H., Ikeda, M., Matsumoto, M., Watanabe, A., Takeuchi, O., Akira, S., Kato, N., Shimotohno, K., and Seya, T. (2010). Hepatitis C Virus Core Protein Abrogates the DDX3 Function That Enhances IPS-1-Mediated IFN–Beta Induction. PLOS ONE 5, e14258.

Paillart, J.C., Marquet, R., Skripkin, E., Ehresmann, B., and Ehresmann, C. (1994). Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. J. Biol. Chem. 269, 27486–27493.

Paillart, J.C., Skripkin, E., Ehresmann, B., Ehresmann, C., and 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. 93, 5572–5577.

Paillart, J.-C., Westhof, E., Ehresmann, C., Ehresmann, B., and Marquet, R. (1997). Non-canonical interactions in a kissing loop complex: the dimerization initiation site of HIV-1 genomic RNA11Edited by J. Karn. J. Mol. Biol. 270, 36–49.

Paillart, J.-C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (2002). In Vitro Evidence for a Long Range Pseudoknot in the 5'-Untranslated and Matrix Coding Regions of HIV-1 Genomic RNA. J. Biol. Chem. 277, 5995–6004.

Paillart, J.-C., Shehu-Xhilaga, M., Marquet, R., and Mak, J. (2004a). Dimerization of retroviral RNA genomes: an inseparable pair. Nat. Rev. Microbiol. 2, 461–472.

Paillart, J.-C., Dettenhofer, M., Yu, X., Ehresmann, C., Ehresmann, B., and Marquet, R. (2004b). First Snapshots of the HIV-1 RNA Structure in Infected Cells and in Virions. J. Biol. Chem. 279, 48397–48403.

Pallesen, J. (2011). Structure of the HIV-1 5′ Untranslated Region Dimer Alone and in Complex with Gold Nanocolloids: Support of a TAR–TAR-Containing 5′ Dimer Linkage Site (DLS) and a 3′ DIS–DIS-Containing DLS. Biochemistry (Mosc.) 50, 6170–6177.

Parkin, N.T., Cohen, E.A., Darveau, A., Rosen, C., Haseltine, W., and 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.

Petrov, A., Tsa, A., and Puglisi, J.D. (2013). Analysis of RNA by Analytical Polyacrylamide Gel Electrophoresis. In Methods in Enzymology, (Elsevier), pp. 301–313.

Plank, T.-D.M., Whitehurst, J.T., and 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.

Pornillos, O., and Ganser-Pornillos, B.K. (2019). Maturation of retroviruses. Curr. Opin. Virol. 36, 47–55.

Putnam, A.A., and 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., and Jankowsky, E. (2015). Division of Labor in an Oligomer of the DEAD-Box RNA Helicase Ded1p. Mol. Cell 59, 541–552.

Raj, S., Bagchi, D., Orero, J.V., Banroques, J., Tanner, N.K., and Croquette, V. (2019). Mechanistic characterization of the DEAD-box RNA helicase Ded1 from yeast as revealed by a novel technique using single-molecule magnetic tweezers. Nucleic Acids Res. 47, 3699–3710.

Rauschendorf, M.-A., Zimmer, J., Ohnmacht, C., and Vogt, P.H. (2014). DDX3X, the X homologue of AZFa gene DDX3Y, expresses a complex pattern of transcript variants only in the male germ line. MHR Basic Sci. Reprod. Med. 20, 1208–1222.

Reyes-Darias, J.A., Sánchez-Luque, F.J., and Berzal-Herranz, A. (2012). HIV RNA dimerisation interference by antisense oligonucleotides targeted to the 5′ UTR structural elements. Virus Res. 169, 63–71.

Rhodes, T., Wargo, H., and Hu, W.-S. (2003). High Rates of Human Immunodeficiency Virus Type 1 Recombination: Near-Random Segregation of Markers One Kilobase Apart in One Round of Viral Replication. J. Virol. 77, 11193–11200.

Riquelme-Barrios, S., Pereira-Montecinos, C., Valiente-Echeverría, F., and Soto-Rifo, R. (2018). Emerging Roles of N6-Methyladenosine on HIV-1 RNA Metabolism and Viral Replication. Front. Microbiol. 9.

Rist, M.J., and Marino, J.P. (2002). Mechanism of Nucleocapsid Protein Catalyzed Structural Isomerization of the Dimerization Initiation Site of HIV-1. Biochemistry (Mosc.) 41, 14762–14770.

Robinson, W.S., Robinson, H.L., and Duesberg, P.H. (1967). Tumor virus RNA's. Proc. Natl. Acad. Sci. U. S. A. 58, 825–834.

Rodamilans, B., and Montoya, G. (2007). Expression, purification, crystallization and preliminary X-ray diffraction analysis of the DDX3 RNA helicase domain. Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun. 63, 283–286.

Rogers, G.W., Lima, W.F., and Merrick, W.C. (2001). Further Characterization of the Helicase Activity of eIF4A SUBSTRATE SPECIFICITY. J. Biol. Chem. 276, 12598–12608.

Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. Cell 169, 1187–1200.

Rudolph, M.G., and Klostermeier, D. (2015). When core competence is not enough: functional interplay of the DEAD-box helicase core with ancillary domains and auxiliary factors in RNA binding and unwinding. Biol. Chem. 396.

Sakuragi, J., Shioda, T., and Panganiban, A.T. (2001). Duplication of the Primary Encapsidation and Dimer Linkage Region of Human Immunodeficiency Virus Type 1 RNA Results in the Appearance of Monomeric RNA in Virions. J. Virol. 75, 2557–2565.

Samatanga, B., and Klostermeier, D. (2014). DEAD-box RNA helicase domains exhibit a continuum between complete functional independence and high thermodynamic coupling in nucleotide and RNA duplex recognition. Nucleic Acids Res. 42, 10644–10654.

Schroeder, R., Barta, A., and Semrad, K. (2004). Strategies for RNA folding and assembly. Nat. Rev. Mol. Cell Biol. 5, 908–918.

Seif, E., Niu, M., and Kleiman, L. (2013). Annealing to sequences within the primer binding site loop promotes an HIV-1 RNA conformation favoring RNA dimerization and packaging. RNA 19, 1384–1393.

Sekiguchi, T., Iida, H., Fukumura, J., and Nishimoto, T. (2004). Human DDX3Y, the Y-encoded isoform of RNA helicase DDX3, rescues a hamster temperature-sensitive ET24 mutant cell line with a DDX3X mutation. Exp. Cell Res. 300, 213–222.

Sekiguchi, T., Kurihara, Y., and Fukumura, J. (2007). Phosphorylation of threonine 204 of DEAD-box RNA helicase DDX3 by cyclin B/cdc2 in vitro. Biochem. Biophys. Res. Commun. 356, 668–673.

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., and Yokoyama, S. (2006). Structural Basis for RNA Unwinding by the DEAD-Box Protein Drosophila Vasa. Cell 125, 287–300.

Sharma, D., and Jankowsky, E. (2014). The Ded1/DDX3 subfamily of DEAD-box RNA helicases. Crit. Rev. Biochem. Mol. Biol. 49, 343–360.

Sharma, A., Yilmaz, A., Marsh, K., Cochrane, A., and 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., Putnam, A.A., and Jankowsky, E. (2017). Biochemical Differences and Similarities between the DEAD-Box Helicase Orthologs DDX3X and Ded1p. J. Mol. Biol. 429, 3730–3742.

Shehu-Xhilaga, M., Kraeusslich, H.G., Pettit, S., Swanstrom, R., Lee, J.Y., Marshall, J.A., Crowe, S.M., and Mak, J. (2001). Proteolytic Processing of the P2/Nucleocapsid Cleavage Site Is Critical for Human Immunodeficiency Virus Type 1 RNA Dimer Maturation. J. Virol. 75, 9156–9164.

Shetty, S., Kim, S., Shimakami, T., Lemon, S.M., and Mihailescu, M.-R. (2010). Hepatitis C virus genomic RNA dimerization is mediated via a kissing complex intermediate. RNA 16, 913–925.

Shih, J.-W., Tsai, T.-Y., Chao, C.-H., and 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.

Silvers, R., Keller, H., Schwalbe, H., and Hengesbach, M. (2015). Differential Scanning Fluorimetry for Monitoring RNA Stability. ChemBioChem 16, 1109–1114.

Sinck, L., Richer, D., Howard, J., Alexander, M., Purcell, D.F.J., Marquet, R., and Paillart, J.-C. (2007). In vitro dimerization of human immunodeficiency virus type 1 (HIV-1) spliced RNAs. RNA 13, 2141–2150.

Skripkin, E., Paillart, J.C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994). Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. Proc. Natl. Acad. Sci. 91, 4945–4949.

Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F., and Haseltine, W. (1985). Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. Science 229, 74–77.

Song, R., Kafaie, J., Yang, L., and Laughrea, M. (2007). HIV-1 Viral RNA Is Selected in the Form of Monomers that Dimerize in a Three-step Protease-dependent Process; the DIS of Stem–Loop 1 Initiates Viral RNA Dimerization. J. Mol. Biol. 371, 1084–1098.

Song, R., Kafaie, J., and Laughrea, M. (2008). Role of the 5′ TAR Stem−Loop and the U5-AUG Duplex in Dimerization of HIV-1 Genomic RNA. Biochemistry (Mosc.) 47, 3283–3293.

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

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

Soto-Rifo, R., Limousin, T., Rubilar, P.S., Ricci, E.P., Décimo, D., Moncorgé, O., Trabaud, M.-A., André, P., Cimarelli, A., and Ohlmann, T. (2012c). 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., and 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.

Soulat, D., Bürckstümmer, T., Westermayer, S., Goncalves, A., Bauch, A., Stefanovic, A., Hantschel, O., Bennett, K.L., Decker, T., and Superti-Furga, G. (2008). The DEAD-box

helicase DDX3X is a critical component of the TANK-binding kinase 1-dependent innate immune response. EMBO J. 27, 2135–2146.

Stunnenberg, M., Geijtenbeek, T.B.H., and Gringhuis, S.I. (2018). DDX3 in HIV-1 infection and sensing: A paradox. Cytokine Growth Factor Rev.

Su, C.-Y., Lin, T.-C., Lin, Y.-F., Chen, M.-H., Lee, C.-H., Wang, H.-Y., Lee, Y.-C., Liu, Y.-P., Chen, C.-L., and Hsiao, M. (2015). DDX3 as a strongest prognosis marker and its downregulation promotes metastasis in colorectal cancer. Oncotarget 6, 18602–18612.

Sun, M., Song, L., Li, Y., Zhou, T., and Jope, R.S. (2008). Identification of an antiapoptotic protein complex at death receptors. Cell Death Differ. 15, 1887–1900.

Sun, X., Zhang, Q., and Al-Hashimi, H.M. (2007). Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg2+ binding: role in the kissing–duplex structural transition. Nucleic Acids Res. 35, 1698–1713.

Sundquist, W.I., and Heaphy, S. (1993). Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA. Proc. Natl. Acad. Sci. 90, 3393–3397.

Tahirov, T.H., Babayeva, N.D., Varzavand, K., Cooper, J.J., Sedore, S.C., and Price, D.H. (2010). Crystal structure of HIV-1 Tat complexed with human P-TEFb. Nature 465, 747–751.

Takahashi, K.I., Baba, S., Chattopadhyay, P., Koyanagi, Y., Yamamoto, N., Takaku, H., and Kawai, G. (2000). Structural requirement for the two-step dimerization of human immunodeficiency virus type 1 genome. RNA 6, 96–102.

Tanner, N.K. (2003). The Newly Identified Q Motif of DEAD Box Heicases IS Involved in Adenine Recognition. Cell Cycle 2, 18–19.

Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG Motif. Mol. Cell 50, 613–623.

Thomson, M.M., Pérez-Álvarez, L., and Nájera, R. (2002). Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. Lancet Infect. Dis. 2, 461–471.

Tirumuru, N., Zhao, B.S., Lu, W., Lu, Z., He, C., and Wu, L. (2016). N6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. ELife 5, e15528.

Toro-Ascuy, D., Rojas-Araya, B., García-de-Gracia, F., Rojas-Fuentes, C., Pereira-Montecinos, C., Gaete-Argel, A., Valiente-Echeverría, F., Ohlmann, T., and Soto-Rifo, R. A Rev–CBP80–eIF4AI complex drives Gag synthesis from the HIV-1 unspliced mRNA. Nucleic Acids Res.

Tran, T., Liu, Y., Marchant, J., Monti, S., Seu, M., Zaki, J., Yang, A.L., Bohn, J., Ramakrishnan, V., Singh, R., et al. (2015). Conserved determinants of lentiviral genome dimerization. Retrovirology 12.

Uhlenbeck, O.C. (1995). Keeping RNA happy. RNA 1, 4–6.

Ulyanov, N.B., Mujeeb, A., Du, Z., Tonelli, M., Parslow, T.G., and James, T.L. (2006). NMR Structure of the Full-length Linear Dimer of Stem-Loop-1 RNA in the HIV-1 Dimer Initiation Site. J. Biol. Chem. 281, 16168–16177.

Valiente-Echeverría, F., Hermoso, M.A., and Soto-Rifo, R. (2015). RNA helicase DDX3: at the crossroad of viral replication and antiviral immunity. Rev. Med. Virol. 25, 286–299.

Webb, J.A., Jones, C.P., Parent, L.J., Rouzina, I., and Musier-Forsyth, K. (2013). Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: Implications for viral genomic RNA packaging. RNA 19, 1078–1088.

Weeks, K.M., and Mauger, D.M. (2011). Exploring RNA Structural Codes with SHAPE Chemistry. Acc. Chem. Res. 44, 1280–1291.

Weill, L., James, L., Ulryck, N., Chamond, N., Herbreteau, C.H., Ohlmann, T., and 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, K.A., Merino, E.J., and Weeks, K.M. (2006). Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. Nat. Protoc. 1, 1610–1616.

Windbichler, N., Werner, M., and Schroeder, R. (2003). Kissing complex-mediated dimerisation of HIV-1 RNA: coupling extended duplex formation to ribozyme cleavage. Nucleic Acids Res. 31, 6419–6427.

Wu, J., and Chen, Z.J. (2014). Innate Immune Sensing and Signaling of Cytosolic Nucleic Acids. Annu. Rev. Immunol. 32, 461–488.

Wu, D.-W., Lee, M.-C., Wang, J., Chen, C.-Y., Cheng, Y.-W., and Lee, H. (2014). DDX3 loss by p53 inactivation promotes tumor malignancy via the MDM2/Slug/E-cadherin pathway and poor patient outcome in non-small-cell lung cancer. Oncogene 33, 1515–1526.

Xiang, N., He, M., Ishaq, M., Gao, Y., Song, F., Guo, L., Ma, L., Sun, G., Liu, D., Guo, D., et al. (2016). The DEAD-Box RNA Helicase DDX3 Interacts with NF-κB Subunit p65 and Suppresses p65-Mediated Transcription. PLOS ONE 11, e0164471.

Yang, Q., and Jankowsky, E. (2005). ATP- and ADP-Dependent Modulation of RNA Unwinding and Strand Annealing Activities by the DEAD-Box Protein DED1 †. Biochemistry (Mosc.) 44, 13591–13601.

Yang, Q., and Jankowsky, E. (2006). The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. Nat. Struct. Mol. Biol. 13, 981–986.

Yang, Q., Del Campo, M., Lambowitz, A.M., and Jankowsky, E. (2007a). DEAD-Box Proteins Unwind Duplexes by Local Strand Separation. Mol. Cell 28, 253–263.

Yang, Q., Fairman, M.E., and Jankowsky, E. (2007b). DEAD-box-protein-assisted RNA Structure Conversion Towards and Against Thermodynamic Equilibrium Values. J. Mol. Biol. 368, 1087–1100.

Yedavalli, V.S.R.K., and Jeang, K.-T. (2010). Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. Proc. Natl. Acad. Sci. 107, 14787–14792.

Yedavalli, V.S.R.K., Neuveut, C., Chi, Y., Kleiman, L., and Jeang, K.-T. (2004). Requirement of DDX3 DEAD Box RNA Helicase for HIV-1 Rev-RRE Export Function. Cell 119, 381–392.

Zhao, L., Mao, Y., Zhou, J., Zhao, Y., Cao, Y., and Chen, X. (2016). Multifunctional DDX3: dual roles in various cancer development and its related signaling pathways. Am. J. Cancer Res. 6, 387–402.