Rôle du facteur d'initiation eIF3 dans la traduction de l'ARNm de l'histone H4

Hassan Alaskari Hayek

▶ **To cite this version:**

Hassan Alaskari Hayek. Rôle du facteur d'initiation eIF3 dans la traduction de l'ARNm de l'histone H4. Génomique, Transcriptomique et Protéomique [q-bio.GN]. Université de Strasbourg, 2019. Français. NNT : 2019STRAJ081 . tel-03081207

HAL Id: tel-03081207

https://theses.hal.science/tel-03081207

Submitted on 18 Dec 2020

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

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

Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

UPR 9002 du CNRS, Architecture et Réactivité de l'arN (ARN)

THÈSE présentée par :

Hassan Alaskari HAYEK

soutenue le : **18 octobre 2019**

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

Discipline : Sciences de la vie et de la santé
Spécialité : Aspects moléculaires et cellulaires de la biologie

Rôle du facteur d'initiation eIF3 dans la traduction de l'ARNm de l'histone H4

THÈSE dirigée par :

Dr. ALLMANG-CURA Christine	Directeur de recherches, Université de Strasbourg
Dr. ERIANI Gilbert	Directeur de recherches, Université de Strasbourg

RAPPORTEURS :

Dr. JALINOT Pierre	Directeur de recherches, ENS Lyon
Dr. MILLEVOI Stefania	Chargée de recherches, Université de Toulouse

---

EXAMINATEUR EXTERNE :

Dr. JEAN-JEAN Olivier	Directeur de recherches, Université Pierre et Marie Curie, Paris

EXAMINATEUR INTERNE :

Pr. DIMITROVA Maria	Professeur, Université de Strasbourg

REMERCIEMENTS

Pour commencer, je tiens à remercier les Dr. Pierre Jalinot et Dr. Stefania Millevoi d'avoir accepté d'évaluer mon travail de thèse. Je remercie également Dr. Olivier Jean-Jean Pr. Maria Dimitrova d'avoir bien voulu examiner mon travail.

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

J'exprime ma gratitude et mes vifs remerciements à Gilbert Eriani pour m'avoir accueilli au sein de son équipe, il y a de cela environ trois ans, pour sa co-direction de ma thèse et pour avoir fourni le cadre amical, compétitif et humain dans lequel ce travail a été effectué.

Je tiens à exprimer ma profonde reconnaissance envers Christine Allmang, ma directrice de thèse. La passion qui l'anime s'est vite révélée contagieuse et son soutien enthousiaste a porté mon travail pendant ces trois années. J'ai apprécié sa disponibilité, son sens de la convivialité et son souci permanent de voir ses étudiants s'épanouir.
J'ai particulièrement apprécié les conseils et les discussions scientifiques avec Christine et Gilbert, mais surtout le fait qu'ils aient été à l'écoute et disponibles. Merci de m'avoir fait confiance pour mener à bien notre projet et pour leur coté humain !

Un grand merci à Franck Martin, pour sa disponibilité, pour les discussions et les réflexions scientifiques, pour la bonne humeur et surtout pour les moments passés à Cracovie avec Christine et Fatima !

Un super méga merci aux personnes avec lesquelles j'ai partagé un coin de paillasse au labo 333 :

Tout d'abord, Laure, 'la grande' pour son aide et ses conseils précieux, pour les moments musicaux et la franche rigolade qui rendent la vie au labo si sympathique.

Fatima ! Ma collègue libanaise, avec qui j'ai passé des longs week-end dans le labo ainsi que dans la ville et surtout pendant les dernières semaines, son amitié m'a été précieuse.

Aurélie Janvier 'la petite', Aurélie Durand et Antonin Tidu, merci pour leur bonne patience, leur gentillesse, leur disponibilité, leur humour et leur bonne humeur.

Je souhaite aussi remercier les personnes ayant quitté le 333, Laurianne, Onaïs, Léa, Manon, Anaïs et Zina pour les bons moments passés avec eux ainsi que pour le soutien qu'ils m'ont apporté.

Un grand merci aux personnes de la salle de culture du 3ème pour leur gentillesse, leur bonne humeur et leur disponibilité.

Je tiens également à remercier tous les gens que je rencontre au hasard dans les couloirs de l'IBMC.

Je souhaite aussi remercier toute l'équipe de la plateforme protéomique : Philippe Hammann, Johana Chicher et Lauriane Kuhn, ainsi que Fréderic Gros du plateforme de cytométrie en flux d'avoir traité mes protéines et mes cellules, je serai très reconnaissant envers eux.

Je remercie également la municipalité d'Adchit pour la confiance qu'elle m'a accordée et le soutien financier au cours de ces trois années de thèse.

J'ai une pensée spéciale pour un ami de 6 mois, mais je sens qu'on se connaît depuis longtemps. Ahmad 'tafach', mon collègue libanais, son amitié m'a été précieuse.

Merci également pour les amis et copains de Strasbourg. Mohamad' slango', Nour 'affoufé', Hussein, Hassan, Hamzah et Ali ainsi que la grande communauté libanaise à Strasbourg et aussi mon ami Mohamad à Paris, pour les soirées et les temps amusants passés côte à côte et pour m'avoir supporté tout au long de mes études.

Enfin, je souhaite remercier ma famille, c'est auprès d'elle que j'ai puisé la force d'avancer : je suis profondément reconnaissant envers mes grands parents, mes oncles et mes frères Ali et Abbass, pour l'exemple de courage et de dévouement qu'ils m'ont donné. Mes parents, Mohamad et Jamila, pour le courage qu'ils m'ont donné, le support pour maintenir les pieds sur terre, la joie de vivre et le sens de l'humour, Cette thèse je vous la dédie toute entière !

SOMMAIRE

SOMMAIRE **1**

ABRÉVIATIONS **3**

INTRODUCTION **7**

1. Initiation de la traduction chez les eucaryotes **7**
1.1. Préambule / Bref aperçu de la synthèse protéique 7
1.2. Initiation de la traduction canonique ou coiffe-dépendante 8
1.2.1. Formation du complexe de pré-initiation 43S 8
1.2.2. Fixation du complexe 43S sur l'ARNm et obtention du complexe 48S 10
1.2.2.1. « Activation » de l'ARNm par le facteur eIF4F et recrutement des particules 43S 10
1.2.2.2. 'Scanning' de l'ARNm et formation du complexe 48S 11
1.2.3. Formation du ribosome 80S 11
1.3. Mécanismes d'initiation de la traduction non canoniques 12
1.3.1. Mécanismes d'initiation de la traduction guidés par des structures de l'ARN 12
1.3.1.1. Mécanismes d'initiation coiffe-indépendant utilisant les IRES 12
1.3.1.1.1. Les IRES virales de type I 14
1.3.1.1.2. Les IRES virales de type II 14
1.3.1.1.3. Les IRES virales de type III 15
1.3.1.1.4. Les IRES virales de type IV 15
1.3.1.1.5. Les IRES dans des ARNm cellulaires 15
1.3.1.1.6. Rôle des facteurs 'IRES Trans-Acting Factors' ou ITAF 16
1.3.1.2. Mécanismes d'initiation coiffe-indépendant chez les virus de plantes 16
1.3.1.3. Les mécanismes d'initiation de la traduction grâce aux modifications des adénosines 17
1.3.1.4. L'initiation de la traduction des ARN circulaires (circARN) 18
1.3.1.5. L'initiation de la traduction d'ARNm à coiffe hyperméthylée 19
1.3.1.6. L'initiation de la traduction de l'histone H4 par 'tethering' du 43S sur l'AUG 21
1.4. Régulation de l'initiation de la traduction 22
1.4.1. Régulation par les facteurs d'initiation (eIFs) 23
1.4.2. Régulation par des facteurs homologues aux facteurs d'initiation 24
1.4.3. Régulation par des éléments structuraux présents dans l'extension 5' non traduite des ARNm 25
1.4.3.1. Les pseudo-nœuds dans l'ARNm de l'interféron g 26
1.4.3.2. Les éléments IRE et le métabolisme du fer 26
1.4.3.3. Les structures de type G-quartet 26
1.4.3.4. Les éléments TIE et IRES des ARNm Hox 27
1.4.4. Régulation de l'initiation de la traduction par le biais de la région 3'UTR 27
1.4.5. Régulation par les uORF 30
1.4.6. Régulation par le ribosome 30
2. Le facteur d'initiation eIF3 ..... **33**
2.1. Généralités 33
2.2. Assemblage et structure du complexe eIF3 33
2.3. Implication d'eIF3 dans les mécanismes d'initiation de la traduction canonique et non-canonique 37
2.3.1. Fonctions d'eIF3 dans l'initiation de la traduction canonique 37
2.3.1.1. Interaction d'eIF3 avec les facteurs d'initiation eIF1 et eIF5 : 38
2.3.1.2. Interaction d'eIF3 avec eIF1A 40

2.3.2. Fonctions d'eIF3 tout au long du cycle de la traduction 41
2.3.3. Rôle d'eIF3 dans la traduction sélective d'ARNm et dans les mécanismes de traduction non-canoniques 43
2.4. Régulation des sous-unités d'eIF3 et cancer 45
3. Les histones 49
3.1. Les histones, constituants du nucléosome 49
3.2. Expression et régulation des gènes d'histones réplication-dépendants 51
3.3. Mécanismes de traduction des ARNm des histones réplication-dépendantes 52

MON PROJET DE THESE 55

RÉSULTATS 59

1. Bases moléculaires de l'interaction entre eIF3 et l'ARNm H4 59
1.1. eIF3 interagit *in vitro* avec l'ARNm de l'histone H4 59
1.2. eIF3 interagit *in vivo* avec les ARNm d'histones 61
1.3. Identification des sous-unités d'eIF3 interagissant avec l'ARNm H4 par pontage *in vitro* 63
1.5. Les quatre sous-unités d'eIF3c, d, e et g isolées interagissent avec l'ARNm H4 *in vitro* 67
2. Etude de rôle de l'interaction d'eIF3 sur la traduction des ARNm des histones 73
2.1. Impact de l'élément régulateur de l'ARNm H4 sur la traduction d'un gène rapporteur 73
2.2. Rôle fonctionnel d'eIF3 sur la synthèse des histones *in vivo* 81
2.2.1. Effet de l'inactivation des sous-unités d'eIF3 c, d, e, et g sur la viabilité cellulaire 81
2.2.2. Effet de l'inactivation des sous-unités d'eIF3 c, d, e, et g sur la néosynthèse des ARNm d'histones in vivo 82

CONCLUSION 92

DISCUSSION ET PERSPECTIVES 95

MATÉRIEL ET METHODES 105

ANNEXE 1 : PUBLICATION 1 123
ANNEXE 2 : PUBLICATION 2 133
ANNEXE 3 : PUBLICATION 3 151

BIBLIOGRAPHIE 169

ABRÉVIATIONS

3'CITE	3' Cap-Independent Translation Enhancer
4E-SE	eIF4E-Sensitive Element
ACTB	Protéine du gène de la β-actine
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
AMV-RT	Avian Myeloblastosis Virus Reverse-Transcriptase
ARN	Acide ribonucléique
ARNm	ARN messager
ARNr	ARN ribosomique
ARNt	ARN de transfert
BS	Binding Site
BTG1	B Cell Translocation Gene 1
CDK1	Cyclin-Dependent Kinase 1
c-JUN	Protéine encodée par le gène JUN
C-terminal	carboxy-terminal
CBP	Cap Binding Pocket
CLIP-SEQ	Crosslinking immunoprécipitation Sequencing
cpm	coups par minute
cryo-EM	cryo-Electron Microscopy
Ctrl	Control
dNTP	désoxyriboNucléotides Tri-Phosphates
DTT	Dithiotherol
eIFs	eukaryotic Initiation Factors
F-luc	Firefly luciferase
GDP	Guanosine DiPhosphate
GST	Glutathion-S-transférase
H1	Histone H1
H2A	Histone H2A
H2B	Histone H2B
H3	Histone H3
H4	Histone H4
HisGST	Hexa-histidine Glutathion S-Transferase
IEF	Isoelectrofocalisation
IP	Immunoprécipitation
IRES	Internal Ribosome Entry Site
kDa	kilo Dalton
m^6A	N6-méthyladénosine
m^7G	coiffe monométhylguanosine
miARN	microARNs
MPN	Mpr1/Pad1 N-terminal
NTD	NH_2 Terminal Domain
NMD	Nonsense Mediated Decay
nt	nucléotides
ORF	Open Reading Frame

PABP	Poly(A)-Binding Protein
PAR-CLIP	PhotoActivavtable Ribonucleoside-enhanced Crosslinking and Immuno-recipitation
PCI	Proteasome, COP9/signalosome, eIF3
PCR	Polymerase Chain Reaction
PIC	Pre-Initiation Complex
R-Luc	Renilla Luciferase
RBPs	RNA-Binding Proteins
RNPs	particules ribonuléo-protéiques
RRL	Rabbit Reticulocytes Lysates
RRM	RNA Recognition Motif
RT-qPCR	Reverse Transcriptase-qPCR
S. cerevisiae	*Saccharmyces cerevisiae*
SDS-PAGE	Sodium Dodécyl Sulfate-PolyAcylamide Gel Electrophoresis
siARN	petit ARN interférent
SL	Stem-Loop
SLBP	Stem-Loop Binding Protein
SLIP1	Stem-Loop Interacting Protein 1
ThioU	Thio Uridine
TWJ	Three Way Junction
U2	ARN nucléaire U2
uORF	upstream Open Reading Frame
UTR	Untranslated Region
WB	Western Blot
WD40	tryptophan-aspartic acid (W-D) domain
WHD	Winged Helix Domain
YLC	Yeast-Like Core

Introduction

INTRODUCTION

1. Initiation de la traduction chez les eucaryotes

1.1. Préambule / Bref aperçu de la synthèse protéique

La traduction de l'ARNm est l'étape finale de l'expression des gènes. Elle fait partie des processus cellulaires les plus complexes, les plus conservés du monde vivant et les plus coûteux énergétiquement. La synthèse d'une seule protéine nécessite l'action concertée d'une centaine de protéines et d'ARN de différents types. Le passage du code à 4 bases des acides nucléiques en code à 20 acides aminés se repliant en architectures complexes aux fonctions diverses et multiples fait appel à quatre catégories cruciales de molécules biologiques : les ARN messagers (ARNm ou mRNA), les ARN de transfert (ARNt ou tRNA), les aminoacyl-tRNA synthétases et les ribosomes.

La traduction peut être divisée en trois phases principales, chacune coordonnée par plusieurs facteurs. La première phase est l'initiation qui consiste en l'assemblage du complexe de traduction. Chez les eucaryotes, elle dépend généralement de la présence d'une coiffe mono-méthylée m^7GpppN située à l'extrémité 5' des ARNm et nécessite au moins 12 facteurs d'initiation (eIFs), les sous-unités ribosomiques 40S et 60S ainsi que l'ARNt initiateur (Met-$ARNt_i^{Met}$) dont la fonction est de s'apparier à l'aide de son anticodon au codon AUG initiateur de l'ARNm dans le site P du ribosome 80S (Sonenberg and Hinnebusch, 2009). Le processus d'initiation est complexe et constitue l'étape limitante de la synthèse protéique. Il a été montré que la dérégulation de cette étape pouvait conduire à des changements métaboliques drastiques et à la formation de cancers et maladies neurodégénératives (Bhat et al., 2015; Chu et al., 2016; Robichaud and Sonenberg, 2017).

L'étape d'initiation est suivie de l'étape d'élongation qui correspond à la synthèse proprement dite du peptide. Le peptidyl-ARNt présent dans le site P du ribosome 80S (ou le Met- $ARNt_i^{Met}$ lors du premier cycle de synthèse) s'attache d'une manière covalente par une liaison peptidique au nouvel acide aminé estérifié sur l'ARNt présent dans le site A du ribosome. L'étape d'élongation nécessite l'intervention de 3 facteurs protéiques. Le facteur eEF1A assure l'apport et le dépôt de l'aminoacyl-ARNt dans le site A du ribosome où il subit la réaction de transfert de la chaine peptidique. Le facteur eEF2 intervient ensuite durant l'étape de translocation du ribosome sur la phase codante de l'ARNm, conduisant au glissement de l'ARNt du site A vers le site P. Enfin le facteur eEF1B recycle eEF1 en y fixant une nouvelle molécule de GTP (Riis et al., 1990). La synthèse du peptide s'arrête lorsque le ribosome atteint la fin de la phase codante et lorsque l'un des codons stop (UAA, UGA, ou

UAG) pénètre dans le site A du ribosome. C'est la phase de terminaison durant laquelle le peptide synthétisé est relâché par les ribosomes grâce aux facteurs de terminaison eRF1 et eRF3 (eukaryotic Release Factor) (Zhouravleva et al., 1995). Le ribosome est ainsi dissocié en sous-unités 40S et 60S qui sont recyclées pour une nouvelle synthèse peptidique (Dever and Green, 2012).

Durant ma thèse, je me suis focalisé sur l'étape d'initiation de la traduction chez les eucaryotes et plus précisément sur l'étude d'un mécanisme d'initiation atypique qui a été identifié dans le cas de l'ARNm de l'histone H4.

1.2. Initiation de la traduction canonique ou coiffe-dépendante

Plusieurs évènements importants ponctuent l'initiation de la traduction. Les complexes de pré-initiation doivent être efficacement recrutés par l'ARNm. Le Met-ARNt$_i^{Met}$ présent dans le site P de ces particules doit alors rechercher le codon AUG initiateur de l'ARNm et s'y hybrider par le biais de son anticodon en vue du démarrage de la traduction. Ces évènements comportent plusieurs étapes particulièrement critiques qui sont décrites ci-dessous **(voir *Figure 1*)**.

1.2.1. Formation du complexe de pré-initiation 43S

Le complexe de pré-initiation 43S est constitué par l'assemblage des facteurs d'initiation libérés d'un précédent cycle d'initiation sur les petites sous-unités ribosomiques 40S. Les facteurs protéiques eIF3, eIF1, eIF1A et eIF5 sont recrutés dans un premier temps et cet assemblage permet l'association du facteur eIF2 lié au GTP et au Met-ARNt$_i^{Met}$ (Jackson et al., 2010). Le complexe ternaire eIF2/GTP/Met-ARNt$_i^{Met}$ est en charge de la reconnaissance de l'AUG initiateur par l'établissement d'un appariement codon-anticodon. C'est un des complexes ribonucléo-protéiques qui illustre dès la phase d'initiation les relations intimes qui relient les protéines et les acides nucléiques tout au long du processus de traduction. Le facteur eIF2 est constitué de 3 unités (α, β et γ) toutes trois nécessaires au bon fonctionnement du complexe ternaire (Sonenberg and Dever, 2003). Une fois formé, le complexe de pré-initiation 43S est recruté sur l'ARNm grâce à l'interaction entre eIF3 et le facteur eIF4G présent dans le complexe eIF4F fixé à la coiffe située à l'extrémité 5' de l'ARNm (Fraser and Doudna, 2007; Hinnebusch and Lorsch, 2012; Jackson et al., 2010). Le facteur eIF3 interagit également avec la Poly A Binding Protein (PABP) présente sur la queue poly (A) située à l'extrémité 3' de la plupart des ARNm cellulaires, ce qui assure la circularisation des ARNm. Tout au long de ces étapes, le facteur eIF3 est un des facteurs clés qui orchestre la formation du complexe 43S (Hinnebusch, 2006). Il joue un rôle anti-

associant vis-à-vis de la grande sous-unité 40S, et en se fixant dès les premiers stades, il assure le bon recrutement des autres facteurs et la stabilisation de la particule 43S finale (***Etape 1, Figure 1***).

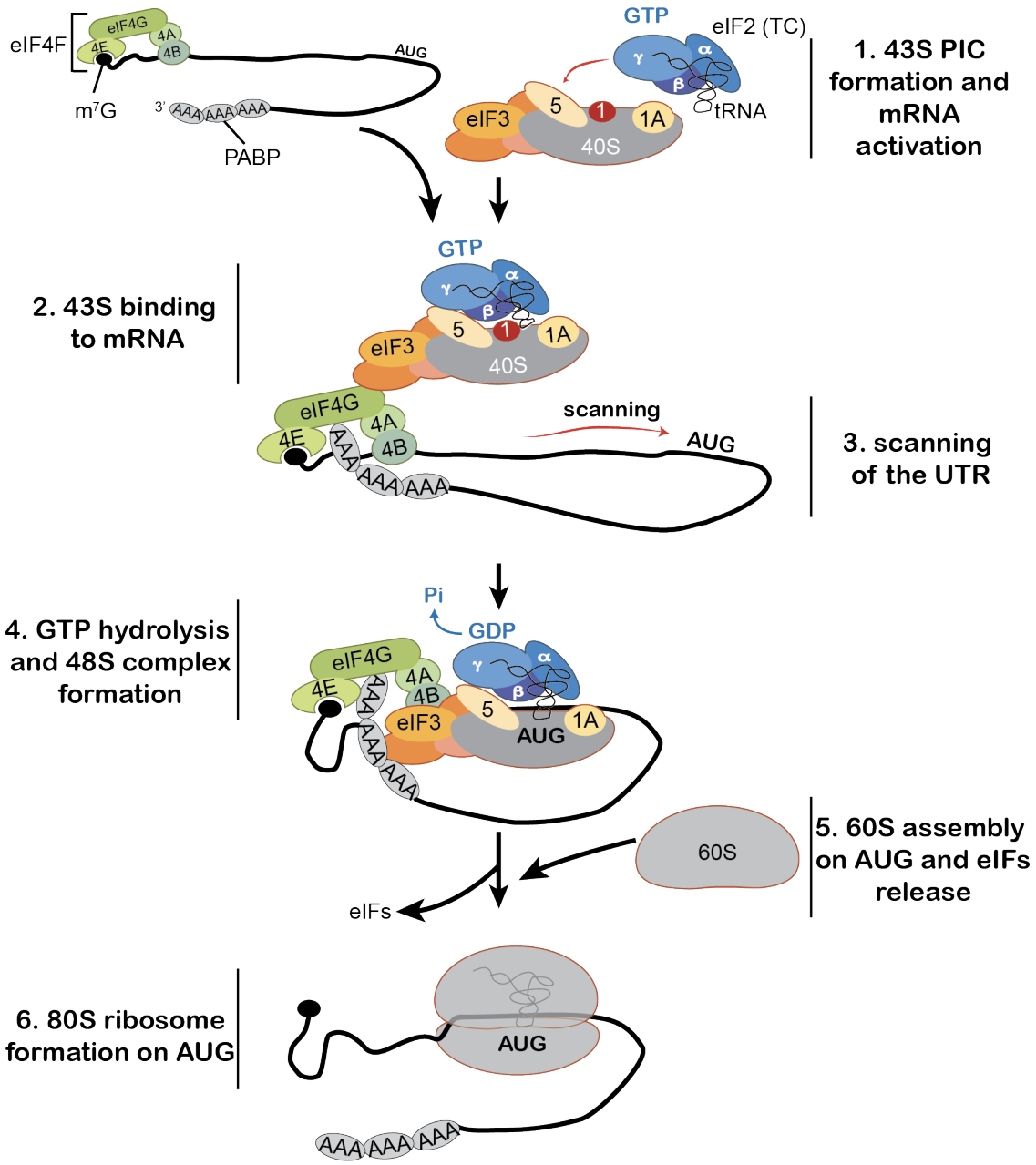


Figure 1. Le mécanisme d'initiation de la traduction canonique « coiffe-dépendant » chez les eucaryotes. Le mécanisme d'initiation canonique dépendant de la coiffe peut être divisé en 6 étapes (étapes **1**-**6**). Le complexe ternaire eIF2-GTP-Met-ARNt$^{Met}_{i}$ est assemblé dans le complexe de pré-initiation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

1.2.2. Fixation du complexe 43S sur l'ARNm et obtention du complexe 48S

Le complexe de pré-initiation 43S se lie à la coiffe présente à l'extrémité 5' de l'ARNm grâce à l'interaction qu'eIF3 établit avec le facteur de liaison à la coiffe eIF4F. Le complexe de pré-initiation 43S ainsi lié va glisser le long de l'extrémité 5' non traduite (5' UTR) de l'ARNm à la recherche du codon d'initiation AUG. Ce processus appelé 'scanning', requiert les facteurs eIF1 et eIF1A. L'interaction codon-anticodon entre l'ARNm et l'ARN_i^{Met} conduit à la formation du complexe d'initiation 48S, prêt pour le démarrage de la traduction (Asano et al., 2000; Chaudhuri et al., 1999).

1.2.2.1. « Activation » de l'ARNm par le facteur eIF4F et recrutement des particules 43S

Les ARNm cellulaires sont coiffés en 5' par un nucléotide particulier modifié, le 7-méthylguanosine ou m^7GpppN. Cette coiffe participe au processus de traduction ; elle assure également la protection contre les RNases cellulaires et intervient dans le transport nucléo-cytoplasmique (Ramanathan et al., 2016). La coiffe m^7GpppN de l'ARNm interagit avec le facteur d'initiation eIF4F. Le facteur eIF4F est lui-même composé de l'assemblage de trois facteurs : le facteur de liaison à la coiffe eIF4E, l'hélicase eIF4A à motif 'DEAD' et le facteur eIF4G servant de plateforme d'amarrage pour ces facteurs. La première étape de l'activation de l'ARNm consiste à fixer le facteur eIF4E sur la coiffe (Sonenberg et al., 1979) en compagnie des autres facteurs eIF4G et eIF4A qui y sont liés au sein du complexe eIF4F (***Etape 2, Figure 1***). Le rôle de l'hélicase eIF4A est de déplier les structures secondaires situées à proximité de la coiffe permettant ainsi à la sous-unité ribosomique 40S de s'y fixer par le sillon dit de l'ARNm qui passe par les sites A, P et E du ribosome (Conroy et al., 1990). Le facteur eIF4G sert de plateforme d'assemblage pour les autres facteurs présents dans le complexe eIF4F mais aussi pour d'autres partenaires (PABP, Mnk1, SLIP1) (Prévôt et al., 2003). Il interagit avec eIF4E via son domaine central et avec eIF4A via deux domaines de liaison distincts. Par ailleurs, les PABP liées à la queue poly (A) des ARNm se fixent à son domaine N-terminal permettant ainsi leur circularisation. Enfin l'ARNm circularisé peut recruter le complexe de pré-initiation 43S grâce à l'interaction entre le domaine C-terminal d'eIF4G et le facteur eIF3 (Lamphear et al., 1995; Mader et al., 1995; Ohlmann et al., 1996). Cette structure circulaire de l'ARNm favorise le recyclage et l'assemblage des complexes d'initiation, stimulant ainsi la traduction (Jackson et al., 2010; Sonenberg & Hinnebusch, 2009).

1.2.2.2. ‘Scanning’ de l’ARNm et formation du complexe 48S

Une fois fixé à proximité de l’extrémité 5’ de l’ARNm, le complexe 43S va parcourir la région 5’ non traduite (5’ UTR) de l’ARNm dans la direction 5’ -> 3’ à la recherche de l’AUG initiateur (***Etape 3, Figure 1***). Ce mécanisme appelé ‘scanning’ (Hinnebusch, 2017; Kozak, 1978) dépend fortement des facteurs eIF1 et eIF1A qui maintiennent le complexe 43S dans une conformation ouverte dite ‘scanning compétente’ ainsi que des hélicases DHX29 et DDX3 plus efficaces qu’eIF4A pour déplier les structures rencontrées dans l’extrémité 5’ UTR de l’ARNm (Hinnebusch, 2017; Pisareva et al., 2008). Durant le scanning, l’établissement de l’interaction codon-anticodon renforce l’interaction eIF1A-40S. Le facteur eIF1, situé dans le site P du ribosome, valide le bon positionnement du complexe 43S en vérifiant la liaison codon-anticodon et dès lors il va diminuer l’énergie du complexe de ‘scanning’ pour en arrêter le mouvement. A ce stade, eIF1 est déplacé du voisinage du site P ce qui entraine le basculement dans une conformation dite fermée sur l’ARNm.

Le facteur eIF5 fixé à eIF2α active alors l’activité GTPasique d’eIF2γ (***Etape 4, Figure 1***). Cette activation intervient une fois qu’eIF1 a été déplacé de son site initial (Guenther et al., 2018; Hinnebusch, 2014). Une fois l’$ARNt_i^{Met}$ libéré dans le site P, le complexe binaire eIF2-GDP est relâché du ribosome avant d’être réactivé par remplacement du GDP en GTP grâce au facteur pentamérique eIF2B (Jennings and Pavitt, 2010).

L’ensemble de ces étapes est favorisé par un contexte de séquence spécifique appelé séquence Kozak. La séquence Kozak présente un consensus de type [$GCCR_{-3}CC$**AUG**G_{+4}] dans lequel la purine en position -3 et la G en position +4 sont particulièrement importantes (Kapp and Lorsch, 2004; Kozak, 1984). Ce contexte permet au complexe 43S de ‘passer’ un certain nombre d’AUG présents dans la 5’ UTR et de favoriser celui de la protéine à traduire.

A l’issue de cette étape de ‘scanning’, on obtient une particule ribosomique interagissant via l’anticodon de l’$ARNt_i^{Met}$ avec l’AUG initiateur de l’ARNm dans le site P. Un tel complexe devient un complexe 48S qui est prêt à lier la grande sous-unité ribosomique 60S et démarrer la traduction.

1.2.3. Formation du ribosome 80S

Après la formation du complexe d’initiation 48S sur l’AUG initiateur, les facteurs eIF1, eIF1A, eIF3, eIF5 et eIF2-GDP sont libérés. Ce départ est provoqué par eIF5B, une GTPase ribosome-dépendante qui permet l’association de la grande sous-unité ribosomique 60S (Jackson et al., 2010; Pestova & Hellen, 2000). L’hydrolyse du GTP d’eIF5B n’est pas requise pour l'association à la sous-unité 60S mais pour son propre départ. Les facteurs eIF5B et eIF2 occupent le même site à l’interface des sous-unités ; l’arrivée d’eIF5B entraine donc le départ d’eIF2 (***Etape 5, Figure 1***). L’interaction d’eIF5B avec eIF1A est aussi requise

pour l'assemblage, ce qui entraine un départ tardif d'eIF1A, tout comme eIF4G et eIF3 qui sont situés sur le côté exposé au solvant. L'hydrolyse du GTP ne se fait qu'après la réunion des sous-unités ribosomique 40S et 60S (Kapp and Lorsch, 2004). Le ribosome 80S ainsi formé, est alors prêt à entrer dans la phase d'élongation marquant la fin du processus d'initiation canonique de la traduction (***Etape 6, Figure 1***).

1.3. Mécanismes d'initiation de la traduction non canoniques

Comme décrit ci-dessus, le processus standard d'initiation de la traduction eucaryote se base sur la présence d'une coiffe à l'extrémité 5' de l'ARNm et fait appel à une douzaine de facteurs d'initiation conduisant à la recherche et la localisation du codon AUG initiateur après une étape de 'scanning' de l'extrémité 5' non traduite. Il existe d'autres populations d'ARNm qui sont traduits par d'autres mécanismes d'initiation. Ces mécanismes sont dits « non conventionnels » ou « non canoniques ». Ces mécanismes alternatifs peuvent faire appel à des éléments structuraux dans l'ARNm qui recrutent directement les ribosomes ou les facteurs et nécessitent un nombre variable et toujours réduit de facteurs d'initiation. Souvent ils découlent d'une absence de coiffe en 5' comme dans le cas des ARN viraux. Les virus sont riches en mécanismes d'initiation alternatifs, en particulier les virus simple brin à polarité positive non coiffés qui sont directement traduits dans la cellule infectée. Certains ARNm cellulaires peuvent également être traduits de manière non conventionnelle grâce à la présence d'éléments structuraux dans les régions non traduites de l'ARN ou plus rarement dans la phase codante. Ces processus alternatifs permettent l'expression de certains ARNm durant des stades particuliers du cycle cellulaire ou bien en réponse à différents stimuli (stress cellulaire, infection virale, etc.). (Lacerda et al., 2019; Shatsky et al., 2010, 2018). Chez les virus, ces structures conduisent généralement à une initiation plus efficace que le processus canonique.

1.3.1. Mécanismes d'initiation de la traduction guidés par des structures de l'ARN

1.3.1.1. Mécanismes d'initiation coiffe-indépendant utilisant les IRES

Au cours de l'évolution, les virus ont mis en place des stratégies pour détourner la machinerie traductionnelle de la cellule hôte en faveur de la traduction de leurs propres ARNm. Ceci se fait au détriment de celle des ARNm cellulaires et permet une multiplication rapide des particules virales (Au and Jan, 2014). De nombreux virus à ARN à polarité positive utilisent des structures d'ARN appelées IRES ou 'Internal Ribosome Entry Site', présentes dans les 5' UTR ou les régions intergéniques dans le cas de virus exprimant

plusieurs cistrons (pour une revue récente (Mailliot and Martin, 2018)). Ces structures d'ARN permettent le recrutement direct en *cis* des particules ribosomiques sans l'aide de facteurs (Jackson and Kaminski, 1995; Pelletier and Sonenberg, 1988) ou avec un nombre réduit de facteurs d'initiation, le facteur de liaison à la coiffe eIF4E étant systématiquement exclu (Terenin et al., 2017). Les IRES sont généralement retrouvées dans les ARNm viraux dépourvus de coiffe en 5' mais on les trouve aussi dans certains ARNm cellulaires où ils leur confèrent la capacité d'être traduits en absence d'un ou plusieurs facteurs d'initiation. Par exemple durant le stress cellulaire induit par l'apoptose, les protéases activées peuvent cliver le facteur eIF4E indispensable à la traduction coiffe-dépendante. Cependant, certains ARNm essentiels pour l'apoptose ont besoin d'être traduits et cela grâce aux IRES contenues dans leurs 5' UTR. Il existe de nombreux types d'IRES dont le point commun est leur indépendance par rapport à la coiffe, ils diffèrent cependant grandement par le nombre des facteurs d'initiation requis et le mode de recrutement des ribosomes (voir ***Figure 2***).

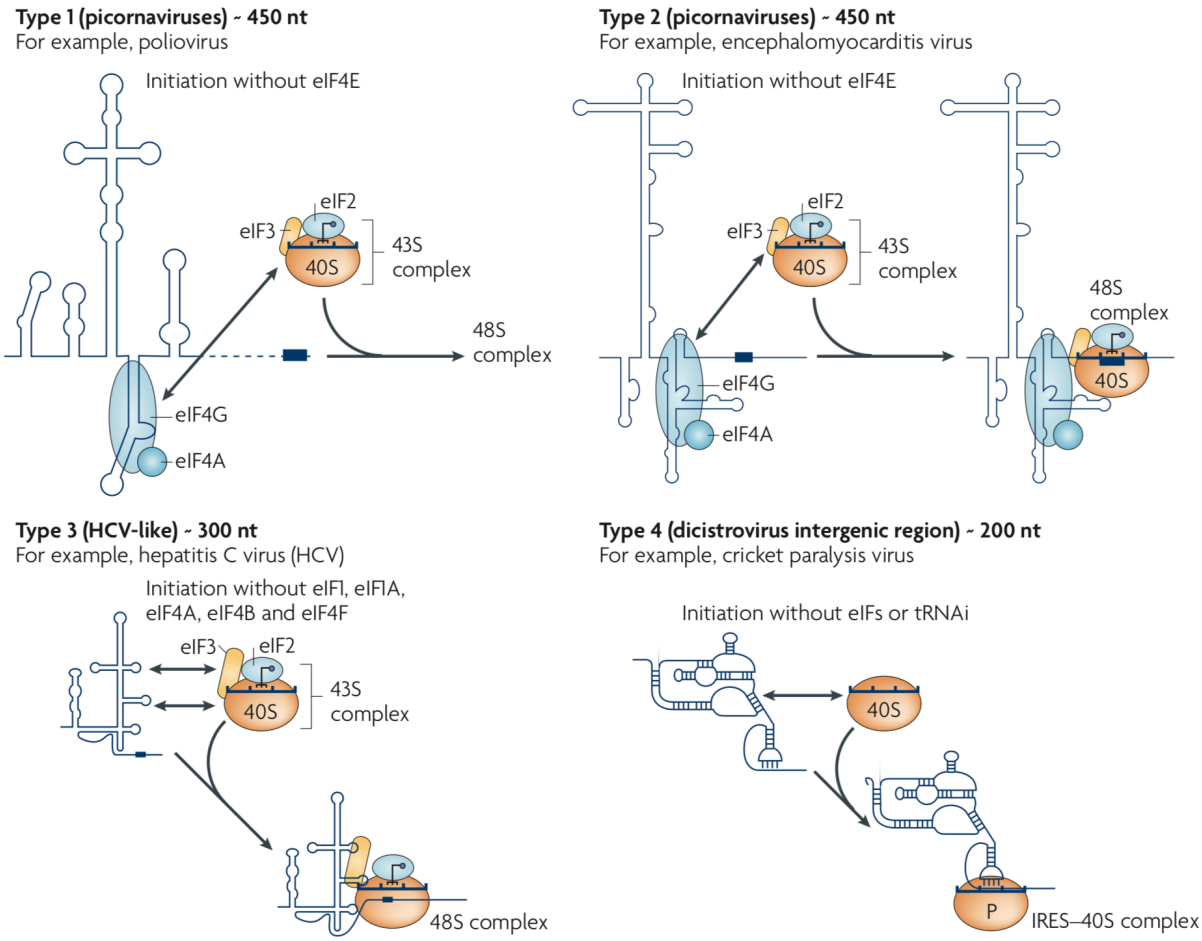


Figure 2. Mécanismes d'initiation de la traduction coiffe-indépendants induits par les IRES (figure issue de (Jackson et al., 2010)). Quatre types d'IRES ont été caractérisées. Les IRES ont un point commun qui est leur capacité à induire l'initiation de la traduction indépendamment de la coiffe. Elles diffèrent grandement par le nombre des facteurs d'initiation avec lesquels ils interagissent et le

mode de recrutement des ribosomes. Les IRES de type **I** nécessitent tous les facteurs d'initiation sauf eIF4E pour initier la traduction. Les IRES de type **II** recrutent directement le complexe de pré-initiation sur l'AUG initiateur sans étape de 'scanning'. Les IRES de type **III** recrutent le complexe de pré-initiation 43S sur le codon AUG initiateur, sans passer par l'étape de 'scanning'. Les deux facteurs d'initiation eIF2 et eIF3 sont suffisants pour l'assemblage du complexe d'initiation 48S. Les IRES de type **IV** possèdent la capacité d'assembler le ribosome 80S sans l'aide d'aucun facteur d'initiation, ni même du complexe ternaire eIF2-GTP-Met-ARNt$_i^{Met}$.

1.3.1.1.1. Les IRES virales de type I

Les IRES de classe I sont principalement présentes dans les génomes des *Picornaviridae*, comme par exemple les Entérovirus A71 (EV-A71), Coxsackievirus (CVB3), Poliovirus (PV) et Rhinovirus Humain (HRV). Il s'agit de virus à ARN de polarité positive dont le génome est traduit dès les premiers stades de l'infection par la machinerie cellulaire (Pelletier & Sonenberg, 1988; Sweeney et al., 2014). Pour positionner le ribosome sur l'AUG initiateur, ces IRES recrutent tous les facteurs canoniques de l'initiation de la traduction à l'exception du facteur eIF4E qui se fixe à la coiffe (Sonenberg and Meerovitch, 1990). Au début de l'infection virale, le Poliovirus synthétise la protéase virale 2A qui dissocie le facteur eIF4F en coupant eIF4G au sein de son domaine de liaison avec eIF4E. Cette dissociation libère un fragment d'eIF4G restant associé avec eIF4E mais ne pouvant plus se fixer à l'extrémité 5' coiffée des ARNm. Le virus utilise cette stratégie pour inactiver l'initiation cellulaire coiffe dépendante, tout en profitant de la machinerie de traduction cellulaire libérée (Gradi et al., 1998; Haghighat et al., 1996). L'ARN viral recrute le complexe 43S en amont de l'AUG initiateur via l'interaction entre le facteur eIF4G tronqué et l'élément IRES. Une fois recruté, le complexe 43S poursuit les étapes classiques du modèle canonique, à savoir la recherche de l'AUG initiateur par le processus de 'scanning' où le ribosome va être assemblé (Sonenberg and Meerovitch, 1990b). Ainsi, les IRES de la classe I nécessitent tous les facteurs d'initiation, sauf eIF4E, pour initier la traduction.

1.3.1.1.2. Les IRES virales de type II

Les IRES de type II sont également retrouvées dans les génomes des *Picornaviridae* (Jang et al., 1988). Dans cette catégorie, les IRES les plus étudiées sont celles de l' 'Encephalo MyoCarditis Virus' (EMCV) et du 'Foot and Mouth Disease Virus' (FMDV). Ce type d'IRES fait appel aux mêmes facteurs d'initiation que les IRES de type I. Les différences majeures qui caractérisent ces IRES sont l'activation de protéines cellulaires qui habituellement répriment la traduction cellulaire : les protéines 4E-BP ('eIF4E-binding protein'). Ces composants normaux de la régulation cellulaire présentent une très forte affinité pour le site d'interaction d'eIF4E avec eIF4G sous leur forme déphosphorylée. Ils empêchent ainsi la formation d'un complexe eIF4F complet et inactivent la synthèse protéique cellulaire (Gingras et al., 1999; Pause et al., 1994). Par ailleurs, les IRES de type II

recrutent directement le complexe de pré-initiation sur l'AUG initiateur sans étape de 'scanning' (Kaminski et al., 1990).

1.3.1.1.3. Les IRES virales de type III

Ces IRES sont retrouvées dans le génome des *Flaviviridae* dont fait partie le virus de l'hépatite C (HCV) (Niepmann, 2013; Reynolds et al., 1995). Comme les IRES de type II, les IRES de type III recrutent le complexe de pré-initiation 43S sur le codon AUG initiateur, sans passer par l'étape de 'scanning'. Les deux facteurs d'initiation eIF2 et eIF3 sont suffisants pour l'assemblage du complexe d'initiation 48S (Pestova et al., 1998). En se fixant sur la petite sous-unité ribosomique 40S, le facteur eIF3 interagit avec des structures présentes dans l'ARN viral et recrute le complexe ternaire eIF2-GTP-Met-ARNt_i^{Met} au niveau de l'AUG initiateur où le ribosome s'assemble (Ji et al., 2004; Kieft et al., 2001; Otto and Puglisi, 2004). Cet assemblage du ribosome fait intervenir les facteurs eIF5B-GTP et eIF5 comme dans le modèle d'initiation canonique (Locker et al., 2007).

1.3.1.1.4. Les IRES virales de type IV

Les IRES de type IV sont trouvées dans les génomes des *Dicistroviridae*. L'exemple le plus étudié et le mieux caractérisé est celui du virus de la paralysie du criquet (CrPV). Ces IRES sont nommées IGR 'Intragenic Region' car localisées entre les deux ORF du génome. Elles possèdent la capacité d'assembler le ribosome 80S sans l'aide d'aucun facteur d'initiation, ni même du complexe ternaire eIF2-GTP-Met-ARNt_i^{Met} (Sasaki and Nakashima, 1999; Wilson et al., 2000) L'initiation de la traduction résulte de l'interaction directe entre les éléments structuraux de l'IRES avec les sous-unités ribosomiques et la traduction démarre sur un codon qui diffère d'un AUG. Plusieurs études structurales et fonctionnelles ont montré que l'IRES du CrPV forme une structure complexe qui mime le bras anticodon de l'ARNt_i^{Met} apparié à un codon viral ce qui lui permet d'entrer dans le site A du ribosome puis d'être transloqué dans le site P (Deniz et al., 2009; Fernández et al., 2014; Muhs et al., 2015; Murray et al., 2016; Pisarev et al., 2005; Sasaki and Nakashima, 1999; Wilson et al., 2000).

1.3.1.1.5. Les IRES dans des ARNm cellulaires

Plusieurs études suggèrent que 15% des ARNm des mammifères contiendraient des IRES cellulaires, notamment les ARNm qui codent pour des facteurs de transcription, de croissance ou des transporteurs. Les IRES cellulaires les plus étudiées sont localisées dans les ARNm des proto-oncogènes c-myc, de la cycline D1, du facteur de croissance VEGF et de hif-1α finement régulée durant le cycle cellulaire. Ces ARNm particuliers utilisent leur IRES pour la traduction dans des conditions de stress cellulaire où l'initiation canonique dépendante de la coiffe en 5' est inactivée (Lacerda et al., 2017; Mokrejs et al., 2010;

Spriggs et al., 2008). Ces IRES adoptent des repliements moins complexes que les IRES virales et leur fonctionnement nécessite des ITAF ou '*IRES Trans-Acting Factors'* notamment pour stabiliser leurs structures. Par ailleurs, l'activité des IRES cellulaires est généralement bien plus faible que celle observée pour les IRES virales (Komar and Hatzoglou, 2011).

1.3.1.1.6. Rôle des facteurs 'IRES Trans-Acting Factors' ou ITAF

Les facteurs cellulaires ITAF ou '*IRES Trans-Acting Factors'* peuvent influencer l'initiation de la traduction par leur action sur les IRES, en la stimulant ou en l'inhibant. Certains ITAF contiennent des domaines de liaison à l'ARNm (Lee et al., 2017; Martínez-Salas et al., 2013). Ces facteurs cellulaires sont impliqués dans divers mécanismes tels la régulation de la transcription, l'épissage, le transport des ARN ou le contrôle de la traduction. Les ITAF sont capables de former des polymères d'ITAF favorisant ainsi la maturation des ARNm et la protection contre les ribonucléases. Ils stabilisent la liaison IRES-ribosome en coopérant avec les facteurs d'initiation (Lewis and Holcik, 2008).

1.3.1.2. Mécanismes d'initiation coiffe-indépendant chez les virus de plantes

La majorité des virus de plantes à ARN ne sont pas coiffés à leurs extrémités 5' ni polyadénylés en 3'. De ce fait, ces virus font appel à des mécanismes non conventionnels pour recruter les ribosomes à l'extrémité 5' de leurs génomes (Newburn and White, 2015). Les virus des familles *Tombusviridae*, *Luteoviridae*, et *Umbraviridae* présentent des éléments d'ARN en *cis* appelés '3' Cap-Independent Translation Element' (CITE) localisés dans les 3' UTR (Nicholson et al., 2010; Simon and Miller, 2013). Les CITE recrutent des facteurs d'initiation, ainsi que les sous-unités ribosomiques et les déposent dans la région 5' de l'ARN viral, augmentant ainsi l'efficacité de la traduction du génome viral. Ces éléments peuvent prendre différentes structures dites en « I », en « Y », en « T », ou de type pseudo-nœud (Simon and Miller, 2013; Wang et al., 2009). La plupart des CITE fonctionnent en recrutant les facteurs d'initiation eIF4E et/ou eIF4G sur des structures de type 'kissing loop' formées entre l'élément CITE situé en 3' et une tige-boucle présente dans la région 5' UTR du génome viral (Treder et al., 2008). D'autres CITE, tel que celui du 'Turnip Crinkle Virus' (TCV), recrutent directement la sous-unité ribosomique 60S au niveau d'une séquence riche en pyrimidines dans la région 5' UTR (McCormack et al., 2008; Stupina et al., 2008). La sous-unité 60S va ensuite recruter et s'associer à la sous-unité 40S pour terminer l'initiation. D'autres travaux sur le virus à ARN (BYDV) 'Barley Yellow Dwarf Virus' ont proposé un modèle dans lequel un élément viral appelé 'BTE' (pour BYDV-like CITE) situé en 3' interagit avec le facteur d'initiation eIF3 afin de favoriser les interactions 3'-5' UTR et le recrutement

de la sous-unité ribosomique 40S avec les facteurs d'initiation associés (Bhardwaj et al., 2019), voir ***Figure 3***. Des études récentes ont étendu le concept de CITE aux ARNm cellulaires. Des CITE cellulaires pourraient fixer certains facteurs d'initiation comme eIF3 afin de recruter les ribosomes sur les ARNm, stimulant ainsi leur traduction (Shatsky et al., 2018).

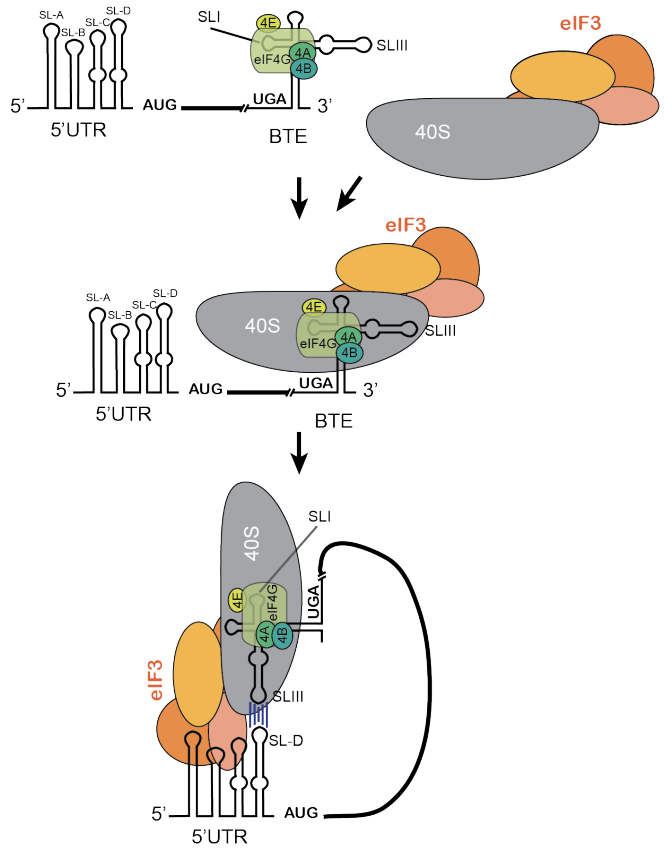


Figure 3. Modèle de recrutement du complexe de pré-initiation dans le 5′ UTR via 3′ BTE. Le complexe eIF4F se lie directement à la tige boucle SLI de l'élément 3′ BTE (élément viral 'BYDV-like CITE') de l'ARN du BYDV (Barley Yellow Dwarf Virus). En présence d'ATP, eIF3 se lie au 3′ BTE afin de recruter le complexe 43S. Par la suite (ou simultanément), une interaction de type 'kissing loop' entre les boucles SLIII (du 3′ BTE) et SL-D (du 5′ UTR) rapproche les UTR dans un alignement approprié permettant à eIF3 de stabiliser l'interaction en se liant directement à l'extrémité 5' UTR. (Bhardwaj et al., 2019; Shatsky et al., 2018).

1.3.1.3. Les mécanismes d'initiation de la traduction grâce aux modifications des adénosines

Les ARNm eucaryotes peuvent contenir dans leurs 5' et 3' UTR des nucléotides N6-méthyladénosine (m^6A) issus de modifications post-transcriptionnelles. Ces modifications impactent la structure et la traduction des ARNm qui les contiennent (Harcourt et al., 2017; Tong et al., 2018; Zhao et al., 2017). Les modifications m^6A favorisent l'assemblage des complexes d'initiation en interagissant avec le facteur d'initiation eIF3 (voir ***Figure 4 A***). La traduction est donc initiée d'une manière coiffe-indépendante et ne nécessite pas la protéine

eIF4E (Meyer et al., 2015). Un tel mécanisme a été décrit pour la protéine HSP70, en réponse à un stress induit par un choc thermique (Zhou et al., 2016).

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

Les ARN circulaires (circARN) sont des boucles d'ARN produites par un grand nombre de gènes exprimés chez les eucaryotes. Ils sont considérés comme des produits de clivage ou d'épissage des introns des précurseurs d'ARNm. Leurs rôles et mécanismes de traduction moléculaires sont encore mal compris (Diallo et al., 2019). Des études récentes montrent qu'ils sont conservés durant l'évolution et abondants dans les cellules. Les circARN présents dans le cytoplasme des cellules sont traduits. Deux mécanismes de traduction sont impliqués dans la traduction des ARN circulaires (voir ***Figure 4 B***), l'un dépendant d'IRES situées en amont du codon initiateur (Leppek et al., 2018; Salzman, 2016), l'autre dépendant de résidus m^6A qui en interagissant avec le facteur d'initiation eIF3 permettent le recrutement du complexe d'initiation (voir paragraphe précédent). Ces mécanismes seraient activés en réponse aux stress cellulaires (Martin et al., 2003).

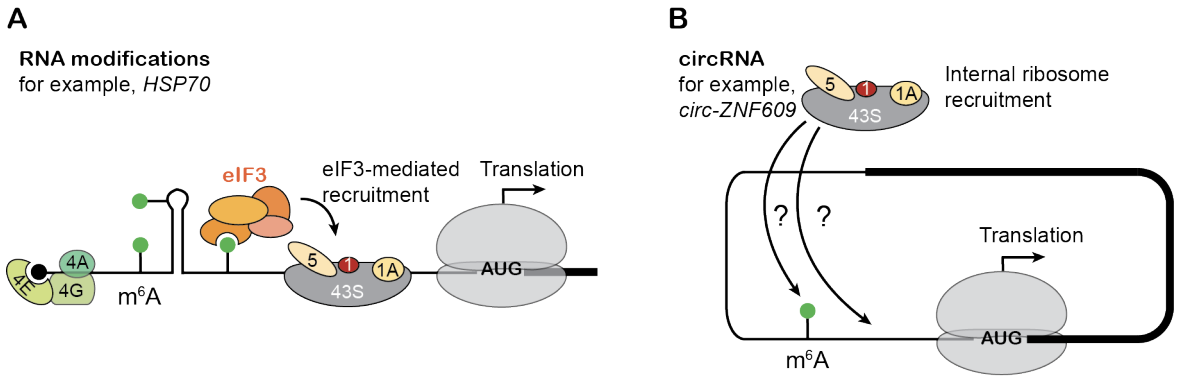


Figure 4. Initiation de la traduction interne grâce aux méthylations des adénosines. (A) En condition de stress cellulaire, la présence de nucléotides modifiés N6-méthyladénosine (m^6A) dans les régions 5' UTR de certains ARNm eucaryotes favorise le recrutement interne du facteur d'initiation eIF3. Ce dernier recrute ensuite le complexe de pré-initiation 43S dans la région 5' UTR de l'ARNm. Ce mécanisme d'initiation coiffe-indépendant a été observé pour l'ARNm codant pour les protéines de réponse aux chocs thermiques HSP70 (Zhou et al., 2016). **(B)** Un mécanisme de recrutement du complexe 43S similaire serait à l'origine de la traduction de certains ARN circulaires (circRNA ou circARN). Chez les eucaryotes les circARN sont des produits d'épissage des pré-ARNm, leurs rôles et leurs mécanismes moléculaires de traduction sont encore mal compris (Diallo et al., 2019).

1.3.1.5. L'initiation de la traduction d'ARNm à coiffe hyperméthylée

Les ARNm eucaryotes se caractérisent généralement par la présence d'une coiffe monométhylée m^7G à leur extrémité 5'. Notre équipe a mis à jour l'existence d'ARNm de mammifères portant une coiffe triméthylée $m_3^{2,2,7}G$, les ARNm de sélénoprotéines (voir ***Figure 5***) (Wurth et al., 2014; Gribling-Burrer et al., 2017). La liaison du facteur d'initiation canonique de la traduction eIF4E est hautement spécifique pour la coiffe m^7G. De façon cohérente, ces ARNm de sélénoprotéines ne sont pas reconnus efficacement par eIF4E, mais sont néanmoins associés aux ribosomes en cours de traduction *in vivo* (Wurth et al. 2014). Une étude au niveau génomique utilisant la torine comme inhibiteur de la traduction dépendante de eIF4E, a identifié des ARNm utilisant des modes d'initiation non canoniques, les ARNm de sélénoprotéines GPx1 et GPx4 à coiffe $m_3^{2,2,7}G$ appartiennent à cette liste d'ARNm (Thoreen et al., 2012). Leur mécanisme d'initiation de la traduction reste cependant incompris. L'ARNg du VIH-1 possède également une coiffe $m_3^{2,2,7}G$ qui est nécessaire à sa traduction (Yedavalli and Jeang, 2010). Dans ce cas, une hélicase ATP dépendante, DDX3 est nécessaire pour une traduction IRES-dépendante de l'ARNg et est capable de se substituer au facteur d'initiation eIF4E dans la liaison à la coiffe. Un complexe de pré-initiation trimérique DDX3/PABP/eIF4G est alors généré qui ne contient ni les protéines de liaison à la coiffe CBP20/CBP80, ni eIF4E (Soto-Rifo et al., 2013).

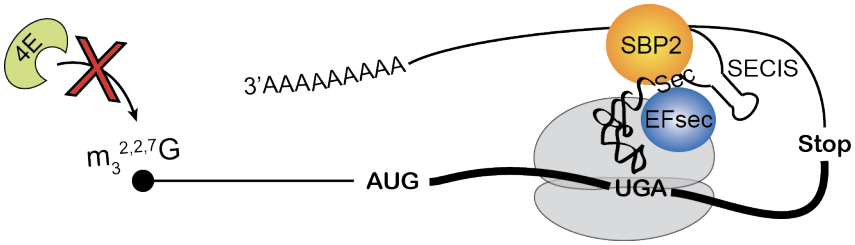


Figure 5. Traduction des ARNm de sélénoproteines à coiffe hyperméthylée. Les ARNm de sélenoproteines **mammifères portant une coiffe triméthylée** $m_3^{2,2,7}G$ ne sont pas reconnus efficacement par eIF4E, et sont néanmoins associés aux ribosomes en cours de traduction. Leur mécanisme d'initiation de la traduction reste cependant incompris (Gribling-Burrer et al., 2017; Wurth et al., 2014). Les ARNm de sélénoprotéines comportent par ailleurs un codon UGA^{Sec} codant une sélénocystéine. Ils sont soumis à un mécanisme de recodage traductionnel. Ce mécanisme implique une structure secondaire en tige-boucle appelée SECIS (SEC Insertion Sequence) situé dans la région 3'UTR des ARNm eucaryotes. La protéine SBP2 reconnaît l'élément SECIS et recrute le facteur d'élongation spécifique eEFSec associé au Sec-$ARNt^{[Ser]Sec}$. Le SECIS et SBP2 sont également indispensables au recrutement de l'enzyme d'hyperméthylation de la coiffe.

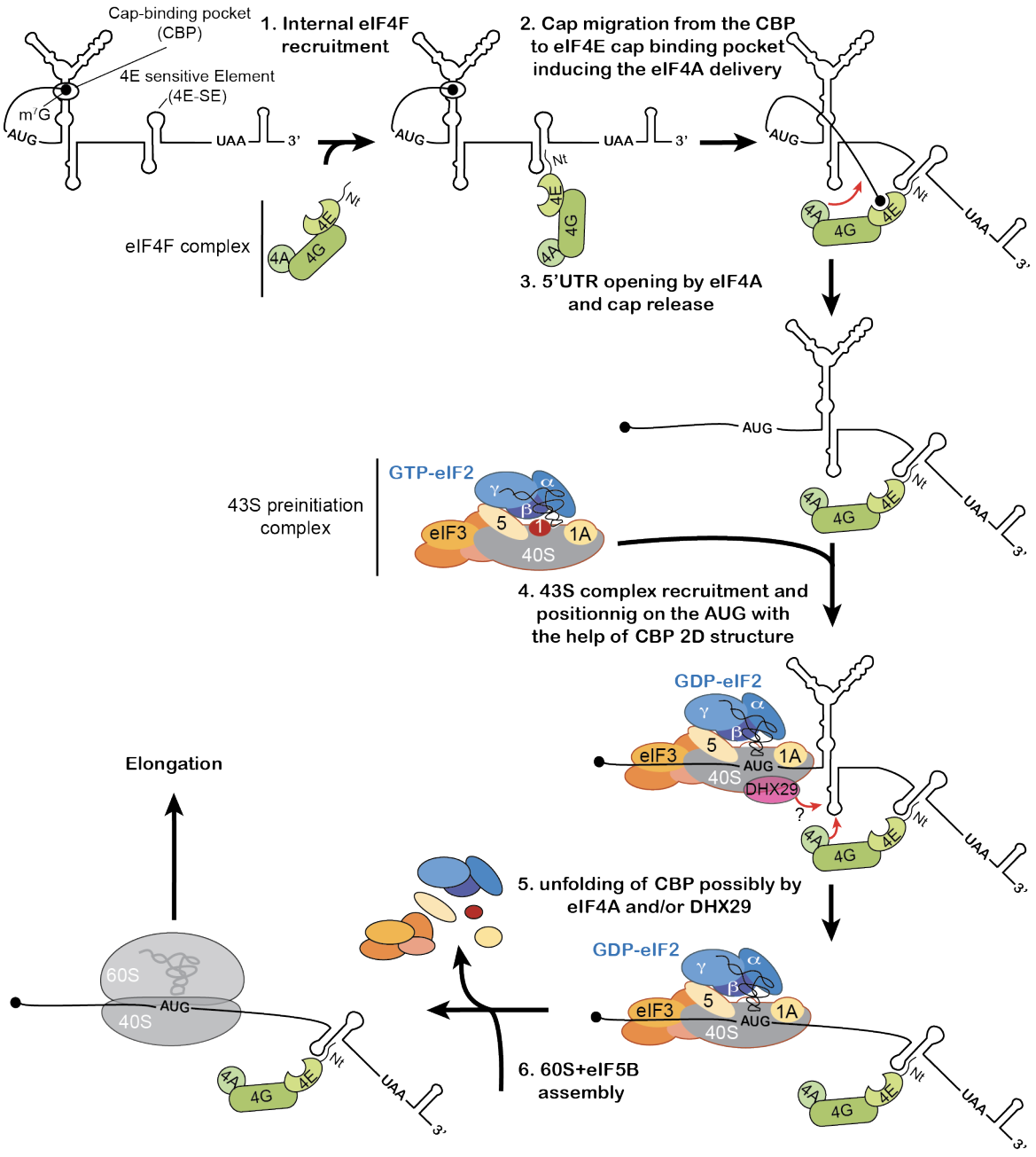


Figure 6. Modèle d'initiation de la traduction coiffe dépendante par 'tethering' du ribosome sur l'ARNm de l'histone H4 de souris. La région codante de l'ARNm d'histone H4 contient deux éléments structuraux impliqués dans le processus d'initiation de la traduction : une jonction à 3 hélices contenant une poche de fixation de la coiffe (CBP) et une double tige-boucle (4E-SE) qui recrute le facteur d'initiation eIF4E à l'intérieur de l'ARNm. La fixation du complexe eIF4F sur l'élément 4E-SE **(1)** permet à la coiffe séquestrée dans la jonction à 3 hélices d'être libérée et d'interagir avec le facteur eIF4E **(2)** induisant un changement de conformation favorable à la fixation du complexe de pré-initiation 43S sur le codon initiateur AUG de l'ARNm **(3-4)** directement sans étape de 'scanning'. L'initiation de la traduction s'achève avec la fixation de la sous-unité 60S et la formation du ribosome 80S **(5-6)** (Martin et al., 2011).

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

Le mécanisme d'initiation de la traduction par 'ribosome tethering' a été décrit dans notre laboratoire pour l'ARNm codant l'histone H4 (Martin et al., 2011). L'ARNm de l'histone H4 de souris (ARNm H4) présente des caractéristiques structurales atypiques qui diffèrent des autres ARNm cellulaires. La région 5' UTR est particulièrement courte avec seulement 9 nucléotides (Meier et al., 1989). De plus, comme tous les ARNm d'histones réplication dépendantes, son extrémité 3' UTR est dépourvue de queue poly (A) et se termine par une structure en tige-boucle (Meier et al., 1989; Piontkivska et al., 2002). Notre laboratoire a montré que la région codante de l'ARNm de l'histone H4 contient des éléments structuraux qui sont impliqués dans le processus d'initiation de la traduction. Le premier, situé 19 nucléotides en aval du codon AUG initiateur, est replié en jonction à trois hélices TWJ 'Three Way Junction'. Cette structure en TWJ séquestre la coiffe de l'ARNm dans une poche appelée 'Cap Binding Pocket' ou CBP empêchant ainsi son interaction avec le facteur d'initiation eIF4E. Le facteur eIF4E se fixe sur un deuxième élément structural situé au milieu de la phase codante, appelé 'eIF4E Sensitive Element' ou 4E-SE, similaire à l'élément 4E-SE retrouvé dans la région 3' UTR de l'ARNm de la cycline D1 (Culjkovic et al., 2006). Le facteur eIF4E interagit avec la structure 4E-SE grâce à son extrémité N-terminale. Après fixation d'eIF4F sur la structure 4E-SE via le facteur eIF4E, l'hélicase eIF4A induit des changements structuraux dans l'élément TWJ conduisant à la libération de la coiffe et à son interaction avec eIF4E. La particule 43S est ensuite directement déposée et positionnée sur le codon AUG, sans étape de 'scanning', grâce à la proximité de la structure TWJ qui est disposée à distance optimale de l'AUG (Martin et al., 2016) (voir ***Figure 6***). Le positionnement du codon AUG dans le site P du ribosome est favorisé par une interaction entre un triplet AGG de l'ARNm H4 situé en amont de la jonction triple (TWJ) et l'hélice h16 de l'ARN ribosomique 18S (Martin et al., 2016) (voir ***Figure 7***). Ce mécanisme, qualifié de 'tethering' des ribosomes sur des structures internes de l'ARNm, assure le démarrage immédiat de la traduction sans étape de 'scanning' coûteuse en temps et en énergie. Il est utilisé durant la phase S du cycle cellulaire afin de produire une quantité massive d'histones nécessaires à la formation des nucléosomes.

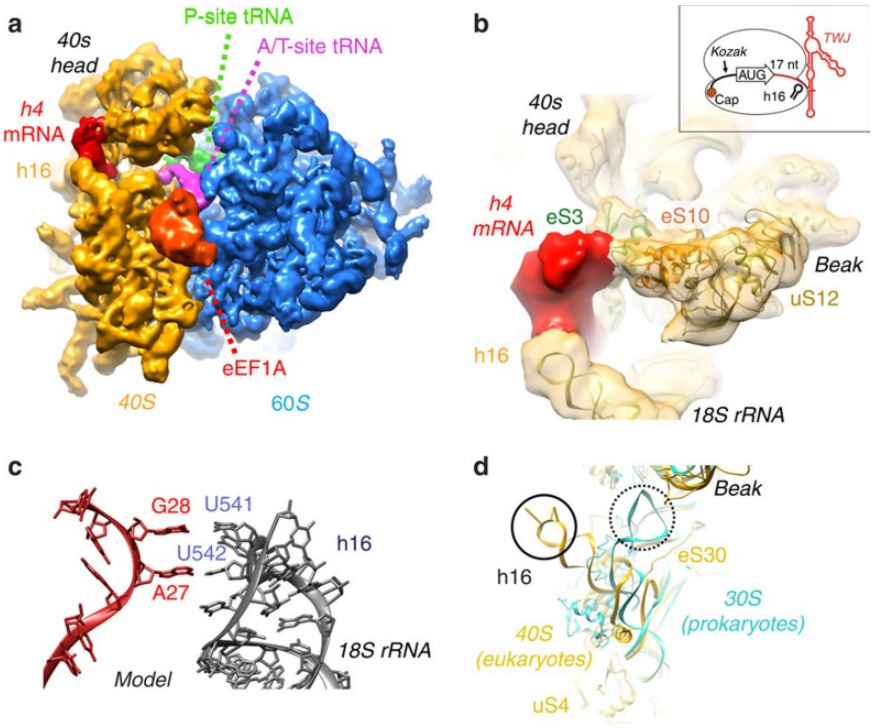


Figure 7. Structure du ribosome 80S complexé à l'ARNm H4 au stade post-initiation. (Figure issue de (Martin et al., 2016)). L'identification du codon initiateur AUG de l'ARN H4 est favorisée par une interaction entre un triplet AGG de l'ARNm localisé en amont de la jonction à 3 hélices et l'hélice h16 de l'ARN ribosomique 18S. **(a)** Modèle structural obtenu par cryo-Microscopie Électronique (cryo-ME) du complexe ARNm H4-80S bloqué au stade post-translocation par la cycloheximide. L'ARNm H4 est représenté en rouge, eEF1A en orange, l'ARNt entrant dans le site A du ribosome encore associé à eEF1A (A/T-site tRNA) est en magenta, l'ARNt présent dans le site P du ribosome est représenté en vert, la sous-unité ribosomique 60S en bleu et la sous-unité ribosomique 40S en jaune. **(b)** Vue rapprochée de la zone de contact entre l'ARNm H4 et le ribosome qui met en évidence la zone de contact de l'ARNm H4 avec la boucle apicale de l'hélice h16 (ARNr 18S) et les protéines eS3 et eS10 du ribosome. **(c)** Modélisation des interactions entre l'ARNm H4 (en rouge) et la boucle apicale de l'hélice h16 de l'ARNr 18S (en gris). **(d)** Superposition des modèles de structure du ribosome eucaryote (en jaune) et du ribosome procaryote (en bleu) mettant en évidence la très grande différence structurale de l'hélice h16. Chez les eucaryotes, l'orientation de la boucle apicale crée un site de liaison à l'ARNm (cercle plein). Chez les procaryotes l'ARNr 16S n'est pas localisé au même endroit (cercle pointillé) et h16 n'est pas orientée de telle sorte à promouvoir une interaction avec l'ARNm.

1.4. Régulation de l'initiation de la traduction

La biosynthèse des protéines est contrôlée par différents mécanismes qui permettent de l'adapter aux besoins de la cellule. L'initiation de traduction impliquant de nombreux facteurs, elle est sujette à plusieurs processus de régulation. La régulation de la traduction offre donc un autre niveau de contrôle de l'expression des gènes autre que celui de la régulation de la transcription. Ci-dessous, sont rappelés les principaux modes de contrôle de la traduction, ciblant les facteurs d'initiation ou impliquant des éléments structuraux présents sur les ARNm eux-mêmes.

1.4.1. Régulation par les facteurs d'initiation (eIFs)

Le contrôle de la traduction est généralement associé à un changement de l'état de phosphorylation des régulateurs qui vont cibler les facteurs d'initiation de la traduction. Le facteur d'initiation eIF2 est l'une des principales cibles des kinases. Dans les conditions normales de traduction, le facteur eIF2 se retrouve dans un état inactif lié au GDP (eIF2-GDP) après son utilisation. Il est ensuite réactivé par le facteur d'initiation eIF2B qui catalyse le remplacement de la molécule de GDP présente sur la sous-unité γ d'eIF2 par une molécule de GTP (Jackson et al., 2010). Dans des conditions de stress cellulaire, les kinases vont phosphoryler la sous-unité α d'eIF2. Cette modification prévient la dissociation du facteur de recyclage eIF2B, ce qui entraîne la séquestration de ces facteurs sous la forme d'un complexe inactif qui empêche la réactivation du facteur eIF2-GDP (voir ***Figure 8***).

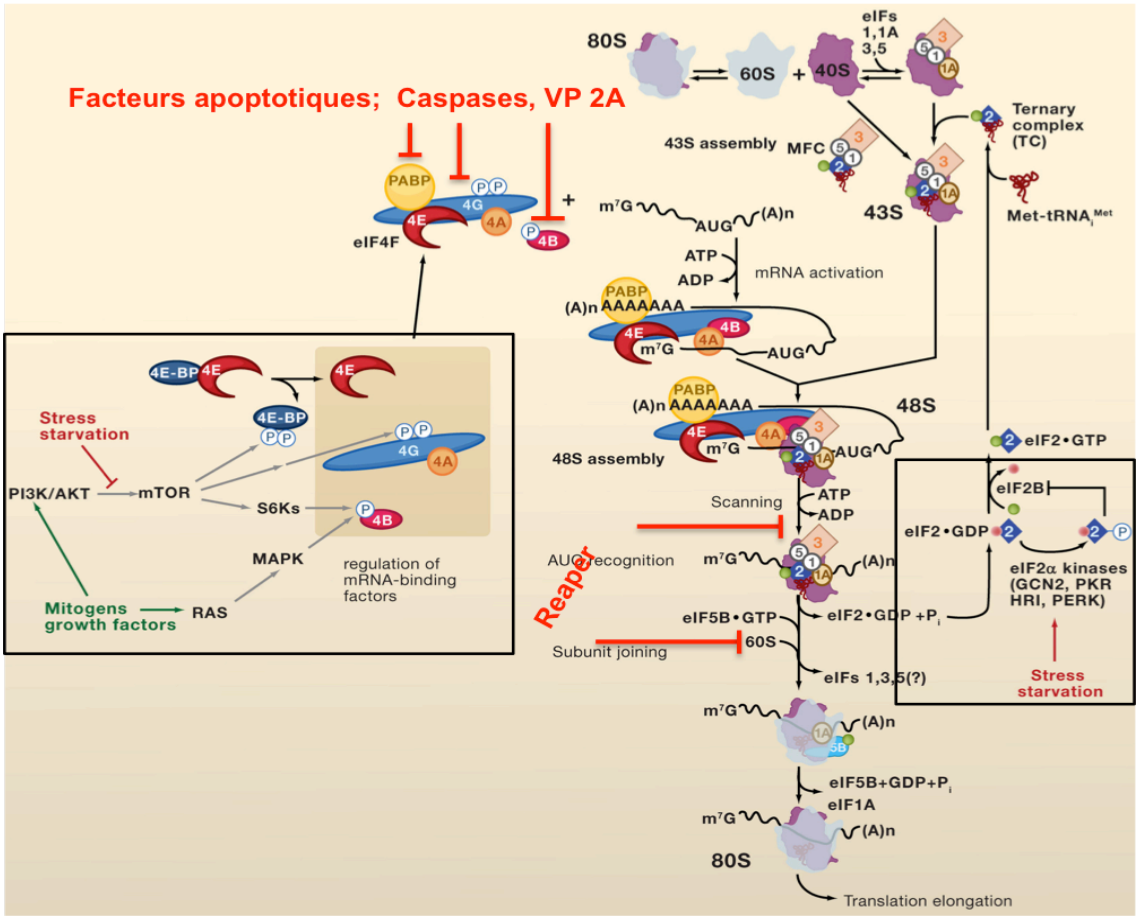


Figure 8. Exemples de régulations de l'initiation de la traduction. Figure adaptée de (Sonenberg and Hinnebusch, 2009). Dans des conditions de stress ou de privation de nutriments, la formation de complexes ternaires est réduite par la phosphorylation de eIF2α (encadré à droite). L'assemblage de eIF4F est bloqué par la liaison de 4E-BP à eIF4E, la phosphorylation de 4E-BP par mTOR dissocie 4E-BP de eIF4E. mTOR favorise également la phosphorylation de eIF4G et eIF4B soit directement, soit via S6K (encadré à gauche). Les mitogènes et les facteurs de croissance favorisent tous ces événements de phosphorylation en activant mTOR via la signalisation PI3K/Akt ou RAS/MAPK (en vert à gauche). D'autres facteurs inhibent la traduction en ciblant soit les facteurs PABP, eIF4G, eIF4B, les facteurs apoptotiques, les caspases, la protéase virale 2A, soit l'étape de scanning et

d'assemblage de la sous-unité 60S. C'est le cas du facteur pro-apoptotique de liaison au ribosome chez la drosophile, Reaper *(rpr)* (au centre de la figure) (Colón-Ramos et al., 2006).

Une autre cible de régulation est le facteur de liaison à la coiffe eIF4E qui interagit spécifiquement avec la protéine plateforme eIF4G via un motif structural de type $YX_4L\ \Phi$ (Φ pour acide aminé hydrophobique) (Mader et al., 1995). Le mécanisme de régulation d'eIF4E fait appel à des protéines inhibitrices présentant des homologies structurales avec ce domaine d'eIF4G s'associant à eIF4E. Ces protéines, appelées 4E-BP (4E-Binding Protein) se lient à eIF4E empêchant ainsi sa liaison avec le facteur eIF4G (Haghighat et al., 1996). Chez l'homme il existe trois 4E-BP appelées 4E-BP1, 4E-BP2 et 4E-BP3. La plus étudiée est la protéine 4E-BP1. Sous sa forme non phosphorylée, 4E-BP1 se lie à eIF4E empêchant la formation d'eIF4F et donc le recrutement du complexe 43S (Hershey, 2010; Hinnebusch and Lorsch, 2012; Jackson et al., 2010; Sonenberg and Hinnebusch, 2009). La phosphorylation de 4E-BP1 par la voie mTOR 'mammalian Target Of Rapamycin' réduit son affinité pour eIF4E et induit un retour au niveau de traduction normal (voir ***Figure 8***) (Hinnebusch and Lorsch, 2012; Sonenberg and Hinnebusch, 2009).

eIF4E peut également être la cible de modifications post-traductionnelles telles que la phosphorylation (Min et al., 2017; Pyronnet et al., 1999). Il a été montré que la phosphorylation du résidu S209 de eIF4E par Mnk1 active de la traduction des ARNm. La phosphorylation du résidu T55 de eIF4E par la kinase MST1 inhibe quant à elle sa capacité de liaison à la coiffe et par conséquent la traduction d'un sous-ensemble d'ARNm dépendant de la coiffe.

D'autres facteurs d'initiation peuvent être la cible des phosphorylations dans des conditions particulières avec des effets très variés : eIF4E, eIF4G,eIF2β, eIF4B, eIF4H, eIF3, eIF1, eIF5, eIF5B (pour une revue récente voir (Roux and Topisirovic, 2018)). Ils n'ont cependant pas la même importance que les protéines 4E-BP et eIF2α décrites ci-dessus qui constituent les acteurs principaux du « contrôle général » de la traduction.

1.4.2. Régulation par des facteurs homologues aux facteurs d'initiation

L'initiation de la traduction de la plupart des ARNm de mammifères est médiée par la liaison du facteur d'initiation eucaryote eIF4E sur la coiffe en 5'. Cependant, l'inactivation de eIF4E n'altère pas la traduction de nombreux ARNm coiffés, ce qui suggère qu'un autre mécanisme inconnu pourrait exister pour la traduction dépendant de la coiffe mais indépendante de eIF4E. Des études récentes ont montrées que le facteur DAP5/eIF4GII (aussi nommé NAT1 et p97), un homologue du facteur d'initiation eIF4GI dépourvu de sites de liaison à eIF4E et PABP, utilise eIF3d pour faciliter la traduction dépendante de la coiffe d'environ 20% des ARNm cellulaires (voir ***Figure 9 A***) (de la Parra et al., 2018). Des

analyses transcriptomiques et protéomiques indiquent que DAP5 est nécessaire à la traduction de nombreux facteurs de transcription et d'ARNm coiffés impliqués dans la survie, la motilité, la différentiation cellulaire et la réparation de l'ADN comme Bcl-2, Apaf-1, cIAP1, CDK1 et p53, ainsi que dans l'initiation de la traduction (Hundsdoerfer et al., 2005; Sugiyama et al., 2017; Weingarten-Gabbay et al., 2014; Yoffe et al., 2016). La traduction de ces ARNm cellulaires est médiée par l'activité IRES-dépendante de DAP5, en cas de stress cellulaire lorsque la traduction eIF4E-dépendante est inhibée (voir ***Figure 9 B***). (Alard et al., 2019; Lee and McCormick, 2006; Marash et al., 2008; Nevins et al., 2003). DAP5 possède des domaines d'interaction avec eIF4A, eIF3 et MNK1/2 (voir ***Figure 9 A***). Les complexes DAP5-eIF3d apparaissent essentiels pour la traduction indépendante d'eIF4E des ARNm coiffés, un mécanisme répandu mais jusqu'alors inconnu (de la Parra et al., 2018).

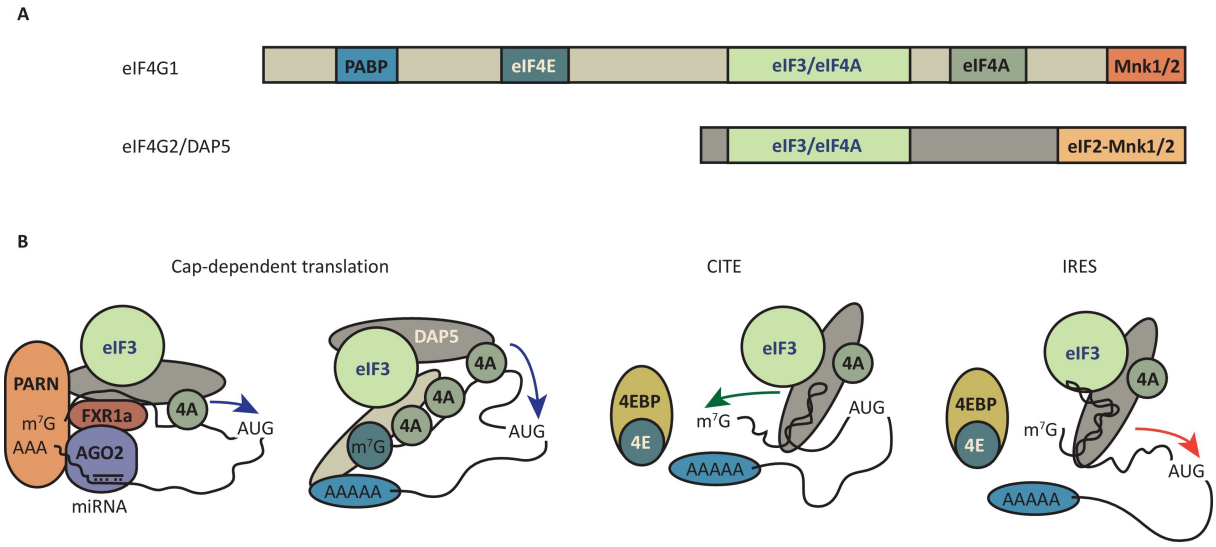


Figure 9. DAP5/eIF4G2 un homologue d'eIF4G1 aux fonctions variées. (A) Représentation schématique d'eIF4G1 et eIF4G2 (DAP5) (selon (Shatsky et al., 2018)). eIF4G2 ne présente pas les domaines de liaison à la PABP et à eIF4E. **(B)** Mécanismes possibles d'action d'eIF4G2. De gauche à droite : PARN se lie à la coiffe en 5' et forme un complexe dépendant des miARN/AGO2, FXR1a et eIF4G2, liant ainsi ce dernier à l'ARNm (Bukhari et al., 2016). eIF4G2 pourrait stimuler la traduction dépendante d'eIF4G1. DAP5 pourrait se lier spécifiquement à des structures de l'ARNm pour favoriser l'initiation de la traduction à l'extrémité 5' dépendante (CITE) ou interne (IRES).

1.4.3. Régulation par des éléments structuraux présents dans l'extension 5' non traduite des ARNm

La régulation de l'initiation de la traduction par les kinases est un processus général qui ne permet pas de cibler des ARNm en particulier. Le contrôle spécifique de l'initiation de la traduction de certains ARNm fait appel à des éléments structuraux présents dans les régions non traduites des ARNm. Compte tenu de la complexité structurale que peuvent

adopter les repliements de l'ARN une grande variété de mécanismes de régulation peut exister.

1.4.3.1. Les pseudo-nœuds dans l'ARNm de l'interféron γ

Les pseudo-nœuds sont des structures formées par des interactions entre la boucle d'une tige-boucle et une région complémentaire en aval. Ces structures ont été décrites dans la région 5' UTR de l'ARNm codant l'interféron γ chez l'homme. Une fois formées après le passage du ribosome sur l'ARNm, ces structures recrutent les protéines kinases pour phosphoryler le facteur eIF2α inhibant ainsi la traduction ultérieure de l'interféron γ (Ben-Asouli et al., 2002; Cohen-Chalamish et al., 2009).

1.4.3.2. Les éléments IRE et le métabolisme du fer

Les éléments IRE (Iron Responsive Elements) ont été décrits dans plusieurs ARNm codant des protéines impliquées dans le métabolisme du fer (Gebauer and Hentze, 2004). Ils forment de longues tiges-boucles qui peuvent être répétées. Dans l'ARNm de la ferritine, une protéine de stockage du fer, il est localisé dans l'extrémité 5' UTR alors que dans le récepteur 1 de la transferrine, une protéine de transport du fer, on le trouve dans l'extrémité 3' UTR. Lorsque le niveau de fer est bas, la protéine régulatrice 'Iron Regulatory Protein' (IRP) est désaturée de son fer et se fixe sur les éléments IRE des ARNm correspondants. Deux effets opposés sont induits: (i) un arrêt de la traduction de l'ARNm de la ferritine (donc diminution du stockage du fer) par empêchement stérique de recrutement des particules 43S (voir ***Figure 10 A***) ; (ii) une augmentation de la synthèse du récepteur de la transferrine par allongement de la demi-vie de l'ARNm correspondant (Wang and Pantopoulos, 2011). De cette façon, le fer reste à l'état libre dans la cellule et sa concentration cellulaire est maintenue constante.

1.4.3.3. Les structures de type G-quartet

Ces structures se forment en 5' UTR de certains ARNm contenant des répétitions de résidus G. Les résidus G interagissent entre eux pour former des plateaux de quatre bases G, les plateaux sont espacés de manière régulière formant un canal occupé par des cations K^+ qui stabilisent les plateaux de base. Selon leur localisation dans les 5' UTR, ces structures répriment ou activent la traduction des ARNm qui les contiennent en inhibant la liaison du complexe de pré-initiation ou en bloquant l'étape de 'scanning' du ribosome pendant l'initiation de la traduction (voir ***Figure 10 B***) (Halder et al., 2009; Kumari et al., 2008).

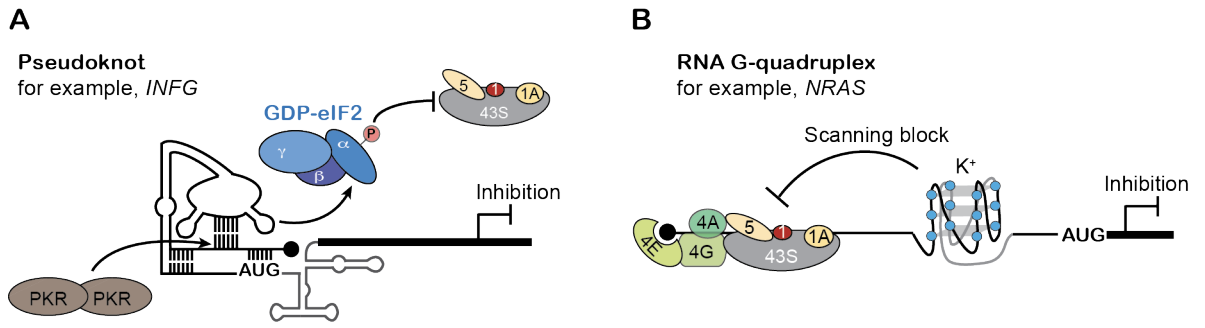


Figure 10. Régulation de l'initiation de la traduction par des structures présentes dans les régions 5' UTR des ARNm. (A) La région 5' UTR de l'ARNm de l'interféron ɣ contient une structure tertiaire de type pseudo-nœud dont les changements de conformation induisent le recrutement de kinases PKR. Les kinases PKR phosphorylent eIF2α, inhibant de ce fait la traduction (Ben-Asouli et al., 2002; Cohen-Chalamish et al., 2009). **(B)** D'autres ARNm possèdent des structures tertiaires de type G-quartet dans leurs 5' UTR. Ces structures sont dépendantes des ions (K^+). Le mécanisme moléculaire d'action des G-quartet est encore mal connu, ces structures joueraient un rôle dans la répression de la traduction (Cammas and Millevoi, 2016).

1.4.3.4. Les éléments TIE et IRES des ARNm Hox

D'autres structures présentes dans les régions 5' UTR des ARNm des gènes homéotiques ou Hox, jouent un rôle dans la régulation de la traduction pendant le développement embryonnaire. En effet, une perturbation de ces mécanismes de régulation engendre de graves conséquences pendant le développement embryonnaire (Kondrashov et al., 2011). Le contrôle de la traduction des gènes Hox se fait grâce à deux éléments structuraux, d'une part des 'Translation Inhibitory Element' (TIE) qui inhibent la traduction 'coiffe-dépendante' et d'autre part des IRES qui guident la traduction de la phase codante des gènes Hox. La régulation de ces gènes se fait par synergie entre ces deux éléments pour réguler précisément le démarrage et l'arrêt de la traduction pendant le développement embryonnaire (Xue et al., 2015). Des résultats suggèrent que les éléments TIE inhibent le recrutement du complexe de pré-initiation ou empêchent l'étape de 'scanning' pendant l'initiation de la traduction coiffe dépendante, en favorisant la traduction IRES-dépendante. Cette coordination entre les deux éléments structuraux permet une régulation spatio-temporelle fine de l'initiation de la traduction « IRES-dépendante » des gènes Hox (Xue et al., 2015).

1.4.4. Régulation de l'initiation de la traduction par le biais de la région 3'UTR

Dans plusieurs études, d'autres facteurs liant eIF4E (en plus que ceux décrits dans le paragraphe 1.4.1. ci-dessus) ont été identifiés comme régulateurs de la traduction d'ARNm spécifiques. Certains de ces facteurs se fixent simultanément à eIF4E et à la région 3'UTR des ARNm. C'est le cas de 4E-T 'eIF4E-Transporter' (également connu sous le nom de

eIF4NIF1) et de ses homologues Cup chez *D. Melanogaster* et Spn-2 (également appelé Pqn-45 ou IFET-1) chez *C. elegans*. Lors des premières étapes du développement, la traduction d'ARNm contenant des éléments de polyadénylation cytoplasmique (CPE) en 3'UTR est réprimée suite au déplacement de eIF4G par 4E-T. Ce dernier est recruté vers l'ARNm par la protéine de liaison à CPE (CPEB) (voir ***Figure 11***) (Fernández-Miranda and Méndez, 2012). C'est un motif $YX_4L\phi$, présent dans 4E-T et similaire à celui présent dans eIF4G qui empêche de manière compétitive la liaison productive de eIF4E à eIF4G, réduisant ainsi la synthèse protéique (Sonenberg and Hinnebusch, 2009).

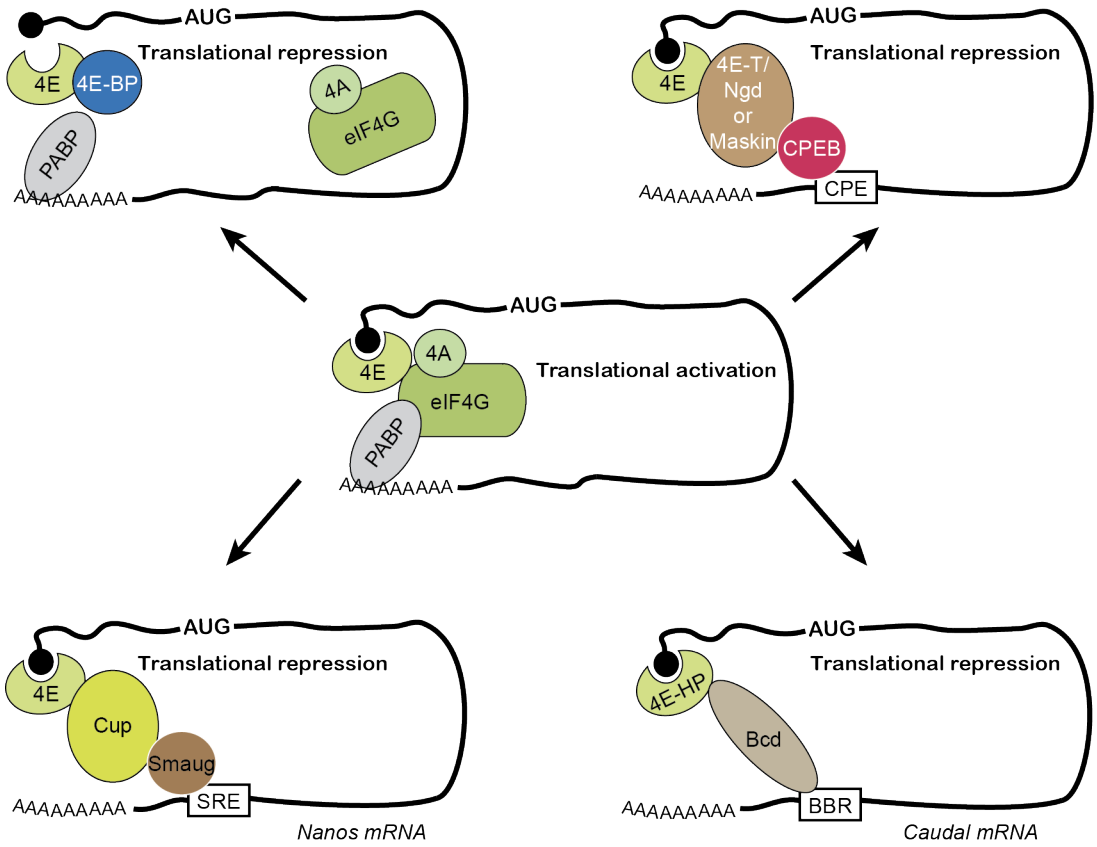


Figure 11. Différents mécanismes de répression de la traduction. Le mécanisme le plus général est engendré par les membres d'une famille de petites protéines, appelées protéines de liaison à 4E (4E-BPs), qui entrent en compétition avec eIF4G pour l'interaction avec eIF4E empêchant la fixation de ce dernier sur la coiffe. La répression de la traduction peut résulter de divers processus visant à empêcher la fixation d'eIF4F sur la coiffe et la circularisation canonique 5'-3' des ARNm (schéma central). Les trois modèles présentés ici se basent sur la présence d'un élément régulateur dans l'ARNm (CPE, SRE ou BBR ci-dessus) et font appel à divers facteurs protéiques homologues d'eIF4E (4E-HP). Par exemple, la traduction d'ARNm contenant un élément de polyadénylation cytoplasmique (CPE) est réprimée par le déplacement d'eIF4G par la Neuroguidin (Ngd) / Maskin/ 4E-T, qui est recrutée sur l'ARNm par le biais de son interaction avec la protéine de liaison au CPE (CPEB). La traduction de *nanos* est inhibée par la fixation d'eIF4E à l'élément de réponse Smaug (SRE) via Smaug et Cup. La protéine Bicoïd (Bcd) inhibe spécifiquement la traduction de *caudal* chez la drosophile en se liant simultanément à l'élément BBR de l'ARNm et à l'homologue d'eIF4E, 4E-HP (Sonenberg and Hinnebusch, 2009).

D'autres mécanismes similaires ont été identifiés. Chez la drosophile, la traduction de l'ARNm *oskar* est inhibée dans la région antérieure de l'embryon par le recrutement de eIF4E vers un élément de séquence BRE (Bruno Response Element) en 3'UTR via son interaction avec les protéines Bruno et Cup (voir ***Figure 11***). La levée d'inhibition pourrait résulter d'une phosphorylation de la protéine Cup à la manière des 4E-BP ou d'un allongement de la queue poly (A) de l'ARNm *oskar* déstabilisant le complexe eIF4E/Cup (Nakamura et al., 2004; Wilhelm and Smibert, 2005). L'ARNm *caudal* est quant à lui n'est exprimé que dans la région postérieure (Johnstone and Lasko, 2001). L'inhibition de sa traduction fait intervener la protéine Bicoid qui se lie simultanément à la region BBR (Bicoid-binding region) en 3'UTR et à la proteine 4EHP, homologue de eIF4E, qui reconnait la coiffe m^7G et entre en compétition avec eIF4E (Cho et al., 2005). Les protéines 4E-T, Cup et Bicoid contrôlent ainsi l'accès des ribosomes à la coiffe en 5' de ces ARNm spécifiques depuis leur site de liaison en 3' UTR (Kamenska et al., 2014a; Rhoads, 2009).

Chez les femelles de drosophiles, la répression traductionnelle de l'ARNm *msl-2* (male-specific lethal 2) est une étape essentielle de la régulation de l'expression des gènes liés au chromosome X. Cette répression est orchestrée par la protéine Sex-lethal (SXL), qui se lie à la region 3'UTR de *msl-2* et recrute la protéine UNR (upstream of N-ras) afin bloquer la liaison du complexe de pré-initiation 43S à l'extrémité 5 'de l'ARNm. SXL bloque également l'étape de scanning des ribosomes dans le 5′UTR (Beckmann et al., 2005). Des travaux récents ont identifié la protéine Hrp48, un nouveau cofacteur de SXL, également capable de se lier en 3'UTR de l'ARNm *msl-2* et d'interagir directement avec la sous-unité eIF3d du complexe d'initiation eIF3 (Szostak et al., 2018).

Enfin, la régulation de la traduction en 3'UTR peut également se faire par le biais de microARNs (miARNs) qui régulent l'expression des gènes à différents niveaux en stimulant la dégradation de leurs cibles mais aussi en réprimant leur traduction. Dans ce cas les miARN ciblent exclusivement les étapes d'initiation de la traduction avant le recrutement de la sous-unités ribosomique 60S (Ricci et al., 2013). Des études ont montré que 4E-T participe à la répression de la traduction médiée par les micro-ARN et était associée à plusieurs sous-unités du complexe CCR4-NOT, un complexe protéique impliqué dans la production et la dégradation des ARNm (Collart, 2016; Kamenska et al., 2014b). CCR4-NOT est un important effecteur de miRISC (Jonas and Izaurralde, 2015). La protéine CNOT1, un des membres du complexe CCR4-NOT, se lie à l'ARN hélicase DDX6 permettant ainsi l'inhibition médiée par les microARN (voir ***Figure 12***) (Jonas and Izaurralde, 2015; Rouya et al., 2014). DDX6 joue un rôle dans la répression de la traduction et stimule l'élimination de la coiffe. C'est aussi un partenaire direct de 4E-T ; ce réseau d'interactions pourrait donc être à la base de la répression de l'initiation de la traduction par miRISC (Kamenska et al., 2016; Minshall et al., 2007; Ozgur et al., 2015).

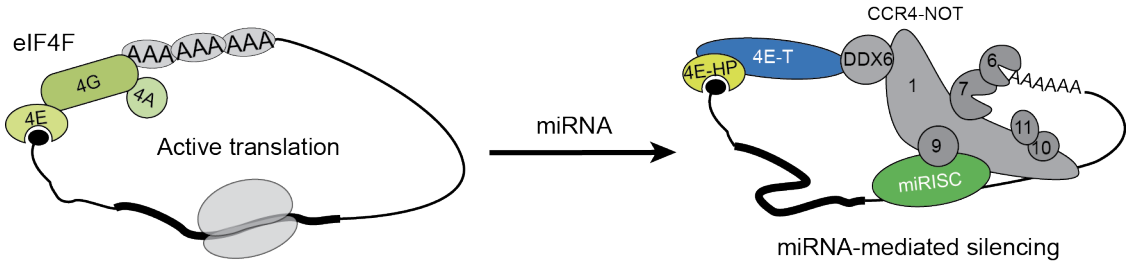


Figure 12. Modèle de répression de la traduction par les miARN par l'intermédiaire de 4E-HP. Le recrutement de 4E-HP sur l'ARNm cible du miARN via CCR4 – NOT / DDX6 / 4E-T favorise sa liaison à la coiffe. L'assemblage de ce complexe est susceptible d'initier la formation d'une structure en boucle fermée (droite) ressemblant à la conformation de l'ARNm en boucle fermée coiffe-queue poly (A) impliquant une interaction eIF4G / PABP (gauche) (Chapat et al., 2017).

1.4.5. Régulation par les uORF

L'initiation de la traduction peut être contrôlée par des séquences situées en amont d'une phase codante (ORF) appelées uORF pour 'upstream Open Reading Frame'. Les uORF sont constituées d'un codon d'initiation et d'un codon stop séparés par au moins un codon sens. Les uORF peuvent chevaucher ou se terminer avant la phase codante principale. Des études bioinformatiques ont permis de détecter jusqu'à 13%, 44% et 49% d'uORF chez la levure, la souris et l'homme respectivement (Calvo et al., 2009). L'utilisation de ces uORF a été validée par des expériences de 'ribosome profiling' confirmant leur traduction *in vivo* (Johnstone et al., 2016) et leur fonction potentielle dans des mécanismes de régulation de la phase codante principale. Certains ARNm sont particulièrement enrichis en uORF, incluant des ARNm de facteurs de croissance, de facteurs de transcription et de proto-oncogènes (Davuluri, 2000).

D'un point de vue mécanistique, dans le modèle classique de 'scanning' de la 5' UTR d'un ARNm, la particule 43S glisse le long de l'ARNm jusqu'à ce qu'elle rencontre un premier codon AUG. Dans la plupart des cas, et en l'absence de stress cellulaire, la traduction des uORF entraine une diminution de la traduction de l'ORF principale par un effet de compétition pour la particule 43S (Calvo et al., 2009). Cependant en conditions de stress cellulaire (Spriggs et al., 2010), on peut observer une augmentation de l'expression de certains ARNm comme GCN4 chez la levure (Hinnebusch, 2005) et de nombreux autres ARNm qui ont été répertoriés dans une banque de données (Wethmar et al., 2014).

1.4.6. Régulation par le ribosome

Il a été proposé que le ribosome lui même pourrait, grâce à sa composition particulière en protéines ribosomiques, réguler l'expression de certains ARNm conduisant à la notion de 'code ribosomique'. Par exemple l'expression des ARNm des protéines Hox décrites ci-dessus est favorisée par la protéine ribosomique RPL38 et son absence entraine des anomalies de développement chez la souris (Kondrashov et al., 2011). Un autre exemple est la protéine ribosomique RACK1 ou 'Receptor for Activated C Kinase1' qui contrôle une catégorie d'ARN possédant des IRES afin de stimuler l'initiation de la traduction IRES-dépendante (voir ***Figure 13***) (Majzoub et al., 2014). Le proto-oncogène c-myc entre également dans cette catégorie d'ARNm. Dans ce cas, la protéine cytoplasmique Kindlin-3 est capable d'interagir avec le ribosome via RACK1 et de réguler l'expression de c-myc nécessaire à la prolifération de cellules de leucémie myéloïde chronique (Qu et al., 2016; Stoneley et al., 1998).

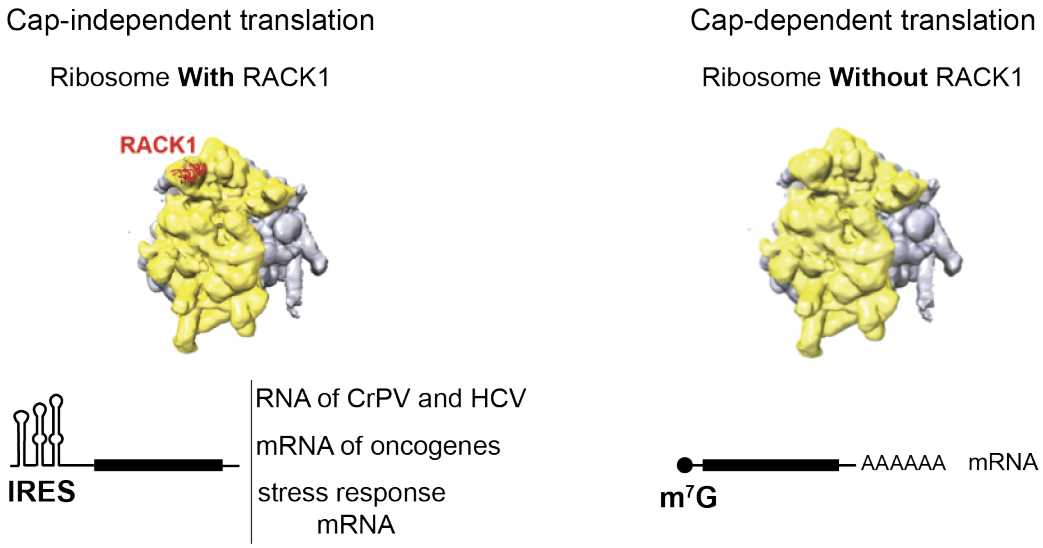


Figure 13. La protéine du ribosome RACK1 participe à la traduction sélective des ARNm en fonction des conditions physiologiques de la cellule. À droite) RACK1 n'influence pas la traduction coiffe-dépendante des ARNm cellulaires qui sont exprimés de manière constitutive dans la cellule. **(À gauche)** En revanche, RACK1 est indispensable à la traduction spécifique des ARNm possédant des IRES (ARN viraux de CrPV, HCV et ARNm cellulaires codant pour des oncogènes) (Majzoub et al., 2014). La traduction de ces ARNm viraux est augmentée lors d'une infection virale. La traduction des ARNm oncogènes est augmentée pendant la propagation des cancers.

H. sapiens			
Subunit	Domains	Named	M.W. (kDa)
eIF3a	PCI, Spectrin	**p170**	166.6
eIF3b	WD40, RRM	**p116**	92.5
eIF3c	PCI	**p110**	105.3
eIF3d	CBP	**p66**	64.0
eIF3e	PCI	**p48**	52.2
eIF3f	MPN	**p47**	37.6
eIF3g	RRM, Zn finger	**p44**	35.6
eIF3h	MPN	**p40**	39.9
eIF3i	WD40	**p36**	36.5
eIF3j	-	**p35**	29.1
eIF3k	PCI	**p28**	25.1
eIF3l	PCI	**p67**	66.7
eIF3m	PCI	**GA17**	42.5

Tableau 1. Le tableau représente les poids moléculaires ainsi que les différents domaines présents dans les sous-unités du complexe eIF3 humain (adapté de (Valášek et al., 2017))

2. Le facteur d'initiation eIF3

2.1. Généralités

Chez la plupart des eucaryotes supérieurs, le facteur eIF3 est le plus grand et le plus complexe des facteurs d'initiation. Il est formé de 13 sous-unités nommées eIF3a à eIF3m de poids moléculaire d'environ 800 kDa (Browning et al., 2001; Smith et al., 2016). eIF3 est impliqué dans l'initiation de la traduction mais également la terminaison, le recyclage des sous-unités ribosomiques et le recodage de codons stop. Lors de l'initiation de la traduction, eIF3 joue plusieurs rôles : il favorise l'assemblage du complexe de pré-initiation 43S (ou PIC, qui comprend la sous-unité ribosomique 40S, le Met-tRNA$_i^{Met}$ et plusieurs eIF), il s'associe avec l'ARNm lié au complexe eIF4F (eIF4E, eIF4G et eIF4A) et permet la mise en place du 'scanning' ainsi que la reconnaissance du codon AUG. En plus de son rôle général dans l'initiation de la traduction, eIF3 agit également comme un activateur ou un répresseur de la traduction en se liant à des structures d'ARN dans les régions 5' UTR et 3' UTR non traduites de certains ARNm spécifiques. eIF3 orchestre la traduction canonique et non-canonique, c'est un acteur majeur et central de la traduction dans les cellules eucaryotes. Par ailleurs, eIF3 fait partie de réseaux protéiques complexes et est impliqué dans différents processus biologiques. La dérégulation de plusieurs sous-unités d'eIF3 a pu être corrélée avec des cancers mais également avec d'autres pathologies.

Mes travaux ont contribué à décortiquer l'implication du facteur eIF3 dans le mécanisme d'initiation de la traduction atypique de l'histone H4.

2.2. Assemblage et structure du complexe eIF3

Chez les eucaryotes supérieurs, les 13 sous-unités (eIF3a-m) qui constituent le facteur eIF3, ne sont pas identiques et varient en taille de 170 à 25 kDa (Browning et al., 2001; Damoc et al., 2007) (voir ***Tableau 1***). Elles s'assemblent d'une manière interdépendante pour former le complexe eIF3 (Smith et al., 2016).

En revanche, chez la levure *S. cerevisiae*, le facteur eIF3 est beaucoup plus simple: il se compose uniquement des orthologues de six sous-unités de mammifère eIF3a, b, c, g, i et j (Phan et al., 1998). Ces dernières sont toutes indispensables à la traduction *in vivo* (Greenberg et al., 1998; Hanachi et al., 1999; Naranda et al., 1997; Valásek et al., 1998) mais semblent jouer des rôles distincts dans les mécanismes de régulation de la traduction notamment en interagissant avec des facteurs d'initiation différents de ceux des eucaryotes multicellulaires (Querol-Audi et al., 2013).

Plusieurs structures de complexes d'initiation eucaryotes incluant eIF3 ont été déterminées par cryo-microscopie électronique (cryo-EM) et cristallographie fournissant des informations essentielles sur l'organisation du complexe à la fois chez la levure et l'homme mais également sur les mécanismes moléculaires et les changements conformationnels des complexes ribosomiques lors la liaison de l'ARNm et la reconnaissance des codons AUG (Aylett et al., 2015; Erzberger et al., 2014; des Georges et al., 2015; Hashem et al., 2013a). Les résultats de ces études sont résumés ci-dessous.

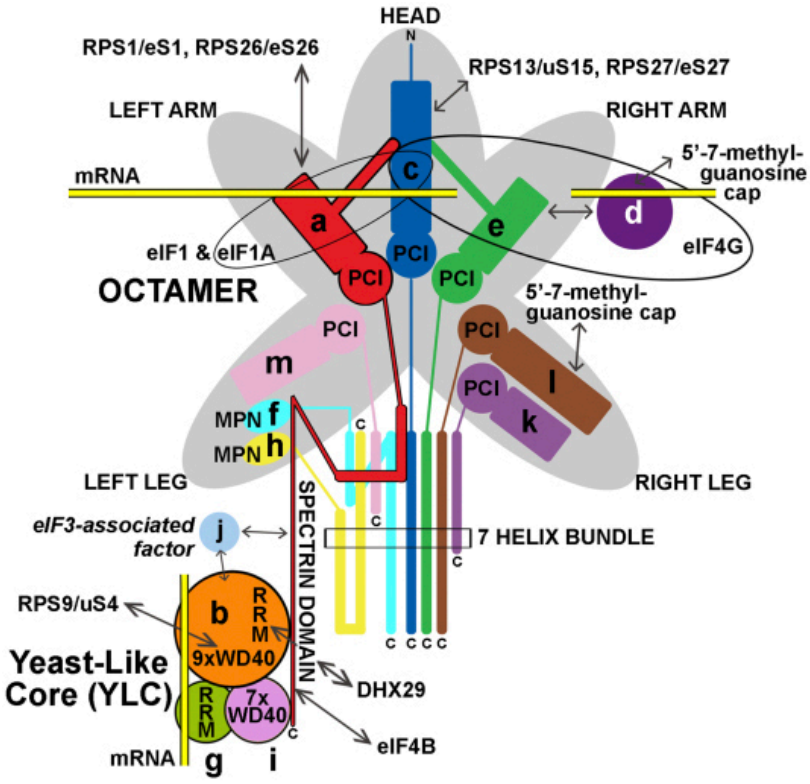


Figure 14. Représentation schématique du facteur d'initiation eIF3 et des interactions entre ses 13 sous-unités. (Figure issue de (Valášek et al., 2017)). Les sous-unités a, c, e, k, l, m, f et h du facteur d'initiation eIF3 constituent le cœur octamérique (entouré en gris). Les sous-unités b, d, g, i et j sont localisées en périphérie du cœur. Elles participent à la liaison du facteur eIF3 à l'ARNm grâce à leurs motifs RRM et WD40. Les sous-unités eIF3a, b, i et g sont rattachées au noyau octamérique eucaryotique *via* l'extrémité C-terminale d'eIF3a. Elles forment le complexe appelé 'Yeast-Like Core' (YLC) qui regroupe quatre des cinq sous-unités constituant le core fonctionnel minimal d'eIF3 chez la levure. La sous-unité eIF3j interagit avec eIF3a et b (représenté par des flèches), elle serait associée de manière transitoire à eIF3 pour stimuler son activité. Le facteur eIF3 interagit avec certains facteurs d'initiation, des protéines ribosomiques et l'hélicase DHX29. Ces interactions sont représentées par des flèches.

Chez les mammifères les 13 sous-unités sont organisées en deux modules : les protéines du cœur structural (a, c, e, f, h, l, k et m) et les sous-unités périphériques (b, d, g, i et j) (voir ***Tableau 1***).

Le cœur structural est un noyau octamérique présent également dans deux autres complexes protéiques : le protéasome 26S et le signalosome COP9 qui régulent la synthèse et la dégradation des protéines (Pick et al., 2009). Ce noyau octamérique structuralement conservé est nommé PCI/MPN pour 'Protéasome, COP9/signalosome, eIF3 et Mpr1/Pad1 N-terminal' (Ellisdon and Stewart, 2012; Pick et al., 2009). Les six sous-unités a, c, e, k, l et m d'eIF3 contiennent les domaines PCI, qui consistent en des répétitions hélicoïdales N-terminales suivies d'un domaine WHD (Winged helix domain) qui permet la polymérisation des domaines PCI (Ellisdon and Stewart, 2012). Les deux autres sous-unités du noyau octamérique (f et h) contiennent des domaines MPN qui sont des tonneaux-β entourés d'hélices-α et des brins-β, elles ont pour fonction de promouvoir l'assemblage d'autres protéines ou sous-unités sur le noyau octamérique (voir ***Figure 14***) (Enchev et al., 2010; Pena et al., 2007). L'ensemble des sous-unités PCI/MPN constituant la structure octamérique adoptent une structure à cinq lobes conservée dans le protéasome 26S et le signalosome (Ellisdon and Stewart, 2012; Pick et al., 2009). Ce module central d'eIF3 ne peut être surproduit et purifié de façon stable que lorsque les 8 sous-unités formant le noyau octamérique sont co-exprimées dans *E. coli* (Sun et al., 2011).

Les quatre sous-unités périphériques (b, d, g et i) sont liées au noyau PCI/MPN de manière flexible (Hashem et al., 2013b, 2013a; Sun et al., 2011). Trois de ces sous-unités, eIF3b, g et i, forment un module séparé qui s'attache au noyau PCI/MPN par l'intermédiaire du domaine C-terminal de la sous unité eIF3a (des Georges et al., 2015; Herrmannová et al., 2012) alors que la sous unité eIF3d interagit avec eIF3e (Wagner et al., 2016). Ce module est appelé 'Yeast-Like Core' (YLC) en référence au complexe eIF3 de levure qui n'est formé que de six sous-unités, dont eIF3a, b, c, g et i constituent le cœur fonctionnel minimal (Valášek et al., 2017) (voir ***Figure 14***). Les sous-unités b et g possèdent des motifs de liaison à l'ARNm appelés 'RNA Recognition Motif' (RRM) (ElAntak et al., 2007). Les sous-unités b et i contiennent également des motifs en hélice-β formés par la répétition des acides aminés tryptophane et acide aspartique (motif WD40), ces derniers sont impliqués dans la coordination de l'assemblage des complexes multi-protéiques. Les sous-unités périphériques d'eIF3 forment un réseau d'interactions encerclant la sous-unité ribosomique 40S et relient l'entrée du canal de l'ARNm à sa sortie (Aitken et al., 2016; des Georges et al., 2015; Hashem et al., 2013a; Sun et al., 2011). Le noyau octamérique et le module YLC du facteur eIF3 sont ainsi présents sur deux faces opposées du ribosome 40S, de telle sorte à l'envelopper presque totalement (voir ***Figure 15***).

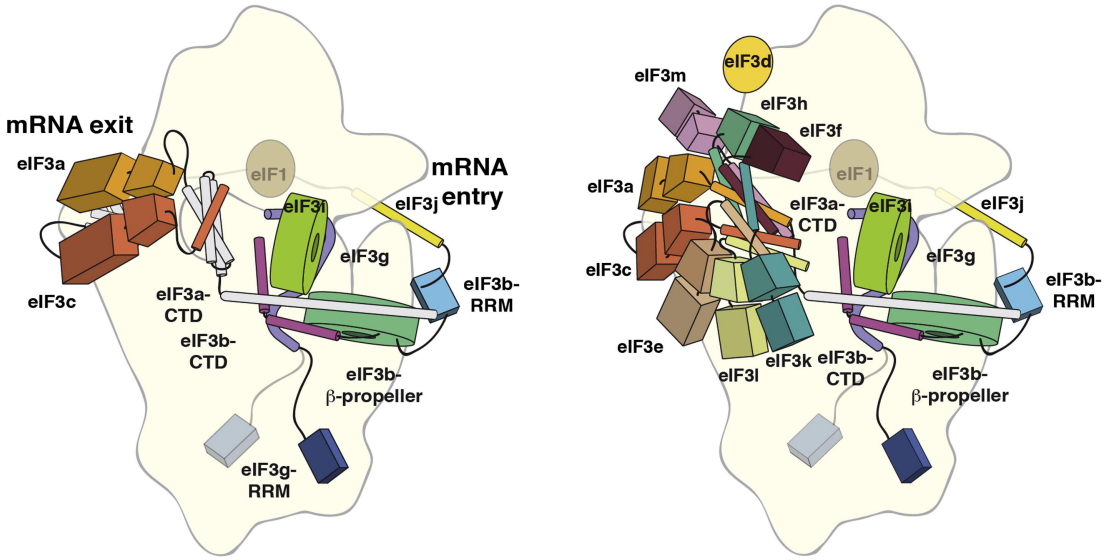


Figure 15. Modèles d'organisation des sous-unités du facteur d'initiation eIF3 autour de la sous-unité 40S du ribosome de deux types d'eucaryotes (figure issue de (Erzberger et al., 2014). Le complexe 43S de levure est représenté à droite et celui des mammifères à gauche. Les treize sous-unités d'eIF3 sont localisées de part et d'autre de la petite sous-unité 40S du ribosome. Les sous-unités périphériques b, d, g, i et j forment un réseau d'interactions encerclant la sous-unité ribosomique 40S et relient l'entrée du canal de l'ARNm à sa sortie. eIF3j occupe le même emplacement que l'ARNm dans le complexe 43S ; elle interagit dans le canal d'entrée de l'ARNm de la sous-unité ribosomique 40S (Aitken et al., 2016; des Georges et al., 2015; Hashem et al., 2013a).

Chez la levure, l'interaction entre les sous-unités eIF3b et eIF3a est essentielle à l'assemblage du complexe eIF3 car elle sert de noyau de nucléation autour duquel s'assemblent les autres sous-unités de manière ordonnée (Khoshnevis et al., 2014; Valásek et al., 2002). Chez l'homme cette interaction est strictement requise pour la formation du complexe *in vivo* (Wagner et al., 2016). En effet, la première étape d'assemblage de l'octamère humain consiste en l'ajout simultané d'eIF3c, f et m (et peut-être eIF3h) au noyau de nucléation eIF3a/3b (Wagner et al., 2014, 2016).

De nombreux travaux suggèrent que la sous-unité eIF3j ne ferait pas partie intégrante du complexe eIF3 chez les levures et les mammifères et doit être considérée comme un facteur associé (Fraser et al., 2004). eIF3j renforcerait les interactions entre les sous-unités périphériques d'eIF3 et les autres facteurs d'initiation ainsi qu'avec la sous-unité ribosomique 40S lors de la formation du complexe de pré-initiation 43S. eIF3j occupe le même emplacement que l'ARNm dans le complexe 43S, elle interagit avec le canal d'entrée de l'ARNm de la sous-unité ribosomique 40S (Fraser et al., 2007) (voir ***Figure 15***), elle interagit également avec eIF1A réduisant ainsi l'affinité de la sous-unité 40S pour l'ARNm. eIF3j joue également un rôle indépendant dans la biogenèse du ribosome 40S (Valásek et al., 2001). Des études récentes suggèrent que eIF3j serait recruté sur le ribosome pendant la phase de terminaison de la traduction et jouerait un rôle dans la dissociation de la sous-unité

ribosomique 60S. Son inactivation par siARN mène à une ré-initiation de la traduction sur l'extrémité 3' UTR de l'ARNm (Young and Guydosh, 2019).

2.3. Implication d'eIF3 dans les mécanismes d'initiation de la traduction canonique et non-canonique

Plusieurs études ont souligné l'importance du facteur eIF3 durant la phase d'initiation de la traduction mais eIF3 contribue aussi à contrôler la synthèse protéique tout au long du cycle de traduction (Valášek et al., 2017).

2.3.1. Fonctions d'eIF3 dans l'initiation de la traduction canonique

Le facteur d'initiation eIF3 est impliqué dans presque toutes les étapes d'initiation de la traduction. Des expériences biochimiques et génétiques ont montré que le rôle principal d'eIF3 est d'orchestrer la formation des complexes de pré-initiation 43S et 48S et de coupler l'initiation de la traduction à la transcription et à l'export des ARNm (Harel-Sharvit et al., 2010). eIF3 empêche l'association des sous-unités ribosomiques (Hashem et al., 2013a; Kolupaeva et al., 2005). Il stimule également le recrutement et la stabilisation du complexe ternaire eIF2-GTP-Met-ARNt_i^{Met} et de l'ARNm vers le complexe 43S (Valásek et al., 2004). Pendant l'attachement du complexe 43S à l'ARNm, eIF3 interagit avec le facteur eIF4G présent dans le complexe eIF4F. Cette interaction implique des contacts entre les trois sous-unités eIF3c, d et e et une région interne d'eIF4G (1015-1118) (Korneeva et al., 2001; Villa et al., 2013). A son tour, eIF4G interagit avec les protéines de liaison à la queue poly (A) de l'ARNm (PABP) afin de circulariser l'ARNm (Kahvejian et al., 2001). Les PABP se lient aux facteurs Paip1 (PABP-interacting protein 1) et eIF4B (Méthot et al., 1996), qui à leur tour se lient avec eIF3 stimulant ainsi la liaison de l'ARNm au complexe de pré-initiation 43S (Bushell et al., 2001; Martineau et al., 2008, 2014). Enfin, eIF3 contrôle également la processivité du processus de scanning ainsi que la fidélité de la sélection du codon d'initiation AUG (Hinnebusch, 2006; Jackson et al., 2010; Pestova and Kolupaeva, 2002; Valásek, 2012). Par ailleurs, il a été montré que les quatre sous-unités d'eIF3 a, b, c, et g sont capables d'interagir directement avec certains ARNm (Lee et al., 2015).

La diversité des rôles d'eIF3 dans l'initiation de la traduction provient probablement du fait qu'il interagit directement avec plusieurs facteurs d'initiation et coordonne leur assemblage fonctionnel à la surface de la petite sous-unité ribosomique (Hinnebusch, 2017; Valásek, 2012). L'ensemble des interactions caractérisées à ce jour est résumé dans le ***Tableau 2*** et la ***Figure 16***. Les interactions majeures sont décrites ci-dessous.

eIF3 subunit	Interacting subunit	Other interactions
eIF3a	eIF3b, eIF3c, eIF3e, eIF3f, eIF3j, eIF3k, eIF3l, eIF3m	40S, eIF1, eIF4G, eIF4B, mRNA, SHC, Raf-1, DHX29, p190A RhoGAP, RRM2
eIF3b	eIF3a, eIF3c, eIF3g, eIF3i, eIF3j	40S, DHX29
eIF3c	eIF3a, eIF3b, eIF3e, eIF3h, eIF3j, eIF3k, eIF3l, eIF3m	40S, eIF1, eIF4G, schwannomin, merlin
eIF3d	eIF3e	eIF4G, 5' cap
eIF3e	eIF3a, eIF3c, eIF3d, eIF3l	eIF4G, HERC2
eIF3f	eIF3a, eIF3h, eIF3m	eIF4G, mTOR, Notch
eIF3g	eIF3b, eIF3i, eIF3k	Pelota, AIF
eIF3h	eIF3c, eIF3f, eIF3m	40S, HERC2
eIF3i	eIF3b, eIF3g	TGF-β, TRIP-1
eIF3j	eIF3a, eIF3b, eIF3c	eIF1A, 40S
eIF3k	eIF3a, eIF3c, eIF3g, eIF3j, eIF3l	cyclin D3, keratin
eIF3l	eIF3a, eIF3c, eIF3e, eIF3k	-
eIF3m	eIF3a, eIF3c, eIF3f, eIF3h	HERC2

Tableau 2. Composition du facteur eIF3 chez les mammifères et liste des interactions établies avec les sous-unités d'eIF3 et les autres constituants de la machinerie de traduction **(issu de (Gomes-Duarte et al., 2018)).**

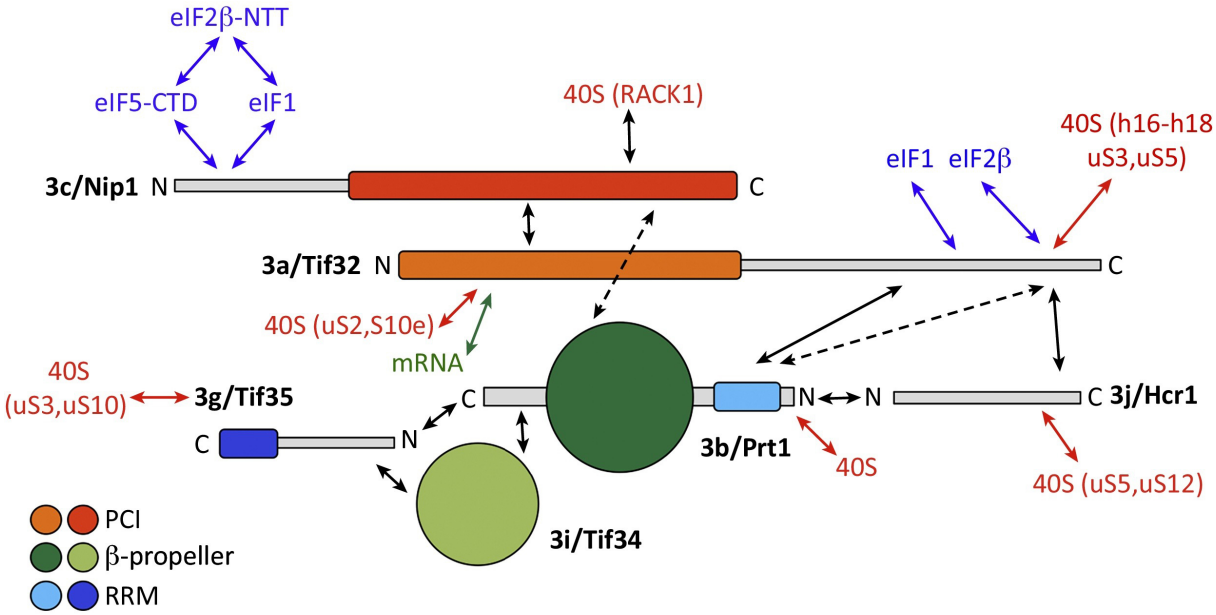


Figure 16. Interactions des sous-unités d'eIF3 de la levure *S. cerevisiae* (figure issue de (Hinnebusch, 2017). Représentations linéaires des sous-unités d'eIF3 interagissant avec d'autres facteurs d'initiation présents dans le complexe MFC 'Multi-Factor Complex' (en bleu), la sous-unité 40S (et les constituants spécifiques du 40S entre parenthèses) (en rouge) et l'ARNm présent au niveau du canal de sortie (en vert). Les domaines (PCI), hélice-β et RRM, au sein du facteur eIF3 sont indiqués par des rectangles ou des cercles colorés (pour une revue des interactions (Valásek, 2012)).

2.3.1.1. Interaction d'eIF3 avec les facteurs d'initiation eIF1 et eIF5 :

eIF3 lie le côté du ribosome 40S accessible au solvant et joue un rôle majeur dans le recrutement des facteurs d'initiation eIF1 et eIF5 vers le centre de décodage (Asano et al., 2000; Hinnebusch, 2017; Hinnebusch et al., 2004; Obayashi et al., 2017; Sokabe et al., 2012). C'est le domaine N-terminal (NTD) d'eIF3c qui est impliqué dans ces interactions. Celui-ci peut être divisé en trois sous domaines: 3c0 (1-58), 3c1 (59-87) et 3c2 (88-163) (voir

Figure 17). La partie N-terminale 3c0 lie fortement eIF5 alors que les domaines 3c1 et 3c2 forment un complexe stœchiométrique avec eIF1. 3c2 fait face à la protéine 40S uS15 / S13 pour ancrer eIF1 au complexe de pré-initiation (PIC) lors de l'étape de scanning. Obayashi et collaborateurs ont proposé que le domaine N-terminal de eIF3c pourrait s'étendre au cœur du centre de décodage, non seulement pour ancrer le facteur eIF1, mais également pour faciliter sa libération lors de la sélection du codon AUG. eIF5 semble jouer un rôle dans ce phénomène (Obayashi et al., 2017). L'altération de l'interaction entre eIF3 et eIF5 affecte la formation du complexe 43S, alors que celle de l'interaction avec eIF1 empêche la reconnaissance du codon AUG initiateur (Karásková et al., 2012; Querol-Audi et al., 2013). Des travaux récents de reconstitution *in vitro* du complexe eIF3 de levure combinés à des expériences de pontage et modélisation 3D suggèrent au contraire une structure très compacte d'eIF3 (Zeman et al., 2019). La liaison d'eIF1 et eIF5 (voir ***Figure 18***) semble même compacter davantage cette structure globulaire. Le site de liaison d'eIF5 a été localisé dans la région des sites P et E de la sous-unité 40S (Zeman et al., 2019). Les auteurs suggèrent qu'eIF3 subit d'importants changements de conformation lors de son interaction avec le 40S qui vont lui permettre de s'enrouler autour de la plateforme de la sous-unité 40S afin d'atteindre le site P et y délivrer eIF1 et eIF5 lors de la formation du complexe de pré-initiation 43S (Zeman et al., 2019).

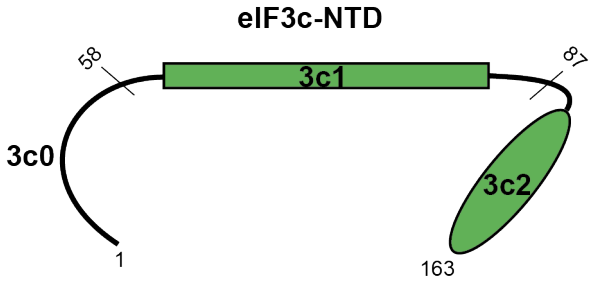


Figure 17. Le domaine N-terminal de eIF3c (adaptée de (Obayashi et al., 2017)). Le domaine N-terminal (NTD) d'eIF3c est divisé en trois sous-domaines 3c0 (1-58), 3c1 (59-87) et 3c2 (88-163).

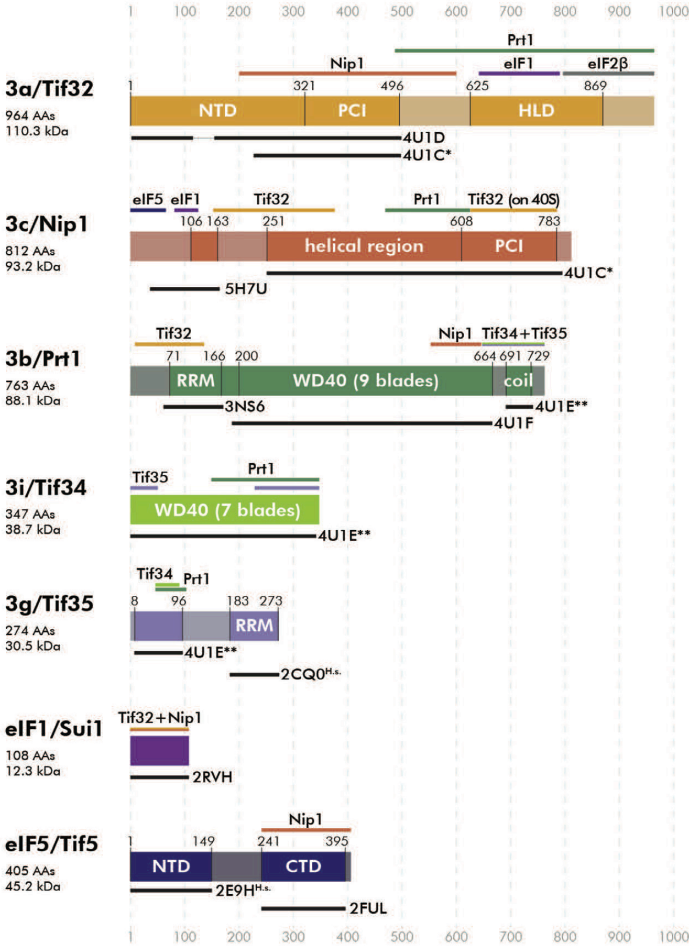


Figure 18. Représentation schématique de toutes les sous-unités d'eIF3 de la levure, eIF1 et eIF5 (figure issue de (Zeman et al., 2019). Les barres de couleur au-dessus de chaque protéine délimitent les régions d'interaction avec les partenaires. Les barres noires sous chaque protéine indiquent les domaines dont les structures 3D sont résolues ainsi que les numéros d'accession à la PDB. Les nuances plus claires représentent des régions de structure inconnue ou supposées non structurées. PCI: Proteasome, COP9, domaine du facteur d'initiation 3; HLD: domaine de type HCR1; RRM: motif de reconnaissance d'ARN; WD40: domaine de l'hélice bêta; NTD: domaine N-terminal; CTD: domaine C-terminal.

2.3.1.2. Interaction d'eIF3 avec eIF1A

eIF3 interagit également avec le facteur eIF1A, qui, avec eIF1 joue un rôle critique dans le processus de scanning et la sélection du codon AUG. Il apparaît que les parties N-terminales d'eIF1A et eIF1 sont capables d'interagir avec les deux sous-unités eIF3a et c (Querol-Audi et al., 2013).

2.3.2. Fonctions d'eIF3 tout au long du cycle de la traduction

Des données récentes suggèrent fortement que le facteur eIF3, qui coordonne la plupart des étapes d'initiation, ne se détache pas après l'assemblage des sous-unités ribosomiques mais reste attaché au ribosome 80S durant les premiers cycles d'élongation (Valášek et al., 2017). Ce maintien permet le redémarrage de la traduction en aval du même ARNm après la lecture du codon stop. Dans le cas de la traduction d'une uORF (upstream Open Reading Frame) suffisamment courte (moins de 30 codons) eIF3, toujours présent, favorise le redémarrage de la traduction de l'ORF qui suit (Hronová et al., 2017; Jackson et al., 2012; Wethmar, 2014). eIF3 joue également un rôle dans les mécanismes de terminaison de la traduction, de recyclage des ribosomes et de recodage traductionnel (Jackson et al., 2012; Mohammad et al., 2017; Young and Guydosh, 2019). Dans certains cas spécifiques, eIF3 est capable de promouvoir la translecture programmée d'un codon stop lorsqu'un aminoacyl-ARNt apparenté (dont l'anticodon présente deux appariements sur trois avec le codon) ou un ARNt suppresseur entrent en compétition avec le facteur de terminaison de la traduction eRF1 pour la lecture du codon stop (Beznosková et al., 2015; Dabrowski et al., 2015; Schueren and Thoms, 2016). Par ailleurs, il a été montré que le complexe eIF3 est également impliqué dans la dégradation des ARNm possédant un codon stop prématuré ou NMD (Nonsense Mediated mRNA Decay) et que la sous-unité eIF3e (INT6) interagit avec CBP80 et le facteur du NMD UPF2 (Hwang et al., 2010; Mocquet et al., 2012; Morris et al., 2007) (voir ***Figure 19***).

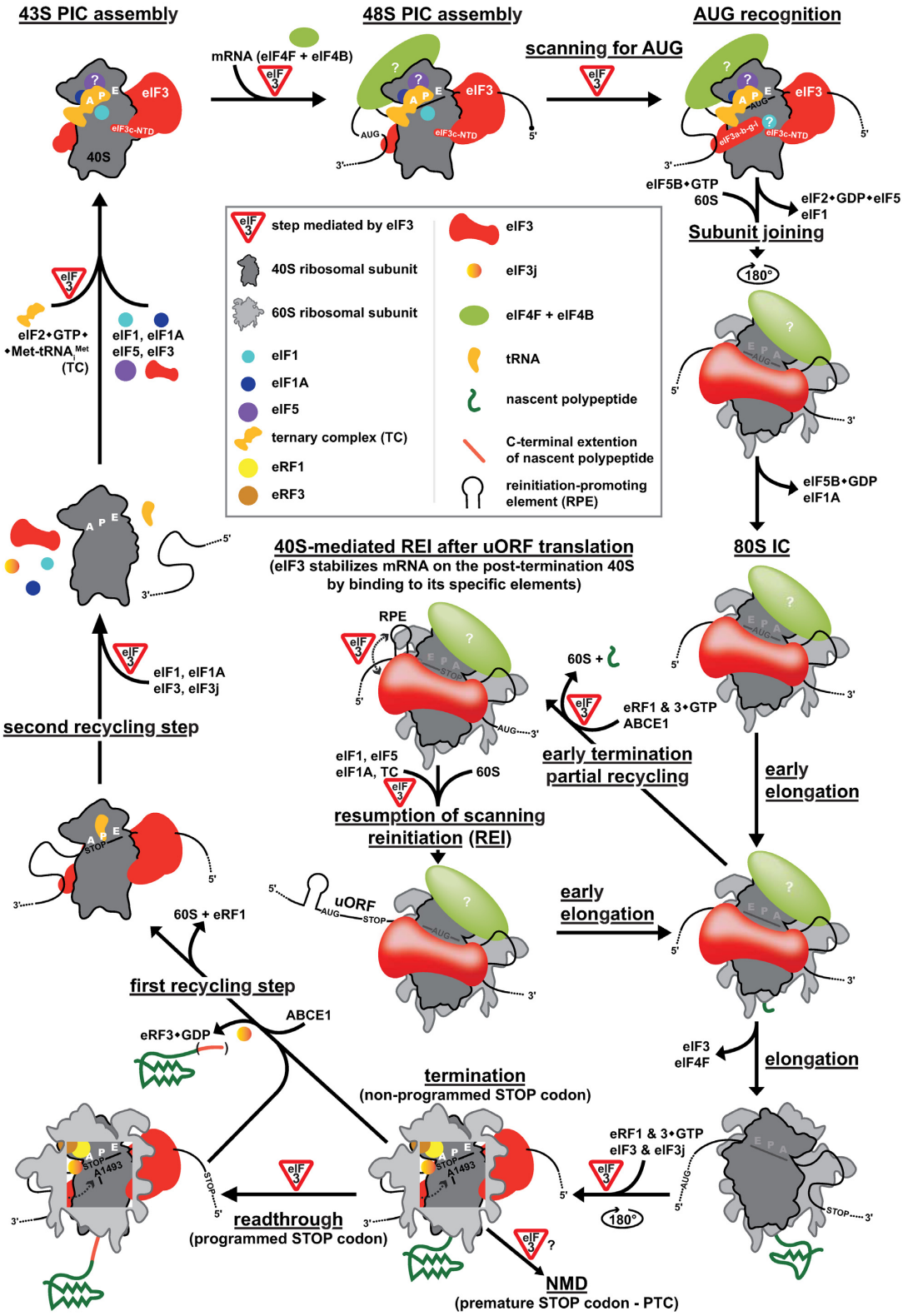


Figure 19. Différentes fonctions d'eIF3 dans la synthèse protéique (figure issue de (Valášek et al., 2017). En dehors de l'initiation canonique, eIF3 intervient lors de la réinitiation de la traduction, la lecture programmée du codon stop et le NMD.

2.3.3. Rôle d'eIF3 dans la traduction sélective d'ARNm et dans les mécanismes de traduction non-canoniques

Des études récentes ont montré que le facteur d'initiation eIF3 pouvait réguler l'initiation de la traduction de certains ARNm. Des expériences de 'PAR-CLIP' (PhotoActivatable Ribonucleoside-enhanced CrossLinking and Immuno-Precipitation) couplées au séquençage haut débit ont montré qu'eIF3 assurait le recrutement sélectif d'environ 500 ARNm cellulaires en se liant à des structures en tige-boucle d'une longueur moyenne de 25 nucléotides présentes dans ces ARN, notamment dans les extrémités 5' UTR et ceci afin de contrôler leur expression (Lee et al., 2015). Ces ARNm ont des fonctions spécifiques et sont impliqués dans le contrôle de la prolifération cellulaire, du cycle cellulaire, de l'apoptose et de la différentiation cellulaire. eIF3 interagit avec ces ARNm par le biais de combinaisons distinctes d'au moins 4 sous-unités : eIF3a, b, d et g (Lee et al., 2015). Étonnamment, l'analyse fonctionnelle de l'interaction entre eIF3 et deux ARNm codant pour les régulateurs de prolifération cellulaire *c-JUN* et *BTG1* a révélé des modes de régulation de la traduction opposés (voir ***Figure 20***).

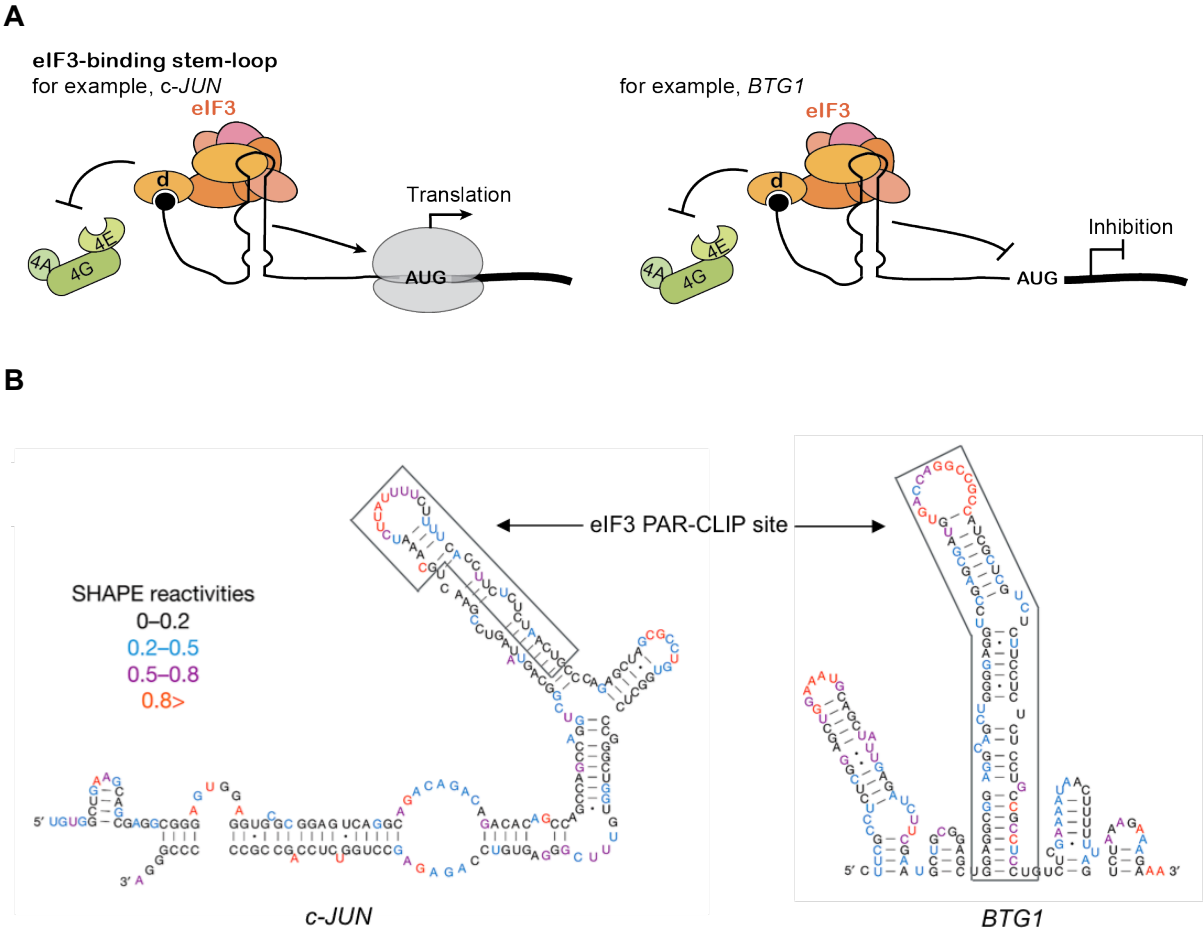


Figure 20. Modèle de structure secondaire des éléments structuraux retrouvés dans les ARNm *c-JUN* et *BTG1* qui interagissent avec eIF3 et conséquence sur la traduction. (A) Deux exemples

de structures secondaires trouvées dans les ARNm *c-JUN* et *BTG1*. Ces ARNm, qui sont impliqués dans la prolifération cellulaire, ont été identifiés par PAR-CLIP (Lee et al., 2015). Les sites d'interaction avec eIF3 (site encadré en noir) sont repliés en longues tige-boucles dont les structures ont été établies par sondage en solution. La fixation du facteur eIF3 à la tige-boucle permet le recrutement de la particule 43S pour initier la traduction. **(B)** Modèles de régulation d'eIF3 lié aux ARNm de *c-JUN* et *BTG1*. En se fixant sur les tiges-boucles distinctes présentes dans leurs extrémités 5' UTR, eIF3 est capable d'activer ou d'inhiber la traduction des ARNm. Dans ces processus, la sous-unité eIF3d se lie à la coiffe des ARNm (Lee et al., 2016).

En se fixant sur des éléments structuraux en tige-boucle distincts dans la région 5' UTR de leurs ARNm, eIF3 est capable d'activer ou d'inhiber la traduction des ARNm *c-JUN* ou *BTG1* respectivement. Bien que l'interaction entre eIF3 et la tige-boucle stimule l'assemblage du complexe d'initiation, la traduction sélective de ces ARNm par eIF3 reste dépendante de la coiffe en 5' des ARNm (Lee et al., 2015). En revanche, l'efficacité de leur traduction n'est pas affectée par l'inactivation d'eIF4E (Gandin et al., 2016; Hsieh et al., 2012; Thoreen et al., 2012), ce qui suggère que la reconnaissance de la coiffe pourrait se faire par un mécanisme non-canonique. Il a été montré que la sous-unité eIF3d était capable de se lier à la coiffe de certains ARNm (Villa et al., 2013), dont le messager *c-JUN* (Lee et al., 2016). Ceci suggère l'existence d'un mécanisme d'initiation de la traduction coiffe-dépendant spécialisé impliquant eIF3d et indépendant du facteur eIF4E (Lee et al., 2016). Le fait qu'eIF3d ne se lie pas à tous les ARNm semble s'expliquer par la présence dans eIF3d d'un domaine 'RNA gate' qui bloque l'accès à la coiffe de certains ARNm. Par ailleurs, l'activation allostérique par des éléments structuraux présents dans certains ARNm spécifiques stimule la fixation d'eIF3d sur la coiffe. C'est le cas de l'ARNm de *c-JUN*, un élément d'ARN inhibiteur capable de bloquer le recrutement d'eIF4E forcerait la reconnaissance alternative de la coiffe par eIF3d (Lee et al., 2016).

Des travaux récents menés chez la drosophile ont montré qu'eIF3d était aussi capable de se lier à la région 5' UTR de l'ARNm *msl-2* et d'en réprimer la traduction (Szostak et al., 2018). Dans ce cas, il a été montré qu'eIF3d interagit aussi avec la protéine de liaison à l'ARNm Hrp48, ce qui aurait pour conséquence d'interférer avec le recrutement efficace du complexe de pré-initiation 43S sur l'ARNm (Szostak et al., 2018).

Un autre mécanisme de régulation sélective de la traduction de certains gènes liés aux métastases impliquant eIF3 et DDX3 a été caractérisé (Chen et al., 2018). En effet, il a été montré que DDX3 est capable de renforcer spécifiquement l'association du complexe de fixation de la coiffe (CBC) avec des ARNm contenant des uORF et de faciliter le recrutement du facteur d'initiation eucaryote 3 (eIF3) pour activer leur traduction (Chen et al., 2018).

La phosphorylation du complexe eIF4F régule également la traduction sélective d'ARNm. (Walsh and Mohr, 2014) ont montré que la sous-unité eIF3e était capable de recruter la kinase Mnk1 et de réguler la phosphorylation de eIF4E.

Par ailleurs, il a été montré qu'eIF3 était capable de se lier à des adénosines méthylées (m^6A) présentes dans la région 5' UTR de certains ARNm et de promouvoir leur traduction indépendamment de la coiffe (Meyer et al., 2015). Divers stress cellulaires sont capables d'induire une augmentation et une redistribution des modifications m_6A dans les régions 5' UTR permettant ainsi de contourner les mécanismes canoniques de reconnaissance de la coiffe pour favoriser la traduction coiffe-indépendante.

En plus de son implication dans la régulation de l'expression génique cellulaire, eIF3 stimule également la traduction des protéines du virus de l'hépatite C (HCV) (Sun et al., 2013) en se fixant directement sur l'IRES virale par l'intermédiaire de ses sous-unités eIF3a, b et c (Cai et al., 2010; Sizova et al., 1998) et non sur la sous-unité ribosomique 40S comme le montre la structure du complexe ribosome-HCV-eIF3 (Hashem et al., 2013b). De façon inattendue, eIF3 peut également agir sur l'inhibition de la réplication du virus de l'immunodéficience humaine (HIV) dans le cadre du mécanisme d'auto-défense antiviral (Jäger et al., 2011).

2.4. Régulation des sous-unités d'eIF3 et cancer

Avec un si large éventail de fonctions, il n'est pas surprenant que des défauts de régulation de l'expression de plusieurs sous-unités d'eIF3 soient impliqués dans l'oncogenèse et le maintien d'états cancéreux. Plusieurs sous-unités d'eIF3 sont sur-exprimées ou sous-exprimées dans un grand nombre de cancers humains (Rajasekhar and Holland, 2004). Par exemple, une surexpression d'eIF3a a été mesurée dans les cancers du sein, du col de l'utérus, de l'œsophage, des poumons et de l'estomac (Pincheira et al., 2001). Les sous-unités eIF3b, c, h, i et m sont également surexprimées dans d'autres types de cancers. Par contre, une sous-expression des sous-unités eIF3e et f a été détectée dans certains cancers du sein, des poumons, du pancréas et les mélanomes (Doldan et al., 2008; Marchetti et al., 2001; Shi et al., 2006) ce qui suggère un lien entre la dérégulation de l'initiation de la traduction et la transformation néoplasique. La contribution des sous-unités d'eIF3 dans les cancers, diffère d'une sous-unité à l'autre comme décrit ci-dessous.

eIF3a :

Un niveau élevé d'eIF3a augmente également le niveau des autres sous-unités du complexe eIF3 pouvant conduire aux effets observés dans les tumeurs (Zhang et al., 2007). L'inhibition d'eIF3a au contraire empêche la formation de complexes eIF3 stables et conduit à une importante diminution de la synthèse protéique (Wagner et al., 2014). De plus, des

altérations d'eIF3a ont des répercussions sur la fonction de plusieurs molécules, dont eIF4B, l'ARNm et les protéines du cytosquelette (MacDonald et al., 1999; Pincheira et al., 2001).

eIF3b :

La surexpression de la sous-unité eIF3b, a été détectée dans les cancers du sein, de la vessie et de la prostate (Lin et al., 2001; Wang et al., 2013a). eIF3b joue le rôle de plateforme d'assemblage du complexe eIF3 en interagissant avec eIF3a, c, g et i (Fraser et al., 2004; Wagner et al., 2014) ainsi qu'avec AUF1 (facteur de liaison aux séquences riches en AU dans la région 3' UTR des ARNm) (Lee et al., 2014). L'inhibition d'eIF3b par siRNA dans certaines cellules cancéreuses de la prostate, vessie et ovaire réduit la prolifération cellulaire ainsi que leurs phénotypes malins et l'aptitude à former des tumeurs chez les souris, mais inhibe seulement 50% de la synthèse des protéines (Wang and Ouyang, 2019; Wang et al., 2013a).

eIF3c :

Un taux élevé d'eIF3c a été détecté dans des tumeurs séminomateuses et méningées ainsi que dans des cellules du cancer du colon (Scoles et al., 2006; Song et al., 2013). eIF3c est également un marqueur tumoral majeur parmi les 15 RBPs (RNA-Binding Proteins) surexprimées dans le cancer du foie. Son expression est associée à la dérégulation de voies de signalisation cellulaires reliées au cancer telles que KRAS, VEGF et Hedgehog (Li et al., 2017). Des études récentes ont montré que l'inhibition par siARN d'eIF3c, inhibe la prolifération cellulaire et stimule l'apoptose de cellules cancéreuses des ovaires (Wen et al., 2019). Ceci s'explique du fait de l'interaction d'eIF3c avec les facteurs d'initiation eIF1 et eIF4G. L'état hautement phosphorylé d'eIF3 durant la croissance exponentielle des cellules suggère un rôle important d'eIF3c dans les mécanismes de régulation cellulaire (Damoc et al., 2007).

eIF3d :

L'implication d'eIF3d a été décrite dans le développement de plusieurs types de cancers. eIF3d a été trouvé surexprimé dans les cancers de l'ovaire et du sein. Son inhibition par siARN réduit la prolifération cellulaire, bloque les cellules dans la phase G2/M en altérant la kinase dépendante des cyclines 'CDK1' (Lin et al., 2016) et inhibe l'expression de la cycline D1, de la β-caténine et de c-Myc dans les cellules cancéreuses du sein (Fan and Guo, 2015). eIF3d est requis pour une association stable du complexe eIF3 (Bandyopadhyay et al., 2002). Récemment, eIF3d a été identifié comme un facteur de liaison à la coiffe des ARNm impliqués dans la prolifération cellulaire. Il stimule la traduction du proto-oncogène *c-JUN* et inhibe celle de la protéine antiprolifératives *BTG1* (Lee et al., 2016; Matsuda et al., 2001; Vleugel et al., 2006). Ceci explique pourquoi sa surexpression est impliquée dans l'oncogenèse et le maintien d'états cancéreux.

eIF3e et eIF3f :

Plusieurs études ont rapporté l'implication d'eIF3e dans l'oncogenèse. eIF3e est essentiel pour la prolifération et la survie des cellules de glioblastomes humains en modulant l'expression des facteurs HIF1α et HIF2α (Hypoxia-Inducible Factors) (Sesen et al., 2014). D'autres études ont montré qu'eIF3e n'est pas essentiel pour la traduction mais joue le rôle de régulateur de la traduction de certains ARNm dans des conditions spécifiques (Grzmil et al., 2010; Kim et al., 2007; Zhou et al., 2005). D'autres études ont qualifié eIF3e de suppresseur de tumeurs, vu que son inhibition est liée à des transitions épithélio-mésenchymateuses dans le cancer des cellules épithéliales du sein (Gillis and Lewis, 2013) et à la trans-différentiation des cellules souches mésenchymateuses en fibroblastes associés au carcinome 'CAF' (Suo et al., 2015). eIF3e forme avec eIF3d un module qui orchestre l'expression d'ARNm spécifiques impliqués dans le contrôle du métabolisme énergétique pouvant être dérégulé dans le cancer (Shah et al., 2016). eIF3e possède donc à la fois des capacités oncogènes et suppressives de tumeurs (Sesen et al., 2017).

Comme pour eIF3e, un taux faible d'eIF3f a été détecté dans plusieurs cancers tels le cancer du sein, du pancréas, de l'ovaire ainsi que dans les mélanomes (Shi et al., 2006). L'inhibition d'eIF3f par siRNA stimule la traduction IRES et cap-dépendante ainsi que la prolifération des cellules épithéliales humaines du pancréas, suggérant qu'eIF3f agit comme un inhibiteur de l'activité du complexe eIF3 et de la synthèse des protéines. Son inhibition sur-active par contre les autres facteurs d'initiation de la traduction (Wen et al., 2012).

eIF3g :

Une expression très élevée de la sous-unité eIF3g a été rapportée dans le cancer colorectal. L'inhibition de son expression dans les cellules HCT116 réduit la prolifération cellulaire en réduisant l'effet kinase de la voie de signalisation mTOR sur les protéines AKT, S6K et 4E-BP1 qui sont impliquées dans la régulation de la prolifération et la croissance cellulaire (Yang et al., 2019). De plus, eIF3g interagit avec des protéines présentes dans le noyau des cellules du cancer du sein, telles la protéine du cytosquelette β-actine, hnRNP U qui se lie aux pré-ARNm et la protéine à doigt de zinc ZNF823 qui forment un complexe pouvant être altéré dans le cancer (Zheng et al., 2016).

eIF3h et eIF3i :

Un taux élevé de la sous-unité eIF3h a été détecté dans les cancers du sein et de prostate. Son inhibition par siARN diminue la prolifération des cellules cancéreuses (Zhang et al., 2008). De même, un taux élevé de la sous-unité eIF3i a été détectée dans le carcinome squameux du cou et dans le cancer du foie (Ahlemann et al., 2006; Huang et al., 2004). Du fait de son interaction avec les protéines kinases, eIF3i est considéré comme un régulateur de la prolifération cellulaire, par exemple, en agissant sur le taux de phosphorylation de la protéine AKT1 (Wang et al., 2013b).

eIF3k et eIF3m :

La sous-unité eIF3k semble participer au développement du cancer. Ceci semble lié à la présence de nombreux domaines d'interaction protéine-protéine du côté C-terminal ainsi qu'à des domaines de liaison à l'ARN du côté N-terminal, suggérant un rôle structural et fonctionnel important de eIF3k (Dong and Zhang, 2006; Wei et al., 2004). La sous-unité eIF3m est impliquée dans le cancer colorectal ainsi que dans plusieurs carcinomes et leucémies (Goh et al., 2011). L'importance de la sous-unité eIF3m est liée à son rôle dans le maintien de l'intégrité du complexe eIF3 et la stabilisation des sous-unités eIF3c, f et h du cœur structural (Zeng et al., 2013). Ceci explique le rôle général d'eIF3m dans la tumorigenèse.

3. Les histones

3.1. Les histones, constituants du nucléosome

Les histones sont des protéines localisées dans le noyau des cellules eucaryotes. Elles constituent le composant majeur de la chromatine. En s'assemblant avec l'ADN sous forme de nucléosomes, elles sont essentielles à la compaction de l'ADN et jouent un rôle fondamental dans la régulation des gènes.

Les histones font partie des protéines les plus conservées au cours de l'évolution et les plus lentes à évoluer (Kamakaka and Biggins, 2005). Ce sont de petites protéines d'environ 100 à 130 résidus particulièrement riches en acides aminés basiques. Elles se caractérisent toutes par un domaine globulaire C-terminal appelé 'histone fold', très conservé depuis les archées jusqu'aux eucaryotes supérieurs, qui est formé de 3 hélices qui leur permettent d'interagir avec l'ADN (Arents and Moudrianakis, 1995). Chez les métazoaires, 5 classes d'histones ont été décrites, à savoir les histones H1, H2A, H2B, H3 et H4 réparties en deux catégories en fonction de leur rôle dans la formation de la chromatine.

Les histones H2A, H2B, H3 et H4 ou histones s'assemblent en octamère pour former le 'cœur' du nucléosome qui contient deux exemplaires de chacune des histones. Au sein de l'octamère, les protéines se regroupent en hétérodimères H2A-H2B et H3-H4 en s'associant grâce à leur domaine 'histone fold' pour former un motif de type 'hand-shake' (Arents et al., 1991). L'ADN est enroulé autour de cet octamère en formant 1,7 tour (soit 145-147 paires de bases) (voir ***Figure 21***) (Arents et al., 1991; Kornberg, 1974; Kornberg and Thonmas, 1974; Luger et al., 1997).

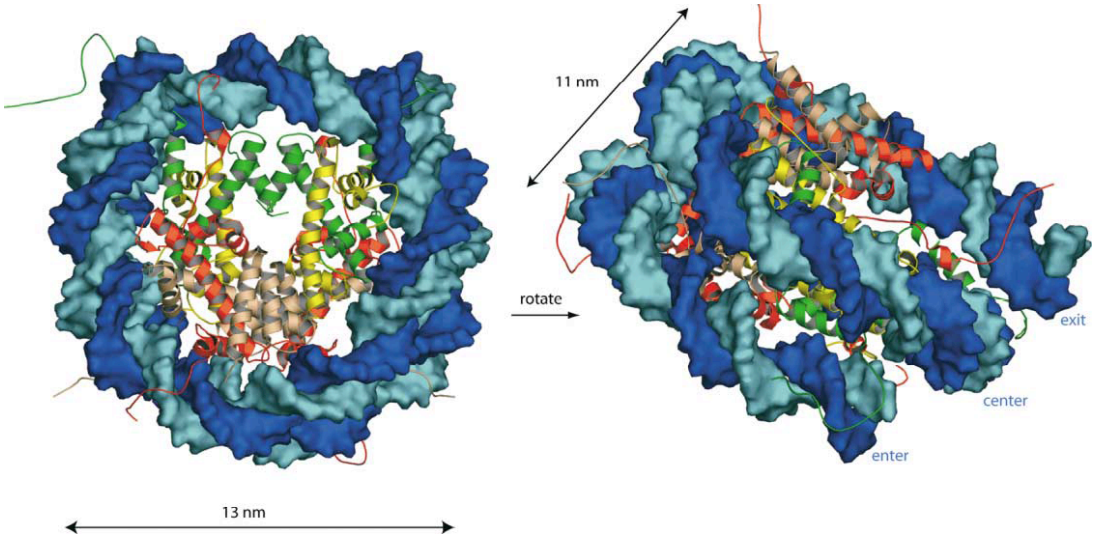


Figure 21. Vue générale de la structure tridimensionnelle du nucléosome (figure issue de (Khorasanizadeh, 2004)). Les 2 brins d'ADN se distinguent par des teintes de bleu différentes. L'ADN

effectue 1,7 tour autour d'un octamère constitué de deux molécules de chacune des histones : H2A (en rouge), H2B (en rose), H3 (en vert) et H4 (en jaune). L'histone H1 est absente du modèle.

Les histones H1, ou histones de liaison, se fixent sur cet octamère afin d'assurer la compaction du nucléosome et de rigidifier la structure de l'ADN. Pour cela, l'histone H1 possède un domaine globulaire central contenant un motif de liaison à l'ADN de type 'winged helix-turn-helix' motif composé d'un repliement de type $\alpha\beta\alpha\alpha\beta\beta$ (Khorasanizadeh, 2004).

Chaque classe d'histones comprend elle même plusieurs familles selon que leur profil d'expression soit couplé ou non à la réplication de l'ADN. La première famille regroupe les histones canoniques, qualifiées de réplication-dépendantes. Leur expression est massive et leur incorporation à l'ADN s'effectue uniquement au cours de la phase S du cycle cellulaire, lors de la réplication de l'ADN (Osley, 1991; Pyronnet et al., 2001). Plusieurs variants d'histones partagent une homologie de séquence avec les histones canoniques (Talbert et al., 2012). Ces variants d'histones, qui constituent la deuxième famille, sont réplication-indépendants. Ils jouent des rôles importants dans le développement (Banaszynski et al., 2010; Maze et al., 2014).

Chez les eucaryotes, la compaction de l'ADN génomique en chromatine permet de réguler la transcription, la réplication, la réparation et la recombinaison (Koyama and Kurumizaka, 2018). Le degré de condensation de l'ADN autour des nucléosomes d'histones est variable au sein de la chromatine (voir ***Figure 22***).

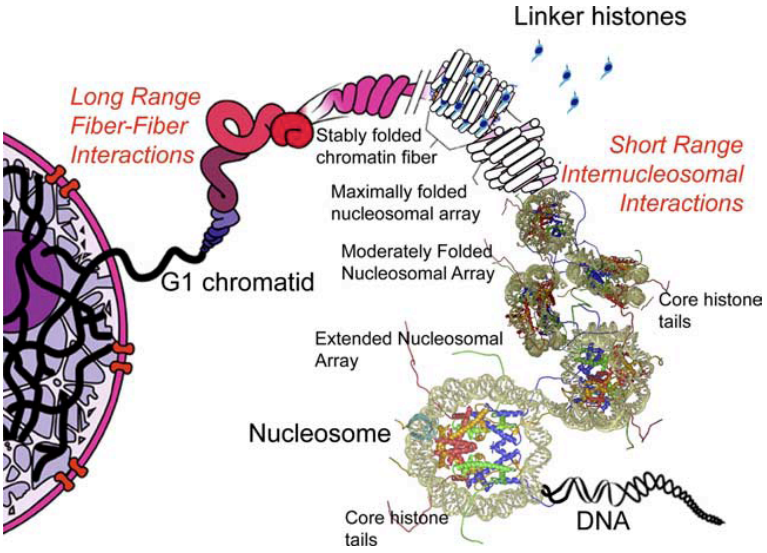


Figure 22. Fonction du nucléosome dans la compaction de l'ADN (figure issue de (Chakravarthy et al., 2005)). Représentation schématique des différents niveaux de compaction de l'ADN chez les eucaryotes. De longs enchaînements de nucléosomes sont compactés par le biais d'interactions à courte ou à longue distance dans des fibres dont l'architecture n'est pas décrite.

Il est faible dans l'euchromatine que l'on dit 'ouverte' et accessible à la machinerie des ARN polymérases. Il est élevé dans l'hétérochromatine, que l'on dit 'fermée' et 'inaccessible' à la machinerie de transcription. Chaque histone possède une extension N-terminale dépassant du nucléosome. Ces queues N-terminales sont accessibles à des enzymes de modification dont l'action permet de réguler la compaction de la chromatine. Ainsi des modifications post-traductionnelles, comme l'acétylation, la méthylation, l'ubiquitinylation et la sumoylation des lysines, la méthylation des arginines, la phosphorylation des résidus sérine et thréonine, l'ADP-ribosylation des résidus glutamates ou aspartates (Bannister and Kouzarides, 2011) ont un impact sur l'expression des gènes et le maintien de la stabilité du génome. Elles ciblent habituellement les queues N- et C-terminales des histones, mais elles peuvent également cibler leur domaine globulaire. Ces modifications agissent sur l'affinité entre l'histone modifiée et l'ADN, induisant une inhibition ou une stimulation de la transcription. Des telles modifications post-traductionnelles ont été détectées dans presque tous les types de cancers, dans les oncogènes, les gènes suppresseurs de tumeur et les facteurs de transcription (Shanmugam et al., 2018).

3.2. Expression et régulation des gènes d'histones réplication-dépendants

Les ARNm des histones réplication-dépendants sont les seuls ARNm cellulaires eucaryotes dépourvus de queue poly(A) à leur extrémité 3' (Piontkivska et al., 2002). L'extrémité 3' se termine par une séquence de 26 nucléotides fortement conservée et contenant une structure en tige-boucle de 16 nt (Dominski and Marzluff, 1999). Cette tige-boucle interagit avec la protéine SLBP ou 'Stem-Loop Binding Protein' (Martin et al., 1997) qui participe à la maturation, au transport nucléo-cytoplasmique, à la régulation de la traduction et la dégradation de l'ARNm (Jaeger et al., 2006; Marzluff et al., 2008). Cette tige boucle, identique dans tous les ARNm d'histones, fixe des protéines impliquées dans les régulations post-transcriptionnelles (Whitfield et al., 2000).

Les pré-ARNm d'histones réplication-dépendants sont également dépourvus d'introns. Ils requièrent une seule étape de maturation par clivage endonucléolytique au niveau de leur région 3' UTR au cours d'un processus impliquant plusieurs protéines et le snARN U7 (Dominski and Marzluff, 1999). Cette étape de maturation requiert la tige-boucle et un élément appelé HDE pour 'Histone Downstream Element', riche en purine et situé 15 nucléotides en aval du site de clivage de l'ARNm qui s'apparie avec la région 5' du snARN U7 (Mowry and Steitz, 1987; Spycher et al., 1994). Cette interaction est stabilisée par SLBP et la protéine ZFP100 'Zinc Finger Protein 100 kDa'. SLBP recrute également une endonucléase CPSF-73 (Dominski et al., 2005) qui catalyse le clivage du pré-ARNm d'histone. Après clivage, la protéine SLBP reste associée à l'ARNm d'histone mature et

l'accompagne vers le cytoplasme où elle participe à la traduction de l'ARNm (Sànchez and Marzluff, 2002). L'efficacité de l'étape de maturation augmente de 8 à 10 fois au cours de la phase S, permettant une accumulation rapide d'ARNm matures (Marzluff and Duronio, 2002). L'expression des ARNm d'histones est également augmentée de 30 à 40 fois durant la phase S du cycle cellulaire. L'augmentation de leur transcription est maximale lorsque les cellules entrent en phase S, puis ces ARNm décroissent rapidement à la fin de cette phase (Harris et al., 1991). Cette régulation est SLBP-dépendante. En effet, le niveau de la protéine SLBP est régulé par le cycle cellulaire et augmente lorsque les cellules entrent en phase S pour permettre une traduction accrue des ARNm d'histones. La dégradation de SLBP à la fin de la phase S empêche l'accumulation ultérieure d'ARNm ou d'histones (Whitfield et al., 2000; Zheng et al., 2003). Dans les cellules HeLa, la phase S dure environ 6 heures. L'analyse par micropuces de cellules HeLa synchronisées a montré une expression séquentielle de groupes de gènes au cours de la phase S. Au début de la phase S, les gènes dédiés à la réplication de l'ADN sont exprimés. Après 3 heures, ce sont les gènes d'histones qui sont exprimés de façon prédominante (van der Meijden et al., 2002). Les ARNm correspondants ne sont détectés qu'en milieu de phase S pendant environ 1 heure. Soixante millions de nucléosomes sont alors synthétisés, ce qui suggère qu'un mécanisme de traduction particulièrement efficace est utilisé pour leur synthèse. Notons que la traduction est strictement dépendante de la coiffe pendant la phase S lorsque les histones sont produites (Pyronnet et al., 2001).

Enfin, il convient de mentionner qu'une étude récente a montré que les ARNm d'histones réplication-dépendants pouvaient être polyadénylés dans des cellules différenciées ne se divisant plus. Ces ARNm servent à produire des histones de remplacement nécessaires au bon maintien de la chromatine (Lyons et al., 2016).

3.3. Mécanismes de traduction des ARNm des histones réplication-dépendantes

Les ARNm d'histones étant dépourvus de queue poly-A, le mécanisme classique d'initiation de la traduction basé sur la circularisation des ARNm n'est *a priori* pas applicable. Plusieurs mécanismes alternatifs ont été proposés permettant de palier à cette absence d'interaction entre la PABP et l'extrémité 5' de l'ARNm.

Il a été montré que la protéine SLBP fixée à la tige-boucle située à l'extrémité 3' des ARNm pouvait se substituer fonctionnellement à la PABP. En effet, l'activité traductionnelle de SLBP résulte de son interaction avec la protéine SLIP1 pour 'SLBP-Interacting Protein 1', une protéine comprenant un domaine MIF4G homologue au facteur d'initiation 4G ('Middle domain of Initiation Factor 4G'). Des études ont montré que les facteurs d'initiation eIF3g et

eIF3e interagissent avec la protéine MIF4GD/SLIP1 et entrainent une traduction efficace des ARNm d'histones (von Moeller et al., 2013; Neusiedler et al., 2012).

Dans d'autres études, il a été montré que la structure de l'ARNm jouait un rôle essentiel pour la traduction de l'histone H4. En effet, les ARNm d'histones présentent généralement des régions 5' UTR relativement courtes (entre 20 nt et 60 nt), l'exemple extrême étant celui de l'histone H4-12 de souris qui ne contient que 9 nt (Meier et al., 1989). Cette taille réduite semble loin d'être optimale pour un recrutement efficace des ribosomes et pourtant ces ARNm sont exprimés de façon très efficace durant la phase S du cycle cellulaire. D'autre part, les régions codantes des ARNm d'histones sont anormalement riches en résidus Guanosine (G) et Cytidine (C) qui y sont retrouvés à des niveaux équivalents (suggérant un grand nombre d'appariements G:C). Les alignements de séquence des histones H4 ont également mis en évidence une conservation en nucléotides exceptionnellement élevée qui va au delà de ce qui est strictement nécessaire pour la conservation des acides aminés. Dès 1992, il avait été proposé que cette conservation au niveau nucléotidique pourrait résulter d'une pression de sélection pour le maintien de structures secondaires dans l'ARNm, elles-mêmes dotées de propriétés fonctionnelles qui restaient à découvrir (Huynen et al., 1992).

Toutes ces particularités structurales, s'additionnant au besoin de synthèse extrêmement élevé en histones ont suggéré l'existence d'un mécanisme de traduction non canonique. C'est pour cette raison que mon laboratoire d'accueil s'est intéressé au mécanisme d'initiation de la traduction de l'ARNm de l'histone H4. Les études réalisées ont permis la caractérisation, dans la région codante de l'ARNm de l'histone H4 de souris, des structures responsables du recrutement interne des ribosomes (voir partie 1.3.1.5. pour les détails du mécanisme d'initiation de la traduction de l'histone H4). D'une manière générale, ces structures jouent un rôle clé au cours du processus d'initiation de la traduction de l'ARNm H4. Une jonction triple-hélice (TWJ) présente dans l'ARNm H4 est responsable du positionnement du codon AUG dans le site P du ribosome tandis qu'un deuxième élément appelé 4E-SE (élément sensible à eIF4E) recrute le complexe de liaison de la coiffe eIF4F contenant eIF4E, 4G et 4A (Martin et al., 2011, 2016). Bien que ces résultats aient souligné le rôle essentiel des structures de l'ARNm dans le mécanisme d'initiation de H4, ils n'excluent pas la participation de facteurs d'initiation non conventionnels comme ceux décrits précédemment (von Moeller et al., 2013; Neusiedler et al., 2012). Néanmoins, il est probable que les éléments structuraux de l'ARNm H4 jouent un rôle essentiel dans ce modèle d'initiation non canonique puisque la plupart des expériences qui ont été réalisées pour l'établir se sont déroulées *in vitro* dans des extraits de traduction de réticulocytes de lapin dépourvus de certains facteurs comme SLBP.

Des travaux plus récents ont montré l'importance d'une autre structure de l'ARNm H4 dans un processus de régulation impliquant eIF3. En effet, eIF3 est capable d'assurer le recrutement de certains ARNm cellulaires, dont celui de l'histone H4, en se liant à des structures en tige-boucle présentes dans ces ARNm (Lee et al., 2015) (voir **partie 2.3.3.**). Dans le cas de l'histone H4, les expériences de sondage en solution réalisées au laboratoire ont confirmé la présence d'une longue tige boucle dans la phase codante de l'ARN au niveau du site d'interaction d'eIF3 identifié par CLIP-SEQ (Lee et al., 2015, 2016).

L'objectif de ma thèse a été de caractériser cette interaction et d'étudier le rôle du facteur d'initiation eIF3 dans la traduction de l'ARNm de l'histone H4.

MON PROJET DE THESE

Le laboratoire s'intéresse aux mécanismes d'initiation de la traduction non conventionnels. L'histone H4 occupe une place de choix dans les projets du laboratoire car l'initiation de la traduction de son ARNm procède par un mécanisme hybride entre traduction canonique dépendante de la coiffe et entrée interne des complexes de préinitiation sans étape de scanning. Dans ce mécanisme, les structures formées par l'ARNm H4 jouent un rôle essentiel et miment la fonction des structures trouvées dans les IRES virales.

L'ARNm de l'histone H4 est particulièrement structuré dans sa phase codante où ont été trouvés les éléments clés impliqués dans le processus d'initiation. Cependant, une structure repliée en longue tige-boucle restait sans fonction décrite. Les expériences de CLIP-SEQ réalisées par le groupe de J. Cate ont permis d'identifier un site d'interaction du facteur d'initiation eIF3 dans cette structure. eIF3 est un imposant complexe multifactoriel interagissant avec nombreux partenaires protéiques (protéines ribosomiques, facteurs d'initiation) et acides nucléiques (ARN ribosomique, IRES virales). Les résultats des expériences de CLIP-SEQ suggèrent l'existence d'interactions avec un groupe de près de 500 ARNm impliqués dans des processus de contrôle de la croissance cellulaire incluant le cycle cellulaire, la différenciation et l'apoptose. Face à une telle variété d'ARNm se pose la question de la spécificité de reconnaissance. Les premières études ont montré qu'eIF3 reconnait des régions double brin situées dans les extrémité 5' et 3' UTR des ARNm. Parmi les cibles d'eIF3 sont retrouvés les ARNm des histones. Dans l'ARNm de l'histone H4, la séquence identifiée par séquençage haut-débit se situe dans la longue tige-boucle située près du codon stop, dont la fonction n'était pas connue jusque là. L'objectif de mon projet de thèse était d'explorer cette interaction, avec l'ARNm de l'histone H4 mais aussi des autres histones, c'est-à-dire de vérifier son existence puis de l'étudier au niveau moléculaire à l'aide d'approches *in vitro* et *in vivo* dans les cellules. Comprendre la fonction de cette interaction a constitué le second objectif majeur de mon projet de thèse, objectif que j'ai conduit par des approches essentiellement cellulaires. L'ensemble de mes travaux nous permet désormais de mieux comprendre le rôle d'eIF3 dans la traduction des histones.

RÉSULTATS

RÉSULTATS

1. Bases moléculaires de l'interaction entre eIF3 et l'ARNm H4

1.1. eIF3 interagit *in vitro* avec l'ARNm de l'histone H4

Il a été précédemment montré que le facteur d'initiation eucaryote eIF3, assurait le recrutement sélectif d'environ 500 ARNm cellulaires en se liant à des structures en tige-boucle afin de contrôler leur expression (Lee et al., 2015). Au cours de leur étude de CLIP-SEQ initiale, Lee et ses collaborateurs ont identifié une séquence d'ARNm H4 de 25 nt interagissant avec eIF3 (Lee et al., 2015). Cette séquence est située dans la phase codante de l'ARNm entre les nucléotides 294 et 319 (voir ***Figure 23***).

Des expériences d'analyses structurales en solution réalisées au laboratoire (D'Orchymont, 2013; Gross, 2018) ont permis de proposer un modèle de repliement de l'ARNm H4 de souris (voir ***Figure 24***) et de définir la position des principaux éléments structuraux et fonctionnels de la phase codante, dont la jonction à trois hélices TWJ (1-135) responsable de la reconnaissance de la coiffe et le site de liaison de eIF4E ou région 4E-SE (189-249). Le site d'interaction du facteur d'initiation eIF3 semble se situer sur l'un des côtés d'une longue structure en tige-boucle appelée eIF3-BS (eIF3 binding site) (250-320) dont le laboratoire a prédit le repliement mais dont la fonction était jusqu'alors inconnue (voir ***Figure 24***).

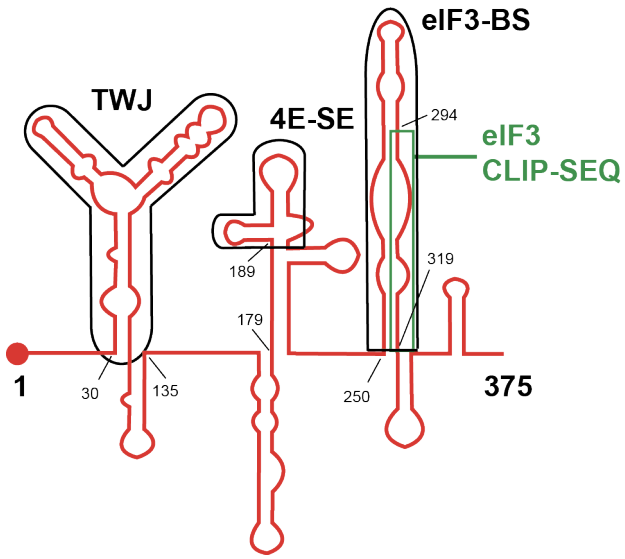


Figure 23. Schéma des différents éléments structuraux présents dans l'ARNm de l'histone H4 de souris. Des expériences de sondage en solution de l'ARNm de l'histone H4 de souris entier (375 nucléotides), ont montré qu'il présentait trois éléments structuraux. La première structure (30-135) appelée TWJ (Three Way Junction) est le site de liaison de la coiffe et de positionnement du ribosome. La deuxième structure (189-249) appelée 4E-SE est le site de liaison du facteur eIF4E. La

troisième structure (250-320) appelée eIF3-BS contient la région de 25 nucléotides en interaction avec eIF3 (294-319) (encadrée en vert) identifiée par les expériences de CLIP-SEQ (Lee et al., 2015).

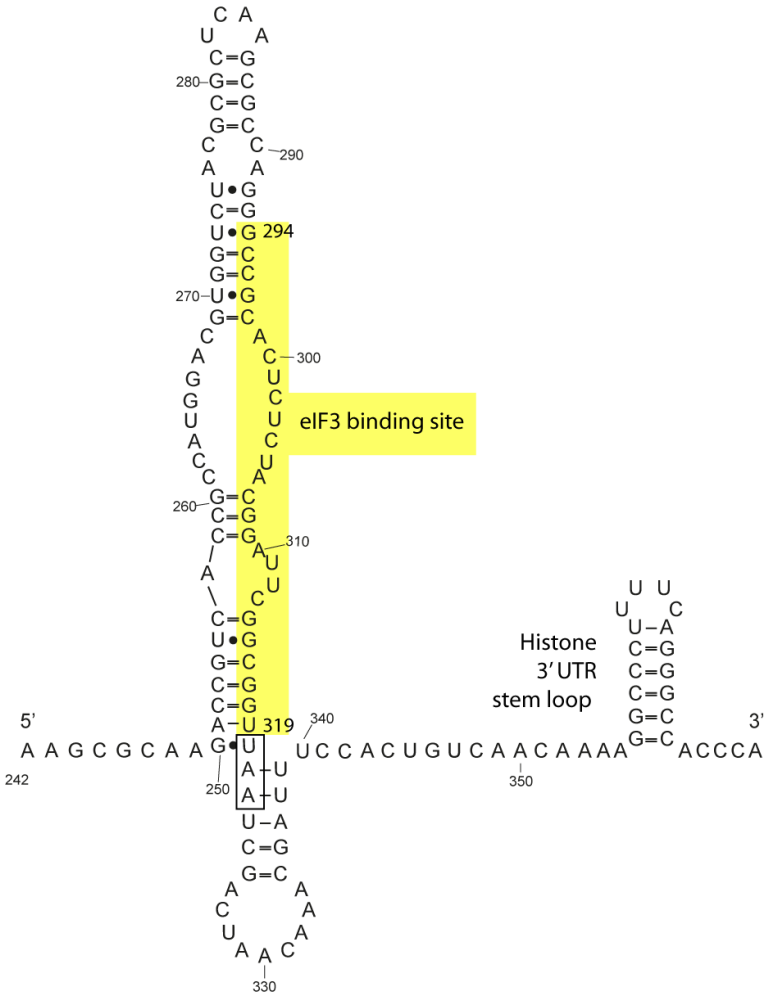


Figure 24. Séquence et modèle de repliement secondaire de la région 242-375 de l'ARNm de l'histone H4. Issu de la thèse (D'Orchymont, 2013). La région identifiée par CLIP-SEQ impliquée dans l'interaction avec eIF3 a été surlignée. Le codon stop est encadré. La structure en tige-boucle située à l'extrémité 3' UTR et caractéristique de tous les ARNm d'histones est indiquée.

Afin de vérifier l'interaction du facteur d'initiation eIF3 avec l'ARNm de l'histone H4 de souris *in vitro* des expériences de retard sur gel ont été réalisées par Laure Schaeffer. Ce test consiste à comparer, en conditions natives, le profil de migration sur gel de l'ARNm H4 isolé ou en complexe avec eIF3. L'interaction ARN-protéine se traduit par un retard de migration du complexe par rapport à l'ARN isolé.

Le facteur eIF3 que nous avons utilisé dans ces expériences a été obtenu par le laboratoire de W.C. Merrick (purifié à partir de réticulocytes de lapin) ; il comprend l'ensemble des 13 sous-unités. L'ARNm H4 entier (1-375) ainsi que deux fragments d'ARN H4 tronqués ($H4_{1\text{-}178}$) et ($H4_{179\text{-}375}$) générés par transcription *in vitro* et marqués de manière interne à l'aide d'[α^{32}P]-ATP ont été testés. Les différents transcrits ont été incubés en présence de concentrations croissantes d'eIF3 (1, 2 et 4 µM) puis séparés sur gel de

polyacrylamide 6% en conditions natives. La **Figure 25** montre qu'un retard sur gel est observé pour l'ARNm H4 entier (pistes 6-8) ainsi que pour le fragment $H4_{179-375}$ (pistes 14-16) dans lequel avait été identifié le site potentiel de liaison à eIF3. Seul un très faible retard sur gel est observé pour le fragment $H4_{1-178}$ en présence de fortes concentrations d'eIF3 (piste 12). Cette région de l'ARN ne semble donc pas contribuer à l'interaction avec eIF3. Aucune différence de mobilité de l'ARNm H4 n'est observée lorsque celui-ci est incubé avec les mêmes concentrations de BSA utilisée comme contrôle négatif (pistes 2-4) ou en l'absence de protéine (pistes 1, 5, 9 et 13).

Ce premier résultat révèle que le facteur d'initiation eIF3 interagit *in vitro* avec l'ARNm de l'histone H4 dans la région (179-375) contenant le site de fixation d'eIF3 identifié par les expériences de CLIP-SEQ. Il est vraisemblable que cette reconnaissance se fasse dans le contexte structural l'élément eIF3-BS, ceci devra être démontré.

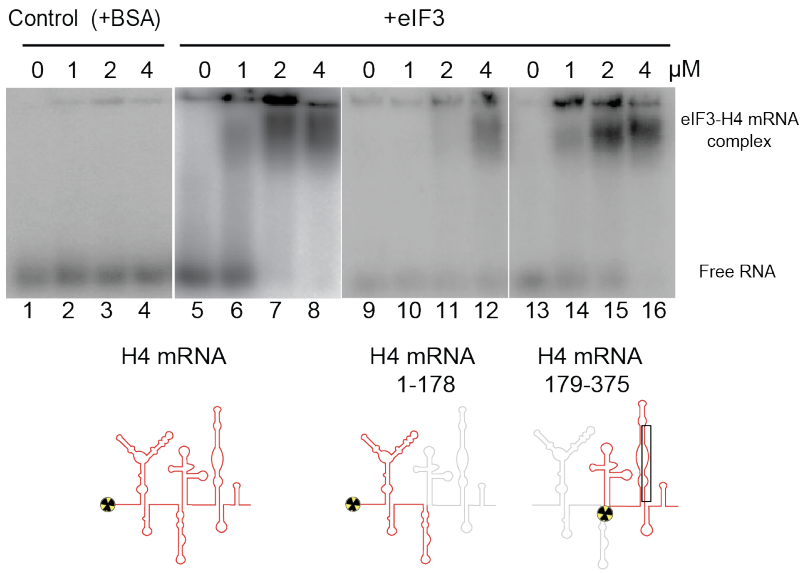


Figure 25. Le facteur eIF3 purifié interagit avec la région (179-375) de l'ARNm de l'histone H4 *in vitro*. Expériences de retard sur gel montrant l'interaction du facteur eIF3 avec des ARNm H4 transcrits *in vitro* et radiomarqués à l'aide d'[α^{32}P]-ATP, entiers (1-375) ou tronqués (1-178 et 179-375). Les transcrits ont été incubés en présence d'une quantité croissante d'eIF3 (1, 2 et 4 µM). eIF3 a été omis dans les pistes 1, 5, 9 et 13. La BSA (1, 2 et 4 µM) a été utilisée comme protéine de contrôle (pistes 2-4). Les complexes ont été séparés sur gel de polyacrylamide 6% en conditions natives.

1.2. eIF3 interagit *in vivo* avec les ARNm d'histones

Afin de déterminer si eIF3 est capable d'interagir avec l'ensemble des ARNm d'histones nous avons immunoprécipité les complexes ARN/eIF3 *in vivo* à partir de cellules HEK293FT. De par sa fonction le complexe eIF3 est majoritairement impliqué dans des interactions transitoires avec les ARNm. Afin de stabiliser ces interactions et de minimiser les

réarrangements de complexes RNP pouvant survenir lors de la lyse cellulaire nous avons optimisé une méthode de pontage à la formaldéhyde pour la préparation de nos extraits ribonucléoprotéiques (voir article sous presse en annexe (Hayek, H., Gross, L., ALHGOUL, F., Martin, F., Eriani, G., Allmang).

Le complexe eIF3 endogène ainsi stabilisé a été immunoprécipité à l'aide d'un anticorps dirigé contre la sous-unité eIF3b (Jivotovskaya et al., 2006; Lee et al., 2015). Les résultats de Western Blot (***Figure 26 A***) montrent que onze des treize sous-unités d'eIF3 (a, b, c, d, e, f, g, h, i, k, l) ont été spécifiquement co-immunoprécipitées par le biais de l'anticorps anti-eIF3b. Les sous-unités j et m n'ont pas été testées. La protéine GAPDH utilisée comme contrôle négatif n'a pas été retenue et aucune rétention aspécifique de sous-unités eIF3 n'a été observée en absence d'anticorps (***Figure 26 A***). Ces résultats montrent que nous avons immunoprécipité spécifiquement l'intégralité du complexe eIF3.

Les ARN associés à eIF3 ont été analysés par RT-qPCR (***Figure 26 B***). Nos résultats ont été comparés et normalisés par rapport à ceux obtenus pour l'ARNm LDHA. Les ARNm c-JUN et BTG1 sont des cibles validées d'eIF3 dans les expériences de CLIP-SEQ (Lee et al., 2015). Il apparaît que les ARNm c-JUN et BTG1 sont efficacement co-immunoprécipités par eIF3, avec un taux de rétention 2 fois plus faible pour BTG1 (conformément aux résultats obtenus précédemment) alors que ce n'est pas le cas de l'ARNm ACTB (Lee et al., 2015). Nos résultats indiquent que tous les ARNm d'histones sont retenus par eIF3. Le taux de rétention de l'ARNm de l'histone H4 (5,5 fois) est le plus élevé et se rapproche de celui de c-JUN (4 fois); celui des histones H1, H2A, H2B et H3 varie entre 1,5 et 1,75 et se rapproche de BTG1 (***Figure 26 B***). Ceci confirme que les ARNm d'histones sont également des cibles privilégiées du facteur eIF3. Les ARNm des gènes de ménages HPRT, GAPDH, LDHA, PGBK1 utilisés comme contrôles négatifs, tout comme ACTB, sont retenus à un taux inférieur ou égal à 1. Ce taux reflète vraisemblablement le rôle général d'eIF3 dans la traduction des ARNm. Aucune rétention du snARN U2 n'est observée.

Nous avons donc montré que l'ARNm H4, mais aussi les autres ARNm d'histones (H1, H2A, H2B, H3) sont capables d'interagir *in vivo* avec le complexe eIF3.

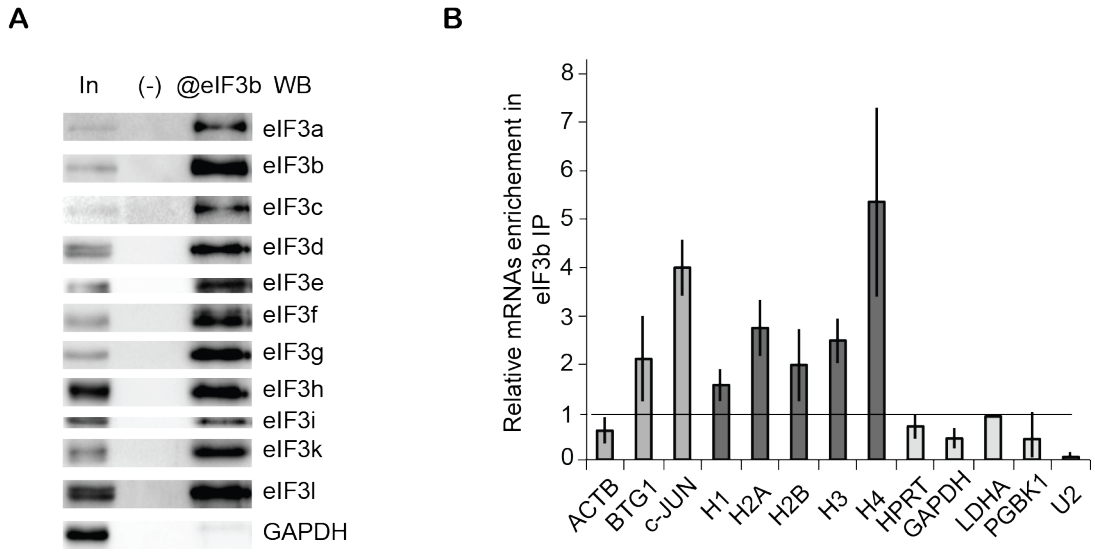


Figure 26. eIF3 interagit *in vivo* avec les ARNm d'histones. Immunoprécipitation du complexe eIF3 endogène et des ARN qui lui sont associés à l'aide de l'anticorps anti-eIF3b à partir d'extraits totaux de cellules HEK293FT. **(A)** Détection par Western Blot des différentes sous-unités d'eIF3 immunoprécipitées avec l'anticorps anti-eIF3b (@eIF3b). In : Input, (-) : Contrôle sans anticorps. GAPDH est utilisé comme contrôle négatif **(B)** Analyse par RT-qPCR des ARN messagers interagissant avec eIF3. Le taux de fixation des ARNm a été normalisé par rapport à celui de LDHA, il est matérialisé par une barre horizontale. Les messagers ACTB, BTG1 et c-JUN sont des cibles de eIF3 d'après Lee et al. (2015). H1, H2A, H2B, H3 et H4 sont les ARNm d'histones. Les ARNm HPRT, GAPDH, LDHA, PGBK1 et le snARN U2 sont utilisés comme des contrôles négatifs. Les barres d'erreurs représentent l'écart-type de trois expériences indépendantes.

1.3. Identification des sous-unités d'eIF3 interagissant avec l'ARNm H4 par pontage *in vitro*

Il a été montré précédemment qu'un nombre restreint de sous-unités du complexe eIF3 établissait des contacts directs avec les ARN. Des motifs de liaison aux ARN ont été identifiés dans les sous-unités eIF3a, b et g (Sun et al., 2011) alors qu'eIF3d est capable de se lier à la coiffe des ARNm par une poche de fixation dédiée (Lee et al., 2016). Seules les sous-unités eIF3a, b et c se fixent sur les éléments de type IRES (Cai et al., 2010; Jackson et al., 2010; Ujino et al., 2012). Les ARNm cibles d'eIF3 identifiés par CLIP-SEQ interagissent avec des combinaisons distinctes des sous-unités eIF3a, b, d et g (Lee et al., 2015). Ceci suggère qu'il puisse y avoir plusieurs modes d'interaction entre eIF3 et les ARNm.

Afin d'identifier précisément les sous-unités du complexe eIF3 en interaction directe avec l'ARNm H4, nous avons réalisé des expériences de pontage aux UV en utilisant un transcrit d'ARNm H4-thioU radio-marqué en présence de complexe eIF3 purifié (***Figure 27***). Après pontage de l'ARNm radioactif, une digestion par la RNase A est réalisée afin de dégrader la partie de l'ARNm non protégé par l'interaction avec les sous-unités d'eIF3. Ceci

permet le transfert d'un petit fragment d'ARNm radioactif vers les sous-unités d'eIF3 en contact direct avec l'ARNm. La **Figure 27 A** montre la séparation des protéines radiomarquées sur gel de polyacrylamide dénaturant en une dimension. L'analyse du gel révèle le marquage radioactif d'au moins 4 sous-unités d'eIF3 différentes et de poids moléculaires apparents de 110, 66, 50 et 44 kDa. Malheureusement, les expériences de western blot préliminaires n'ont pas permis d'identifier de façon formelle ces sous-unités car plusieurs présentent des poids moléculaires très proches (voir ***Introduction Tableau 1***). C'est le cas par exemple des sous-unités eIF3e, f et g (45 kDa) ; eIF3d et l (65 kDa) ainsi que eIF3a, b et c (110 kDa).

Pour résoudre définitivement les ambiguïtés d'attribution liées à la taille des sous-unités, nous avons séparé les produits des pontages par électrophorèse bidimensionnelle (2D) (***Figure 27 B et 27 C***). Les protéines ont été séparées dans la première dimension selon leur point isoélectrique par isoélectrofocalisation (IEF) puis selon leur poids moléculaire dans la seconde dimension (***Figure 27 B***). Les sous-unités eIF3 pontées au fragment d'ARN radiomarqué ont été identifiées à l'aide d'anticorps spécifiques dirigés contre 11 des 13 sous-unités après superposition des signaux de western blot (panneaux de gauche) et d'autoradiographie (panneaux droite) (***Figure 27 C et 27 D***). Nos résultats montrent clairement que les signaux de pontage correspondent aux sous-unités eIF3**c**, **d**, **e** et **g** (***Figure 27 D***). Ils permettent également d'exclure toute interaction des sous-unités eIF3 b, f, h, i, k et l avec l'ARNm H4 car les signaux de western blot ne se superposent pas avec les signaux de radioactivité (***Figure 27 C***). Nos résultats diffèrent partiellement de ceux publiés précédemment montrant que l'ARNm de H4 interagit avec eIF3b et d (Lee et al., 2015). Dans ce cas, les complexes eIF3/ARNm avaient été élués à partir de gels monodimensionnels qui ne permettaient pas de séparer les protéines eIF3b et c. Il est donc probable qu'il y ait eu co-migration des sous-unités eIF3 c et b pontées aux ARN ce qui aurait pu entrainer l'erreur lors de la confection des banques d'ARN pour le séquençage à haut débit.

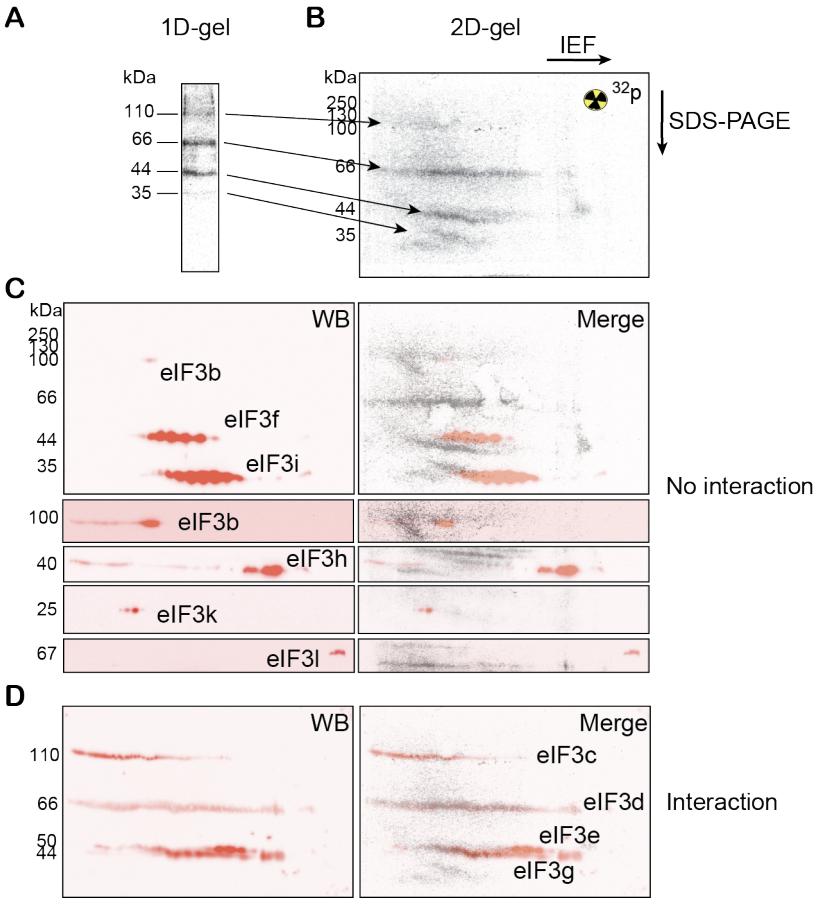


Figure 27. Identification des sous-unités de eIF3 en interaction avec l'ARNm de l'histone H4 par pontage aux UV et western blot. (A) Expérience de pontage aux $UV_{365\ nm}$ de l'ARNm thioU-H4 marqué au ^{32}P-α-ATP en présence d'eIF3 purifié. Après l'étape d'irradiation aux UV, l'ARNm a été digéré par la RNase A, permettant le marquage radioactif des sous-unités eIF3 et leur révélation par autoradiographie après séparation sur gel monodimensionnel. Le poids moléculaire apparent des sous-unités pontées est indiqué. **(B)** Séparation des produits de pontage sur gel bidimensionnel (2D). Dans la première dimension, les protéines sont séparées selon leur point isoélectrique par isoélectrofocalisation (IEF) entre les pH 4 à 7. Dans la seconde dimension les protéines sont séparées par SDS-PAGE. Les protéines pontées à un fragment d'ARN sont révélées par autoradiographie. **(C et D)** Identification des protéines pontées par western blot après transfert sur membrane de PVDF à l'aide d'anticorps dirigés contre les sous-unités individuelles d'eIF3 (signaux rouges, panneaux de gauche). Les anticorps individuels ont été utilisés successivement. La superposition des signaux radioactifs et des signaux de WB (Merge) issus de la même membrane de PVDF a conduit à l'identification des protéines pontées (panneaux de droite). Aucune interaction entre les sous-unités eIF3 b, f, h, i, k et l et l'ARNm H4 n'est détectée **(C)**. La superposition des signaux de WB et de radioactivité indique que les sous-unités d'eIF3 (c, d, e et g) interagissent avec l'ARN H4 **(D).**

Ici il est intéressant de commenter les profils de séparation des gels gel 2D. *In vivo*, les sous-unités d'eIF3 subissent de nombreuses modifications post-traductionnelles (Damoc et al., 2007). Après la migration en 2D, ces modifications conduisent à un profil de migration sous forme de plusieurs points contigus ne différant que par une charge. Ce profil est clairement modifié lorsque les protéines eIF3c, d, e et g sont pontées aux fragments d'ARN. Ces fragments d'ARN résultant de la digestion partielle par la RNase A ont des tailles diverses et entrainent un étalement de la migration sur une plus grande étendue de charges

(***Figure 28***). L'observation d'une différence des profils entre sous-unités pontées et non pontées conforte donc l'identification des 4 sous-unités pontées eIF3c, d, e et g (***Figure 28***).

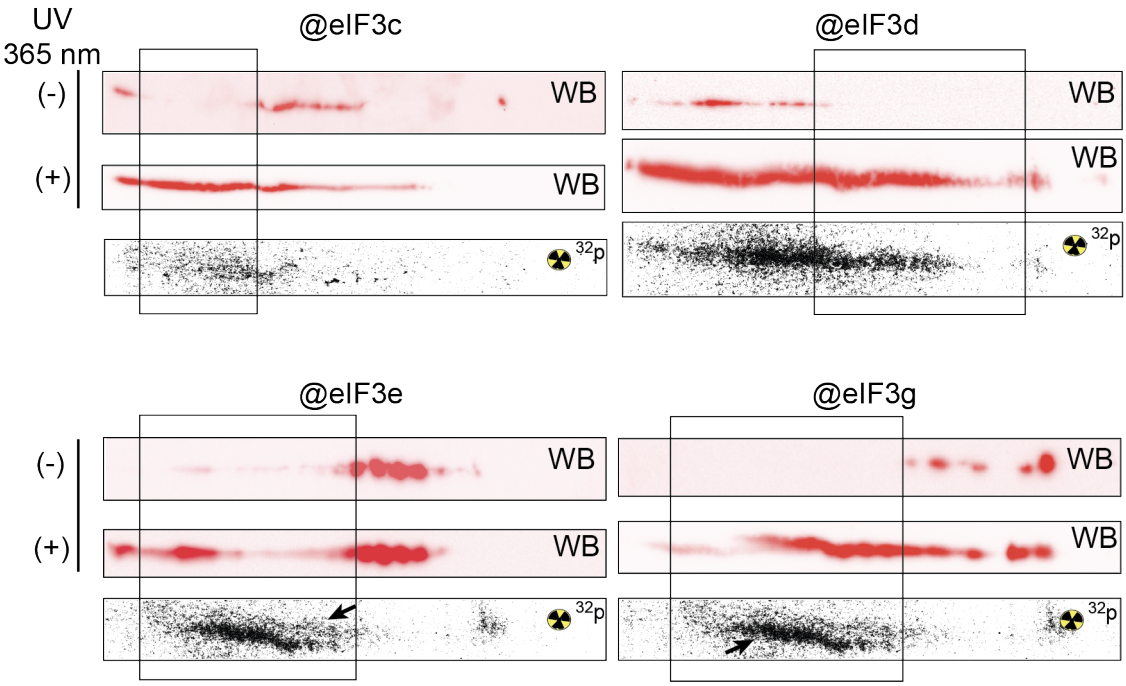


Figure 28. Comparaison des profils de migrations des protéines eIF3c, d e, et g sur gel 2D en présence et absence de pontage avec l'ARNm H4. Les expériences de pontage aux UV, séparations des échantillons sur gels 2D et analyses western blot ont été réalisées comme décrit dans la **Figure 27.** Les zones encadrées montrent les différences entre les profils de migration des protéines avant (-) et après pontage (+) à l'ARNm H4. Les flèches précisent la position du signal radioactif correspondant au pontage des sous-unités eIF3e et g.

1.5. Les quatre sous-unités d'eIF3c, d, e et g isolées interagissent avec l'ARNm H4 *in vitro*

Les quatre sous-unités d'eIF3c, d, e et g que nous avons identifiées en interaction avec l'ARNm H4 jouent des rôles importants dans la formation et le positionnement du complexe eIF3 dans le contexte du ribosome 80S. Parmi ces protéines, seule la sous-unité eIF3g possède des motifs de liaison à l'ARN de type 'RNA Recognition Motif' (RRM) (ElAntak et al., 2007) et eIF3d est capable de se lier à la coiffe de certains ARNm. Les sous-unités eIF3g et d font partie du module périphérique et s'attachent au cœur structural du facteur eIF3 d'une manière flexible notamment par le biais d'interactions entre eIF3d et la sous-unité eIF3e. Les sous-unités d'eIF3c et e appartiennent au cœur structural d'eIF3 (Ellisdon and Stewart, 2012; Enchev et al., 2010; Obayashi et al., 2017; Pena et al., 2007).

Afin de tester la capacité de ces 4 sous-unités à interagir directement et indépendamment du complexe eIF3 avec les ARNm d'histones, nous avons effectué des expériences de GST pull-down couplées à la détection des ARNm par qRT-PCR (***Figures 29 – 31***). Pour cette étude des protéines eIF3c, d, e et g recombinantes fusionnées à une double étiquette HisGST (l'hexa-histidine et la glutathion S-transférase) ont été exprimées dans *E. coli* et purifiées au laboratoire sur colonne de nickel Ni-NTA par Aurélie Janvier. Celles-ci sont représentées schématiquement dans la **Figure 29 A et B**.

Si les protéines eIF3d, e, et g entières (***Figure 29 B***) ont pu être purifiées sans problème de solubilité ou de dégradation majeure, il n'en va pas de même pour la protéine eIF3c pour laquelle de nombreux produits de protéolyse ont été obtenus (données non montrées). Nous nous sommes donc orientés vers la purification de versions tronquées HisGST-eIF3c 1-318 et HisGST-eIF3c 319-913 dont les extrémités avaient été définies lors d'études antérieures (Cai et al., 2010) (***Figure 29 A***). L'analyse par spectrométrie de masse Nano LC-MS/MS (plateforme de protéomique de l'IBMC) a par ailleurs révélé la présence d'un site de protéolyse majeur après le résidu Glycine 190. La présence de ce dernier révèle vraisemblablement la présence d'une région flexible entre 2 domaines structurés. En effet, alors que le domaine C-terminal est impliqué dans des interactions avec les protéines du cœur structural d'eIF3, le domaine N-terminal contient trois sous-domaines (entre les acides aminés 1-163) impliqués dans des interactions avec d'autres facteurs d'initiation tels eIF1 et eIF5 chez la levure (voir **Introduction 2.3.1.**). Nous avons également produit et purifié les protéines HisGST-eIF3c 1-190 et 191-318 (***Figure 29 A***).

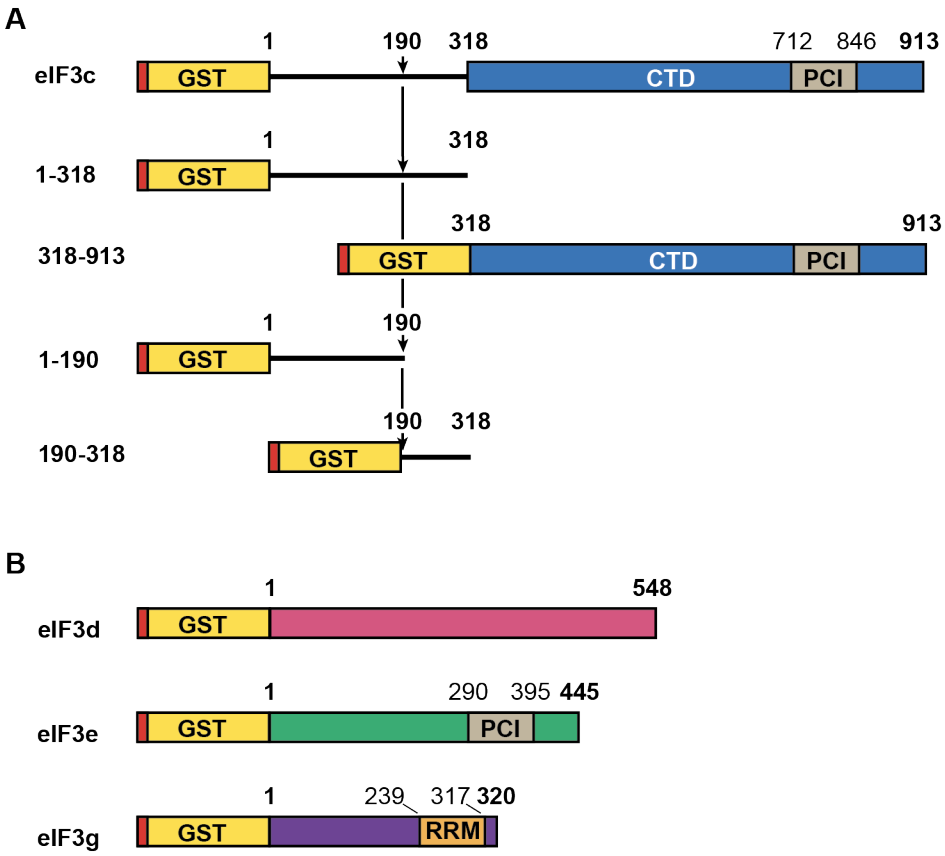


Figure 29. Représentation schématique des différentes protéines eIF3c, d, e et g recombinates
Les protéines eIF3c, d, e et g ont été fusionnées en N-terminal à une double étiquette : l'hexa-histidine et la glutathion S-transférase (HisGST) afin de permettre leur purification et leur utilisation dans les expériences de GSTpulldown. **(A)** La protéine d'eIF3c entière (1-913) contient dans sa région C-terminale le domaine PCI (712-846) pour 'Protéasome, COP9/signalosome, eIF3' ainsi qu'un domaine N-terminal qui s'étend dans le centre de décodage de la particule 40S et interagit avec d'autres facteurs d'initiation. Trois protéines tronquées recouvrent la région N-terminale (1-318) ; (1-190) ; (190-318) d'eIF3c et une protéine correspond à la région C-terminale (318-913). **(B)** Du haut, la protéine entière eIF3d (1-548) ; la protéine eIF3e entière (1-445) qui contient dans sa région C-terminale le domaine PCI (290-395) et en bas, la protéine eIF3g entière (1-320) qui contient dans sa région C-terminale le domaine RRM (239-317) pour 'RNA Recognition Motif'.

Pour les expériences de GST pull-down, les protéines recombinantes ont été incubées en présence d'ARN totaux extraits de cellules HEK293FT. La rétention des protéines recombinantes sur les billes d'agarose anti-GST a été validée par western blot dans chaque expérience (***Figure 30 et 31***). Les ARNm en interaction avec les protéines ont été analysés et quantifiés par RT-PCR (***Figure 30 et 31***). Comme décrit précédemment, les ARNm de c-JUN et BTG1 sont utilisés comme contrôles positifs. La fixation de l'ensemble des ARNm des histones (H1, H2A, H2B, H3 et H4) a été analysée. Enfin, les ARNm de gènes de ménage tels que la β-actine (ACTB), de la GAPDH, de la HPRT, de la LDHA et le snARN U2 ont été utilisés comme contrôles. Une rétention aspécifique des ARN sur l'étiquette 'HisGST'

de 1 % en moyenne a été mesurée et soustraite des valeurs obtenues pour les différentes ARNm retenus dans l'ensemble de nos expériences.

Les résultats des expériences de GST pull-down concernant les protéines eIF3c tronquées sont représentés dans la **Figure 29 A**. Ils montrent que la protéine HisGST-eIF3c 1-318 interagit fortement avec les ARNm des histones H1, H2A, H3, et H4 avec un taux de fixation de 38%, 44%, 42% et 44% respectivement. Ces valeurs sont environ 3 fois plus élevées que celle observée pour l'ARNm de c-JUN qui a un taux de fixation de 15 %. La fixation de l'ARNm de BTG1 par la protéine HisGST-eIF3c 1-318 est similaire aux ARNm contrôles qui ont un taux de fixation d'environ 10%. On constate néanmoins que le taux de fixation de l'ARNm de l'histone H2B se situe au même niveau que ceux retrouvés pour les ARNm contrôles, indiquant que l'ARNm de l'histone H2B est le seul ARNm d'histone à ne pas être reconnu préférentiellement par la région 1-318 d'eIF3c. Globalement, le taux d'ARNm fixé par la protéine HisGST-eIF3c 1-318 est considérablement plus important que celui fixé par la protéine HisGST-eIF3c 319-913 (***Figure 30 A***). En effet, le taux de fixation pour la construction HisGST-eIF3c 1-318 est en moyenne de 40 % alors qu'il n'est que de 3 % pour HisGSTeIF3c 319-913. Ces résultats indiquent que le domaine de liaison aux ARNm d'eIF3c se situe entre les acides aminés 1-318. L'ensemble de ces expériences montre que la sous-unité eIF3c isolée est capable d'interagir avec les ARNm d'histones et notamment l'ARNm H4, confirmant les résultats de pontages obtenus précédemment. Par ailleurs, eIF3c interagit préférentiellement avec les ARNm d'histones H1, H2A, H3, et H4 *in vitro* et ceci par l'intermédiaire de sa région 1-318. La capacité des protéines HisGST-eIF3c 1-190 et 191-318 à interagir avec les ARNm a été testée afin d'affiner les bornes du domaine d'interaction (***Figure 30 B***). On observe que le taux des ARNm fixé par la protéine HisGST-eIF3c 1-190 n'est que de 1% en moyenne, et est inférieur aux ~5 % d'ARNm fixé pour HisGST-eIF3c 191-318. Les ARNm des histones H1, H3 et H4 restent toutefois préférentiellement fixés. Ceci suggère que le domaine de reconnaissance des ARN est altéré dans ces fragments et que la totalité de la région 1-318 est nécessaire à l'interaction avec les ARNm. La région 191-318 d'eIF3c contiendrait néanmoins la majeure partie du domaine de reconnaissance des ARN.

Les résultats des expériences de GST pull-down concernant les protéines eIF3d, e et g sont représentés dans la **Figure 31**. Il apparaît que la protéine HisGSTeIF3d interagit très distinctement avec l'ensemble des ARNm d'histones avec un taux de fixation d'environ 5% pour H1 et H4, de 3% pour H2A et H3 et de 1% pour H2B (***Figure 31 A***). Bien que modestes ces taux de rétention sont 10-50 fois plus élevés que ceux observés pour les ARNm de ACTB, HPRT, GAPDH, LDHA et du snARN U2 qui sont de 0,1% en moyenne. Le niveau de rétention d'H2B est par ailleurs comparable à celui de c-JUN qui n'est que faiblement retenu alors que BTG1 ne l'est pas.

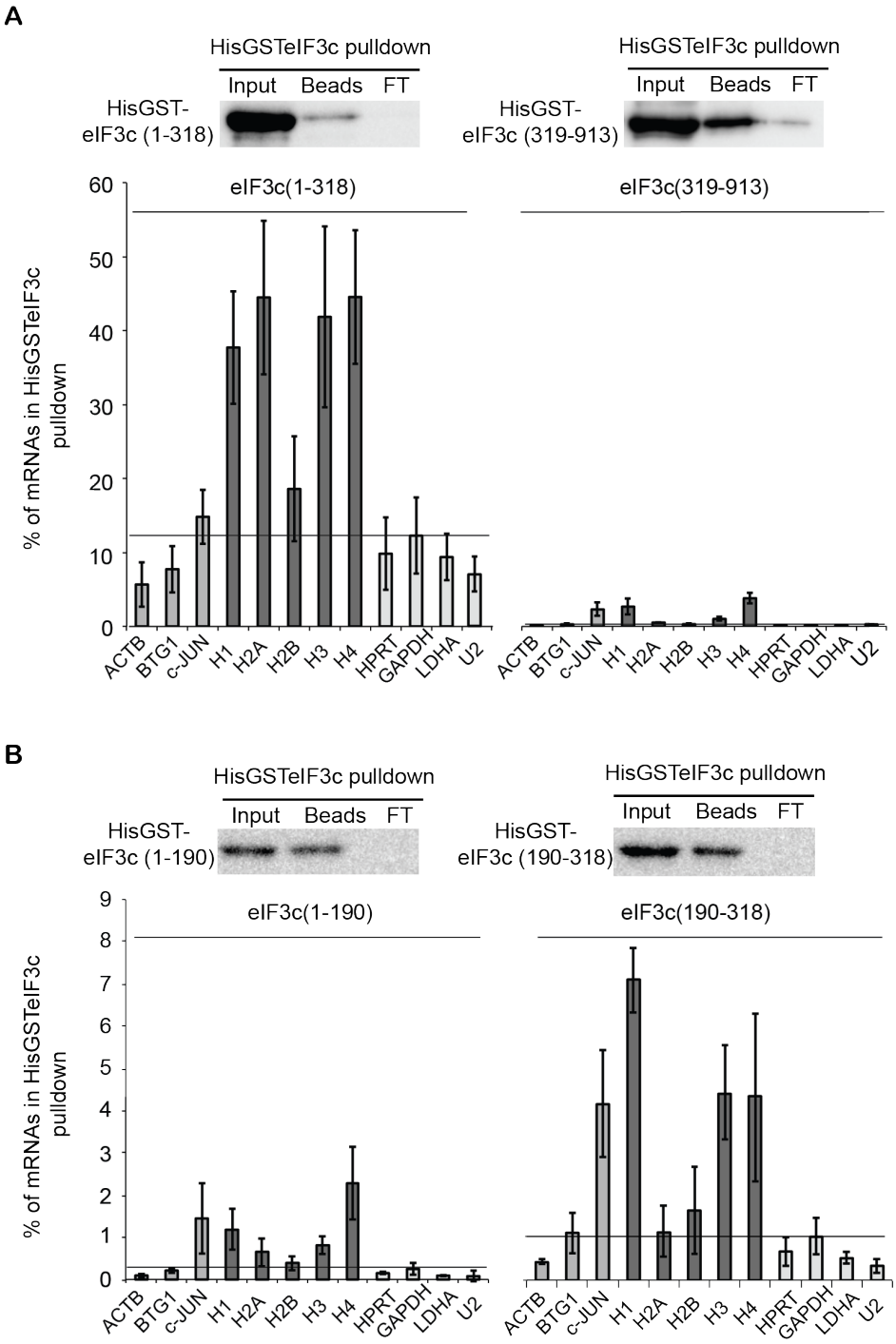


Figure 30. L'extension N-terminale d'eIF3c interagit avec les ARNm des histones. Résultats des expériences de GST pull-down effectuées à l'aide des protéines HisGST-eIF3c 1-318, 319-913 et HisGST **(A)** et les protéines HisGST-eIF3c 1-190, 191-318 **(B)** purifiées. Pour chaque expérience de GST pull-down la rétention des protéines HisGST sur les billes anti-GST a été vérifiée par western blot (encarts au dessus des histogrammes). Input : protéine purifiée. Beads : protéine eIF3c retenue sur les billes anti-GST. FT : effluent. Les histogrammes **(A et B)** représentent le pourcentage d'ARNm lié aux différentes protéines tronquées d'eIF3c et quantifiés par RT-qPCR. Les messagers BTG1 et c-JUN sont utilisés comme contrôles positifs. H1, H2A, H2B, H3 et H4 sont les ARNm des histones. Les ARNm de la β-actine (ACTB), de HPRT, GAPDH, LDHA et du snARNU2 sont utilisés comme des contrôles négatifs. Les barres d'erreurs représentent l'écart-type de trois réplicats biologiques. La ligne horizontale indique le taux de fixation pour l'ARNm GAPDH (contrôle négatif).

Le profil d'interaction de la protéine HisGSTeIF3g (***Figure 31 C***) révèle que c'est l'ARNm H1 qui est le plus fortement retenu avec un taux de 17% alors que les autres ARNm d'histones le sont à un taux d'environ 5% assez proche de celui de c-JUN (8%) contre 2% pour les ARNm de contrôle. Enfin, nos résultats montrent que la sous-unité isolée HisGSTeIF3e interagit avec l'ARNm de l'histone H1 (14%) mais pas avec H2A, H2B et H4. Elle n'interagit pas non plus avec les ARNm de c-JUN et BTG1 pour lesquels un taux de rétention similaire à celui des ARNm de contrôle HPRT, GAPDH et LDHA a été observé (***Figure 31 B***).

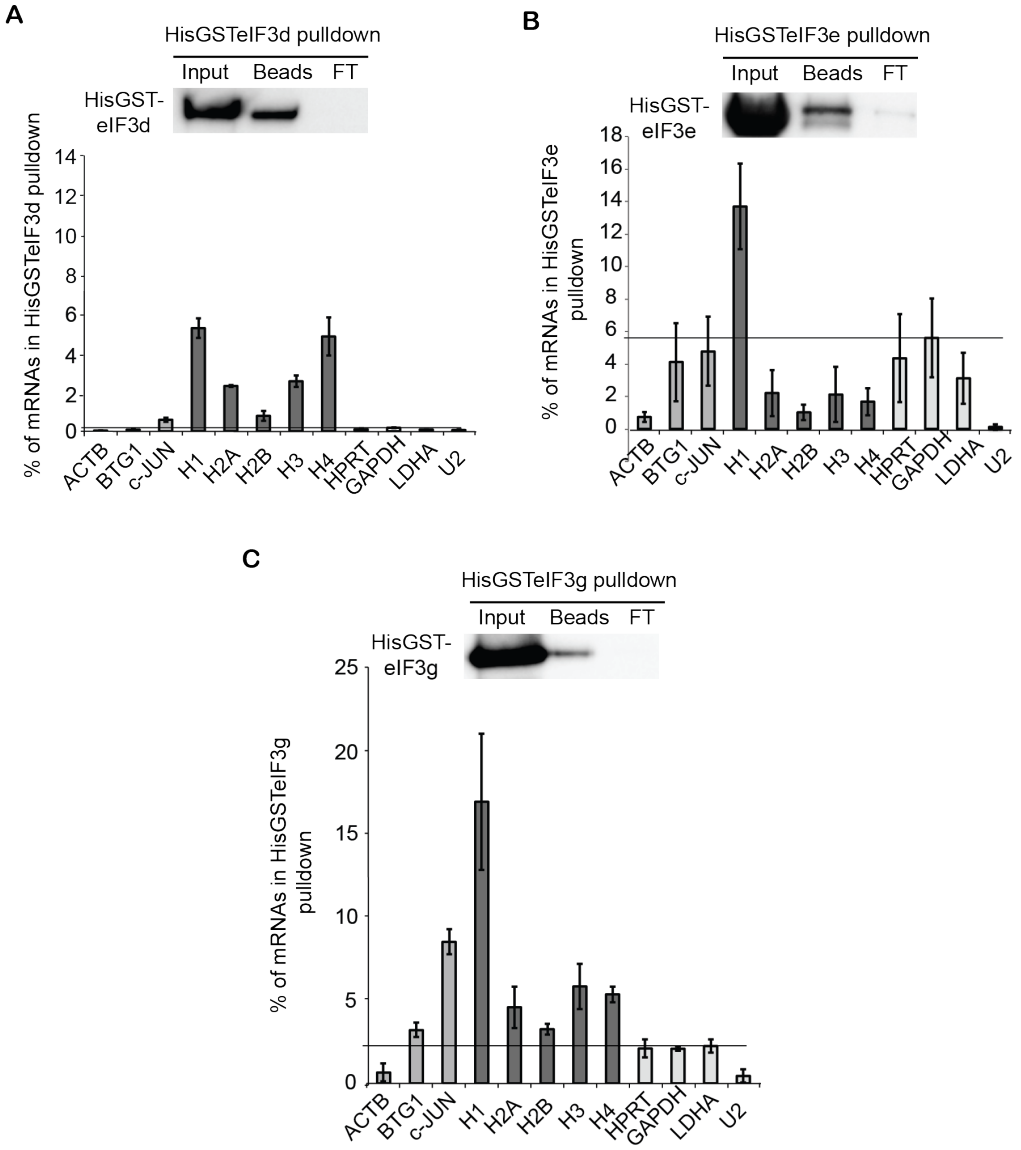


Figure 31. Interaction des ARNm d'histones avec les protéines recombinantes HisGSTeIF3 d, e et g. Résultats des expériences de GST pull-down effectuées à l'aide des protéines HisGST-eIF3d **(A)**, HisGST-eIF3e **(B)** et HisGST-eIF3g **(C)**. L'efficacité de l'expérience de GST pull-down a été vérifiée par western blot. Le pourcentage d'ARNm lié aux protéines tronquées d'eIF3d, g et e a été quantifié par RT-qPCR L'ensemble des résultats est représenté comme décrit dans la **Figure 30**.

En conclusion, nous avons montré que les trois sous-unités eIF3 c, d et g sont capables d'interagir avec les ARNm d'histones H1, H2A, H2B, H3 et H4 *in vitro* indépendamment du complexe eIF3. eIF3e en revanche interagit uniquement avec l'ARNm de l'histone H1 et pas avec l'ARNm H4 comme le suggéraient nos expériences de pontage (voir **récapitulatif *Tableau 3***). Il est possible que l'interaction d'eIF3e avec l'ARNm H4 ne puisse avoir lieu que dans le contexte du complexe eIF3 entier ou que le pontage H4/eIF3e résulte d'une proximité structurale de la sous-unité eIF3e avec l'ARNm. Dans l'ensemble, nos résultats révèlent l'existence de schémas d'interaction différents entre le complexe eIF3 et les différents ARNm d'histone comme suggéré également pour d'autres ARNm cibles de eIF3 (Lee et al., 2015)

mRNA	eIF3c	eIF3d	eIF3e	eIF3g
H1	✓	✓	✓	✓
H2A	✓	✓	X	✓
H2B	+/-	✓	X	+/-
H3	✓	✓	X	✓
H4	✓	✓	X	✓
BTG1	X	X	+/-	+/-
c-JUN	✓	✓	+/-	✓

Tableau 3. Récapitulatif des interactions des différentes protéines recombinantes HisGSTeIF3c, d, e et g avec les ARNm d'histones ainsi qu'avec c-JUN et BTG1, obtenues dans les expériences de HisGST pulldown.

2. Etude de rôle de l'interaction d'eIF3 sur la traduction des ARNm des histones

Dans la première section des résultats, nous avons établi le schéma d'interaction entre les sous-unités d'eIF3 et les ARNm d'histones. Dans cette nouvelle section, nous avons cherché à comprendre la fonction de ces interactions.

2.1. Impact de l'élément régulateur de l'ARNm H4 sur la traduction d'un gène rapporteur

Dans une première approche nous avons tenté de reconstituer l'interaction ARNm d'histone-eIF3 dans un système utilisant un gène rapporteur. Disposer d'un tel système nous aurait permis d'aborder plus facilement les études structurales et fonctionnelles. A cette fin, l'élément structural de l'ARNm H4 identifié par séquençage haut débit lors des expériences de PAR-CLIP-SEQ (Lee et al., 2015) a été placé dans les régions non traduites du gène de la Renilla Luciférase (R-Luc ci-après) afin d'observer l'effet sur la traduction de celle-ci.

Les travaux de séquençage de Lee et coll. (2015) ont montré que le facteur eIF3 se fixait le plus souvent sur des éléments structuraux d'ARNm localisés dans les régions 5' UTR (environ 70% des cas) (Lee et al., 2015). Les 30% des cas restants se répartissent entre les régions 3' UTR et les phases codantes des ARNm. Concernant l'ARNm de l'histone H4 (de 375 nucléotides de longueur), l'élément fixant le facteur eIF3 est une séquence de 25 nucléotides, située entre les nucléotides 294 et 319 dans la phase codante (Lee et al., 2015). Cette séquence est située dans une longue structure en tige-boucle identifiée par sondage en solution dans notre laboratoire (***Figure 32 A, à gauche et Figure 24***) (D'Orchymont, 2013).

Afin de tester l'effet de cette tige-boucle sur la traduction, nous avons généré différentes constructions dans lesquelles le gène rapporteur R-Luc était tantôt placé en 5' ou en 3' de la tige-boucle de H4.

D'un point de vue expérimental, l'approche choisie a été la suivante : des fragments d'ADN présentant les assemblages désirés ont été produits dans un premier temps par amplification PCR. Ces assemblages présentaient tous en amont le promoteur de l'ARN polymérase du phage T7 pour la production d'ARN et le gène rapporteur R-Luc pour la détection de l'activité de traduction. Pour insérer l'élément régulateur de H4 en 5' ou 3' de l'ORF de R-Luc, sa séquence était respectivement contenue dans l'oligonucléotide de PCR 'Forward' ou Reverse' (***Tableau M5***). Grâce à la réaction de PCR réalisée sur matrice d'ADN d'un gène R-Luc présent dans un plasmide 'pUC19 a3 Glo R-Luc' nous avons ainsi pu obtenir les constructions attendues qui ont été transcrites par la T7 ARN polymérase *in vitro* (***Figure 32 & Matériel et Méthodes***). Les ARN produits et purifiés ont ensuite été coiffés en 5' par la coiffe m^7GpppG à l'aide du kit 'ScriptCap m7G Capping System' (Cellscript).

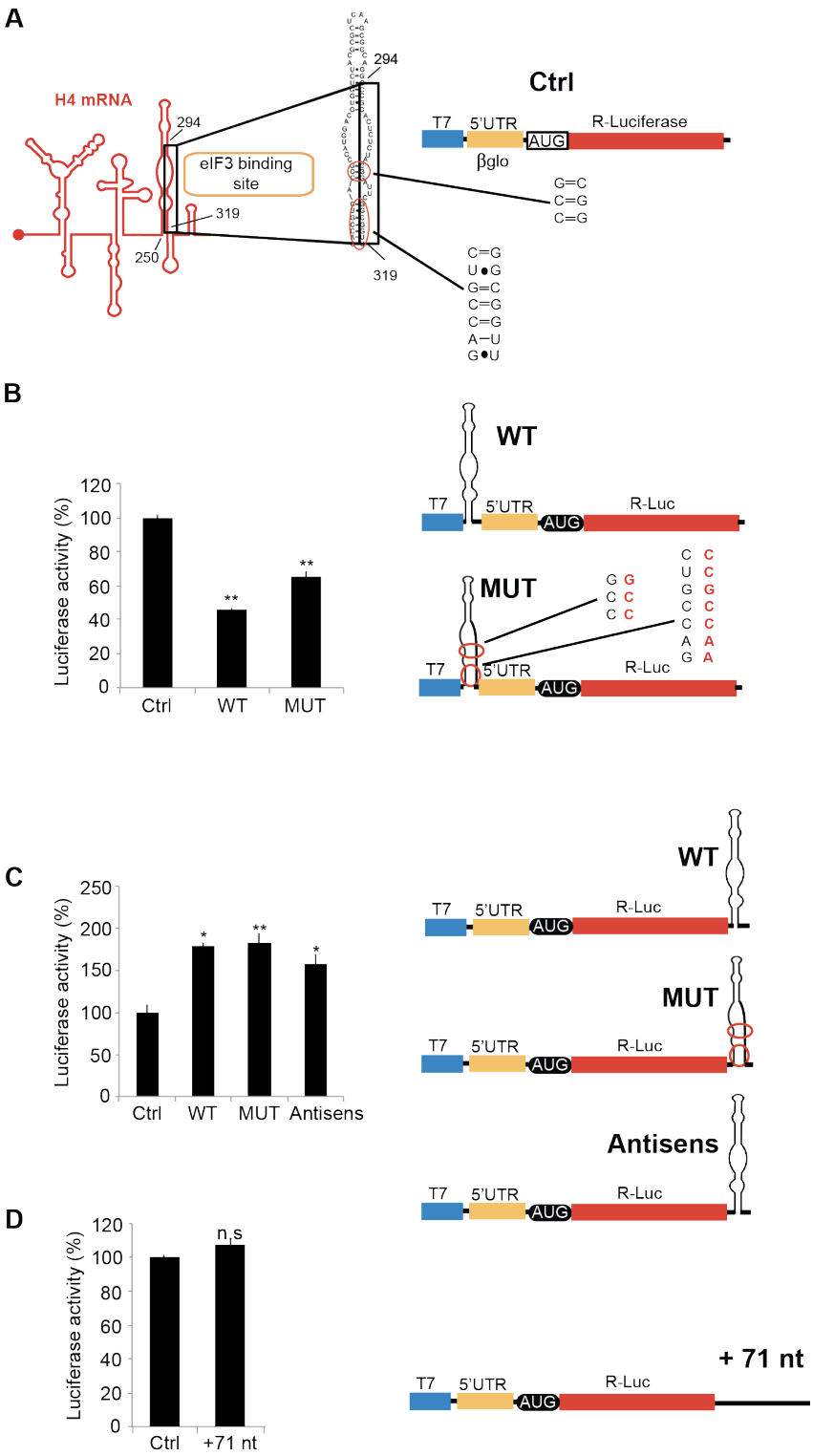


Figure 32. Effet de la tige-boucle (250-320) de l'ARNm de l'histone H4 sur la traduction du gène rapporteur R-Luc. (A) A gauche : Structure schématique du repliement de l'ARNm de l'histone H4. La longue tige-boucle contenant le site d'interaction d'eIF3 (25 nucléotides, résidus 294 à 319) est encadrée en noir et la séquence est présentée dans l'agrandissement. A droite : construction contrôle 'Ctrl' utilisée durant les expériences. Le fragment d'ADN qui a été transcrit présentait de 5' vers 3' un promoteur pour la T7 ARN polymérase, l'extrémité 5' UTR de l'ARNm de la b-globine, et la séquence de la Renilla luciférase (R-Luc). **(B)** A droite : constructions dérivées du fragment précédent par insertion de la tige-boucle de H4 fixant eIF3 en amont de la séquence 5' UTR de la β-globine. La construction sauvage 'WT' contient les nucléotides natifs de l'ARNm H4 ($_{306}$CGG$_{308}$ et $_{313}$GGCGGTT$_{319}$, cerclés de rouge) alors que dans la construction 'MUT', les nucléotides sont remplacés par des complémentaires (GCC et CCGCCAA). A gauche sont représentées les activités R-Luc, normalisées par rapport au 'Ctrl'. **(C)** A droite : constructions présentant la tige-boucle de H4

insérée en aval de la séquence codante de R-Luc. Les tige-boucles 'WT' et 'Mut' présentent les mêmes séquences qu'en **(B)**. La construction 'Antisens' présente une tige-boucle dont les nucléotides ont été entièrement remplacés par leurs complémentaires (selon Watson Crick). A gauche sont représentées les activités R-Luc, normalisées par rapport au 'Ctrl'. **(D)** Construction présentant à l'extrémité 3' un fragment de 71 nt de l'extrémité 3' UTR de la Renilla luciférase. Ce fragment présente la même longueur que la tige-boucle de H4. A gauche sont représentées les activités R-Luc, normalisées par rapport au 'Ctrl'. Toutes ces constructions ont été transcrites par la T7 ARN polymérase et les ARNm à la concentration finale de 2 μM ont été traduits dans les lysats de réticulocytes de lapin (RRL).

La réaction de traduction a été effectuée dans des lysats de réticulocytes de lapin (RRL ci-après) (***2.15 de Matériel et Méthodes***). La quantité de luciférase produite par traduction *in vitro* est déterminée par mesure de l'activité luciférase à l'aide d'un kit 'Dual-Glo® Luciferase Assay System' (Promega).

Notons ici que dans toutes les constructions la région 5' UTR de la β-globine (βglo) a été placée en amont de l'ORF de R-Luc (***Figure 32***). Cette séquence est non structurée et particulièrement adaptée à la traduction dans les RRL puisque la β-globine est très fortement produite dans ces cellules.

Une première série d'expériences a montré que lorsque l'élément en tige-boucle de H4 était localisé en amont du rapporteur R-Luc, une chute de 54% de l'activité R-Luc était observée par rapport à l'ARNm témoin qui en était dépourvu (construction Ctrl dans la ***Figure 32 A***). Une telle inhibition pourrait être due à un ralentissement de la vitesse de 'scanning' due à la présence de la tige-boucle à déplier (Kozak, 1990), mais aussi résulter d'un effet inhibiteur lié à la fixation d'eIF3 comme observé dans le cas de BTG1 (Lee et al., 2015). Nous avons donc construit un variant dans lequel les nucléotides (250-320) du fragment identifié par CLIP-SEQ ont été mutés de façon à les rendre anti-complémentaires pour leurs vis-à-vis dans la tige-boucle (construction 'MUT' dans la ***Figure 32 B***). Par conséquent, cette construction ne devrait plus présenter ni la structure en tige-boucle originelle, ni la séquence d'interaction avec eIF3. Après traduction, l'activité R-Luc de la construction MUT est supérieure à celle de la construction WT mais reste inhibée à 35% par rapport au Ctrl. Ce résultat montre que la traduction reste diminuée malgré l'absence de la séquence de fixation à eIF3 et ceci avec une structure partiellement dépliée au niveau de la base de la tige-boucle. Cela suggère que l'effet inhibiteur résiduel pourrait être indépendant d'eIF3 et pourrait résulter de la présence de la partie supérieure inchangée de la tige-boucle. Ce constat laisse penser que l'effet observé résulte d'un simple effet de ralentissement du scanning et donc de la traduction au niveau de la tige-boucle et n'est pas lié à la fixation d'eIF3.

Dans une deuxième série de constructions, nous avons testé l'effet de l'insertion de la tige-boucle de H4 en aval de l'ORF de R-Luc (***Figure 32 C***). En effet, dans l'ARNm de H4 la séquence de la tige-boucle est localisée dans la dernière partie de la phase codante de

l'histone. Elle pourrait donc agir à distance sur le recrutement de l'appareil d'initiation. Pour mimer cette situation, nous avons inséré la tige-boucle immédiatement après le codon stop de la séquence de la R-Luc. Nous ne l'avons cependant pas inséré dans la phase codante par crainte de modifier l'activité enzymatique de la R-Luc. De cette façon, nous escomptions reproduire l'effet à distance du recrutement d'eIF3 sur la tige-boucle.

Comme précédemment une tige-boucle mutée (MUT dans la ***Figure 32 C***) a été insérée en aval de la R-Luc, ainsi qu'une tige-boucle antisens (complémentaire à la séquence WT) ('Antisens' dans la ***Figure 32 C***) dont toutes les paires de bases ont été inversées ce qui doit en principe conduire à un élément structural dont la séquence et le repliement sont différents de ceux de l'élément originel. Après traduction des produits de transcription l'activité luciférase résultant des constructions WT, MUT et Antisens augmente respectivement de 78, 82 et 57% par rapport au Ctrl (***Figure 32 C***). Cette stimulation de l'activité avec les 3 ARNm, présentant ou non un site de liaison intègre pour eIF3, suggère fortement qu'il ne s'agit pas d'un effet attribuable à la fixation d'eIF3. Suspectant un effet de stabilisation de l'ARNm lié à l'ajout des séquences à l'extrémité 3', nous avons modifié la construction Ctrl dépourvue de 3' UTR en y insérant une séquence de 71 nt trouvée dans la région 3' UTR de la R-Luc. Cependant, cette construction rallongée d'une longueur équivalente à la tige-boucle de H4 ne présentait pas la même stimulation de traduction et l'activité luciférase restait très semblable à celle de la construction Ctrl de la R-Luc non rallongée (***Figure 32 D***).

En conclusion, ces derniers résultats suggèrent que la tige-boucle de H4 est capable de stimuler la traduction lorsqu'elle est placée au niveau de l'extrémité 3'. Cependant, l'ouverture partielle de la tige (construction MUT) ou l'inversion complète de la séquence (construction Antisens) qui en principe devraient détruire le site de liaison à eIF3, ont le même effet ce qui suggère que ce résultat est indépendant de la fixation d'eIF3. L'effet stimulant pourrait résulter de la présence des tige-boucles, qui dans les 3' UTR pourraient agir sur le repliement ou la stabilité des ARN (éventuellement via des facteurs additionnels), les rendant plus ou moins aptes à être traduits.

Face à la difficulté à normaliser les données issues de ce genre d'expériences où les UTR sont différentes et les résultats obtenus difficilement comparables, nous avons décidé d'utiliser le test du bicistron sur la tige-boucle d'intérêt de H4. Le système du bicistron présente l'avantage d'une normalisation plus facile, mais il être la cause d'artéfacts que nous redoutions et qui nous avaient conduits à privilégier les mono-cistrons *in vitro* comme décrit ci-dessus. Ici nous avons choisi le test du bicistron *in vivo* dans lequel l'ADN correspondant est transfecté dans des cellules eucaryotes. Celles-ci transcrivent l'ADN en ARN, produisant

un ARN bicistronique traduit dans la cellule. La préparation d'extraits cellulaires protéiques permet ensuite de mesurer l'activité des 2 luciférases du bicistron. Le test du bicistron a été développé pour mesurer les activités IRES (Pelletier and Sonenberg, 1988). Dans ce test, l'IRES ou la séquence à tester, est placée entre 2 cistrons, dans notre cas la Firefly luciferase (F-Luc) en 5' et la R-Luc en 3'. La traduction du premier cistron est dépendante de la coiffe tandis que l'expression du second dépendra uniquement de la capacité de la séquence placée entre les 2 cistrons à recruter les ribosomes pour initier sur le second cistron. En effet, contrairement à la situation des procaryotes, les ribosomes eukaryotes ne peuvent ré-initier la traduction sur un second cistron (exception faite des uORF où la ré-initiation peut survenir à courte distance du codon stop). Dans ce système, notre attente était de provoquer le recrutement d'eIF3 sur la séquence régulatrice de H4 placée dans la région intergénique ce qui entrainerait le recrutement des particules 43S pour une initiation sur l'AUG de la R-Luc.

Pour réaliser ces expériences, nous avons utilisé le vecteur eucaryote pcDNA3.1/Zeo(+) dans lequel était déjà construit un bicistron exprimable *in vivo* grâce au promoteur CMV (don du Dr. Annie Cahour, CERVI, laboratoire de virologie de la Pitié, Paris). Dans ce vecteur bicistronique, la F-Luc est placée en amont de la R-Luc, la région intergénique étant séparée par des sites de restriction dans lesquels ont été clonés les différents fragments testés.

Comme contrôle positif du bon fonctionnement du biscistron nous avons utilisé une construction contenant l'IRES de HCV dans la région intergénique. Le contrôle négatif était constitué du bicistron « vide » sans insert dans la région intergénique. Une autre construction contenait la séquence entière de l'ARNm de l'histone H4 et enfin nous avons fragmenté la séquence de H4 en 3 fragments (1-138 ; 136-225 ; 244-339) dans lesquels avaient été identifiées des structures fonctionnelles auparavant. Le fragment 1-138 de H4 contient la structure repliée en jonction triple ('TWJ' sur la ***Figure 33 A***) séquestrant la coiffe et servant au positionnement du ribosome sur l'AUG initiateur (Martin et al., 2016). Le fragment 136-225 contient le site de liaison à eIF4E appelé '4E-SE' (eIF4-Sensitive Element) (Martin et al., 2011). Enfin le fragment 244-339 présente la longue tige-boucle fixant eIF3 (Lee et al., 2015) qui nous intéresse tout particulièrement. La construction de ces bicistrons a été réalisée par « clonage HiFi » des fragments de gène d'H4 préalablement amplifiés par PCR. Ce travail a été réalisé par A. Janvier à l'aide du 'NEBuilder® HiFi DNA Assembly Cloning Kit'.

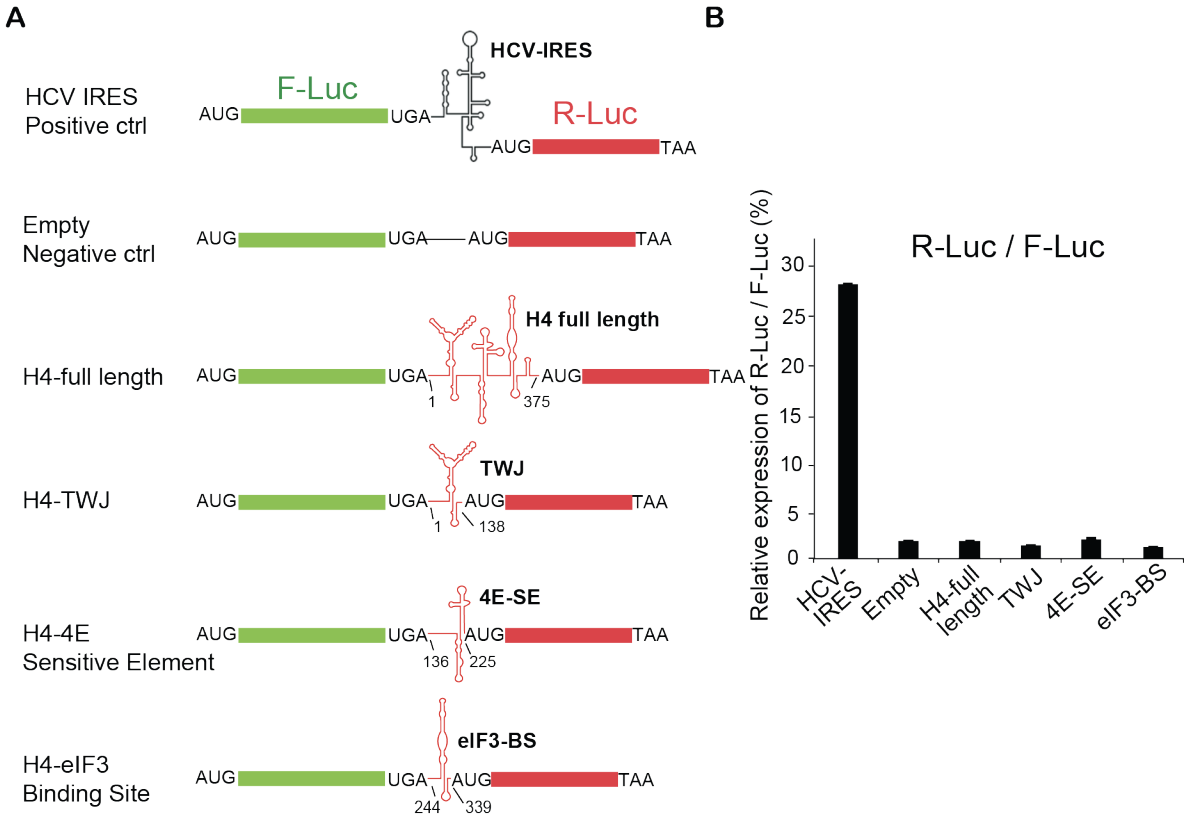


Figure 33. Mesure de l'activité de recrutement interne des ribosomes par le test du bicistron. **(A)** Structures schématiques des différentes constructions testées. Tous les bicistrons construits présentent le gène de la firefly luciferase en amont et de la renilla luciferase en aval. Les différents fragments d'ARN à tester ont été insérés dans la région intergénique entre les 2 luciférases. La construction 'HCV IRES' est utilisée comme témoin positif et le bicistron vide comme témoin négatif ('Empty'). La construction 'H4 full length' contient l'intégralité de la séquence de l'ARNm de l'histone H4 tandis que les 3 suivantes contiennent des fragments présentant des éléments structuraux distincts. Les chiffres réfèrent à la position des nucléotides dans l'ARNm H4. **(B)** Résultats des activités luciférases mesurées après transfection de cellules HEK293FT par les plasmides exprimant les bicistrons. Les résultats sont exprimés en rapports R/F-Luc et normalisés par rapport à l'IRES de HCV.

Les cellules HEK293FT ont été transfectées par les différents plasmides pcDNA/Zeo(+) à l'aide de lipofectamine, après 48 heures les 2 activités luciférase ont été mesurées à l'aide d'un luminomètre à partir d'extraits protéiques natifs comme décrit dans Matériel et Méthodes. Les résultats de la **Figure 33 B** sont représentés sous forme de rapports des activités R-Luc / F-Luc. Ce rapport permet de normaliser les résultats et d'effacer les variations d'efficacité de transfection, de préparation des extraits protéiques ou de stabilité des différents ARNm bicistroniques. Le fait que le rapporteur témoin (F-Luc) soit intégré au bicistron est un avantage majeur du système du bicistron qui permet de normaliser les résultats d'une construction à l'autre. De toutes nos constructions, seule celle contenant l'IRES de HCV présentaient une forte activité R-Luc ce qui confirme que l'IRES est bien capable de recruter les ribosomes de façon interne et d'initier la traduction sur l'AUG

du second cistron. Certaines autres constructions ont parfois montré une faible activité R-Luc, cependant celle-ci était « gommée » après l'étape de normalisation vis-à-vis de F-Luc (***Figure 33 B***).

De ces expériences menées *in vitro* sous forme de monocistron ou *in vivo* dans un bicistron, nous pouvons conclure que nous n'avons pas pu reconstituer une interaction eIF3-ARNm H4 détectable par un système rapporteur. L'interaction pourrait être trop faible ou trop transitoire pour être reconstituée hors de son contexte moléculaire. Dans les deux tests utilisés, nous avons clivé l'ARNm H4 en fragments qui ont été fusionnés à l'ARNm du rapporteur. Des problèmes d'encombrement stérique ou de repliement inadéquat ont très bien pu agir à l'encontre du résultat espéré. Dans des expériences similaires, Lee et collaborateurs avaient pu montrer que les structures fixant eIF3 sur les ARNm c-JUN et BTG1 était respectivement capables de stimuler et d'inhiber la traduction d'un gène rapporteur monocistronique (Lee et al., 2015). Nous n'avons cependant pas obtenu une pareille réussite, ce qui nous a conduit à focaliser notre travail sur d'autres approches décrites ci-après.

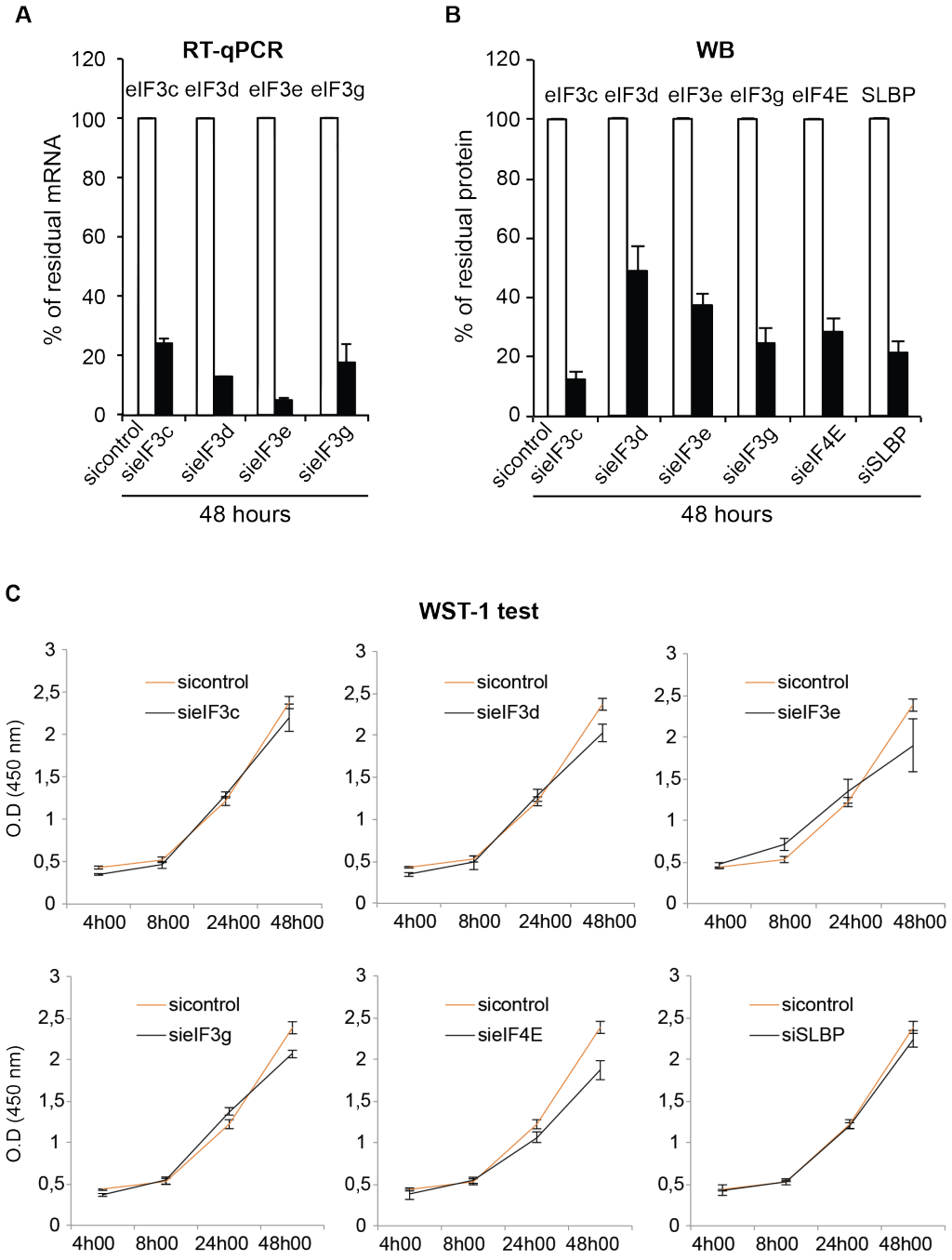


Figure 34. Analyse de l'efficacité de l'inactivation des sous-unités d'eIF3 c, d, e, et g et de son impact sur la viabilité cellulaire. Les ARNm des protéines cibles sont inactivés par siARN pendant 48 heures. **(A)** Niveaux des ARNm quantifiés par RT-qPCR. Les barres blanches représentent les conditions de contrôle (sicontrol) **(B)** Quantification des protéines résiduelles par western blot (WB) pour les 4 sous-unités d'eIF3c, d, e et g ainsi que pour les deux contrôles eIF4E et SLBP. **(C)** Test de la prolifération cellulaire : mesure de l'activité des déshydrogénases mitochondriales à l'aide du réactif WST-1 après 4, 8, 24 et 48 heures d'inactivation des 4 sous-unités d'eIF3 (c, d, e et g) ainsi que des deux protéines de contrôle eIF4E et SLBP. Les résultats sont comparés à ceux obtenus avec des siARN de contrôle.

2.2. Rôle fonctionnel d'eIF3 sur la synthèse des histones *in vivo*

Dans une seconde approche nous avons choisi d'analyser le rôle fonctionnel du complexe eIF3 *in vivo* en étudiant l'impact de l'inhibition des sous-unités eIF3 **c**, **d**, **e** et **g** par siARN sur la traduction des histones.

Avant de mener à bien cette expérience nous avons dû tout d'abord nous assurer de sa faisabilité par une série de mises au point expérimentales. En effet, en raison de la fonction essentielle du complexe eIF3 dans diverses étapes de la traduction la mise en œuvre d'une telle approche soulevait d'emblée de nombreuses questions. En particulier, quel sera l'effet général de l'inactivation des sous-unités d'eIF3 sur la traduction protéique, la viabilité cellulaire, l'intégrité du complexe eIF3. Mais également comment suivre sélectivement l'effet de l'inactivation des sous-unités d'eIF3 sur la traduction des histones.

2.2.1. Effet de l'inactivation des sous-unités d'eIF3 c, d, e, et g sur la viabilité cellulaire

Des études antérieures d'inhibition de diverses sous-unités d'eIF3 ont pu être réalisées avec succès dans des cellules HeLa et HEK293 (Neusiedler et al., 2012; Wagner et al., 2014, 2016) et ont montré des effets variables sur la viabilité cellulaire, l'intégrité et le fonctionnement du facteur eIF3. Nous avons donc inhibé individuellement chacune des sous-unités eIF3 c, d, e et g dans les cellules HEK293FT à l'aide d'un mélange de quatre siARN dirigés contre l'ARNm cible (ON-TARGET Plus SMART Pool, DharmaconTM) (***Tableau M6***). Ces siARN sont 2'-O-methylés au niveau des riboses afin de limiter leur hybridation aspécifique hors cible sans affecter celle avec l'ARNm ciblé. Ces derniers ont été transfectés dans les cellules à l'aide de lipofectamine et les concentrations de siARN ont été optimisées afin d'atteindre un niveau d'inhibition de l'ARNm cible maximal (données non montrées). En guise de contrôles nous avons également inactivé l'expression du facteur de liaison à la coiffe eIF4E (Sonenberg et al., 1979) qui devrait affecter de façon importante la traduction protéique coiffe-dépendante ainsi que la protéine de liaison à la tige-boucle en 3' des ARNm d'histones (SLBP) requise spécifiquement pour la maturation des ARNm des histones (Martin et al., 1997). Le niveau des ARNm a été quantifié par RT-qPCR (***Figure 34 A***) et celui des protéines par western blot (***Figure 34 B et Figure 39 A***). Nos résultats montrent qu'après 48h d'inhibition le niveau des ARNm est réduit à 24, 12, 5 et 18% pour eIF3c, d, e et g respectivement et le niveau de protéine correspondante à 12, 49, 37 et 25% par comparaison aux résultats obtenus pour des siARN contrôles sans cible. Ceci correspond en moyenne à une diminution des protéines de l'ordre de 70%. Ces résultats sont comparables à ceux obtenus par (Wagner et al., 2014) montrant une diminution de 75% à 60% de l'expression de ces protéines dans les cellules HeLa et HEK293 respectivement (Wagner et

al., 2014). Les niveaux résiduels après inactivation des protéines de contrôle eIF4E et SLBP sont similaires (28 et 21% respectivement).

L'effet de l'inhibition des sous-unités d'eIF3 sur la prolifération des cellules HEK293FT a été testé en parallèle à l'aide d'un test WST-1 (voir ***Matériel et Méthodes 2.3.***) destiné à mesurer l'activité des déshydrogénases mitochondriales et qui reflète directement le nombre de cellules métaboliquement actives (***Figure 34 C***). Il apparaît que l'inactivation des sous-unités d'eIF3 ou de SLBP n'a pas d'effet majeur sur la prolifération cellulaire, à l'exception de celles d'eIF3e et eIF4E qui réduisent l'activité des déshydrogénases de 20%.

En conclusion, nos conditions expérimentales ont permis d'abaisser suffisamment le niveau d'expression des sous-unités d'eIF3 ciblées sans pour autant altérer significativement la viabilité des cellules permettant de ce fait d'analyser l'impact de cette inhibition sur la néosynthèse des histones.

2.2.2. Effet de l'inactivation des sous-unités d'eIF3 c, d, e, et g sur la néosynthèse des ARNm d'histones in vivo

L'impact de l'inactivation des sous-unités eIF3c, d, e et g sur la traduction des histones a été mesuré par marquage métabolique des protéines néo-synthétisées à la [^{35}S]-Met puis extraction sélective des histones en condition acide (Voir ***partie 2.2. du Matériel et Méthodes***). Le protocole que nous avons utilisé est schématisé sur la ***Figure 35.***

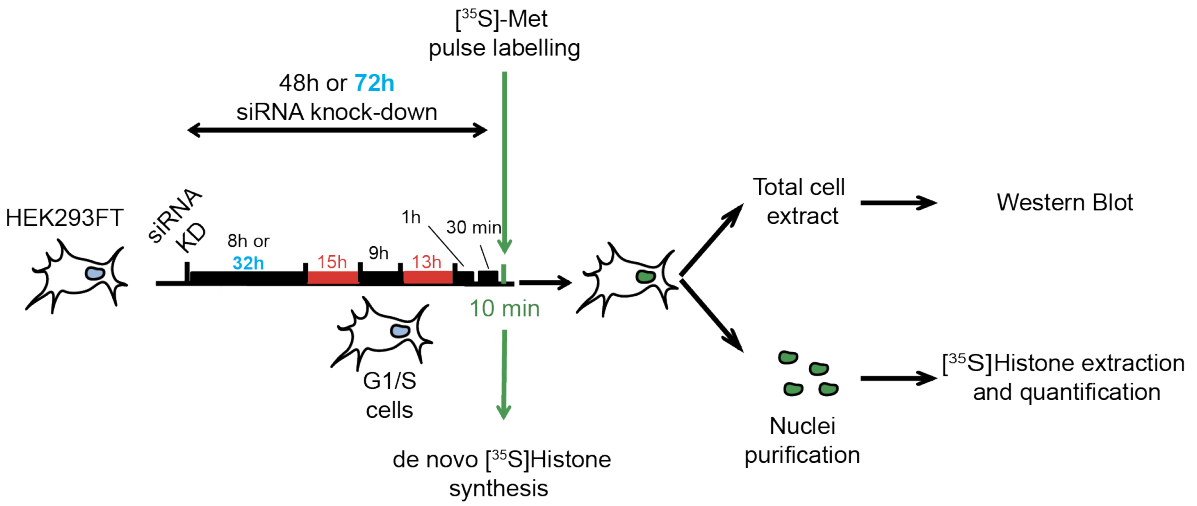


Figure 35. Représentation schématique du protocole utilisé pour étudier l'impact de l'inhibition des sous-unités d'eIF3 sur la néosynthèse des histones. L'expression des protéines cibles est inactivée par siARN (siRNA KD) dans des cellules HEK293FT. Pendant la phase d'inactivation les cellules sont synchronisées en phase G1/S par un double blocage à la thymidine pendant deux périodes de 15h et 13h (marquées en rouge) séparées par un intervalle de 9h. Cette synchronisation est démarrée soit 8 heures (siARN 48h) soit 32h (siARN 72h) après la transfection. Les cellules sont libérées du dernier blocage pendant 1h puis incubées pendant 30 min dans du milieu dépourvu de méthionine. L'effet de l'inactivation sur la néosynthèse des histones est mesuré après marquage métabolique à la [^{35}S]-méthionine pendant 10 min. Les cellules sont lavées 3 fois avec du PBS froid

afin d'éliminer la [^{35}S]-méthionine non incorporée. Des extraits cellulaires totaux sont préparés à partir de 25% des cellules afin d'analyser par western blot l'efficacité de l'inhibition par siARN. Les noyaux issus des 75% de cellules restantes sont purifiés et les histones sont sélectivement extraites par l'acide sulfurique (comme décrit par (Cakmakci et al., 2008) et dans la ***Figure 37*** (voir ***partie 2.2. dans Matériel et Méthodes***)).

Les histones étant fortement et massivement exprimées durant la phase S du cycle cellulaire (Osley, 1991; Pyronnet et al., 2001), nous avons synchronisé les cellules HEK293FT pendant la phase d'inactivation par les siARN (***Figure 35***). Nous avons utilisé pour cela la méthode de double blocage à la thymidine (Bostock et al., 1971), L'excès de thymidine a pour effet de bloquer le cycle cellulaire à la transition G1/S en inhibant la ribonucléotide réductase, empêchant ainsi la production de nucléotides précurseurs du dCTP. Ainsi, les cellules qui se trouvent en phase S ne peuvent l'achever, alors que celles qui sont en G1 ne peuvent pas amorcer la synthèse d'ADN. La synchronisation des cellules a été réalisée pendant deux périodes de 15h et 13h séparées par un intervalle de 9h. Les cellules ont alors été libérées du blocage pendant 1h (voir ***Matériel et Méthodes 2.2.1.***). La synchronisation a été contrôlée par cytométrie de flux (Plateforme de cytométrie en flux de l'IBMC). Les résultats sont représentés dans la ***Figure 36***. Ils révèlent une distribution prédominante des cellules synchronisées en phase S par comparaison aux cellules non synchronisées.

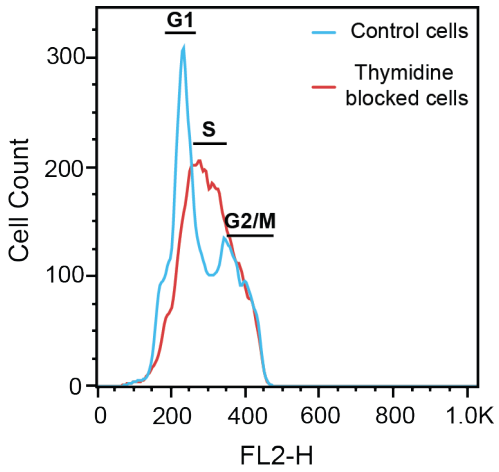


Figure 36. Vérification de la synchronisation des cellules HEK293FT en phase de transition G1/S du cycle cellulaire. L'histogramme montre les résultats de cytométrie en flux obtenus pour des cellules HEK293FT bloquées à la thymidine (en rouge) par rapport aux cellules non traitées (en bleu). L'ADN cellulaire est marqué à l'aide de l'iodure de propidium, intercalant fluorescent qui permet de quantifier les cellules en fonction de leur stade au sein du cycle cellulaire (G1, S et G2/M). FL2-H représente l'intensité du signal de fluorescence à 617 nm.

Nous avons réalisé un marquage métabolique des cellules en présence de [^{35}S]-méthionine 1h30 après la levée du blocage. Ceci en vue d'analyser spécifiquement les

protéines néosynthétisées, ayant incorporé de la [^{35}S]-méthionine, et de déterminer l'impact de l'inactivation par siARN des facteurs d'initiation sur leur traduction.

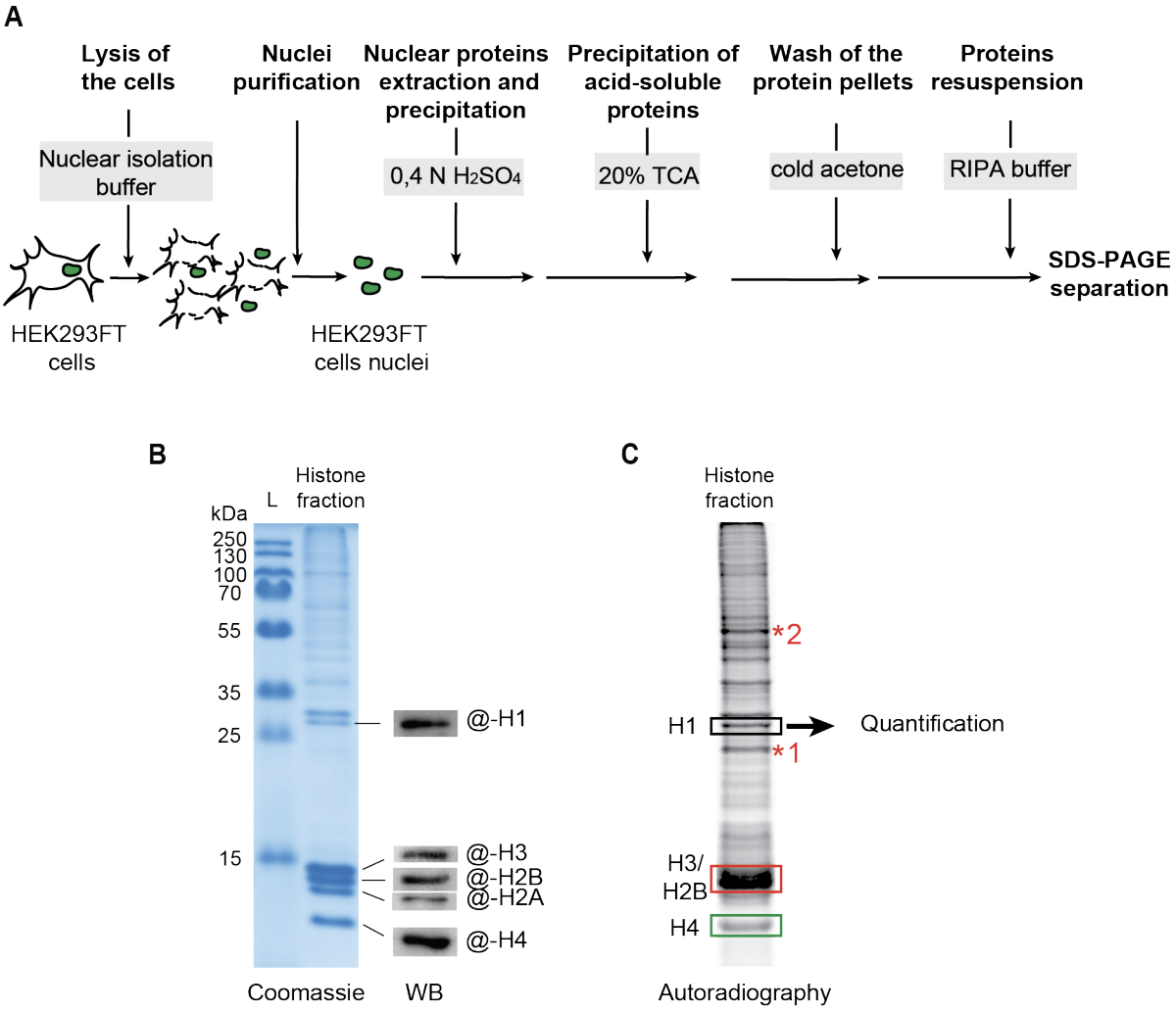


Figure 37. Enrichissement des histones néosynthétisées à partir des noyaux cellulaires de cellules HEK293FT. (A) Schéma expérimental de la méthode d'extraction acide des histones. Les noyaux des cellules sont isolés après lyse des cellules dans un tampon d'extraction nucléaire. Les histones sont extraites dans 0,4 N H_2SO_4 (les histones sont solubles dans l'acide sulfurique) puis précipitées par ajout de 20% d'acide trichloroacétique (TCA). Les culots contenant les protéines sont lavés avec de l'acétone froid et mis en suspension dans du tampon RIPA avant d'être analysés sur gel SDS-PAGE. **(B)** Séparation des fractions d'histones sur gel SDS-PAGE 15% et coloration au bleu de Coomassie (à gauche). L'identité des protéines a été validée par western blot à l'aide d'anticorps spécifiques (à droite). **(C)** Autoradiographie du gel SDS-PAGE permettant de révéler les histones néosynthétisées marquées au [^{35}S]. Les bandes encadrées correspondant aux histones ont été quantifiées à l'aide du logiciel ImageQuant. Le résultat des quantifications a été normalisé par rapport au signal d'une bande 'non histone' désignée par un astérisque (*1). Le rapport entre deux bandes de protéines 'non histones' désignées les astérisques (*1 et *2) est resté constant (± 5%). Les histones H2B et H3 co-migrent et sont quantifiées ensembles.

Afin de pouvoir différencier les [^{35}S]-histones des autres [^{35}S]-protéines néosynthétisées, ces dernières ont été isolées des noyaux cellulaires par extraction à l'acide sulfurique (d'après (Cakmakci et al., 2008), (***voir Figure 37 A et Matériel et Méthodes 2.2.2.***). Les histones très basiques, restent solubles en milieu acide alors que la plupart des autres protéines précipitent. Ceci permet de les séparer sélectivement des autres protéines cellulaires par centrifugation. Les histones ainsi isolées peuvent être fractionnées sur SDS-PAGE, puis colorées au bleu de Coomassie (***Figure 37 A et B***). Des expériences de western blot effectuées en utilisant des anticorps dirigés contre les protéines histones ont permis d'identifier sans ambiguïté la position des histones H1, H2B, H2B, H3 et H4 (***Figure 37 B***). L'incorporation de [^{35}S]-Met dans les histones néosynthétisées lors du marquage métabolique permet leur visualisation par autoradiographie et leur quantification (***Figure 37 C***). Les histones possèdent entre 1 et 3 résidus Met (1 Met dans les histones H1 et H2A, 2 dans H2B et H4, et 3 dans H3) et incorporent de ce fait des quantités faibles de [^{35}S]-Met. Les histones H1 et H2A n'incorporent qu'un seul résidu Met au niveau du codon initiateur. Celle-ci est probablement clivée comme la plupart des Met initiatrices. Nous n'avons pas tenté de quantifier H2A sur les autoradiogrammes. D'autre part, les histones H2B et H3, de poids moléculaire très proches, ont dû être quantifiées simultanément (***Figure 37 C***). Le signal obtenu pour chaque bande d'histone a été normalisé par rapport à une bande protéique discrète servant de contrôle interne (***Figure 37 C***). Des exemples de résultats bruts sont représentés **Figure 38**.

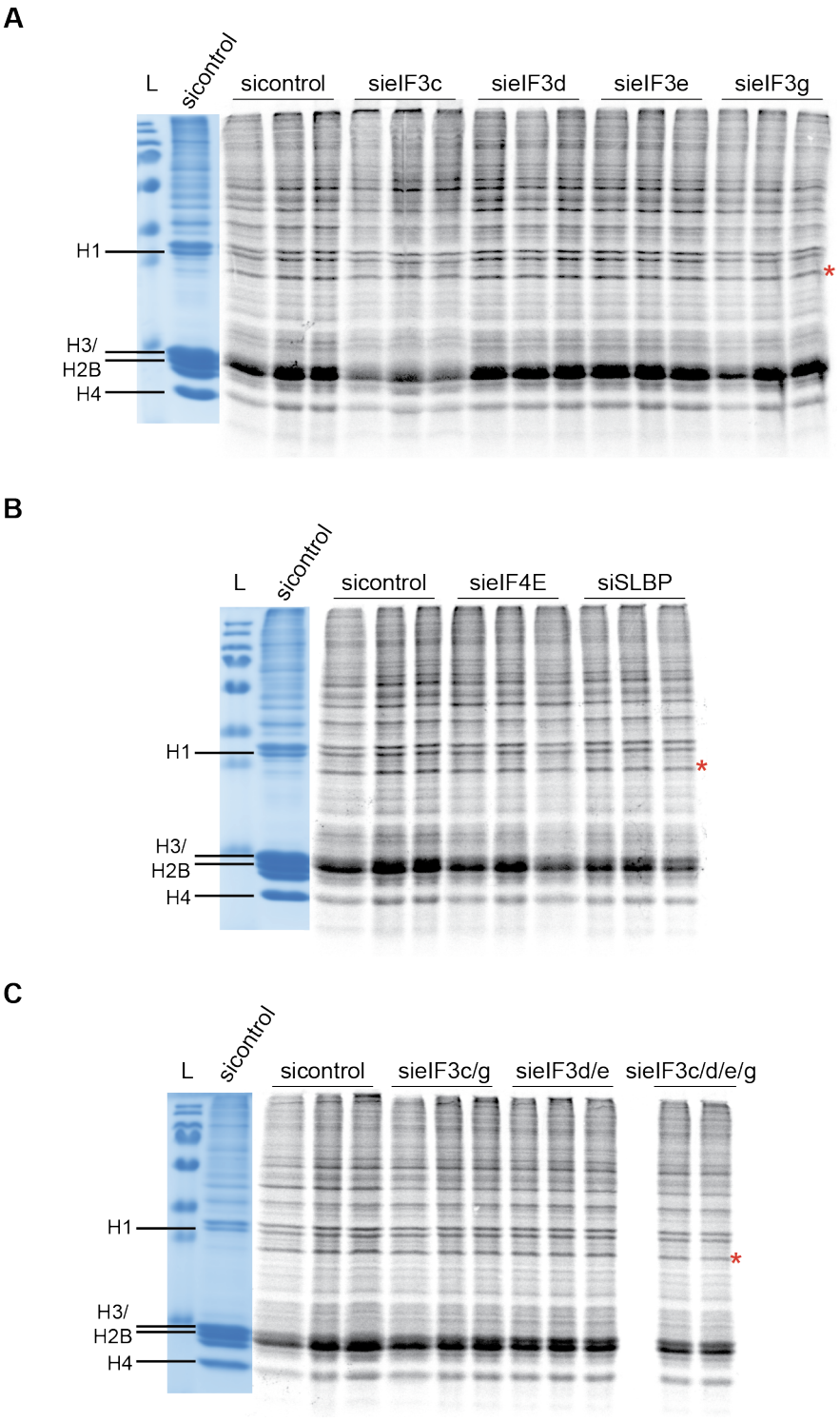


Figure 38. Examples d'autoradiographies permettant de visualiser la synthèse *de novo* des [^{35}S]-histones. Séparation des protéines radiomarquées sur gel SDS-PAGE 15% après inactivation des sous-unités eIF3c, d, e et g **(A)**, des protéines contrôles eIF4E et SLBP **(B),** et inactivation double des sous-unités eIF3 (c/g), (d/e) ou quadruples (c/d/e/g) **(C)**. Les protéines radiomarquées sont détectées par autoradiographie. La première piste de chaque gel coloré au bleu de Coomassie est montrée à gauche ainsi que l'échelle de poids moléculaire (L). Le signal de la bande de radioactivité correspondante à chaque histone est quantifié à l'aide du logiciel ImageQuant et normalisé par rapport au signal de la bande 'non histone' désignée par l'astérisque.

Les résultats de l'inhibition individuelle des 4 sous-unités eIF3c, d, e et g sur la synthèse *de novo* des histones sont représentés dans la ***Figure 39***. Les résultats de western blot (***Figure 39 A et B***) sont ceux discutés dans le **paragraphe précédent 2.2.1.**, ils révèlent qu'après 48h d'inhibition par siARN les niveaux des sous-unités eIF3 c, d, e et g ont pu être abaissés efficacement à 12%, 49%, 37% et 25% respectivement et celui des protéines eIF4E et SLBP à 28% et 21% respectivement, alors que le niveau de la protéine de contrôle GAPDH reste inchangé (***Figure 39***).

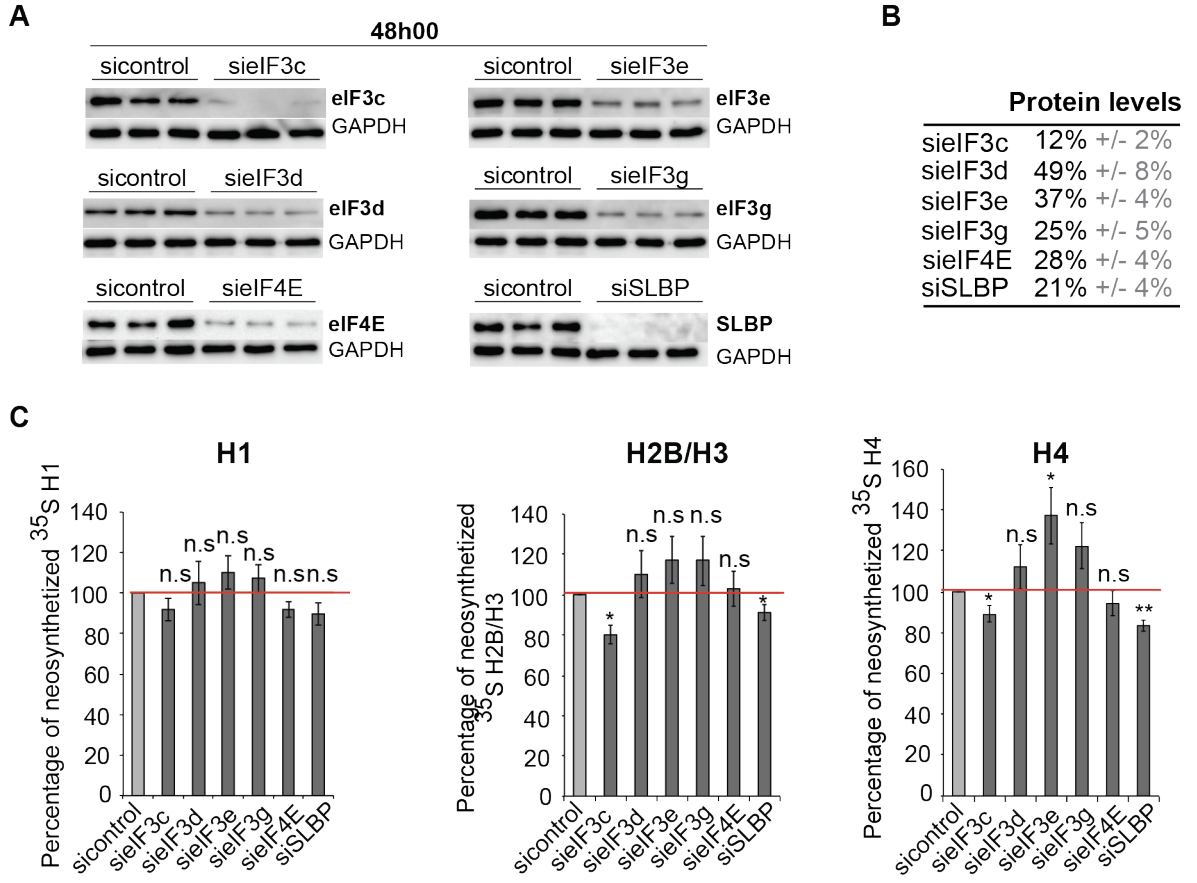


Figure 39. Impact d'une inhibition des sous-unités d'eIF3 (c, d, e et g) sur la traduction des histones après 48 heures. (A) Analyse de l'efficacité de l'inhibition par siARN des sous-unités d'eIF3 (c, d, e et g), eIF4E et SLBP par western blot. **(B)** Quantification des protéines ciblées après inhibition de leur expression par siARN. Les résultats sont normalisés par rapport à la protéine endogène GAPDH. **(C)** Quantification des signaux de radioactivité des histones H1, H2B/H3 et H4 néo-synthétisées comme décrit dans la légende de la ***Figure 37***. L'ensemble des résultats et les barres d'erreurs sont issus de triplicats biologiques. Les astérisques représentent des différences significatives par rapport au contrôle (sicontrol) : * P<0,05 et ** P<0,005 selon le test de Student. n.s : différence non significative

Il apparaît que l'inactivation de SLBP, facteur requis pour la traduction des histones conduit à une diminution de 15-20% de la traduction des histones H1, H2B/H3 et H4 après 48h. L'effet sur la synthèse de H4, statistiquement significatif, est le plus marqué (***Figure 39***). Ces résultats sont en accord avec ceux de (Mejlvang et al., 2014) qui révèlent que l'inactivation de SLBP dans des cellules U2OS conduisent à une baisse plus rapide du niveau des

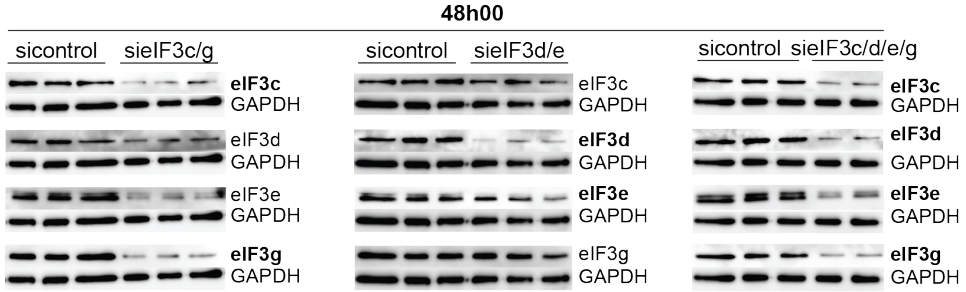
protéines H4 et H3 alors que les histones H2B sont moins fortement affectées, validant notre approche expérimentale. En revanche, l'inhibition d'eIF4E n'a pas d'effet significatif sur la traduction des histones dont le niveau diminue toutefois de l'ordre de 10% pour H1 et H4. Dans une précédente étude, il avait été obtenu, avec les mêmes siARN dirigés contre eIF4E et dans les mêmes conditions, environ 25% d'inhibition de la traduction d'un gène rapporteur (Wurth et al., 2014). Nos présents résultats montrent donc une inhibition de la traduction atténuée qui est peut-être à relier au mécanisme particulier d'initiation des histones qui pourrait être moins sensible à une carence cellulaire en eIF4E (Gross et al., 2018; Martin et al., 2011). Aucun effet n'est observé par ailleurs pour la condition contrôle (sicontrol). Les mêmes résultats ont été obtenus lorsque d'autres protéines internes de référence ont été utilisées pour la normalisation des résultats (données non montrées).

Dans ces conditions, seule l'inhibition de la sous-unité eIF3c réduit significativement de 20 et 15% le niveau d'expression des histones H2B/H3 et H4 respectivement, alors que le niveau de H1 n'est pas significativement abaissé. L'inhibition des sous-unités d, e et g, au contraire, augmente faiblement le niveau de traduction de H1, H2B/H3 et H4 de l'ordre de 10 à 20%. De façon remarquable, seule l'inhibition d'eIF3e conduit à une augmentation importante et significative du niveau d'expression de H4 de 40 % (***Figure 39***). Wagner et coll. (2014) ont montré que l'inhibition des deux sous-unités eIF3c et eIF3e conduisait à la formation de sous-complexes d'eIF3 (a, b, g, i) capables d'interagir avec la sous-unité ribosomique 40S et les ARNm mais conduisant à des défauts d'initiation de la traduction. De même, Neusieldler et coll. (2012) ont montré que l'inhibition d'eIF3e par siARN conduisait, à l'inverse de nos résultats, à une diminution de la traduction des ARNm des histones H4, H2B et H3 (Neusiedler et al., 2012). De manière générale, les effets modestes que nous avons mesurés sur la traduction de la majorité des histones après l'inactivation des sous-unités isolées laissent supposer qu'une quantité suffisante de complexe eIF3 intègre est toujours présente dans la cellule et permet de maintenir un niveau élevé de traduction des histones. Il est probable que les ARNm d'histones pourraient détourner les complexes eIF3 intègres restants pour leur propre usage au détriment des autres ARNm cellulaires. L'interaction directe d'eIF3 avec les ARNm d'histones leur offrirait donc un avantage certain en situation de pénurie des composants de l'appareil de traduction. L'augmentation du temps d'inactivation des sous-unités isolées à 72h n'a pourtant pas permis d'amplifier ces effets (données non montrées).

Dans un second temps, nous avons donc choisi d'inactiver simultanément 2 sous-unités d'eIF3 (c/g) ou (d/e) ou les 4 sous-unités (c/d/e/g) et d'analyser l'impact de cette inhibition sur la synthèse *de novo* des histones après 48h et 72h (***Figures 40 et 41*** respectivement). Les mêmes conditions de synchronisation et de traitement des cellules ont été utilisées que pour l'expérience précédente (voir ***Figure 35***).

Les résultats de western blot montrent qu'au bout de 48h d'inhibition conjointe les taux d'expression des sous-unités ciblées sont réduits en moyenne à 35 % (à l'exception d'eIF3e dans le cas de sieIF3d/e), ce qui est comparable aux résultats obtenus pour l'inhibition des sous-unités isolées. Comme décrit précédemment (Wagner et al., 2014), nous avons également observé un impact de l'inhibition des sous-unités c/g sur le niveau d'expression d'autres sous-unités du complexe, ici eIF3d et eIF3e qui voient leur niveau abaissé à 64 et 48 % respectivement (***Figure 40 A et B).*** L'effet de l'ensemble de ces inactivations multiples sur la néosynthèse des histones (***Figure 40 C***) n'est cependant pas différent de celui observé pour l'inactivation des sous-unités isolées. On observe dans tous les cas une augmentation significative d'environ 20% de l'expression des histones par rapport aux contrôles (***Figure 40***). La proportion de complexe eIF3 intègre semble toujours être suffisante pour permettre une traduction efficace des histones H1, H2B/H3 et H4.

A



B

	sieIF3c/g	sieIF3d/e	sieIF3c/d/e/g
eIF3c	38% +/- 12%	94% +/- 16%	47% +/- 4%
eIF3d	64% +/- 9%	38% +/- 5%	31% +/- 2%
eIF3e	48% +/- 13%	81% +/- 8%	38% +/- 5%
eIF3g	34% +/- 11%	96% +/- 28%	31% +/- 2%

C

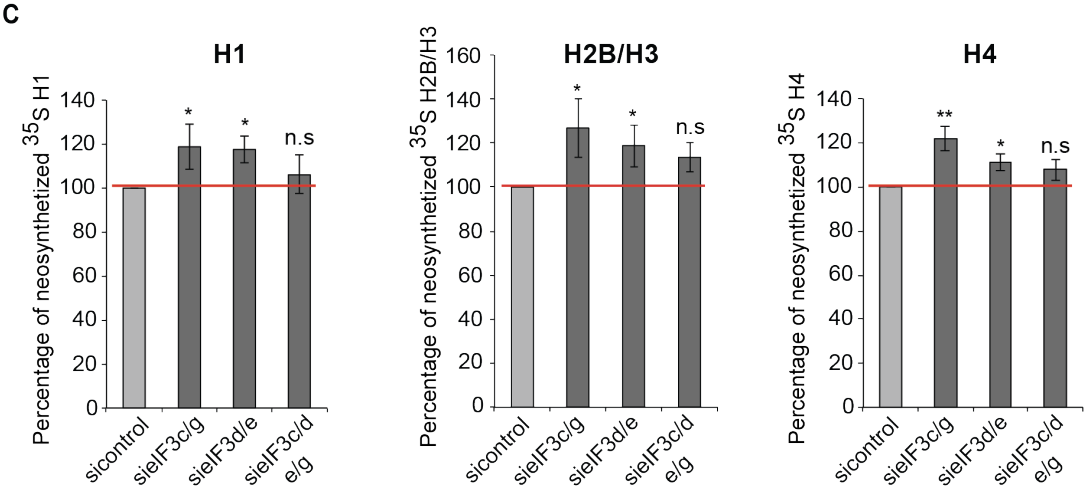
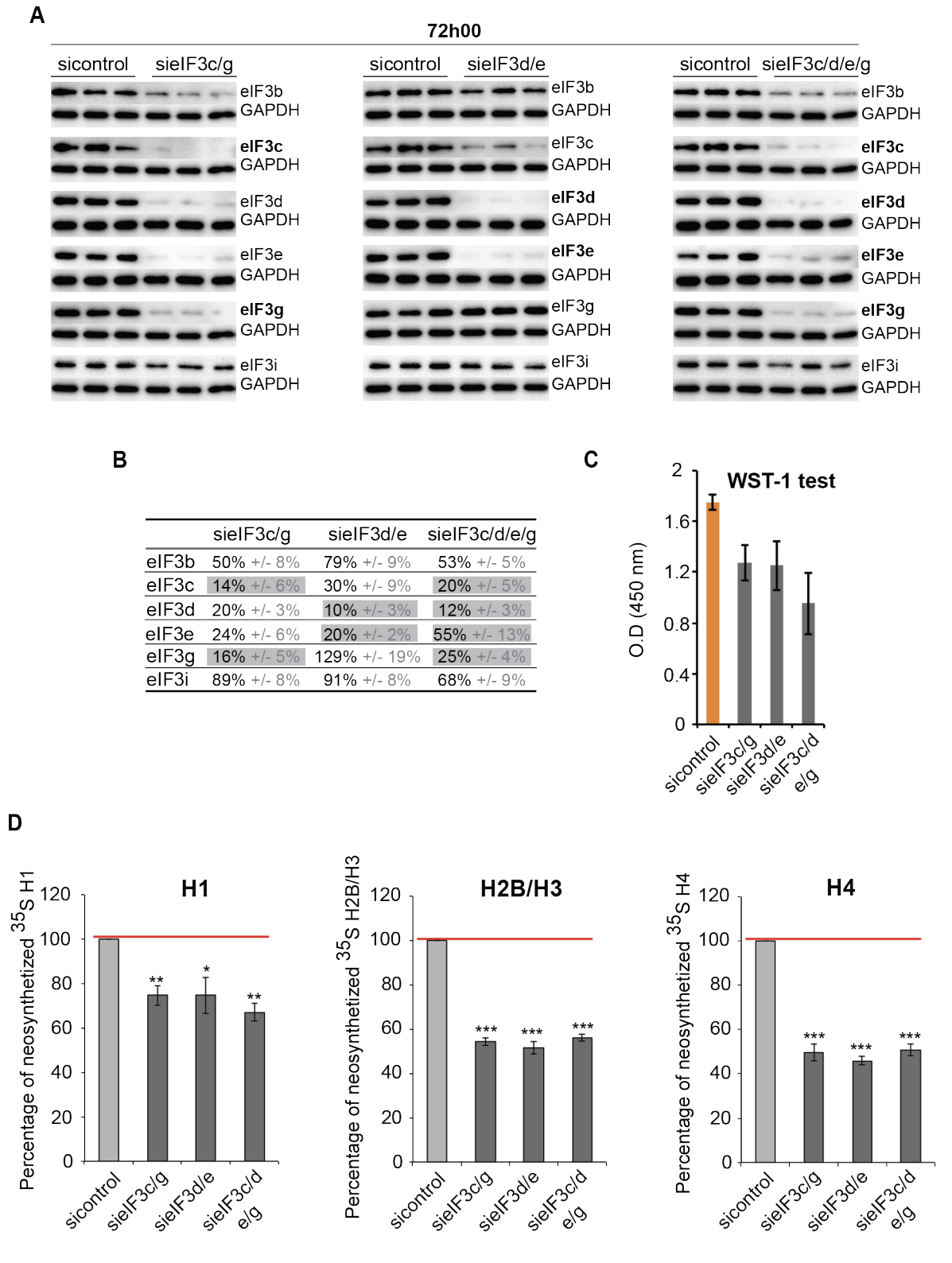


Figure 40. Impact de l'inhibition simultanée des sous-unités d'eIF3 (c/g), (d/e) et (c/d/e/g) pendant 48 heures sur la traduction des histones. L'ensemble des résultats est représenté comme décrit dans la ***Figure 39*****.**



	sieIF3c/g	sieIF3d/e	sieIF3c/d/e/g
eIF3b	50% +/- 8%	79% +/- 9%	53% +/- 5%
eIF3c	14% +/- 6%	30% +/- 9%	20% +/- 5%
eIF3d	20% +/- 3%	10% +/- 3%	12% +/- 3%
eIF3e	24% +/- 6%	20% +/- 2%	55% +/- 13%
eIF3g	16% +/- 5%	129% +/- 19%	25% +/- 4%
eIF3i	89% +/- 8%	91% +/- 8%	68% +/- 9%

Figure 41. Impact de l'inhibition simultanée des sous-unités d'eIF3 (c/g), (d/e) et (c/d/e/g) pendant 72 heures sur la traduction des histones. L'ensemble des résultats est représenté comme décrit dans la ***Figure 39***. **(C)** Test de la prolifération cellulaire à l'aide du réactif WST-1 après 72h d'inactivation double des sous-unités eIF3 (c/g), (d/e) ou quadruple (c/d/e/g)**.** Les astérisques représentent des différences significatives par rapport au contrôle (sicontrol) : * P<0,05, ** P<0,005 et ***P<0,0005 selon le test de Student.

Après 72h d'inhibition simultanée des sous-unités d'eIF3 les résultats obtenus sont en revanche très différents. Dans ce cas, le niveau d'expression des protéines résiduelles a pu être réduit à 15% en moyenne pour sieIF3c/g et sieIF3d/e (***Figure 41 A et 41 B***) et entre 12 et 55% pour les sieIF3c/d/e/g. Dans tous les cas, l'inhibition de ces sous-unités a également abaissé le niveau d'expression d'autres sous-unités du complexe telles que eIF3b, d, e ou eIF3i (***Figure 41 B***). On peut considérer que le complexe eIF3 entier a été inactivé dans ces conditions. Nos tests de prolifération cellulaire indiquent que l'activité métabolique des cellules est réduite de 30 à 40% pour les doubles ou quadruples inactivations respectivement (***Figure 41 C***). Au bout de 72h, l'impact de ces inhibitions sur la traduction des histones est particulièrement marqué et conduit à une diminution très significative du taux d'histone H1 de 30 % et de celui des histones H2B/H3 et H4 néo-synthétisées de 50 % par comparaison avec la néo-synthèse de la protéine non-histone ayant servi de contrôle interne qui reste stable (***Figure 38***). Cette stabilité s'observe aussi pour la protéine GAPDH dont l'expression n'est pas altérée après 72h d'inhibition comme l'atteste l'ensemble de nos western blots pour (***Figure 41 A***). Il apparaît donc que l'inhibition du complexe eIF3 affecte différemment la traduction des histones par rapport à celle d'autres protéines. Il est possible que dans ces conditions, la cellule ne contienne plus suffisamment de molécules d'eIF3 fonctionnelles et structuralement intègres pour favoriser l'expression des ARNm contenant un site d'interaction.

CONCLUSION

L'objectif initial de mon projet de thèse était de démontrer et de caractériser l'interaction entre l'ARNm de l'histone H4 (et autres histones) et le facteur d'initiation eIF3 puis à analyser sa fonction. Cet objectif a été atteint dans une large mesure. Dans un premier temps j'ai pu visualiser l'interaction *in vitro* en utilisant des transcrits d'ARN et le facteur eIF3 purifié. Puis j'ai montré que des interactions semblables existaient *in vivo* par des expériences d'immunoprécipitation des ARN. J'ai ensuite identifié les sous-unités d'eIF3 impliquées dans l'interaction avec les ARNm d'histones par des expériences de pontage de l'ARN et électrophorèse en 2 dimensions. Les 4 sous-unités d'eIF3 (c, d, e et g) identifiées ont été exprimées dans *E. coli* sous forme de protéines recombinantes étiquetées qui ont été utilisées pour décortiquer le schéma d'interaction avec les différents ARNm d'histones. Puis j'ai entrepris l'étude fonctionnelle en essayant dans un premier temps de reconstituer un système de traduction utilisant un gène rapporteur placé sous l'influence du fragment d'ARNm de H4 liant eIF3. Je n'ai cependant pas pu observer de variation concluante de l'efficacité de traduction du gène rapporteur qui aurait pu être attribuée à eIF3 et par manque de temps, je n'ai pu tester les fragments identifiés dans les ARNm des autres histones. A ce stade, je me suis focalisé sur une analyse plus cellulaire dans laquelle j'ai inhibé l'expression des 4 sous-unités d'eIF3 par siARN. L'effet de cette inhibition a ensuite été mesuré sur la synthèse des histones durant la phase S du cycle cellulaire. Les résultats ont montré que des inhibitions modérées et individuelles des sous-unités d'eIF3 entrainaient une stimulation de la synthèse des histones. Au contraire, des inhibitions plus marquées et simultanées de plusieurs sous-unités ont entrainé une chute de la synthèse des histones mais cet effet était accompagné d'une diminution globale de l'activité métabolique cellulaire. Ceci suggère que l'inhibition de la synthèse des histones pourrait être un effet secondaire lié à l'effondrement du métabolisme. A ce stade de l'étude, mes résultats suggèrent que la fonction de l'interaction avec eIF3 serait de réprimer la traduction des histones, du moins après 1h30 de progression en phase S. Dans des expériences futures, il sera intéressant de suivre cet effet durant la totalité de cette phase et au delà afin de voir si cet effet est constant ou variable comme récemment observé pour un autre ARNm cible d'eIF3 : l'ARNm PTBP1 de la Polypyrimidine Tract-Binding Protein1.

DISCUSSION ET PERSPECTIVES

DISCUSSION ET PERSPECTIVES

La traduction de l'ARNm est l'étape finale de l'expression des gènes, elle est divisée en trois phases principales, l'initiation, l'élongation et la terminaison, chacune coordonnée par plusieurs facteurs. Chez les eucaryotes, la phase d'initiation est complexe et constitue l'étape limitante de la synthèse protéique. Elle nécessite la coordination entre plusieurs facteurs qui forment plusieurs complexes aboutissant à la fixation sur la coiffe située à l'extrémité 5' des ARNm de la particule ribosomique 43S, puis au 'scanning' en vue de l'identification de l'AUG initiateur et de l'assemblage du ribosome 80S (Hinnebusch, 2017; Kozak, 1978; Sonenberg and Hinnebusch, 2009). Les virus ont développé des alternatives à l'initiation cellulaire : les éléments IRES qui sont plus efficaces et qui leur permettent de détourner la machinerie traductionnelle de la cellule hôte à leur profit (Jackson et al., 2010). Un certain nombre d'éléments IRES ont été trouvés dans des ARNm cellulaires, ainsi que des systèmes d'initiation non canoniques dans certaines classes d'ARNm comme les ARNm TISU présentant de très courtes extrémités 5' UTR (Elfakess et al., 2011; Haimov et al., 2017). L'ARNm de l'histone H4 est un autre exemple de messager présentant une très courte extrémité 5' UTR. Son mécanisme d'initiation unique a été décrit par notre laboratoire (Martin et al., 2011). Dans cet ARN, c'est la région codante de l'histone H4 qui remplace l'extrémité 5' UTR dans sa fonction de recrutement des facteurs et des ribosomes. En effet, des éléments structuraux de l'ARN fixent les facteurs d'initiation et les ribosomes, à la façon des éléments IRES mais avec cependant une dépendance stricte pour la coiffe. Un premier élément d'ARN présente un repliement contenant une jonction à 3 hélices (TWJ). Il joue successivement le rôle de site de liaison interne de la coiffe puis d'élément facilitant le positionnement du ribosome à la distance optimale du codon AUG (Gross et al., 2018; Martin et al., 2011). Un deuxième élément appelé 4E-SE (élément sensible à eIF4E) recrute le complexe eIF4F (eIF4E+4G+4A) lui-même recrutant les particules ribosomiques 43S (via l'interaction eIF3-eIF4G) pour les déposer sur l'AUG initiateur de l'ARNm H4 (Martin et al., 2016).

Des expériences de sondage en solution réalisées il y a quelques années sur l'ARNm H4 avaient permis d'identifier un élément structural supplémentaire dans la phase codante de l'ARNm de l'histone H4, formant une longue tige-boucle repliée juste en amont du codon stop TAA (D'Orchymont, 2013). Cependant, la fonction de cette structure restait inconnue jusqu'aux expériences de CLIP-SEQ de Lee et coll. qui ont révélé la présence d'un site de liaison putatif pour le facteur d'initiation eIF3 (Lee et al., 2015) appelé eIF3-BS (eIF3 binding site) dans notre étude. Mon travail de thèse a consisté à démontrer et caractériser

l'interaction entre l'ARNm H4 et le facteur d'initiation eIF3 et à analyser sa fonction. Par extension, j'ai également étudié les ARNm des autres histones (H1, H2A, H2B et H3).

Dans une première série d'expériences, nous avons cherché à visualiser l'interaction entre le facteur eIF3 et l'ARNm H4 par des expériences de retard sur gel. Nous avons ainsi pu observer qu'eIF3 purifié est capable d'interagir avec l'ARNm H4 entier ($H4_{1-375}$) mais également avec un fragment constitué de la seconde moitié de l'ARNm ($H4_{179-375}$) dans lequel a été localisé le site de pontage d'eIF3 par CLIP-SEQ (Lee et al., 2015). En revanche, seule une très faible fixation est observée avec la première moitié de l'ARNm ($H4_{1-178}$) en présence de fortes concentrations d'eIF3 (piste 12) (***Figure 25***). Cette fixation pourrait être due aux domaines RRM (domaine de reconnaissance non spécifiques des ARN) qui sont trouvés dans les sous-unités eIF3a, b et g. Ce premier résultat suggèrait fortement que l'interaction du facteur eIF3 pourrait bien se produire avec la séquence de 25 nt identifiée par CLIP-SEQ entre les positions 294 et 319.

Dans un second temps, nous avons analysé l'interaction *in vivo*, dans des cellules HEK293FT. Nous avons élargi la recherche aux autres ARNm d'histones. Par des expériences d'immunoprécipitations des particules ribonucléoprotéiques (RNP IP) à l'aide d'anticorps dirigés contre eIF3b nous avons démontré que l'ARNm H4, mais aussi les autres ARNm d'histones (H1, H2A, H2B, H3) étaient capables d'interagir *in vivo* avec le complexe eIF3 (***Figure 26***). Afin de stabiliser ces interactions et de minimiser les réarrangements de complexes RNP pouvant survenir lors de la lyse cellulaire nous avons optimisé une méthode de pontage à la formaldéhyde pour la préparation des extraits ribonucléoprotéiques (voir article sous presse en annexe (Hayek, H., Gross, L., ALHGOUL, F., Martin, F., Eriani, G., Allmang).

Pour identifier plus précisément les sous-unités d'eIF3 responsables de ces interactions, nous avons réalisé des expériences de pontage de l'ARNm H4 radio-marqué au [^{32}P] en présence de complexe eIF3 purifié. Après fractionnement sur gel bidimensionnel (2D), nous avons identifié sans ambiguïté les quatre sous-unités d'eIF3 impliquées dans l'interaction directe avec l'ARN H4. Il s'agit des sous-unités eIF3c, d, e et g. Nous n'avons pas obtenu de pontage avec la sous-unité eIF3b précédemment identifiée par CLIP-SEQ (Lee et al., 2015). Au contraire, nous avons montré une interaction avec la sous-unité eIF3c dont le poids moléculaire est voisin d'eIF3b. Une hypothèse pour expliquer cette différence pourrait être que lors de la constitution des banques d'ARN destinées au séquençage haut-débit, les 2 protéines, de poids moléculaires voisins, auraient pu co-migrer (d'autant plus qu'à la protéine se rajoute un fragment d'ARN de taille variable) de sorte que les fragments d'ARN, élués du gel puis séquencés, auraient pu être faussement attribués à un pontage à eIF3b au lieu d'eIF3c. Dans nos expériences de fractionnement en gel 2D, la différence de

taille et de charge des sous-unités d'eIF3 pontées aux fragments d'ARNm de H4 est clairement visible, ce qui conforte l'identification des 4 sous-unités eIF3c, d, e et g (***Figure 28***) et explique la probable erreur d'attribution du pontage à eIF3b lors du séquençage à haut débit (Lee et al., 2015).

Par ailleurs, nos résultats révèlent également que le schéma d'interaction d'eIF3 avec l'ARNm H4 est très différent de celui avec les éléments de type IRES qui fixent uniquement les sous-unités eIF3a, b et c (Cai et al., 2010; Jackson et al., 2010; Ujino et al., 2012). Ceci confirme qu'il puisse y avoir plusieurs modes d'interaction entre eIF3 et les ARNm qu'il contrôle.

En lien direct avec les expériences d'identification des sous-unités, nous avons réalisé des expériences de 'GST-pulldown' à l'aide des protéines eIF3c, d, e et g recombinantes fusionnées à une double étiquette HisGST couplées à la détection des ARNm par qRT-PCR. Nous avons montré que les trois sous-unités eIF3c, d et g étaient capables d'interagir séparément avec les ARNm d'histones H1, H2A, H2B, H3 et H4 *in vitro* en dehors du complexe eIF3. En revanche, eIF3e interagit uniquement avec l'ARNm de l'histone H1 et pas avec l'ARNm H4 comme le suggéraient nos expériences de pontage (voir récapitulatif ***Tableau 3***). Il est possible que l'interaction d'eIF3e avec l'ARNm H4 ne puisse se produire que dans le contexte du complexe eIF3 entier ou que le repliement de la protéine de fusion HisGST-eIF3e ne soit pas optimal pour l'interaction avec les ARNm des histones (exception faite de l'histone H1). Mentionnons ici que la séquence de la sous-unité eIF3e est la seule qui ait nécessité une optimisation pour pouvoir être exprimée dans *E. coli*, ce qui suggère des problèmes sous-jacents de repliement de la protéine. Pour l'optimiser, nous avons utilisé la séquence décrite par Sun et coll. (2011) qui est très proche de la séquence proposée par les programmes d'optimisation de notre fournisseur habituel (Integrated DNA Technologies). Il conviendrait de vérifier le bon repliement de la protéine par des expériences de dichroïsme circulaire ou de DLS avant de conclure définitivement à l'absence d'interaction avec 4 des ARNm d'histones.

Concernant la sous-unité eIF3c nous avons montré que le domaine N-terminal de la protéine (région 1-318) est responsable de la fixation des ARNm d'histones. Cette région de la protéine forme une extension qui est repliée indépendamment des 2/3 restants de la protéine que l'on retrouve dans le noyau PCI/MPN d'eIF3. Chez la levure, il a été montré que l'équivalent de cette extension se situe près du canal de sortie de l'ARNm (Aylett et al., 2015; Erzberger et al., 2014) où elle joue un rôle lors de l'amarrage initial d'eIF1 et eIF5 à leurs emplacements respectifs sous les sites E et P de la petite sous-unité (Obayashi et al., 2017). Cependant le domaine N-terminal d'eIF3c de la levure ne fait que 230 acides aminés contrairement aux 319 résidus de l'enzyme humaine. Les alignements de séquence de protéines eIF3c eucaryotes réalisés au laboratoire montrent de très grandes variations dans

les domaines de liaison d'eIF1 et eIF5 de la levure et l'homme. De plus, eIF3c humaine présente 89 résidus supplémentaires par rapport à la levure (données non présentées). Ce constat, combiné au fait que l'extension N-terminale humaine fixe très fortement les ARN (le plus fortement des 4 sous-unités testées durant ma thèse), inspire de futures études dans le but de délimiter plus précisément ce domaine de liaison à l'ARN et de localiser les domaines d'interaction pour eIF1 et eIF5. Des délétions supplémentaires du domaine N-terminal seront réalisées suivies d'expériences de 'GST-pull-down' d'eIF1 et eIF5 en vue de localiser les sites d'interaction de ces partenaires.

Dans ce contexte, il est aussi intéressant de noter que la traduction de l'histone H4 est étonnamment résistante à l'ajout d'un excès d'eIF1 recombinant (Franck Martin, communication personnelle). En effet, il a été montré qu'un excès d'eIF1 induit sur la plupart des ARNm une déstabilisation du complexe de pré-initiation lors de l'étape de 'scanning' ce qui entraine une diminution de l'initiation sur le premier site de démarrage (Barth-Baus et al., 2013). Les IRES virales qui ne 'scannent' pas l'ARNm sont résistantes à l'inhibition par eIF1, de même que l'ARNm de l'histone H4. Une hypothèse à tester serait que l'extension N-terminale d'eIF3c pourrait se substituer à eIF1 ce qui expliquerait la résistance de H4 à l'inhibition par eIF1. Il faudrait réaliser des tests de traduction en présence de concentrations croissantes de fragments N-terminaux d'eIF3c et observer l'effet sur la traduction de l'histone.

Concernant la sous-unité eIF3d, il serait intéressant de tester la capacité de ce facteur à lier la coiffe des ARNm des histones. En effet il a été montré que ce facteur contient un site de liaison à la coiffe qui lui permettrait de se substituer à eIF4E dans des conditions où la traduction coiffe-dépendante canonique est inhibée (Lee et al., 2016). Il serait intéressant de réaliser des tests de pontage de la coiffe avec le complexe eIF3 purifé comme décrit dans l'article Gross et coll. (2018) et des expériences de compétition en présence d'analogues de la coiffe. Ceci permettrait de déterminer précisément la nature des contacts protéiques à proximité de la coiffe et de tester si eIF3d utilise sa poche de fixation à la coiffe pour interagir avec les ARNm d'histones.

Concernant la sous-unité eIF3g qui contient un domaine de liaison à l'ARN (RRM) dans sa moitié C-terminale, il conviendrait dans un premier temps de muter ce domaine et d'observer l'effet sur la fixation des ARNm d'histones. Si le domaine RRM est réellement impliqué dans la fixation des ARNm d'histones, il sera intéressant de rechercher dans la protéine les peptides générant la spécificité d'interaction aux ARN des histones.

Une seconde partie importante de mon travail de thèse a consisté à rechercher la fonction cellulaire de l'interaction d'eIF3 avec les ARNm d'histones. Précédemment, il a été montré que la fixation d'eIF3 pouvait stimuler ou bien réprimer la traduction selon les ARNm

(Arake de Tacca et al., 2019; Lee et al., 2015). Dans une première approche, nous avons testé l'effet de la tige-boucle de l'ARNm H4 interagissant avec eIF3 (eIF3-BS) sur la traduction d'un gène rapporteur R-Luc, *in vitro* à l'aide d'ARNm monocistroniques purifiés, ou *in vivo* dans un système de bicistrons (voir partie ***résultats 2.1.***). Nous avons observé des effets tantôt stimulateurs, tantôt inhibiteurs selon la localisation de l'élément eIF3-BS en 5' ou 3' de la phase codante de la R-Luc. Cependant, l'utilisation de mutants affectant les appariements de la tige boucle ou sa structure n'a pas permis de conclure qu'il s'agissait d'effets spécifiques liés à la fixation d'eIF3. Les effets observés étaient sans doute liés à une modification de la stabilité ou de la conformation des ARNm. Ceci a confirmé l'importance de tester des variants mutés du site d'interaction pour s'assurer que l'effet observé est bien spécifique d'eIF3 et n'est pas un effet secondaire lié à la modification de structure de l'ARNm rapporteur initial.

Dans une seconde approche, nous avons choisi d'analyser le rôle fonctionnel du complexe eIF3 *in vivo* en étudiant l'impact de l'inhibition des sous-unités eIF3 c, d, e et g par siARN sur la traduction des histones. Ces quatre sous-unités d'eIF3 que nous avons identifiées en interaction avec l'ARNm H4 jouent des rôles importants dans la formation et le positionnement du complexe eIF3 dans le contexte du ribosome 80S. Seule la sous-unité eIF3g possède des motifs de liaison à l'ARN de type 'RNA Recognition Motif' (RRM) (ElAntak et al., 2007) et eIF3d est capable de se lier à la coiffe de certains ARNm. La sous-unité eIF3c se fixe sur des éléments de type IRES (Cai et al., 2010; Jackson et al., 2010; Ujino et al., 2012). Par ailleurs eIF3e forme avec eIF3d un module qui orchestre l'expression d'ARNm spécifiques impliqués dans le contrôle du métabolisme énergétique cellulaire (Shah et al., 2016). En plus, les sous-unités eIF3g et eIF3e interagissent avec la protéine MIF4GD/SLIP1 pour la traduction des ARNm d'histones (von Moeller et al., 2013; Neusiedler et al., 2012). Dans nos expériences, nous avons inactivé les 4 sous-unités d'eIF3 individuellement, ou par paire c/g et d/e, ou les 4 simultanément durant 48 et 72 heures. Les histones étant fortement et massivement exprimées durant la phase S du cycle cellulaire (Osley, 1991; Pyronnet et al., 2001), nous avons synchronisé les cellules HEK293FT pendant la phase d'inactivation par les siARN afin d'analyser l'impact de l'inactivation d'eIF3 spécifiquement sur les histones.

Des études antérieures d'inhibition des sous-unités eIF3a, eIF3c, eIF3e et eIF3j (Neusiedler et al., 2012; Wagner et al., 2014, 2016) ont montré des effets variables sur l'intégrité et l'activité du facteur eIF3. Par exemple, l'inhibition des deux principales sous-unités du noyau octamérique d'eIF3 (eIF3a et eIF3c) induit des effets différents. Alors que l'inhibition par siARN d'eIF3a dissocie entièrement le complexe eIF3 *in vivo*, l'inhibition d'eIF3c entraine l'apparition d'un sous-complexe eIF3a, b, g et i (rappelant le module minimal d'eIF3 'Yeast Like Core' de la levure) qui conserve une affinité élevée pour la sous-

unité ribosomique 40S mais avec une relaxation de spécificité de reconnaissance pour l'AUG initiateur (Wagner et al. 2014). D'autres études ont montré que le facteur eIF3e (INT6) interagissait avec la protéine MIF4GD/SLIP1 (protéine de liaison à la SLBP comprenant un domaine MIF4G homologue au facteur eIF4G) et que son inhibition par siARN entrainait une diminution de la synthèse des histones H4, H2B et H3, démontrant que cette interaction entre INT6 et MIF4GD/SLIP1 est nécessaire pour une traduction efficace des ARNm d'histones (Neusiedler et al., 2012). De même, l'interaction entre la sous-unité eIF3g et la protéine MIF4GD/SLIP1 est nécessaire pour une traduction efficace des ARNm d'histones (von Moeller et al., 2013). Nos résultats ont montré qu'après inhibition individuelle des sous-unités d'eIF3c, d, e et g, la traduction des histones augmentait sensiblement mais n'était réellement significative que pour l'histone H4 après 48h (***Figure 39 C***) et 72h d'inhibition. Ces résultats contrastent singulièrement avec ceux obtenus précédemment après inhibition d'eIF3e qui montraient une nette diminution de la synthèse des histones H4, H2B et H3 (Neusiedler et al., 2012). La différence avec nos expériences réside principalement dans les cellules utilisées qui étaient des HeLa contre des HEK293FT dans notre cas. Cet effet de stimulation est également observé lors de l'inactivation simultanée des sous-unités eIF3 c/g et d/e mais pas c/d/e/g après 48h d'inactivation. Néanmoins, après 72h d'inhibition simultanée des sous-unités d'eIF3, le niveau de synthèse des histones baisse très fortement en même temps que l'activité métabolique cellulaire diminue. Dans ces conditions, la quantité disponible et l'intégrité du complexe eIF3 est fortement compromise comme le montrent les western blots des **Figures 40 et 41** ainsi que les expériences préliminaires d'immunoprécipitation du complexe qui sont encore en cours de réalisation. Collectivement ces résultats suggèrent que la diminution de la quantité d'eIF3 disponible favorise dans un premier temps la synthèse des histones, peut-être grâce au site de liaison dans leurs ARNm qui favorise la liaison aux molécules d'eIF3 résiduelles. Cependant, lorsque le complexe eIF3 intègre vient à manquer (après 72h de doubles et quadruples inhibitions), un effet général sur la traduction cellulaire se manifeste, entrainant la diminution de la synthèse des histones. Une hypothèse alternative pour expliquer l'augmentation de synthèse des histones ici observée serait qu'eIF3 serait un répresseur de l'expression des ARNm d'histones. L'inhibition des sous-unités d'eIF3 lèverait l'effet inhibiteur et provoquerait l'augmentation de la traduction des histones. Néanmoins, cet effet régulateur négatif reste modeste dans nos conditions expérimentales (en moyenne 20% d'augmentation de synthèse, et un maximum de 40% pour H4 en conditions d'inhibition modérées de 48h, inhibition simple d'eIF3e) en phase S lorsque la demande maximale d'histones est attendue. Il n'est pas exclu que l'importance de l'effet répresseur soit plus critique à la fin de la phase S du cycle cellulaire lorsque la synthèse des histones cesse brusquement afin d'éviter une surexpression toxique pour la cellule (Meeks-Wagner and Hartwell, 1986). Dans un tel modèle, eIF3 permettrait

d'accentuer le phénomène d'extinction très rapide de la synthèse des histones, avec un effet bénéfique pour la cellule supérieur à l'effet négatif induit sur leur synthèse. Il serait intéressant de tester une telle hypothèse en conduisant une série d'expériences de mesure de la synthèse des histones tout au long de la phase S jusqu'à sa sortie en phase G2.

ANNEXES

MATÉRIEL ET METHODES

1. Matériel

1.1. Lignées cellulaires eucaryotes

HEK293FT : lignée cellulaire immortalisée issue de cellules embryonnaires humaines de rein. Ces cellules possèdent le gène précoce E1A de l'adénovirus permettant l'activation de certains promoteurs viraux et donc la surproduction de protéines.

1.2. Souches bactériennes

Escherichia coli BL21 Rosetta DE3 (Novagen) : souche bactérienne utilisée pour la surproduction des protéines de fusion HisGST-eIF3 (c, d, e et g). Génotype : F- *ompT hsdSB* (rB- mB-) *gal dcm* (DE3) pRARE (CamR). Cette souche contient le gène codant l'ARN polymérase T7 qui est sous la dépendance d'un promoteur inductible lacUV5. Elle permet l'expression de protéines d'origine eucaryote à partir des plasmides pET. Cette souche est résistante au chloramphénicol et exprime des ARNt isoaccepteurs : arginine (AGG, AGA et CGA), isoleucine (AUA), leucine (CUA) et proline (CCC). Ces codons, rarement utilisés par *E. coli*, posent un problème lors de l'expression des gènes eucaryotiques qui en sont riches. La surexpression de ces ARNt rares simultanément à la protéine d'intérêt eucaryotique permet ainsi d'améliorer son expression.

1.3. Plasmides

pUC19 a3 glo R-Luc: plasmide contenant la séquence 5'UTR ß-Globine en amont du gène rapporteur Renilla Luciferase.

pUC19 H4-12 : plasmide pUC19 permettant la transcription de l'ARNm H4.

pET His6 GST TEV LIC (2G-T) : plasmide d'expression procaryote dans lequel ont été clonés les cDNA des sous-unités d'eIF3 en fusion N-terminale avec une double étiquette : l'hexa-histidine et la glutathion S-transférase (HisGST). Ce plasmide contient le gène de résistance à l'ampicilline.

pcDNA/Zeo(+) : plasmide permettant l'expression des constructions bicistroniques *in vivo* et la mesure de l'activité enzymatique du produit synthétisé en utilisant la luciférase.

Tableau M1 : liste des plasmides

Nom du plasmide	Insert	Etiquette	Utilisation
pUC19 a3 Glo R-Luc	5' UTR β-globin-Renilla-Luciferase	-	Transcription d'ARN chimères
pUC19 H4-12	H4-12	-	Transcription d'ARNm H4
pcDNA3.1/Zeo(+)	Bicistron F-Luc-R-Luc	-	Expression des constructions bicistroniques *in vivo*
p2GT-IF3c 1-913	eIF3c 1-913	HisGST	Vecteur d'expression procaryote de type pHis6 GST TEV LIC (2G-T) permettant la surproduction d'eIF3c entière
p2GT-IF3c 1-318	eIF3c 1-318	HisGST	Surproduction d'eIF3c tronquée
p2GT-IF3c 319-913	eIF3c 319-913	HisGST	Surproduction d'eIF3c tronquée
p2GT-IF3c 1-190	eIF3c 1-190	HisGST	Surproduction d'eIF3c tronquée
p2GT-IF3c 191-318	eIF3c 191-318	HisGST	Surproduction d'eIF3c tronquée
p2GT-IF3d	eIF3d	HisGST	Surproduction d'eIF3d
p2GT-IF3e	eIF3e	HisGST	Surproduction d'eIF3e
p2GT-IF3g	eIF3g	HisGST	Surproduction d'eIF3g

1.4. Anticorps

Tableau M2 : Liste des anticorps primaires et secondaires utilisées pour détecter les différentes sous-unités d'eIF3, les protéines d'histones ainsi que les contrôles GAPDH et GST. Le tableau présente les anticorps primaires utilisés, soit dans des expériences de Western Blot, soit d'immunoprécipitation. Les anticorps secondaires sont conjugués à la peroxydase de raifort « Horseradish Peroxydase » (HRP).

Nom de l'anticorps	Classe / Origine	Fournisseur / Ref.	Type
Anti-eIF3a	Polyclonal / Lapin	Novus Biological / NBP1-18891	Primaire
Anti-eIF3b	Polyclonal / Lapin	Bethyl / A301-761A	Primaire
Anti-eIF3c	Polyclonal / Lapin	Bethyl / A300-377A-T	Primaire
Anti-eIF3d	Polyclonal / Lapin	Bethyl / A301-758A	Primaire
Anti-eIF3e	Polyclonal / Lapin	Bethyl / A302-985A	Primaire
Anti-eIF3f	Polyclonal / Lapin	Bethyl / A303-005A	Primaire
Anti-eIF3g	Polyclonal / Lapin	Bethyl / A301-757A	Primaire
Anti-eIF3h	Polyclonal / Lapin	Bethyl / A301-754A	Primaire
Anti-eIF3i	Monoclonal / Souris	BioLegend / 646702	Primaire
Anti-eIF3k	Polyclonal / Lapin	Novus Biological / NB100-93304	Primaire
Anti-eIF3l	Polyclonal / Lapin	GeneTex / GTX120119	Primaire
Anti-GAPDH	Monoclonal / Souris	GeneTex / GTX627408	Primaire
Anti-H1	Polyclonal / Lapin	Neobiotech / NB-22-3513	Primaire
Anti-H2A	Polyclonal / Lapin	Neobiotech / NB-22_3516	Primaire
Anti-H2B	Polyclonal / Lapin	Neobiotech / NB-22-3519	Primaire
Anti-H3	Polyclonal / Lapin	Neobiotech / NB-22-3524	Primaire

Anti-H4	Polyclonal / Lapin	Neobiotech / NB-22-3534	Primaire
Anti-GST- HRP	Polyclonal / Lapin	GE Healthcare / RPN1236V	Primaire
Anti-eIF4E	Monoclonal / Souris	Santa Cruz / sc-271480	Primaire
Anti-SLBP	Polyclonal / Lapin	Bethyl / A303-968A	Primaire
anti-IgG de lapin HRP	Polyclonal / Chèvre	Bethyl / A120-101P	Secondaire
anti-IgG de souris HRP	Polyclonal / Chèvre	BIO-RAD / 170-6516	Secondaire

1.5. Oligonucléotides

1.5.1. Oligonucléotides utilisés pour les expériences de PCR quantitative (RT-qPCR)

Les oligonucléotides utilisés dans les expériences de RT-qPCR ciblent les jonctions exon-exon, ils sont indiqués dans le tableau ci-dessous.

Tableau M3 : Liste des oligonucléotides utilisés en PCR quantitative (RT-qPCR). Le tableau représente les amorces sens (Fw) et antisens (Rev) utilisées en RT-qPCR, orientés dans la direction 5'->3'.

Cibles	Séquence des oligonucléotides (5'-3')
H1	Fw : CGCTACCCCGAAGAAAAGCA Rev : ACCTTTTTCGCACTCTTGGC
H2A	Fw : CGCAACGACGAGGAGCTAA Rev : AAGAGTCTCGTTTTACTTGCCC
H2B	Fw : GTGACCAAGGCGCAGAAGAAGGAC Rev : TTTAGAGGAGATGCCGGTGTCGGG
H3	Fw : CGTGAAAAAGCCTCACCGTT Rev : TCCGAATCAGCAACTCGGTC
H4	Fw : AGGTGCTGCGGGACAATATC Rev : GCCGAAACCATAAAGGGTGC
ACTB	Fw : CCTTCTACAATGAGCTGCGT Rev : CTCCTTAATGTCACGCACGAT
C-JUN	Fw : TGACTGCAAAGATGGAAACG Rev : CAGGGTCATGCTCTGTTTCA
BTG1	Fw : CACTGGTTCCCAGAAAAGC Rev : CTACCATTTGCACGTTGGTG
LDHA	Fw : TGGCAGCCTTTTCCTTAGAACA Rev : ACGGCTTTCTCCCTCTTGCTGA
GAPDH	Fw : CTTTGGTATCGTGGAAGGACT Rev : CCAGTGAGCTTCCCTTTTCAG
snARN U2	Fw : TTCTCGGCCTTTTGGCTAAG Rev : CTCCCTGCTCCAAAAATCCA
eIF3c	Fw : ACCAAGAGAGTTGTCCGCAGTG Rev : TCATGGCATTACGGATGGTCC

eIF3d	Fw : CTGGAGGAGGGCAAATACCT Rev : CTCGGTGGAAGGACAAACTC
eIF3e	Fw : TTCTTCAATCACCCCAAAGG Rev : TAGAACCTGCCGACGTTTTC
eIF3g	Fw : TCTCGTTTTCGCCCGCTTC Rev : GTCTCCAGTAGGCATCGCAA
dN9 (Random)	NNNNNNNNN

1.5.2. Oligonucléotides utilisés pour générer des matrices de transcription d'ARN *in vitro par PCR*

Tableau M4 : Oligonucléotides utilisés pour la production des matrices de transcription in vitro des ARN messagers d'histone H4. Les séquences d'oligonucléotides utilisés en 5' sont flanquées de la séquence du promoteur T7 (marqué en vert) afin de permettre la transcription des produits d'H4-12 après amplification par PCR.

Cibles	Séquences des oligonucléotides (5'-3')
H4 entier	Fw : TAATACGACTCACTATAGGTCATAACCATGTCTGGACG Rev : CGGAATTCAAGCTTCGGTAAGTGAACAAC

1.5.3. Ultramers utilisés pour réaliser les matrices PCR, utilisées dans les tests de traduction *in vitro* des rapporteurs luciférase

Tableau M5 : Séquence des ultramers utilisés pour la production des matrices de transcription d'ARN chimères pour les expériences de traduction *in vitro*. Les ultramers sens utilisés sont constitués de la séquence du promoteur T7 (marquée en vert) fusionné au gène H4 (250-320) sauvage ou muté (des séquences complémentaires aux séquences sauvages sont en gras soulignées dans le tableau) et à celui de la β-globine (β Glo, en bleu). Les ultramers anti-sens comportent une séquence complémentaire au gène rapporteur de la luciférase (marqué en orange gras).

Séquences des ultramères et d'oligonucléotides	Séquences des ultramères (5'-3')
T7 - HP - H4 WT (250-320) – β Glo	Fw : ATATTAATACGACTCACTATAGGCCCTTTTCAGGGCCACCCAGAC CGTCACCGCCATGGACGTGGTCTACGCGCTCAAGCGCCAGGGCC GCACTCTCTA**CGG**ATTC**GGCGGTT**ACATTTGCTTCTGACACAACT GTGTTCACTAGCAACCTCAAACAGACACC
T7 - HP - H4 Mut (250-320) - β Glo	Fw : ATATTAATACGACTCACTATAGGCCCTTTTCAGGGCCACCCAGAC CGTCACCGCCATGGACGTGGTCTACGCGCTCAAGCGCCAGGGCC GCACTCTCTA**GCC**ATTC**CCGCCAA**ACATTTGCTTCTGACACAACT GTGTTCACTAGCAACCTCAAACAGACACC
	Fw :

T7 - β Glo	ATATTAATACGACTCACTATAACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACC
T7 - H4 (250-269)	Fw : ATATTAATACGACTCACTATAGGACCGTCACCGCCATGGACG
Luc - H4 WT (250-320)	Rev : **AACCGCC**GAAT**CCG**TAGAGAGTGCGGCCCTGGCGCTTGAGCGCGTAGACCACGTCCATGGCGGTGACGGTCAGTAA**TTATTGTTCATTTTTGAG**
Luc - H4 Mut (250-320)	Rev : **TTGGCGG**GAAT**GGC**TAGAGAGTGCGGCCCTGGCGCTTGAGCGCGTAGACCACGTCCATGGCGGTGACGGTCAGTAA**TTATTGTTCATTTTTGAG**
Luc – H4 antisens (250-320)	Rev : TTGGCGGCTTAGGCATCTCTCACGCCGGGACCGCGAACTCGCGCATCTGGTGCAGGTACCGCCACTGCCAGAGTAA**TTATTGTTCATTTTTGAG**
Luc ++71 nt	Rev : CAGTTAAAATAAAATTGTTGAAAATGACATTTATATTATTAAACCCGGGAAAAATGTAAATAAAAAACCAAAGTAA**TTATTGTTCATTTTTGAG**
Renilla Luciferase HindIII	Rev : CAAAAGCTTTTATTGTTCATTTTTGAGAACTCG

1.5.4. Séquence des siARN

Un ensemble de siARN « On-TARGET Plus SMART Pool » (Dharmacon™) dirigé contre les quatre sous unités d'eIF3 (c, d, e et g), le facteur d'initiation eIF4E et la protéine de liaison à la tige-boucle en 3' des ARN d'histones (SLBP) a été utilisé dans les expériences d'inactivation de gènes par ARN interférence. Pour chaque gène, un mélange de 4 siARN ciblant des séquences différentes sur l'ARN cible est utilisé simultanément pour une meilleure spécificité d'inhibition. Des siARN « ON-TARGET plus Non-targeting Pool » sans cible parmi les gènes humains ont été utilisés comme contrôles.

Tableau M6 : Liste des siARN.

Cible	Gene ID	Séquence NCBI	Séquence (5'-3')
EIF3C	8663	NM_001199142	GCGGAGUGCCGUUGGUUAA GGAGAGGGCGUCAUUGUCA CAACUGCUGGUUCAGAUUG GCAAGUACAUCUACGCCAA
EIF3D	8664	NM_003753	GUUCAAGCCUAAUGAGUUU CAAGAUAAGAGGUACACAA AGACAAAGAUCGUCGGAAC CGAAUGAGAUUUGCCCAGA
EIF3E	3646	NM_001568	UGGCUUGUCUUGAGGAUUU

			GGAUCGGCAUCUAGUCUUU GGGUAACAAUGCAGUCUCA AAAGGUCGCGAUAAUAUUA
EIF3G	8666	NM_003755	GGUCAUCAACGGAAACAUA GUCCGCACCUUCAGGAUUG GGACAAGACCACUGGCCAA GACGAUGUCUCUAUGACGU
EIF4E	1977	NM_001968	GACGAUGGCUAAUUACAUU CAUAUCCAGUUGUCUAGUA GUGAUAAGAUAGCAAUAUG GAGACGAAGUGACCUCGAU
SLBP	7884	NM_001306075	GAAAUGAGGACCAGAGUUA GAAGUAUAGUCGACGUUCA CGGCUGACUUUGAGACAGA GACAGAAGCAGAUCAACUA
sicontrol	-	-	UGGUUUACAUGUCGACUAA UGGUUUACAUGUUGUGUGA UGGUUUACAUGUUUUCUGA UGGUUUACAUGUUUUCCUA

2. Méthodes

2.1. Culture cellulaire : lignée HEK293FT

2.1.1. Maintien des cellules

Les cellules HEK293FT sont cultivées dans le milieu de culture DMEM (Dulbecco's Modified Eagle's Medium GlutaMAX, GIBCO), supplémenté de 10% de sérum de veau fœtal (FBS), 1% pénicilline/ streptavidine (GIBCO, 10^4 U/mL pénicilline, 10^5 µg/mL streptavidine) et 1% généticine (GIBCO, 50 mg/mL). Les cellules HEK293FT (3.10^6 cellules) sont ensemencées dans 7 mL du milieu de culture décrit ci-dessus (dans une boîte de culture de 25 cm^2), incubées dans l'étuve à 37°C sous 5% CO_2 et hydratation contrôlée pour environ 24 heures. Les cellules arrivées à confluence sont transférées dans des boîtes de 75 cm^2. Pour cela, le milieu de culture est d'abord éliminé et les cellules sont délicatement lavées puis dissociées dans 1 mL de trypsine-EDTA 0,25% (GIBCO) ; afin de stopper l'action de la trypsine les cellules sont ensuite reprises dans 10 mL de milieu de culture. Environ 2 mL de cellules sont ensemencés dans des nouvelles boîtes de 75 cm^2 en vue de leur utilisation. Elles sont maintenues en culture pendant environ 25 passages successifs.

2.1.2. Transfection de cellules et inactivation des gènes par ARN d'interférence

Les cellules HEK293FT sont transfectées par différents siARN dirigés contre les gènes d'intérêt à l'aide de la lipofectamine 2000 (Invitrogen) comme décrit par le fournisseur. 48 et 72 heures après transfection, les cellules sont récoltées et traitées selon les différentes méthodes décrites ci-dessous.

2.1.3. Préparation d'extraits cellulaires totaux contenant des complexes ribonucléoprotéiques (RNP) pontés à la formaldéhyde

Les cellules HEK293FT sont ensemencées dans les boîtes de pétri (10 cm^2) pendant 24 h. En vue de leur récupération, les cellules sont lavées deux fois avec 1 mL de DPBS (137 mM NaCl ; 26,8 mM KCl ; 40,2 mM Na_2HPO_4 ; 17,6 mM KH_2PO_4 ; pH 7,4). Les cellules sont ensuite récupérées dans 1mL de DPBS (Dulbecco's Phosphate Buffer Saline) puis centrifugées à 1000 rpm (90 g) pendant 10 min à 4 °C. Le culot cellulaire est repris dans un volume de formaldéhyde à 0,2% afin de stabiliser les complexes RNP, puis la réaction est arrêtée par l'ajout de 0,15 M de glycine. Les cellules sont centrifugées et reprises dans un volume de tampon RNP (10 mM HEPES-NaOH pH 7 ; 100 mM KCl ; 5 mM $MgCl_2$; 0,5 % NP40 ; 1 mM DTT ; 100 U /ml RNasin (Promega) ; 400 µM VRC (Vanadyl Ribonucleotide Complex, Sigma) ; 1X cocktail d'inhibiteurs de protéases (Sigma) et incubées pendant 5 min dans la glace. Pour préparer des extraits natifs sans pontage les cellules lavées au DPBS sont directement mises en suspension dans un volume de tampon RNP. Les extraits ainsi récupérés pourront être directement utilisés dans des expériences d'immunoprécipitation des complexes ARN/protéines.

2.1.4. Extraction des ARN cellulaire totaux

Les cellules sont ensemencées pendant 24 h avant l'extraction des ARN dans des boîtes de pétri de 10 cm^2 (environ 3.10^6 cellules/ boîte de pétri à 10 cm^2) et incubées dans l'étuve à 37 °C sous 5% CO_2. Après deux lavages avec du DPBS1X, les cellules sont repris dans 1 mL de TRIReagent™ (Invitrogen), incubées 10 min à température ambiante puis additionnées de 200 µL de chloroforme. Les extraits sont vortexés et incubés 10 min à température ambiante, puis centrifugées pendant 15 min à 14000 rpm (18000 g) à 4 °C. La phase aqueuse est récupérée et précipitée par 1 volume d'isopropanol pendant 10 min à température ambiante. Après centrifugation à 14000 rpm, le culot est lavé à l'éthanol 80% puis repris dans 50 µL d'eau milliQ.

2.2. Transfection des cellules HEK293FT par les différents siARN et analyse de l'expression *de novo* des histones par autoradiographie et Western Blot

Les cellules HEK293FT ont été utilisées pour analyser l'effet de l'inhibition des quatre sous-unités d'eIF3 (c, d, e et g) ainsi qu'eIF4E et SLBP sur l'expression *de novo* des protéines d'histones H1, H2A, H2B, H3 et H4. 24h avant la transfection, 4.10^5 cellules sont mises en culture dans des boîtes de six puits dans 2 mL de milieu de culture DMEM supplémenté en 10% FBS, sans ajout d'antibiotiques. Lorsque les cellules ont atteint environ 70% de confluence, chaque cible est inactivée par 100 pmol de siARN (On-TARGET Plus SMART Pool, Dharmacon™, voir ***Tableau M6***). Les transfections sont faites à l'aide de la lipofectamine 2000 (Invitrogen) comme décrit par le fournisseur. L'inactivation simultanée de deux (eIF3 c/g ; eIF3 d/e) ou quatre cibles (eIF3 c /d /e/g) a été réalisée en présence de 100 pmol de chaque siRNA. Chaque siARN est lui même un mélange de quatre ARN dirigés contre l'ARNm cible. Ces siARN sont 2'-O-methylés au niveau des riboses afin de limiter les hybridations aspécifiques hors cibles. Pour chaque condition d'inhibition 3 répliquas biologiques sont réalisés. L'analyse de l'effet de l'inhibition est analysée après 48 et 72 heures.

2.2.1. Analyse de la traduction *de novo* de protéines d'histones par marquage métabolique à la [^{35}S] méthionine

Les histones sont massivement exprimées durant la phase S du cycle cellulaire. Afin d'étudier de façon optimale l'impact de l'inhibition des sous-unités d'eIF3 par siRNA sur la traduction des histones, les cellules ont été préalablement synchronisées. 32 h après leur transfection, les cellules HEK293FT sont synchronisées par la méthode de double blocage à la thymidine telle que décrite par (Bostock et al., 1971). L'excès de thymidine a pour effet de bloquer le cycle cellulaire à la transition G1/S en inhibant la ribonucléotide réductase, empêchant ainsi la production de nucléotides précurseurs au dCTP. Ainsi, les cellules qui se trouvent en phase S ne peuvent l'achever, alors que celles qui sont en G1 ne peuvent pas amorcer la synthèse d'ADN. Le blocage est effectué par ajout de 2 mM de thymidine (Sigma-Aldrich). Après 15 h les cellules sont rincées avec du milieu DMEM/10% FBS et incubées avec 1 mL de ce même milieu pendant 9 h. Les cellules sont bloquées une seconde fois avec 2 mM de thymidine pendant 13 h en vue d'obtenir une population cellulaire homogène en transition G1/S. Le niveau de synchronisation des cellules est vérifié par cytométrie de flux (FACS). Les cellules sont libérées du blocage et incubées pendant 1 h dans le milieu DMEM/10% FBS. Le milieu de culture est retiré et les cellules sont incubées pendant 30 min dans du milieu dépourvu de méthionine ou « starvation medium » (DMEM Glutamax Gibco dépourvu des acides aminés méthionine, glutamine et cystine, complété par 10% FBS, 3,97

mM de glutamine, 0,153 mM cystine, 1 mM pyruvate de sodium, 0,25 mM HEPES pH 7,5). Ce milieu est alors additionné de 10 µCi de [^{35}S] méthionine pendant 10 min. Les cellules sont lavées 3 fois avec du PBS froid. Afin de préparer des extraits cellulaires totaux contenant l'intégralité des complexes nucléoprotéiques (RNP), 25% des cellules sont traitées comme décrit dans le paragraphe (***2.1.***) ci-dessus. Les histones néosynthétisées sont extraites des 75% d'extraits restants par la méthode décrite ci-dessous.

2.2.2. Extraction spécifique des histones (d'après (Cakmakci et al., 2008)

Pour extraire spécifiquement les histones, les noyaux des cellules sont isolés après lyse du culot de cellules dans 100 µL de tampon d'extraction nucléaire (250 mM sucrose ; 1 mM $CaCl_2$; 2 mM $MgCl_2$; 1% Triton X-100 ; 10 mM Tris-HCl pH 8) pendant 1 h sur la glace. Les cellules sont centrifugées à 5000 rpm pendant 5 min à 4 °C. Le culot contenant les noyaux cellulaires est repris dans 0,4 N H_2SO_4 (les histones sont solubles dans l'acide sulfurique) et incubé dans la glace pendant 30 min. Les protéines solubles dans l'acide sulfurique sont récupérées dans le surnageant après centrifugation des noyaux à 14000 rpm (18000 g) pendant 30 min à 4 °C. Le surnageant contenant les histones est récupéré puis précipité par ajout de 20% d'acide trichloroacétique (TCA) final à -20°C. Le mélange est ensuite centrifugé à 14000 rpm (18000 g) pendant 30 min à 4°C et les culots d'histones sont lavés 3 fois avec de l'acétone froid et repris dans 20 µL de tampon de lyse RIPA (150 mM NaCl, 5 mM EDTA pH 8 ; 50 mM Tris-HCl pH 8 ; 1% NP-40 ; 0,5% sodium déoxycholate ; 0,1% SDS) supplémenté d'un cocktail 1X d'inhibiteurs de protéases (Sigma). Pour achever la dissolution des culots, les extraits sont soniqués 2 fois pendant 3 secondes à l'aide d'un sonicateur (Vibra Cell^{TM} 75022, Bioblock Scientific).

2.2.3. Quantification de la radioactivité par scintillation et séparation sur gel SDS-PAGE

1 µL de protéines radioactives est déposé sur un papier « Wathman », puis plongé 3 fois dans des bains d'acide trichloroacétique (TCA) 5% pendant 10 min afin de précipiter les protéines sur les papiers puis éliminer l'excédent de [^{35}S] méthionine libre. Les papiers sont lavés 2 fois dans des bains d'éthanol 100% pendant 5 min, séchés sous des lampes chauffantes pendant 5 min puis plongés dans 2 mL de liquide de scintillation afin de quantifier la radioactivité grâce au compteur à scintillation (LS-6500, Beckman Coulter^{TM}). Des quantités équivalentes de protéines radioactives sont déposées et séparées sur gel polyacrylamide SDS-PAGE 15%. Le gel est coloré au bleu de Coomassie (50% éthanol ; 10% acide acétique ; 0,03% bleu de coomassie), séché, puis exposé pendant 24 h et révélé par autoradiographie. La bande correspondante à chaque protéine d'histone est quantifiée à

l'aide du logiciel ImageQuant, et normalisée par rapport à la quantité de radioactivité totale déposée dans chaque puits.

2.3. Analyse de l'état métabolique des cellules HEK293FT après la transfection par les différentes siARN à l'aide du réactif de prolifération WST-1

Afin de tester l'effet de l'inhibition des différents siARN sur la prolifération des cellules HEK293FT, les expériences d'inactivation sont reproduites sur 2.10^4 cellules dans des plaques à 96 puits à l'aide de 4 pmol de siARN comme décrit précédemment (***2.2.***). L'activité proliférative des cellules est mesurée à l'aide du sel de tétrazolium WST-1. La prolifération cellulaire induit une augmentation de l'activité des déshydrogénases mitochondriales, qui clive le sel de tétrazolium WST-1 en formazan ce qui reflète directement le nombre des cellules métaboliquement actives dans la culture. La détection colorimétrique du formazan est réalisée à des différents temps après la transfection (4h, 8h, 24h, 48h et 72h) en mesurant l'absorbance de la solution de formazan à 450 nm à l'aide d'un spectrophotomètre à lecteur multi puits (ELISA reader).

2.4. Immunoprécipitation des particules ribonucléoproteiques (RNP)

L'immunoprécipitation de RNP permet de tester l'interaction entre le facteur d'initiation eIF3 et différents ARNm *in vivo*. Ces expériences d'immunoprécipitation sont réalisées à l'aide d'extraits RNP de cellules HEK293FT et d'un anticorps dirigé contre eIF3b (Bethyl, A301-761A) en utilisant les billes magnétiques protéine A µMACS™ (Miltenyi Biotec). 300 µL d'extraits RNP sont incubés en présence de 2 µg d'anticorps anti-eIF3b et de 100 µL de billes protéine A dans 1 mL de tampon de lyse (1% Igepal® CA-630 (NP-40), 50 mM Tris HCl pH 8) pendant 30 min sur la glace. Des complexes incubés en absence d'anticorps sont utilisés comme contrôles. Les complexes ARN-protéines sont chargés sur des colonnes µMACS™ (Miltenyi) placées sur un support magnétique et préalablement lavées avec 200 µL du tampon de lyse. Les effluents non retenus sont conservés. Les complexes ARN-eIF3 retenus sur les colonnes sont lavés 4X avec le tampon de lavage 1, puis 1X avec le tampon de lavage 2 (Miltenyi Biotec) avant d'être élués dans 50 µL de tampon d'élution chauffé à 95 °C pendant 5 min. Les protéines immunoprécipitées (10 µL) sont analysées sur gel SDS-PAGE et immuno-détectectées par Western Blot. Les ARN liés à eIF3b et les ARN libres du surnageant sont récupérés par extraction au phénol-chloroforme. Les échantillons sont additionnés d'un volume de phénol-chloroforme (phénol/chloroforme/alcool isoamylique 50/46/4 en % v/v) puis agités vigoureusement sur un « vortex » avant d'être centrifugés 2 min à 16000 g. La phase aqueuse contenant les ARN est récupérée et les ARN sont

précipités toute la nuit à -20 °C par ajout de 3 volumes d'éthanol 100 % et 250 mM final de NaCl. Après centrifugation 10 min à 16000 g le culot d'ARN est lavé à l'éthanol 80 % et dissous dans un volume adéquat d'eau milliQ stérile.

2.5. Détection des ARN par transcription inverse et PCR quantitative (qRT-PCR)

Les ARN associés au facteur eIF3 et récupérés après immunoprécipitation, sont traités à la DNase I, soumis à une transcription inverse et analysés par qRT-PCR.

2.5.1. Traitement des ARN à la DNase I

Afin d'éliminer l'ADN génomique résiduel qui interfèrerait avec la synthèse des ADNc, 2 µg d'ARN sont incubés en présence de 2,5 U de DNase I RNase free (Fermentas) et de tampon DNase (40 mM Tris-HCl pH 7,9 ; 10 mM NaCl ; 6 mM $MgCl_2$; 1 mM $CaCl_2$) dans un volume final de 20 µL pendant 15 min à 25 °C. La réaction est arrêtée par ajout de 8 mM d'EDTA final et incubation pendant 10 min à 70 °C.

2.5.2. Transcription inverse

Des copies d'ADN simple-brin complémentaires (ADNc) de l'ensemble des ARN présents dans l'échantillon sont générées par transcription inverse. Pour ce faire, 10 µL des échantillons traités à la DNase I sont dénaturés 5 min à 70 °C en présence de 100 pmol d'oligonucléotides dégénérés dN9 (Sigma) et 10 nmol de dNTP, puis mis dans la glace. La transcription inverse s'effectue avec 5 U de « reverse transcriptase » AMV (MP Biomedicals) dans le tampon AMV (50 mM Tris-HCl pH 8,3 ; 6 mM $MgCl_2$; 60 mM NaCl ; 100 mM DTT) et en présence de 40 U de RNasin. L'hybridation est réalisée à 25°C pendant 15 min et l'élongation 1 h à 42 °C. Les ADNc sont stockés à -20 °C.

2.5.3. Quantification des ADNc par qPCR

Les ADNc présents dans l'échantillon sont quantifiés par qPCR en temps réel à l'aide du kit Master Mix Maxima SYBR® Green qPCR (ThermoFischer), contenant du SYBR® Green et une Hot Start Taq ADN polymérase. L'intensité du signal émis, proportionnelle à la quantité d'ADN, est mesurée en temps réel par le thermocycleur (BioRad CFX96). 4 µL des échantillons d'ADNc dilués au 1/20ème sont additionnés de 12,5 µL de réactif Master Mix 2X et de 0,25 µM final de chaque oligonucléotide sens et antisens dans un volume final de 25 µL. Les échantillons sont dénaturés pendant 10 min à 95 °C puis soumis à 40 cycles de qPCR : 15 sec à 95 °C (dénaturation), 30 sec à 58 °C (hybridation) et 30 sec à 72 °C (élongation). Une courbe de fusion des produits de PCR finaux est réalisée dans une

dernière étape de qPCR : 1 min à 95 °C, 30 sec à 55 °C et 30 sec à 95 °C. Les données sont traitées à l'aide du logiciel CFX Manager (BioRad).

2.6. Western Blot

Les protéines sont fractionnées sur gel 10 % polyacrylamide (acrylamide : bisacrylamide 19 :1) SDS-PAGE à 100 V pendant 2 h, puis transférées sur une membrane Immobilon-P (Millipore). Le transfert est effectué à 10 Volts pendant 1 h à l'aide d'un appareil de transfert (semi-dry Trans-Blot® SD) sur une membrane de PVDF (PolyVinyliDene Fluoride) préalablement activée pendant 2 sec dans du méthanol 100%, dans le tampon de transfert pH 8 (25 mM Tris ; 200 mM glycine ; 20% éthanol). Après transfert, la membrane est saturée pendant 2 h dans le tampon de blocage (5% lait ; 0,5% Tween-20 ; PBS1X). Les anticorps primaires ciblant les différentes sous-unités du complexe eIF3 sont dilués selon les instructions du fournisseur dans le tampon de blocage et hybridés sur la membrane pendant la nuit à 4 °C. La membrane est ensuite lavée 3 fois avec le tampon de lavage (PBS1X ; 0,5 %Tween-20) pour enlever l'excès d'anticorps primaire. Suite aux lavages, la membrane est hybridée avec l'anticorps secondaire couplé à la HRP pendant 1 h puis lavée 3 fois avec le tampon de lavage à température ambiante. Le signal lumineux émis par la réaction entre l'HRP et le réactif ECL (Kit ECL Plus Western Blotting Detection System, GE Healthcare) est révélé par chimioluminescence à l'aide d'une station d'imagerie ChemiDoc (BioRad).

2.7. Réaction de polymérisation en chaîne (PCR)

Les matrices de transcription des ARNm sont générées par amplification PCR. Les oligonucléotides sens utilisés ont été souvent rallongés en 5' de la séquence du promoteur T7 afin de permettre la transcription des différentes matrices. 20 ng de plasmide sont mélangés avec 100 µM d'oligonucléotides sens et antisens, en présence de 0,5 mM dNTPs, 20 µL de tampon HF 5X (F-518, Thermo Scientific) 3 µL de DMSO dans 100 µL de milieu réactionnel final. La réaction d'amplification est réalisée à l'aide de la polymérase Phusion (0,006 mg/mL final). Les échantillons sont dénaturés pendant 10 min à 95 °C puis soumis à 35 cycles de PCR. Chaque cycle commence par une dénaturation pendant 30 sec à 95 °C, suivie de l'hybridation des oligonucléotides pendant 30 sec à 55 °C, l'élongation des produits PCR pendant 1,5 min à 72 °C puis 5 min à 72 °C pour une dernière phase d'élongation. 5 µL de produit PCR additionnés de 3 µL de bleu de Sarkosyl sont analysés sur gel d'agarose 1%.

2.8. Transcription *in vitro* d'ARNm

2.8.1. Transcription *in vitro* des ARNm d'histone H4-ThioU radio-marqué à l'[α-^{32}P]-ATP

La synthèse d'ARN contenant des nucléosides modifiés 4-Thiouridine (ThioU) peut être réalisée par transcription à l'aide de l'ARN polymérase T7. Les réactions de transcription sont réalisées dans 100 µL final à partir d'une matrice de transcription générée par PCR (***Tableau M4***, et comme décrit ci-dessus) avec 5 mM de (CTP, GTP et ThioU (Jena Bioscience), 2,5 mM UTP et 10 µL de l'[α-^{32}P]-ATP à 6000 Ci/mmol dans le tampon TMSDT (40 mM Tris-HCl pH 8,1 ; 22 mM $MgCl_2$; 1 mM spermidine ; 5 mM DTT ; 0,01% Triton X-100) en présence de 40 U de RNasin. La réaction est réalisée à l'aide de l'ARN polymérase T7 (0,2 mg/mL final) à 37 °C. De l'ATP froid est rajouté 5 fois toutes les 10 min (1 µL à 50 mM) de façon à ce que la concentration finale devienne 5 mM après 1 h d'incubation à 37°C. Après 1h d'incubation à 37 °C, 0,01 mg/mL de pyrophosphatase (Roche) est ajoutée afin de digérer le pyrophosphate libre et déplacer l'équilibre vers la synthèse d'ARN et l'incubation est poursuivie pendant 30 min. La DNase I est ensuite ajoutée (20 U/mL final) afin de dégrader les brins d'ADN matrices pendant 1 h. La réaction est arrêtée par deux extractions des ARN au phénol-chloroforme et précipitation à l'éthanol (comme décrit dans le paragraphe ***2.4.***).

2.8.2. Ajout co-transcriptionnel de la coiffe ^{7}mGpppG

Les matrices de transcription ont été réalisées par amplification PCR à partir des plasmides recensés dans le ***Tableau M5***. Les réactions de transcription ont été réalisées dans 100 µL final à partir de 20 µL de matrice d'ADN (1 réaction de PCR), 14,3 µL tampon TMSDT 7X (40 mM Tris-HCl pH 8,1 ; 22 mM $MgCl_2$; 1 mM spermidine ; 5 mM DTT ; 0,01% Triton X-100), 5 mM de rNTP (ATP ; CTP et UTP), 0,5 mM m^{7}GpppG, 0,4 U/µL des RNasin et 0,5 mg/mL de T7 ARN polymérase. Après 10 min d'incubation à 37 °C, 1 mM de GTP est ajouté, suivi de 4 autres ajouts à des intervalles de 10 min. Ceci permet d'ajouter la coiffe m^{7}GpppG au début des ARN en cours de transcription. Après 1 h d'incubation à 37 °C, 0,01 mg/mL de pyrophosphatase (Roche) est ajoutée et l'incubation est poursuivie pour 30 min afin de digérer le pyrophosphate libre et déplacer l'équilibre vers la synthèse d'ARN. La DNase I est ensuite ajoutée (20 U/mL final) afin de digérer les brins d'ADN matrices pendant 1 h. La réaction est arrêtée par extraction des ARN au phénol-chloroforme et précipitation à l'éthanol 100%.

2.9. Purification des transcrits sur gel de polyacrylamide dénaturant

Les culots d'ARN sont repris dans 40 µL de tampon Formamide (95% formamide, 0,025% bleu de bromophénol, 0,025% bleu de xylène cyanol), dénaturés pendant 5 min à 95 °C puis séparés par électrophorèse sur gel de polyacrylamide (4% acrylamide ; 8 M urée; tampon TBE (89 mM Tris-Base ; 89 mM acide borique ; 2,5 mM EDTA)). Les ARN radioactifs sont révélés par autoradiographie (Blue X-Ray film) et la bande correspondante à l'ARN d'intérêt est découpée et mise à éluer pendant une nuit dans le tampon d'élution (Cracking buffer : 0,3 M NaCl ; 0,5 mM EDTA ; 10 mM Tris-HCl pH 7,5) à 4 °C sous rotation. Les ARN élués sont précipités à l'éthanol et l'ARN radiomarqué est quantifié à l'aide d'un compteur à scintillation (LS-6500, Beckman Coulter^{TM}).

2.10. Pontage au rayonnement UV 365 nm des ARNm-ThioU radioactifs de l'histone H4 avec le facteur eIF3

50000 cpm d'ARNm ThioU-H4 radiomarqué sont incubés en présence de 5 µM final du complexe protéique eIF3 purifié (fourni par le laboratoire du Prof. W.C. Merrick, Case Western Reserve University, Cleveland, USA) dans un volume final de 4 µL pendant 30 min à 25°C (100 mM KCl ; 20 mM Tris 7.5 ; 1 mM DTT ; 0.1 mM EDTA ; 10% glycérole). Le pontage ARN-protéines est réalisé par la photo-oxidation du ThioU par des rayonnements UV 365 nm, pendant 30 min à 365 nm. Un traitement à la RNAse A (Roche) est réalisé pendant 30 min à 37 °C afin de dégrader les brins d'ARN non protégés par l'interaction avec les sous-unités du facteur eIF3. Les fragments d'ARNm radioactifs pontés aux sous-unités d'eIF3 restent liés aux protéines.

2.11. Expériences de séparation des complexes eIF3-[ARNm-ThioU] radioactifs par électrophorèse bidimensionnelle

L'électrophorèse 2D est une méthode de choix pour séparer des protéines de propriétés physico-chimique proches en fonction de deux propriétés différentes : la charge électrique et le poids moléculaire. Dans la première dimension, les protéines sont soumises à une électrophorèse dans un gel présentant un gradient de pH continu grâce à la présence d'ampholytes. Au cours de cette étape, appelée isoélectrofocalisation (IEF), les protéines migrent dans le gel jusqu'à une position où la valeur du pH est égale à celle de leur point isoélectrique (pI). Dans la deuxième dimension les protéines sont séparées par la technique SDS-PAGE. Cette méthode a été utilisée pour déterminer sans ambiguïté la nature des sous-unités du complexe eIF3 en interaction avec les ARNm-ThioU après pontage.

Pour ce faire, quatre réactions de pontage préparées comme décrit ci-dessus sont rassemblées puis précipitées par 5 volumes d'une solution d'acétate d'ammonium (0,1 M $AcNH_4$; 100% méthanol) pendant une nuit à -20 °C. Les protéines précipitées sont centrifugées à 13000 rpm (15500 g) pendant 15 min à 4 °C, puis lavées 2X avec une solution de (0,1 M $AcNH_4$; 80% méthanol) et centrifugées à 13000 rpm (15500 g) pendant 15 min à 4 °C. Les culots sont séchés pendant 5 min, puis mis en suspension dans 125 µL de tampon UTCT (7 M urée, 2 M thio-urée, 4% CHAPS, 50 mM DTT, 0,2% ampholytes 3/10). Les échantillons sont incubés en présence de gels d'IEF en barrettes IPG 7cm/ pH 4-7 (ReadyStrip^{TM} IPG Strips, BioRad). La migration des échantillons au sein des barrettes est assurée grâce au générateur PROTEAN IEF Cell (BioRad) selon le programme suivant : 0 à 250 Volts pendant 15 min, puis 250 à 4000 Volts pendant 1 h et enfin de 10 à 20000 V-heure (50 µA/ IPG). Les barrettes sont alors équilibrées dans le tampon d'équilibrage 1 (6 M urée ; 0,375 Tris-HCl pH 8,8 ; 2% SDS ; 20% glycérol ; 2% DTT) pour casser les groupes sulfhydriles pendant 10 min sous une faible agitation puis transférées dans le tampon d'équilibrage 2 (6 M urée ; 0,375 Tris-HCl pH 8,8 ; 2% SDS ; 20% glycérol ; 2,5% iodoacétamide) afin d'ajouter des groupes alkyles sur les groupes sulfhydriles réduits. Puis les barrettes sont plongées dans le tampon TGS (25 mM Tris ; 200 mM glycine ; 0,1% SDS) avant d'être incluses dans un gel 10% polyacrylamide SDS-PAGE en vue de la séparation des protéines suivant la seconde dimension. Cette séparation est faite pendant 2 h à 100 V puis les protéines sont transférées sur une membrane Immobilon-P (Millipore) comme décrit dans le paragraphe (***2.6.***). La membrane est ainsi exposée 24 h et révélée par autoradiographie.

2.12. Expériences de retard sur gel

L'interaction entre les ARN H4 et le facteur eIF3 est testée *in vitro* par les expériences de retard sur gel. D'abord, l'ARNm entier (1-375) ainsi que des fragments tronqués (1-178 et 179-375) ont été générés par transcription *in vitro* et marqués de manière interne à l'aide d'[α^{32}P]-ATP. 30000 cpm d'ARN marqué au [^{32}P] a été incubé en présence de 15 µg de tRNA et de 1,5 µL de tampon PBSD5X pH 7,4 (0,7 M NaCl ; 13,5 mM KCl ; 50 mM Na_2HPO_4 ; 9 mM KH_2PO_4 ; 10 mM DTT) en présence de concentrations croissantes d'eIF3 (1 ; 2 et 4 µM) dans un volume final de 8 µL, pendant 30 min à 30°C. Les complexes ARN-eIF3 formés sont ensuite séparés sur un gel de polyacrylamide 6% non-dénaturant (6% polyacrylamide (acrylamide : bisacrylamide 19 :1) ; 2,5 mM Tris base ; 2,5 mM Glycine), dans un tampon d'électrophorèse 0,5X TBE, 5% glycérol. Le gel est séché, exposé pendant 30 min et révélé par autoradiographie.

2.13. Transformation de bactéries compétentes pour la surproduction de protéines

Les bactéries BL21 Rosetta DE3 sont rendues compétentes par la méthode au chlorure de calcium. 50 mL de milieu LB (10 g/L peptone ; 5 g/L d'extrait de levure ; 10 g/L NaCl ; 1 mM NaOH) sont ensemencés avec 0,5 mL de préculture et incubés à 37°C jusqu'à ce que l'absorbance à 600 nm atteigne 0,4 (culture en phase exponentielle). Les cellules sont alors centrifugées à 5000 g durant 5 min et repris par 10 mL d'une solution de 50 mM $CaCl_2$, 10 mM Tris-HCl, pH 8. Au bout d'une heure d'incubation dans la glace, les cellules sont recentrifugées et le culot est resuspendu dans 2 mL de la même solution de $CaCl_2$. 200 µl de bactéries compétentes sont transformées avec 1 µg de plasmide codant les différentes protéines recombinantes eIF3c, d, e et g. Elles sont incubées pendant 1 h à 4 °C en présence du plasmide et le choc thermique est réalisé par 3 min d'incubation à 37 °C. Les bactéries transformées sont ensuite incubées pendant 20 min à 37 °C dans 1 mL de milieu LB pour la régénération de la membrane. Enfin, les bactéries sont étalées sur milieu LB solide (LB + 20 g/L Agar-Agar) additionné de 100 µg/µL d'ampicilline et 30 µg/µL de chloramphénicol et incubées une nuit (O/N) à 37 °C. Une colonie est ensuite utilisée pour ensemencer les cultures liquides.

2.13.1. Surproduction et purification de protéines recombinantes

Les bactéries BL21 Rosetta DE3 transformées sont ensemencées dans 1 L de milieu LB contenant 100 µg/µL d'ampicilline et 30 µg/µL de chloramphénicol. Lorsque l'absorbance à 600 nm atteint 0,6, l'expression de la protéine est induite par ajout de 0,2 mM final d'IPTG. Les bactéries sont incubées pendant 20 h à 20 °C sous agitation pour la surproduction de protéines. Les bactéries sont centrifugées 10 min à 5000 g et le culot est resuspendu dans 8 mL de tampon de lyse (20 mM Tris HCl pH 7,5 ; 300 mM KCl ; 25 mM imidazole ; 1 pastille de cocktail d'inhibiteurs de protéases (Sigma) ; 10 % glycérol). Les bactéries sont lysées à 4 °C par sonication (4x 1 min). Le lysat est clarifié des débris cellulaires par une première centrifugation de 15 min à 8000 g à 4 °C. Le surnageant est récupéré pour une seconde ultracentrifugation de 1 h 30 à 105.000 g à 4 °C. Le surnageant est ensuite incubé avec la résine Ni-NTA, préalablement équilibrée avec le tampon de lyse, pendant 1 h à 4 °C sous rotation. Après sédimentation de la résine dans une colonne, l'effluent est collecté et la résine est lavée par le tampon de lyse jusqu'à ce que l'absorbance atteigne le niveau de base. Les protéines possédant l'étiquette 6xHis sont enfin éluées avec le tampon d'élution (20 mM Tris HCl pH 7,5 ; 300 mM KCl ; 0,5 M imidazole ; 10 % glycérol) et collectées en une dizaine de fractions. Les protéines purifiées sont analysées par électrophorèse sur gel de polyacrylamide 10 % en conditions dénaturantes (SDS-PAGE), puis colorées au bleu de Coomassie (50 % éthanol ; 10 % acide acétique ; 0,03 % bleu de Coomassie) pendant 20

min. Le gel est enfin décoloré par lavages successifs par une solution (7 % éthanol ; 10 % acide acétique).

2.13.2. Dialyse de protéines purifiées

La dialyse est effectuée afin d'éliminer l'imidazole des échantillons de protéines et de les transférer dans leur tampon de stockage. Les fractions contenant les protéines purifiées, identifiées par gel SDS-PAGE, sont rassemblées et transférées dans des boudins de dialyse (limite d'exclusion 6-8000 Da). Les fractions sont dialysées à deux reprises contre le tampon de dialyse (20 mM Tris HCl pH 7,5 ; 300 mM KCl ; 1 mM DTT ; 10 % glycérol) pendant 6 h, puis O/N à 4 °C dans le même tampon contenant 50 % glycérol. La concentration des protéines purifiées est ensuite déterminée par dosage colorimétrique grâce à la méthode de Bradford.

2.14. Expériences de GST pull-down

L'interaction entre les ARN messagers d'histones endogènes (H1, H2A, H2B, H3 et H4) et les sous unités d'eIF3 est analysée *in vitro* par des expériences de GST pull-down à l'aide de protéines recombinantes et d'extraits d'ARN totaux de cellules HEK293FT.

Pour chaque expérience de GST pull-down 50 µL de billes d'agarose couplées à des anticorps anti-GST (GST-TRAP agarose, Chromotek), préalablement lavées 3X avec 500 µL de tampon de lavage froid (20 mM tris HCl pH 7,5 ; 100 mM KCl ; 0,1 mM EDTA : 1 mM DTT ; 1X cocktail d'inhibiteurs de protéases (Sigma) ; 10 % glycérol) sont incubées pendant 1 h à 4 °C sous rotation avec 40 µg de protéine recombinante HisGST-eIF3c purifiée. Par la suite, les billes sont lavées 3X avec 500 µL de tampon de lavage froid et sont incubées pendant 30 min à 4 °C sous rotation avec 60 µg d'ARN totaux extraits de cellules HEK293FT (voir paragraphe ***2.1.***). Le surnageant, correspondant à la portion d'ARN non fixés, est conservé pour l'analyse des données par RT-qPCR. Les billes sont lavées 3X avec 500 µL puis resuspendues dans 500 µL de tampon de lavage froid. Une aliquote de 100 µL de billes est conservée afin de vérifier la fixation de la protéine recombinante par Western Blot.

2.15. Tests de traduction *in vitro* à l'aide des extraits de réticulocytes du lapin

Afin de tester l'effet de la tige-boucle (250-320) de l'histone H4 sur la traduction, des matrices de transcription d'ARN chimères sont générées par PCR, puis transcrites en ARNm (***Tableau M5***). Puis 2 µM d'ARNm transcrits sont incubés 1 h à 30 °C en présence de 10 µL de lysat de réticulocytes de lapins (RRL), en présence de 0,22 mM d'un mélange d'acide aminés sans méthionine, 0,125 µCi/µL de [^{35}S] méthionine, 1 U/µL de RNasin, 75 mM de KCl

et 0,5 mM de $MgCl_2$ dans un volume final de 20 µL. Les protéines sont ensuite séparées par SDS-PAGE puis le gel est coloré au bleu de Coomassie, puis décoloré (comme décrit au paragraphe ***2.2.***) et enfin séché et exposé pendant 24 h avant d'être révélé par autoradiographie.

Alternativement, les produits de traduction peuvent être analysés par mesure de l'activité enzymatique du produit synthétisé. Dans les expériences utilisant la luciférase, 8 µL des produits de traduction étaient placés dans une plaque opaque à 96 puits afin de détecter la luminescence émise par la protéine luciférase, en utilisant les réactifs comme décrit par le fournisseur (Dual-Glo™ Luciferase Assay System, Promega). Le signal est détecté à l'aide d'un luminomètre (GloMax Multi Plus Detection System, Promega).

2.16. Tests de traduction *in vivo* dans les cellules HEK293FT

Les cellules HEK293FT ont été transfectées par les différents plasmides pcDNA/Zeo(+) à l'aide de lipofectamine (comme décrit par le fournisseur). 48 heures après transfection, les cellules ont été récoltées et traitées dans le tampon PLB 'Passive Lysis Buffer' comme décrit par le fournisseur (Dual-Glo™ Luciferase Assay System, Promega) de façon à obtenir des extraits protéiques natifs permettant de mesurer les 2 activités luciférase. Les mesures des 2 activités luciférase ont été réalisées par ajout de 50 µL d'extrait protéique cellulaire total à 100 µL de réactif du kit Dual-Glo™ Luciferase Assay System (Promega). Le signal a été mesuré à l'aide d'un luminomètre (GloMax Multi Plus Detection System, Promega).

ANNEXE 1 : PUBLICATION 1

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

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

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

Methods 137 (2018) 3–10

Contents lists available at ScienceDirect

Methods

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

Tracking the m^{7}G-cap during translation initiation by crosslinking methods

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

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

ARTICLE INFO

Article history:
Received 15 October 2017
Received in revised form 20 December 2017
Accepted 22 December 2017
Available online 4 January 2018

Keywords:
Cap-dependent translation
UV crosslinking
Chemical crosslinking
Ribosome
Histone H4 mRNA
eIF4E

ABSTRACT

In eukaryotes, cap-dependent translation initiation is a sophisticated process that requires numerous *trans*-acting factors, the eukaryotic Initiation Factors (eIFs). Their main function is to assist the ribosome for accurate AUG start codon recognition. The whole process requires a 5′-3′ scanning step and is therefore highly dynamic. Therefore translation requires a complex interplay between eIFs through assembly/release cycles. Here, we describe an original approach to assess the dynamic features of translation initiation. The principle is to use the m^{7}Gcap located at the 5′ extremity of mRNAs as a tracker to monitor RNA and protein components that are in its vicinity. Cap-binding molecules are trapped by chemical and UV crosslinking. The combination of cap crosslinking methods in cell-free translation systems with the use of specific translation inhibitors for different steps such as edeine, GMP-PNP or cycloheximide allowed assessing the cap fate during eukaryotic translation. Here, we followed the position of the cap in the histone H4 mRNA and the cap binding proteins during H4 mRNA translation.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Translation initiation leads to the assembly of an elongation-competent 80S ribosome on the genuine AUG start codon. In eukaryotes, this is achieved through assembly of a 43S pre-initiation complex containing the small ribosomal subunit 40S on the 5′ m^{7}Gcap of mRNA. Next, the so-called 43S slides on the 5′UTR, by a 5′-3′ nucleotide-by-nucleotide inspection that is an ATP-consuming mechanism termed scanning, until the AUG start codon is found [1]. Then, the large ribosomal subunit 60S joins in order to assemble a complete 80S ribosome on the start codon. This sophisticated process is assisted by numerous *trans*-acting factors called eukaryotic Initiation Factors (eIF) [2]. The whole process is highly dynamic with many transient complexes that are difficult to trap. The interplay between the eIF is rather complex and methods to assess the numerous assembly/release cycles of individual players are rare.

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

The position of the m^{7}G-cap during the whole translation initiation process is of special interest. Indeed, numerous studies have demonstrated that the cap first initiates the scanning process by recruiting the whole machinery at the 5′ extremity of mRNA. However, the positioning of the cap during the scanning process, after 80S formation on the AUG and later on during elongation is so far unexplored. Here, we describe two crosslinking methods using the cap as a bait to trap partner molecules in order to assess the cap positioning and its fate during the whole translation process. In this study, we used the mouse histone H4 mRNA as a model to follow the cap fate during the whole initiation process. This particular mRNA is part of the cell cycle-dependent histone mRNAs which are

Abbreviation: U, units.

* Corresponding authors.
E-mail addresses: g.eriani@ibmc-cnrs.unistra.fr (G. Eriani), f.martin@ibmc-cnrs.unistra.fr (F. Martin).

https://doi.org/10.1016/j.ymeth.2017.12.019
1046-2023/© 2017 Elsevier Inc. All rights reserved.

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

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

2. Material and methods

2.1. In vitro transcription

Synthesis of RNAs containing thio-modified nucleosides can be performed with the T7 RNA polymerase system. The gene encoding mouse histone H4 (HIST1H4C) and the human β-globin mRNA have been previously cloned into pUC19 and YpGlo respectively. Transcription templates are synthesized by PCR amplification from the plasmids. The 5′ primers contain the T7 promoter sequence and the 3′ primers are designed to promote *in vitro* run-off transcription at the desired position. The PCR approach is also used to synthesize transcripts of H4 mRNA with truncated 5′ sequences. For that, new 5′ primers containing the T7 promoter sequence connected with the desired H4 sequences are designed to synthesize the PCR fragments with variable sizes. *In vitro* transcription is performed as follows. PCR DNA template (about 20 μg) is mixed with transcription buffer (40 mM Tris-HCl, pH 8.1, 20 mM $MgCl_2$, 5 mM DTT), 1 mM each ATP, CTP, GTP, UTP, 40 U of RNase inhibitor (RNasin-Promega), 0.2 mg/mL of T7 RNA polymerase, and milliQ water to 100 μL. The reaction mixture is incubated at 37 °C for 1 h. Then, 2 μL of pyrophosphatase (10108987001-Sigma) are added and the mix is further incubated for 30 min. Then, the DNA template is degraded by addition of 2 μL of DNase I RNase-free (04716728001-Sigma) and incubation for 1 h at 37 °C. Usually, to check RNA integrity, an aliquot is mixed with Formamide Dye and loaded on urea-denaturing 4% PAGE and visualized under UV light after ethidium bromide staining. To eliminate unincorporated nucleotides, the remaining RNA sample is loaded on a G25 column. Transcripts are further purified by separation on denaturing 4% PAGE and electro-elution from gel slices using a Biotrap apparatus (Schleicher and Schuell). Purified RNA samples are then phenol extracted to eliminate proteins from the synthesized transcripts. The concentration of purified RNA sample is determined by absorbance measurement at 260 nm. Before use, transcripts are folded in water by incubating at 80 °C for 2 min followed by slow cooling to 35 °C and kept on ice.

2.2. Capping of RNA transcripts

RNA transcripts can be m^7G-capped co-transcriptionally by T7 RNA polymerase in the presence of m^7GpppG cap analogue (New England Biolabs). More efficient capping is obtained with the vaccine-capping enzyme (VCE) using the ScriptCap™ m^7G Capping System (Epicentre Biotechnologies) which caps and methylates the 5′ end of mRNAs to nearly completion (Fig. 1A). The ScriptCap™ m^7G Capping System requires purified mRNA (purified by electro-elution, see 2.1.). When the GTP from the kit is substituted by s^6GTP (Jena Biosciences, 1 mM), a m^7s^6G-capped mRNA suitable for photo-crosslinking studies is synthesized by VCE after a 1-h incubation at 37 °C (Fig. 1B). After capping, enzymes used for the reaction are eliminated by phenol-chloroform extraction and the m^7s^6G-capped mRNAs are precipitated with ethanol in the presence of 250 mM NaCl. The capped mRNAs are recovered by centrifugation; the pellets are dried and then resuspended in autoclaved milliQ water.

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

Two pmoles of *in vitro* synthesized and capped m^7s^6G-histone H4 mRNA are dissolved in water, refolded at 95 °C for 2 min and then incubated at 0 °C, then are irradiated with 312 nm for 30 min at a distance of 8 cm from the bulb using a Bio-Link BLX 312 on ice as previously described [6]. Importantly, previous experiments with histone H4 mRNA demonstrated that it has to be refolded in water to favour the optimal conformation to visualize the UV-crosslink of the cap in its cap binding pocket. However, for other mRNAs, it might be critical to first set-up an optimal refolding protocol that may require Mg^{2+} to reach the optimal conformation. Then, crosslinks are detected by primer extension using fluorescently labelled H4-specific primers by reverse transcriptase. Reverse transcription is performed in final volume of 20 μL. First, 2 pmol of cross-linked mRNA are mixed with 2 pmol of fluorescently (VIC or NED, from Integrated DNA Technologies) labelled primer (H4-rev160: 5′-TGAGGCCGGAGATGCGCTTC-3′) and denatured first at 95 °C for 2 min followed by primer annealing at 65 °C for 5 min and incubation on ice for 2 min. Then, 0.2 U AMV reverse transcriptase (Life Sciences Advanced Technologies Inc.) in its buffer as recommended by the manufacturer is added and supplemented with 1.5 mM of each dNTP. Reverse Transcription extension parameters are: 42 °C for 20 min, 50 °C for 30 min, 60 °C for 10 min. Sequencing reactions are performed in parallel in similar conditions, but containing 0.5 mM of each dNTP and 16.6 μM of ddCTP. Reactions are stopped by phenol-chloroform extraction and ethanol precipitation. Samples are resuspended in 10 μL deionized formamide and loaded on a 96-well plate for sequencing on an Applied Biosystems 3130xl genetic analyzer. The resulting electropherograms are analysed using QuSHAPE [8], which aligns signal within and across capillaries, as well as to the dideoxy references, and corrects for signal decay. Normalized reactivities range from 0 to ∼2, with 1.0–1.2 being the average reactivity for highly reactive positions. The average reactivity (∗.shape file) is calculated for each H4 mRNA transcript from three independent experiments with Pearson correlation coefficients ≥0.9.

2.4. 5′-end radioactive labelling of RNA transcripts

For chemical crosslinking studies, mRNA transcripts (80 pmol-10 μg) are first 5′-end-labelled. For that, 50 μCi of [α-^{32}P]-GTP (6000 Ci/mmol, 10 mCi/mL, Hartmann Analytic) is used instead of unlabelled GTP during the capping reaction by the ScriptCap™ m^7G Capping System (Fig. 1C). After radioactive capping, capped mRNA transcripts are separated from unincorporated GTP by denaturing PAGE (4%) and eluted overnight from gel slice at 4 °C in

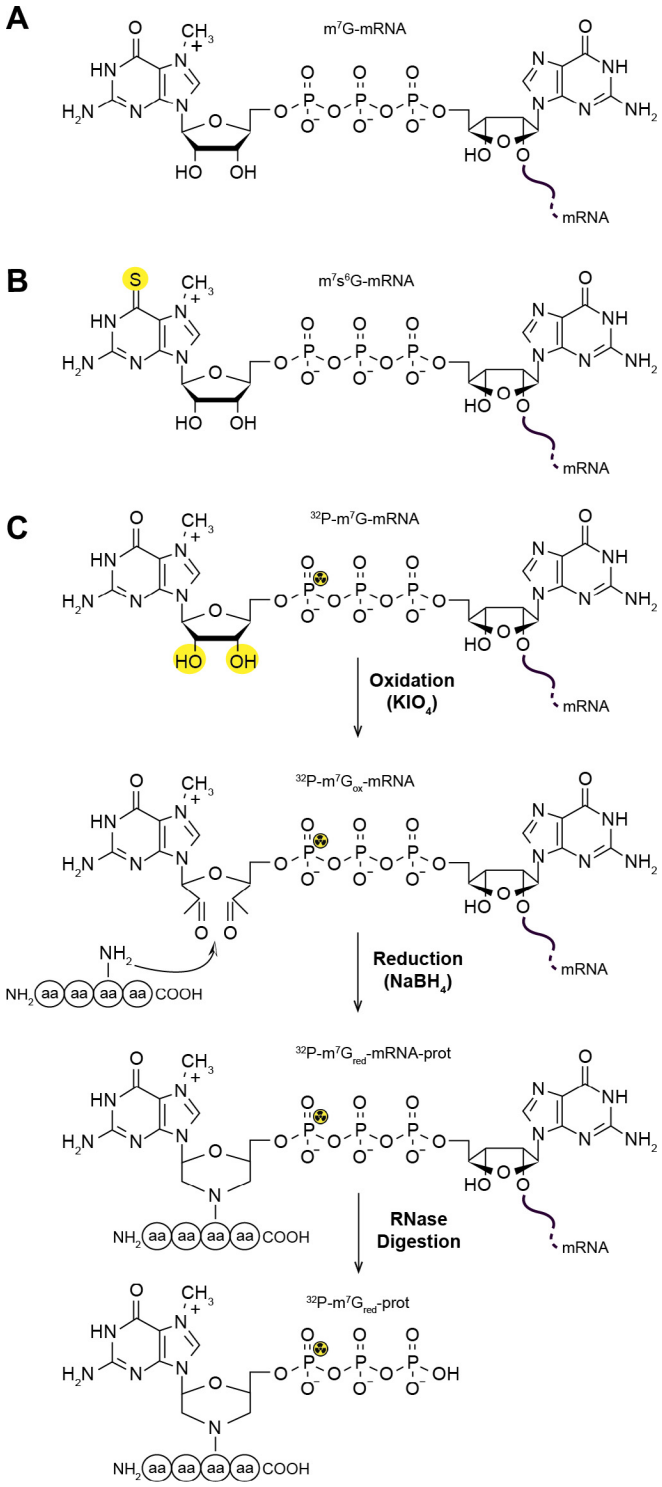


Fig. 1. Capped mRNAs used in this study. A. Structure of native m^7G-capped mRNA. B. Structure of m^7s^6G-mRNA that is site-selectively labelled with 6-thioguanosine (in yellow) within the cap structure. C. Structure of radiolabelled m^7G-capped mRNA that has been use for chemical crosslinking. The 2′ and 3′ OH (in yellow) are first oxidised by KIO_4 and then reduced by $NaBH_4$. This enables a covalent link between a NH_2 group from the lateral chain of a protein that is in a close proximity. The RNA moiety is then degraded by RNase treatment leaving a ^{32}P labelled protein that is further analysed by SDS-PAGE. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

buffer (0.3 M NaCl, 0.5 mM EDTA, 10 mM Tris-HCl pH 7.5). After passive elution, the radioactive m^7G-capped mRNA are precipitated with ethanol in the presence of 250 mM NaCl. The radioactive capped mRNAs are recovered by centrifugation, the pellets are dried and then resuspended in autoclaved milliQ water. Global radioactivity of the mRNA samples are measured by Cerenkov counting and adjusted to 50,000 cpm per μL.

2.5. Chemical crosslinking of the 5′ cap to ribosome initiation complexes

Chemical crosslinking of the 5′ cap of mRNAs is a powerful approach to study RNA binding proteins such as initiation factors. The method was originally used to identify the cap-binding protein eIF4E [9]. The target mRNA containing a radioactively labelled cap, $m^7[^{32}P]G(5')pppG$, is oxidised with sodium periodate to convert the 2′,3′-cis-diol of the 5′-terminal m^7G to a reactive dialdehyde (it is worth noting here that the ribose at the 3′ end of the mRNA is similarly oxidised in a dialdehyde) (Fig. 1C). Then, the oxidised mRNA is incubated in cell-free protein-synthesis systems from rabbit reticulocyte lysate (RRL), or wheat germ extract (WGE) or alternatively with purified components. The resulting mRNA-ribosome initiation complexes are treated with $NaBH_4$ to reduce and stabilize the Schiff bases between mRNA 5′ termini and amino groups of neighbouring proteins, yielding to covalently linked protein-RNA conjugates. The mRNA is then further trimmed by RNase A (Roche Diagnostics) thereby leaving the radiolabelled cap covalently linked to its protein partner(s). Cell-free extracts or protein components are then separated by SDS-PAGE and radioactive crosslinked proteins are revealed by autoradiography. Usually, only polypeptides positioned near the 5′ end are crosslinked to the mRNA, making this approach very useful for studying the spatial relationships between molecules in ribosomal complexes or other ribonucleoprotein (RNP) complexes that are close to the cap. In the example of histone H4 mRNA shown here, the periodate-oxidised mRNA is assembled with initiation complexes in RRL in the presence of various translation inhibitors [10]. Edeine is binding in the mRNA tunnel of the small ribosomal subunit and interferes with the codon-anticodon interaction between initiator $Met\text{-}tRNA_i^{Met}$ and the AUG start codon in the P site. This means that edeine treatment impairs formation of the 48S complex. In other words, edeine treatment enable the accumulation of translation initiation complex prior AUG codon recognition. GMP-PNP is a non-hydrolysable GTP analogue that prevents ribosomal subunit joining by inhibiting the GTPase activity of eIF5B subunit-joining factor [11]. The use of GMP-PNP in cell-free extracts allows the accumulation of 48S particles (small ribosomal subunits with a $Met\text{-}tRNA_i^{Met}$ assembled on the AUG initiation codon). Finally, cycloheximide induces accumulation of 80S particles by blocking translation at the translocation step. In fact, cycloheximide allows one complete round of translocation to proceed before inhibiting further elongation. The inhibitory effect results from binding of both cycloheximide and deacylated tRNA to the E-site, which leads to the arrest of the ribosome on the second codon [12]. Cap-specific crosslinks are validated by competition with an excess of m^7GDP. Experimentally, 5 μg of purified and radiolabelled (250,000 cpm) capped H4 mRNA are incubated for 2–3 h at 0 °C in 250 μL of 100 mM sodium acetate, pH 5.3, 10 mM EDTA, 0.2 mM sodium periodate. Then, glycerol is added to 2% (v/v) final concentration for neutralisation. After 10 min incubation at room temperature, the mixture is phenol extracted twice and ethanol precipitated. The RNA pellet is dissolved in 10 μL of water. One μL of oxidised RNA is incubated with 2.5 μL of RRL, in 10 mM HEPES-KOH, pH7.6, 1 mM ATP, 75 mM KCl, 1 mM DTT, 1 mM $Mg(Ac)_2$ and 1 mg/mL cycloheximide, or 2 mM GMP-PNP or 10 μM edeine or 1 mM m^7GDP in a final volume of 10 μL. After 10 min incubation at 30 °C, 1 μL of 0.2 M $NaBH_4$ is added and incubation is pursued for 2–3 h at 0 °C. Then, RNA is digested by 1 μL of RNase A (10 mg/mL, ThermoFisher Scientific-EN0531) for 30 min at 37 °C and samples are fractionated by SDS-PAGE.

Radioactive bands correspond to proteins crosslinked to the 5′ m^7G-cap of the mRNA are revealed by autoradiography. The specificity of the crosslink is further confirmed by adding the cap analogue m^7GDP that competes with the m^7G-cap and clears the cross-link. Protein identification can be further performed by immuno-precipitation and/or mass spectrometry analysis.

2.6. Sucrose gradient analysis of crosslinked initiation complexes

The cap crosslinking capacity in RRL was further analysed after sucrose gradient fractionation. The protocol used to prepare the complexes with 5′-labelled capped mRNA was slightly modified compared to the one described in Section 2.5. Prior the formation of 80S/H4 initiation complexes, 30 μL of untreated RRL were incubated for 5 min at 30 °C and 20 min in ice in a buffer containing 1 μL of RNase inhibitor (40U)(RNasin, Promega-N2511), 75 mM KCl and 0.5 mM $MgCl_2$. In order to lock the ribosome at translation initiation, RRL were pre-incubated for 3 min at 30 °C and 20 min at 0 °C in the presence of 10 μM edeine or 2 mM GMP-PNP or 1 mg/mL cycloheximide or without inhibitor. Finally, formation of initiation complexes was obtained by adding the oxidised histone H4 mRNA to a final concentration of 500 nM and incubating for 5 min at 30 °C. Then $NaBH_4$ was added to a final concentration of 20 mM and reduction was performed at 0 °C for 1 h. The resulting RRL programed with H4 mRNA were separated on 7–47% linear sucrose gradients in buffer (25 mM Tris-HCl, pH 7.4, 75 mM KCl, 0.5 mM $MgCl_2$, 1 mM DTT, 1 mM cycloheximide). The reactions were loaded on the gradients and centrifuged (37,000 rpm for 2.5 h at 4 °C) in a SW41Ti rotor. Gradients are fractionated on a Biocomp gradient fractionator and fractions were analysed by Cerenkov counting in order to identify the fractions containing translation initiation complexes. Indeed, we demonstrated that the stability of translation initiation complexes is not affected by the 1 h-incubation at 0 °C after reduction as attested by the radioactive profiles. Then the fractions (300 μL) corresponding to RNPs, 80S and polysomes were RNase A treated as described in Section 2.5 for 30 min at 37 °C before being precipitated by TCA (12.5% final concentration) for 30 min at 0 °C. Precipitated proteins were recovered by 15 min centrifugation at 15000 × g, washed by 20 μL acetone, dried and resuspended in 10 μL of loading blue (400 mM Tris-HCl pH 8.8, 4% SDS, 4% βME, 30% glycerol, 0,05% bromophenol blue). Proteins were fractionated by SDS-PAGE and radioactive bands were detected by photostimulable phosphor plate (Fuji Imagine plates). For reproducibility, we noticed that the use of freshly oxidised mRNA is critical.

3. Results and discussion

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

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

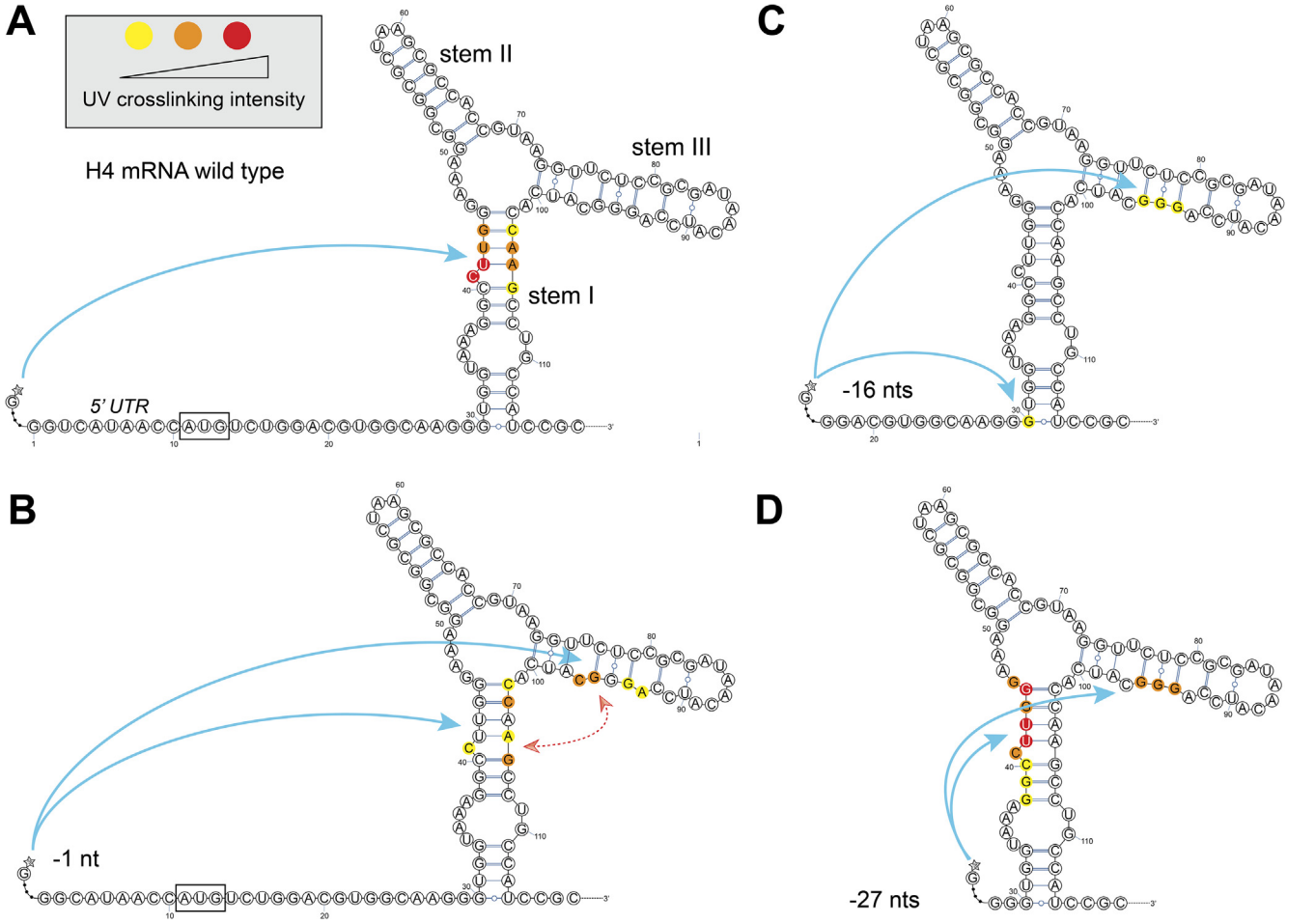


Fig. 2. UV-crosslinking of m^7s^6G-H4 mRNA. A. Secondary structure of the 114 first nts of histone H4 mRNA and interaction with the m^7s^6G cap structure. The structure contains three helical stems (I-III) connected by a three-way helix junction (TWJ). The initiation codon is boxed. Crosslink experiments on m^7s^6G-capped wild-type H4 mRNA. Crosslinks were performed at 360 nm and revealed by reverse transcription. Reverse transcription stops are indicated by colours according to the intensity scale shown in the inset. B-D. Similar experiments performed on mutants deleted of 1, 16 and 27 first nucleotides.

thionucleotides ^{6}sG and ^{4}sU are preferred due to their simple molecular structure, relative stability, and high reactivity. In addition, ^{6}sG and ^{4}sU only differ from the corresponding G and C nucleotides by a single atomic substitution, the replacement of a nucleobase oxygen by sulphur, thus reducing the risk for perturbation of the RNA structure. Exposure of the thionucleotides to UV light produces a sulphur radical that can react very efficiently with functional groups that are in proximity (in proteins or nucleic acids). To increase crosslinking specificity, low energy (360 nm) irradiation is preferred, yielding "zero-length" crosslinks [13] that represent direct contacts (in a 2-Å sphere) of the nucleotide of interest and protein or nucleic acid. Thionucleotides can be incorporated by either random incorporation or site-specific substitution. *In vitro* transcription is a convenient way to incorporate statistically and randomly ^{6}sG and ^{4}sU. This way of incorporation allows a rapid screening of the potential interactions over the entire RNA transcript. Site-specific incorporation by transcription or RNA ligation eliminates the ambiguity of the origin of crosslinking and generally improves the experimental signal since all the RNA molecules are containing the photo agent at a single or very few positions. It allows accurate probing of any position, and generally increases the experimental signal since all the RNA molecules contain a photo agent at a single position rather than spread throughout the molecule, however it is time- and labour-consuming since it requires the synthesis of site-specifically modified RNA. A simplified variant of this approach consists in the synthesis of short RNA fragments containing a single G or U residue that is substituted by ^{6}sG or ^{4}sU during *in vitro* transcription. For instance a set of mRNAs that contained single ^{4}sU at unique positions between −26 and +11, relative to the A of the initiation codon, were synthesized to explore mRNA contacts in the ribosome mRNA channel during translation initiation [14]. Similarly, TISU mRNA fragments containing single unique ^{4}sU between nts −4 and +8 were synthesized to identify ribosomal proteins interacting with TISU sequence [15].

Here we describe another method of site-specific incorporation of photo agent in the m^{7}G cap at the 5′ end of any mRNA. The thionucleotide ^{6}sG has been used for mapping the cap-binding pocket in the mouse histone H4 mRNA [5]. Indeed, the m^{7}G-cap of eukaryotic mRNAs is a modified guanine nucleotide connected to the mRNA via an unusual 5′ to 5′ triphosphate linkage. *In vivo*, it is added co-transcriptionally to the mRNA. *In vitro*, the modified cap is added post-transcriptionally by the Vaccine Capping Enzyme (VCE) using the ScriptCap™ m^{7}G Capping System (Epicentre Biotechnologies) (see Section 2.2). During this process, ^{6}sGTP is substituted to the usual GTP, it is subsequently modified to form a m^{7}s^{6}G-cap structure. In that form, the m^{7}s^{6}G-cap structure is an effective structural probe that combines high crosslinking efficiency, site-specific incorporation with high level of incorporation. Then, the modified mRNA is subjected to UV at 312 nm exposure and detection of the crosslinks is subsequently performed by primer extension with [^{32}P] labelled oligonucleotides and polyacrylamide gel separation as described previously [6,5] or by a an alternative protocol that we are here describing in Section 2.3.

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

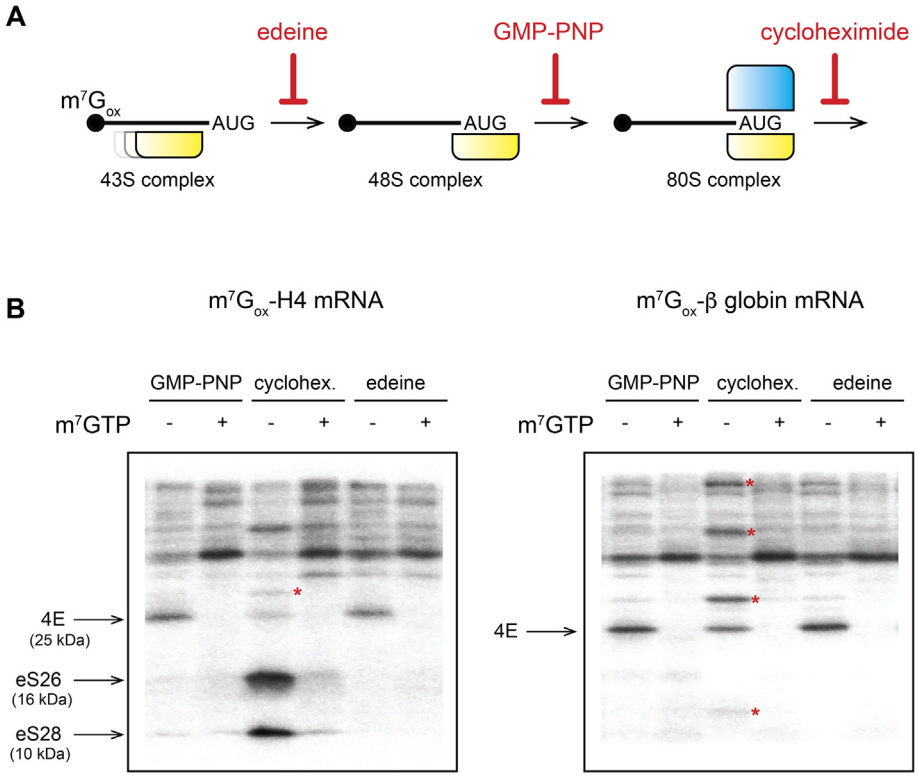


Fig. 3. Chemical crosslinking of oxidised histone H4 and β-globin mRNA. A. Cap-dependent translation initiation is inhibited by specific inhibitors (shown in red). Edeine prevents codon-anticodon interaction of Met-tRNA$_{i}^{Met}$ on AUG codon and thereby leads to 43S accumulation. GMP-PNP, is a non-hydrolysable GTP analogue that blocks eIF5B GTPase activity, it leads to 48S accumulation on the AUG start codon. Cycloheximide inhibits translocation after 60S joining and triggers 80S formation on the AUG start codon. B. Chemical crosslinking of ^{32}P-labelled oxidised histone H4 mRNA and β-globin mRNA in RRL in the presence of GMP-PNP, cycloheximide, edeine with and without m^{7}GDP. Characterized cap-specific crosslinks are indicated by arrows. Uncharacterized cap-specific crosslinks are indicated by red asterisks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the 5′ UTR and checked whether this changed the cap-crosslinking capacity. The truncated mRNA fragments were synthesized starting from shortened truncated H4 genes and capped as the wild-type mRNA with a m^7s^6G-cap structure using the ScriptCap™ m^7G Capping System. One nt deletion induced a severe effect with a near disappearance of the crosslinks of nts 41–44 and new crosslinks at nts 93–94, 96–97 whereas crosslinks at nts 102–103 and 105–106 are maintained. This suggests that nts 93–97 are spatially close of the stem I in the TWJ and can both react with mutant −1 nt of H4 mRNA. Deletions −2 to −8 nts show that the strong crosslink at positions 41–42 is definitively lost and confirm the vicinity of nts 93–97 and nts 103–106 (Supplementary Fig. 1).

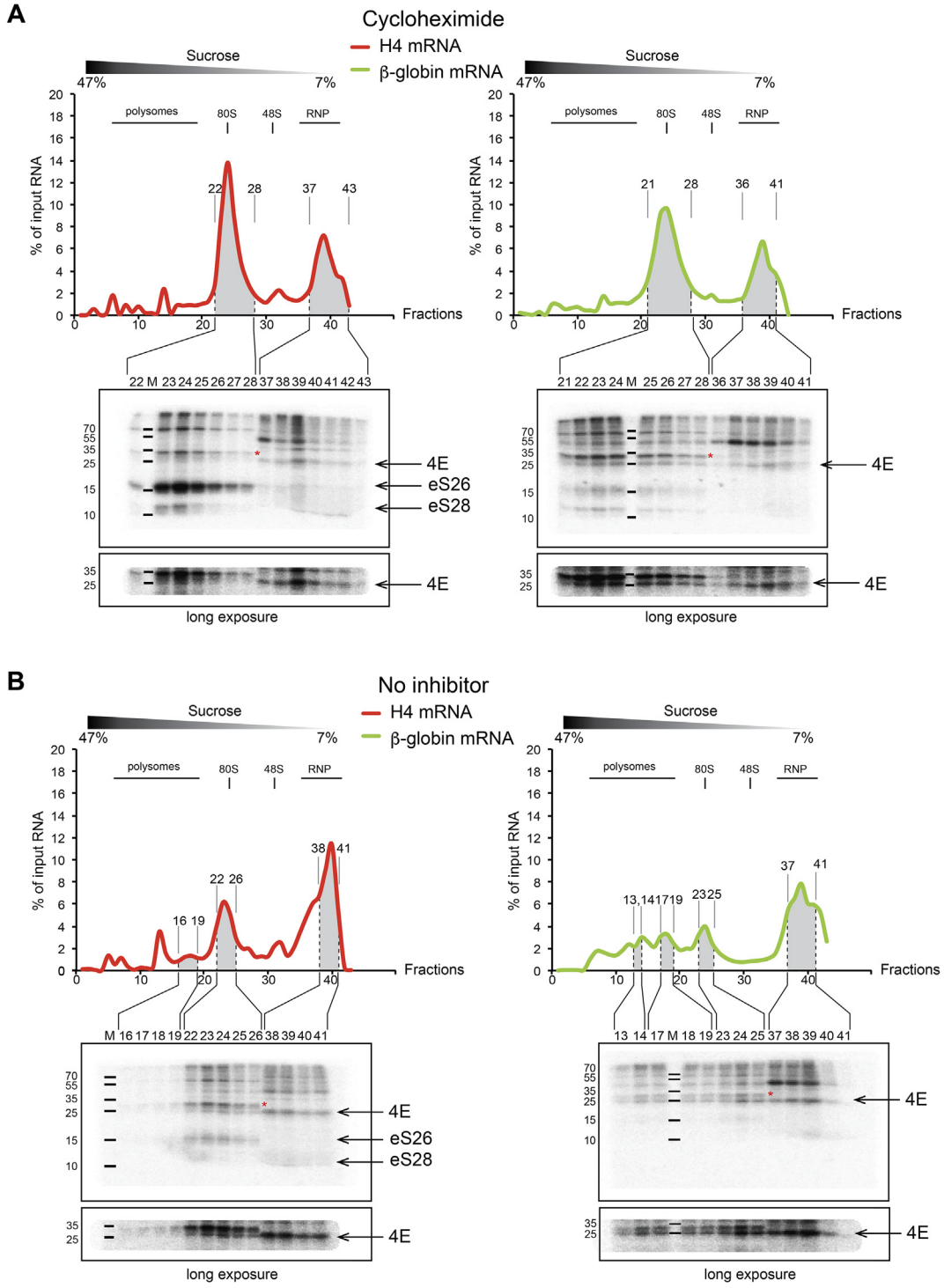


Fig. 4. Chemical crosslinking of oxidised histone H4 and β-globin mRNA and sucrose fractionation. A. Sucrose gradient fractionation of cycloheximide-blocked translation in RRL. The percentage of input mRNA is measured by Cerenkov counting of the sucrose gradient fractions. Proteins contained in fractions that are shown in grey are precipitated and analysed by SDS-PAGE. Characterized cap-specific crosslinks are indicated by arrows; red asterisks indicate uncharacterized cap-specific crosslinks. B. Sucrose gradient fractionation of translation in RRL. The percentage of input mRNA is measured by Cerenkov counting of the sucrose gradient fractions. Proteins contained in fractions that are shown in grey are precipitated and analysed by SDS-PAGE. Characterized cap-specific crosslinks are indicated by black arrows, uncharacterized cap-specific crosslinks are indicated by red asterisks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Most of these mutants exhibit strong crosslink reactivity around nts 104–105. After a 3 nt-deletion, the mutants exhibit different reactivities spread over nts 13 and 30 that indicate that short UTRs can react with closer regions of the mRNA structure (Supplementary Fig. 1). Following these progressive deletions of the UTR, we performed six additional deletions that brought the reactive cap at positions 16, 22, 27, 29, 31 and 37 of H4 mRNA. Except deletion 27, all the mutants lose their ability to react with the original cap-binding site in the pocket 41–44 and 104–106. With these mutants, new positions are crosslinked, providing new data on their position versus the modified cap. Surprisingly, mutant −27 which has the cap at the bottom of stem I, recovered the strong crosslinking activity of nts 41–46 but no reactivity on the other strand of the stem. This indicates that the cap of mutant −27 does not bind the CBP in the native manner since it does not react as the wild-type mRNA does with nts 103–106. The low energy crosslinking experiments might be useful to detect close interactions such as the interaction of the m^7Gcap in the cap-binding pocket of H4 mRNA. In the H4 case, the specificity of the binding pocket was demonstrated by performing a competition assay with m^7GpppG cap analogue [5]. Deletion experiments of nts of the 5′ UTR revealed the spatial proximity of the CBP of stem I and stem III where recurrent crosslinks were found. One major drawback of the deletion approach may concern either the alteration of the folding of the mRNA structure or the severe reduction of reactivity observed with several mutants. For instance, we repetitively observed low reactivity of several independent preparations of mutants −5, −16 and −22 nts. Whether this is due to an unusual high flexibility of the RNA sequence harbouring the 5′ cap resulting in a higher exposure to the solvent side remains to be explored.

3.2. Chemical crosslinking of the cap in RRL

The next issue we addressed was the position of the cap in H4 mRNA during the translation initiation process. For that purpose, we performed chemical cross-linking experiments using mRNA that contain a radiolabelled m^7Gcap that has been oxidised. This histone H4 mRNA transcript was then incubated in RRL in the presence of edeine, GMP-PNP or cycloheximide and then further reduced to induce covalent bonds with proteins that are in contact with the cap. As a control, we performed the same experiment with radioactive oxidised capped β-globin mRNA. Crosslinking experiments in the presence of an excess of m^7GDP as a competitor allowed us to discriminate cap-specific crosslinks that are abrogated versus unspecific bonds that are still present with the competitor (Fig. 3). In the case of H4 mRNA, we detected specific crosslinks of eIF4E, eS26 and eS28 that were identified with specific antibodies as previously described [7]. The crosslinks of eIF4E are detected with edeine and GMP-PNP, but are drastically decreased with cycloheximide. This suggests that eIF4E is bound on the cap on the 43S and 48S complexes but is released from the 80S complex. On the contrary, eIF4E is bound on the cap on β-globin mRNA and remained bound after 80S assembly. Concerning eS26 and eS28, no specific crosslinks are detected on β-globin mRNA. Altogether, these results indicate that in the case of H4 mRNA translation, eIF4E is bound on the cap prior and after 48S complex formation, then eIF4E is released from the cap and is replaced by eS26 and eS28 in the 80S complex. In contrast, eIF4E stayed bound on the cap of β-globin mRNA during the whole translation initiation process. However, the major drawback of this experiment is that the whole input oxidised mRNA is most probably not fully assembled in translation initiation complexes suggesting that a part of the oxidised mRNA is in fact crosslinked with soluble proteins that are also visualised here as attested by the radioactive count profiles (Fig. 4). In other words, the detected crosslinks could also be due to RNPs components in contact with the cap rather than ribosomal complex proteins.

3.3. Translation initiation step-resolved chemical crosslinking of the cap

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

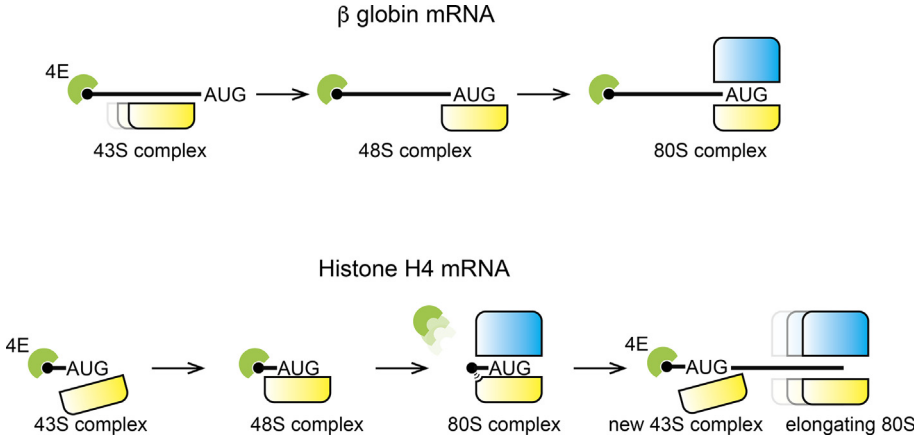


Fig. 5. Model of the cap fate during translation of histone H4 mRNA and β-globin mRNA. During β-globin mRNA translation, eIF4E remains bound on the cap during the whole process. On the contrary, H4 mRNA translation requires eIF4E binding/release cycles to allow efficient ribosome assembly.

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

4. Conclusion

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

Acknowledgments

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

Appendix A. Supplementary data

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

References

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

ANNEXE 2 : PUBLICATION 2

Immunoprecipitation methods to isolate messenger ribonucleoprotein complexes (mRNP).

Hayek, H., Gross, L., Alghoul, F., Martin, F., Eriani, G., and Allmang, C.

***Book chapter** : In Advanced Technologies for Protein Complex Production and Characterization. Editors: Vega, M. Cristina Springer International Publishing Switzerland **(in press**)*

Chapter 2. Immunoprecipitation methods to isolate messenger ribonucleoprotein complexes (mRNP)

Hassan Hayek, Lauriane Gross, Fatima Alghoul, Franck Martin, Gilbert Eriani and Christine Allmang*

Architecture et Réactivité de l'ARN, Université de Strasbourg, Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 67084 Strasbourg, France

*Correspondence to c.allmang@unistra.fr

Running title: IP methods to isolate mRNA complexes

Abstract Throughout their life cycle, messenger RNAs (mRNAs) associate with proteins to form ribonucleoproteins (mRNPs). Each mRNA is part of multiple successive mRNP complexes that participate in their biogenesis, cellular localization, translation and decay. The dynamic composition of mRNP complexes and their structural remodelling play crucial roles in the control of gene expression. Studying the endogenous composition of different mRNP complexes is a major challenge. In this chapter, we describe the variety of protein-centric immunoprecipitation methods available for the identification of mRNP complexes and the requirements for their experimental settings.

Keywords Protein-RNA interactions, RNA immunoprecipitation, Crosslinking, messenger ribonucleoprotein complexes.

2.1 Introduction

Expression of protein-coding genes in eukaryotes is a complex and coordinated mechanism that involves many steps and relies on multiple factors and several molecular machineries. Cellular RNAs are associated with RNA-binding proteins (RBPs) to form ribonucleoprotein (RNP) complexes of highly dynamic compositions [1–2]. Throughout their life cycle, mRNAs are part of multiple successive mRNP complexes that play critical roles in their biogenesis, cellular localization, translation and decay [3–4]. The dynamic composition of mRNP complexes allows both post-transcriptional and specific translational regulation mechanisms to take place [5–6]. Some trans-acting factors can interact stably with specific mRNPs but structural remodelling of mRNPs also involves transient interactions with protein and RNA chaperone complexes and numerous regulatory factors that are difficult to trap [7–9]. Studying the endogenous composition of different mRNP complexes is therefore a major challenge. The advent of new and high-throughput approaches has enabled the genome-wide determination of mRNP composition using techniques such as affinity purification-mass spectrometry (MS) as well as protein–RNA UV crosslinking approaches combined with RNA deep-sequencing (RIP-seq) [10–12]. Two types of approaches to study RNA–protein complexes have emerged: protein-centric and RNA-centric approaches. In protein-centric approaches, mRNAs associated with RBPs are identified after the immunoprecipitation of the RBPs by microarray analysis or RNA-seq [13–15]. In RNA-centric approaches, RBPs are recovered by RNA pull-down methods and subsequently identified by mass spectrometry (MS). Several independent laboratories have developed breakthrough methods for the

mRNA interactome discovery. Based on protein-mRNA interactome capture protocols these studies now allow genome wide cartography of mRNPs in yeast, mammalian, invertebrates or plant cells, under various conditions [16–20]. These unbiased methods have emerged as powerful tools to identify new RBPs; they nevertheless need to be complemented by individual mRNA/protein detection methods in order to map pairwise interactions between the RBPs and their mRNA targets and to identify RBPs with overlapping target specificity. Some mRNA centric methods are limited to the study of polyadenylated mRNAs and thus exclude mRNA categories such as histone mRNAs that lack poly-A tails. Global analysis can also oversee low abundance mRNAs or mRNAs containing non-canonical codons subjected to translation recoding. Targeted mRNA immunoprecipitation experiments are particularly useful in these cases. In this chapter, we will focus on standard protein-centric immunoprecipitation methods that allow the identification of native mRNPs containing specific mRNAs of interest. These methods can be used as initial targeted approaches or to confirm and consolidate high throughput or global analysis results. We have applied the panel of immunoprecipitation methods described herein to analyse the implication of various chaperone complexes and modification enzymes that act transiently in the assembly and biogenesis mechanisms of selenoprotein mRNAs as well as the role of translation initiation complexes in histone H4 mRNA translation.

2.2 Endogenous mRNP purification using RNA immunoprecipitation methods

mRNA form complexes with partner proteins in ribonucleoprotein particles (mRNP) from the onset of their transcription. Their assembly, transport and subcellular localization and degradation involve interactions with numerous intermediate protein complexes. Protein-centric immunoprecipitation methods allow to unravel the fate of mRNAs and their uptake by different known protein complexes. Various methods have been developed for the recovery of endogenous mRNPs under native physiological conditions or after stabilization of the interactions by chemical or UV cross-linking (Table 2.1).

2.2.1 Native immunopurification

Native RNA immunoprecipitation (RIP) allows the identification of RNA–protein interactions under physiological conditions by using a protein-specific antibody and the detection of interacting RNAs. mRNA interactants can be identified by targeted-RNA detection methods like quantitative Real Time-PCR (qRT-PCR), northern blot but also by genome-wide methods such as microarray (RIP-chip) or deep-sequencing (RIP-seq). Typical RIP experiments for the study of eukaryotic mRNPs are outlined in Figure 2.1A, mainly applying to HEK293 or HeLa cells. These methods preserve native complexes present in the cells and most often reflect *in vivo* associations. Nevertheless it was demonstrated that reassociation of RNA-binding proteins can occur after cell lysis [21]. Interactions that were prevented *in vivo* due to differential compartmentalization might occur during the lysis and immunoprecipitation experiment. Several protocols have therefore been optimized to minimize RNP rearrangements during the process of immunoprecipitation [13, 22]. For instance, purification of cytoplasmic mRNPs requires mild lysis conditions in order to leave nuclei intact [23–25]. Specific conditions for nuclear mRNP have also been developed [26, 27]. Another important issue is the specificity of the interactions detected by these methods. Indeed, specific interactions that occur with low abundance mRNAs *in vivo* could be masked by the non-specific interactions of abundant transcripts. Several precautions need to be taken in the experimental procedures to avoid this (see

below).

2.2.1.1 Key requirements for RIP

To prepare high-quality extracts that contain intact pre-mRNPs or mature mRNPs several factors need to be optimized. In particular extracts have to be kept cold and snap-frozen immediately after lysis and centrifugation. To limit degradation, inhibitors of proteases and ribonucleases (RNases) have to be added at all steps. Extracts should be diluted only after thawing and during the immunoprecipitation steps. The total amount of cell extract used in RIP experiments needs to be adapted based on the abundance of the RNA-binding protein or complex targeted by the antibodies as well as the subsequent method of RNA detection. Typically, extracts generated from 3-5.10^6 mammalian cells will be required when the RNA is detected by qRT-PCR or 5-20.10^6 when using microarrays.

It is essential to limit potential nonspecific interactions of the mRNAs of interest with the immunoprecipitation matrix. Various affinity matrices and supports can be used successfully for the recovery of mRNP complexes after incubation of the extracts with a specific antibody. These include protein A-sepharose, agarose but also wide range of magnetic beads. It is advisable to compare different matrices because the nonspecific binding of a given mRNA can vary depending on the immunoprecipitation support. It is also recommended to pre-clear the extracts in contact with non-coated beads before immunoprecipitation and to dilute the lysate in order to remove nonspecific binders. When the targeted mRNA will subsequently be detected by methods such as qRT-PCR, pre-coating of the beads with purified BSA and yeast total tRNA can considerably reduce nonspecific background (Figure 2.1). For standard RIP, extensive washing of the immunoprecipitated pellet is required and the stringency of the washing buffers needs to be adapted to the stability of the RNP complex analysed. The mRNP is ultimately released and dissociated into RNA and proteins. This is mostly performed under denaturing conditions. mRNP affinity purification can also be performed using anti-peptide antibodies against the target protein. In this case bound mRNP complexes can be eluted in native conditions by the peptide used to generate the antibodies. This is particularly interesting when studying low abundance mRNPs. This method allowed us to co-immunoprecipitate selenoprotein mRNAs and protein complexes associated to SECIS binding protein 2 (SBP2), a key protein that interacts with all selenoprotein mRNAs and recruits translation and assembly factors to the mRNP [28, 29]. In any case, validation steps of the RIP will include verification of optimal protein pull-down by Western blot and evaluation of the integrity of protein complexes associated to the mRNP by standard mass spectrometry.

2.2.1.2 mRNA identification and validation of the RIP experiment

After release of the RNP components, the ultimate goal is to isolate and characterize the mRNA from the immunoprecipitated pellet (Figure 2.1). For proper interpretation of the data it is essential to compare the RIP result to a negative control (IP with pre-immune serum or non-RBP protein target). For standard RIP, qRT-PCR remains the method of choice to measure RNA levels as it allows to normalize the immunoprecipitated mRNA to its abundance in total lysate [30]. To determine if the interaction of given RBP can be generalized to an entire family of mRNAs the microarray analysis of the RIP content is a useful strategy. Result differences can be observed between the two detection methods, most likely because qRT-PCR is more sensitive and specific than microarrays for the detection of the endogenous population of mRNAs [31]. Deep-sequencing analysis of RIP experiments allows the genome-wide characterization of RNAs bound to a given RBP and the mapping of protein-RNA regulatory networks using appropriate bioinformatics tools [32, 33]. mRNA interactants identified using

native purification methods often require further experimental validation using multiple distinct experimental approaches. A way to confirm the validity of the mRNA interactant is to compare the RIP results obtained against an endogenous protein under normal and RNAi inhibition of the target protein [34]. Alternative immunoprecipitation approaches based on epitope tagging or cross-linking methods can be used, and they will be discussed below. Functional validation of the mRNA-protein interactions unveiled by the RIP experiment will have to include a wide panel of methodologies such as subcellular localization experiments (FISH) [35], yeast-two hybrid interaction tests (Y2H) [36], electrophoretic mobility shift assays (EMSA) [37, 38], and more.

2.2.2 Epitope tagging and RIP

The use of an epitope-tagged RBP for mRNP isolation is a useful complementary approach to confirm endogenous RIP data. It can also be used when the endogenous target protein is not detected by standard RIP because of its low cellular abundance or the inaccessibility of the epitope to the immunoprecipitating antibody (Figure 2.1B). After transfection and expression of the epitope tagged protein in cells, usually HEK293 or HeLa, mRNAs are coprecipitated using antibodies directed against the epitope tag. Some of the most commonly used epitope tags include FLAG, HA, His, Myc, GST, GFP and V5 [39, 40]. Limitations linked to the use of epitope tags are the possibility for the tag to impact the folding of the fusion protein, to disrupt functional protein complexes and to interact non-specifically with cellular proteins. Large globular tags may also affect the function and subcellular localization of the fusion protein. Among various tags the green fluorescent protein (GFP) has proven particularly efficient for affinity purification of RNP complexes and proteomic analysis, despite its size, because it shows minimal nonspecific binding to mammalian cell proteins [41]. We have used epitope tagging successfully to analyse the contribution of different chaperone complexes and modification enzymes in selenoprotein mRNA assembly and translation. This allowed us to demonstrate interactions between selenoprotein mRNAs and components of the HSP90 chaperone complex [28]. Our findings also establish that a FLAG-tagged version of trimethylguanosine synthase 1 (Tgs1) interacts with selenoprotein mRNAs for cap hypermethylation [31]. We also showed that the ubiquitous Survival of MotoNeurons (SMN) chaperone and the methylosome complex devoted to sn- and snoRNP maturation contribute to recruiting Tgs1 to selenoprotein mRNPs [29]. In this case we used GFP-tagging as a method of choice to decipher the role of individual SMN and methylosome components in selenoprotein mRNP assembly and translation.

2.2.3 Formaldehyde cross-linking and RIP

When studying RNA-protein interactions *in vivo* it is essential to minimize RNP reassortments that can occur during cell disruption. Chemical treatments can be employed for this purpose in order to cross-link proteins with their associated mRNAs prior to the cell lysis and immunoprecipitation steps (Figure 2.2A). Formaldehyde is a useful reversible crosslinking agent capable of capturing interactions between various macromolecules mainly by formation of covalent bonds between amino groups that are in close proximity (around 2 Å) [42]. It has been used extensively as a fixative to maintain the structural integrity of cells and in chromatin immunoprecipitation (ChIP) experiments to study DNA-protein interactions and cellular networks [43, 44], but also to delineate RNA-protein and protein-protein interactions [45]. Formaldehyde cross-linking is particularly useful to detect transitory interactions and to stabilize indirect RNA-binding factors that are part of higher-order RNPs. It was used successfully in various studies including the identification of RNAs binding to the CTD of RNA polymerase II,

the binding of the histone acetyltransferase Elongator to nascent mRNAs [46, 47] as well as the interaction of eukaryotic Translation Initiation Factor 3 (eIF3) with various mRNAs [48]. The success of the RIP assay depends on the degree of cross-linking. The cross-linking procedure therefore needs to be optimized to avoid false positive interactions and to achieve optimal RNA-protein complex recovery. The concentration of formaldehyde and the duration of cross-linking need to be adjusted [49]. Optimal formaldehyde concentration typically ranges between 0.1 to 1.0% and the duration of fixation varies between 5 min and 1 h. Cross-linking reactions are quenched by the addition of glycine (pH 7). Interestingly formaldehyde crosslinking can be reversed by incubation of the IP pellets at 70 °C for the characterization of immunoprecipitated components.

2.2.4 UV cross-linking and RIP

Unlike formaldehyde cross-linking, which is able to capture interactions with protein complexes, UV cross-linking exclusively identifies direct RNA-protein interactions. UV irradiation at 254 nm generates short-lived radicals that react with nucleic acids and amino acids such as lysine, cysteine, phenylalanine, tryptophan and tyrosine located in close proximity (zero distance) forming covalent bonds [50]. *In vivo* UV cross-linking thus allows trapping functional protein-RNAs interactions. Interestingly it can be associated to denaturing purification methods for the identification of the interacting RNA and has given rise to the cross-linking and immunoprecipitation (CLIP) technique [11, 12, 51, 52]. This technique can be used for targeted RNA identification but has mainly been associated to high-throughput sequencing of cDNA library (HITS-CLIP). Several powerful variants of the methodology exist such as PhotoActivatable Ribonucleoside enhanced CLIP (PAR-CLIP) and individual-nucleotide resolution CLIP (iCLIP) that allow the identification of cross-linking sites at a single-nucleotide resolution [14, 15, 53]. PAR-CLIP employs the nucleotide analogue 4-thiouridine (4-thioU), which can be added to the growth medium and is taken up by cultured cells and incorporated into newly-synthesized RNAs [15]. 4-thioU is activated by 365-nm UV irradiation, instead of UV 254 nm used in CLIP, to generate covalent crosslinks between proteins and RNAs. In both CLIP and PAR-CLIP experiments the RNA is fragmented by RNase treatment and immunoprecipitation allows the recovery of co-precipitated RNA fragments. Other variations make use of double epitope-tagged RBPs in cross-linking experiments followed by purification of the RNPs by tandem affinity purification (iCLAP) or immunoprecipitation coupled to immobilized metal-ion affinity chromatography (IMAC) in the CRAC method [54, 55]. Some *in vivo* methods combine UV cross-linking, GFP-based immunoprecipitation and quantification of co-isolated polyadenylated [poly(A)] RNAs with fluorescent oligo(DT) probes [56]. The main high-throughput methods are summarized in Figure 2.2B and Table 2.1.

2.3 Cell-free extract mRNP immunoprecipitation methods

2.3.1 GST pull-down experiments using total RNA from cell extracts

An alternative approach to endogenous pull-down is the exogenous expression of epitope-tagged recombinant proteins that can be incubated with cell extracts for the reconstitution of RNA-protein interactions. Purified recombinant proteins carrying a glutathione-S-transferase (GST) tag have been widely used as a screening technique for the identification of protein-protein interactions. GST-

tagged proteins are also efficient tools to uncover protein-mRNA interactions when incubated with total RNAs isolated from cell extracts, native extracts or rabbit reticulocyte lysates. The GST-tagged proteins can be efficiently recovered from cell extracts using either affinity matrices such as Glutathione sepharose, Glutathione magnetic beads or anti-GST antibodies. The co-precipitated RNA can be subsequently analyzed by qRT-PCR to detect the presence of specific target mRNAs using gene-specific primer pairs (Figure 2.3A and Table 2.1). Glutathione-S-transferase (GST) pull-down procedures have been developed for the purification of eukaryotic mRNAs using a mutant version of the mRNA 5′ cap-binding protein (eIF4E) with increased affinity for the m^{7}GTP moiety of the cap [57]. Using this method we could demonstrate that several selenoprotein mRNAs are not recognized efficiently by translation initiation factor eIF4E and have an alternative cap structure [31].

2.3.2 Crosslinking of total cell extracts with *in vitro* transcribed 4-thioU or 6-thioG labelled mRNAs

The 4-thiouridine PAR-CLIP approach can be adapted to study RNA-protein interactions using an exogenous *in vitro* transcribed ^{32}P-labelled mRNA containing 4-thioU nucleotides that is incubated in the presence of total cell extracts (Figure 2.3B). After UV-cross linking at 365 nm, the crosslinked RNAs is digested by RNase T1 or A [58]. Alternatively *in vitro* transcribed ^{32}P-labelled mRNA devoid of 4-thioU can also be UV cross-linked directly to proteins from cell extracts at 254 nm. Subsequent immunoprecipitation and separation by denaturing gel electrophoresis allows resolving RNase-protected ^{32}P-labelled mRNA fragments cross-linked to specific RNA binding proteins. Overlay of Western-blot and autoradiography analysis of the SDS–PAGE gel can allow the detection of pairwise protein-RNA interactions. Critical to the success of these experiments is the preparation of the *in vitro* 4-thioU labelled RNA, in particular the optimization of the 4-thioUTP:UTP ratio during transcription [59, 60]. UV crosslinking using mRNAs containing 4-thiouridine have allowed the characterization RNA–protein interactions between various mRNAs and ribosomal proteins as well as numerous translation initiation factors including eIF3, eIF4GII [61–63]. These experiments have also revealed the mRNA binding function of Gemin5, a component of the SMN chaperone complex, and its role in translation [64]. We have recently developed UV and chemical crosslinking methods specifically dedicated to track proteins bound to the 5’ cap structure of an mRNA of interest during the translation process in cell-free extracts [65]. Site-specific incorporation of 6-thioG can be achieved after *in vitro* transcription of the mRNA at the m^{7}G capping step. The modified cap is added post-transcriptionally by the Vaccine Capping Enzyme (VCE). At this stage 6-thioGTP (s6G) is substituted to the usual GTP, to generate an m^{7}s6G-cap structure. UV cross-linking can be achieved at 365 nm (Figure 2.3B). Chemical crosslinking of the 5’ cap of mRNAs is an alternative approach to study RNA binding proteins such as initiation factors, it was initially used to identify the cap-binding protein eIF4E [66]. The target mRNA transcripts containing a radioactively labelled m^7[^{32}P]G(5’)pppG cap, is oxidized with sodium periodate and incubated in cell-free extracts leading to the formation of Schiff bases between the 5’ cap and amino groups of interacting proteins [65]. Subsequent RNase A digestion allows to leave the radiolabelled cap covalently linked to its protein partners that can be identified directly by Western blot or after immunoprecipitation. These methods allowed us to follow the histone H4 mRNA cap binding proteins during H4 mRNA translation with the use of specific translation inhibitors for different steps [65, 67].

2.4 Direct mRNA IP: Anti-m^{6}A, m^{5}C and anti-TMG IP

While mRNP assembly and remodelling is essential to gene expression control, mRNA post-transcriptional modifications have emerged as an additional layer of gene expression regulation [68]. Different types of nucleotide modification have been documented in eukaryotic mRNAs. These include internal mRNA modifications such as 5-methylcytosine (m^5C) [69], and N^6-methyladenosine (m^6A) [70], that play regulatory roles, but also m^7G and $m_3^{2,2,7}G$ 5' cap modifications. Mammalian mRNAs, synthesized by RNA polymerase II (polII), are generally characterized by the presence of a 7-methylguanosine (m^7G) cap structure at their 5' end [71]. The m^7G cap required for mRNA processing, translation initiation, mRNA transport, splicing and degradation [72–77]. We showed that several selenoprotein mRNAs bear a trimethylated $m_3^{2,2,7}G$ cap (Figure 2.4A) and are thus subjected to a non-conventional translation initiation mechanism [31]. Only a limited number of mRNA immunoprecipitation methods allow to directly recover endogenous mRNAs using antibodies directed against specific nucleotide modification; these include detection of $m_3^{2,2,7}G$-cap, m^6A and m^5C (Table 2.1).

2.4.1 m^6A Me and m^5C RIP

N^6-methyladenosine (m^6A) is the most prevalent internal (non-cap) modification present in the messenger RNA of higher eukaryotes [78–80]. This RNA methylation is reversible and may dynamically control mRNA metabolism. RIP methods called anti-m^6A Me-RIP [81] or m^6A-seq [82], based on antibody-mediated capture followed by either qRT-PCR or massively parallel sequencing have been developed. They allowed the identification of m^6A modified mRNAs [83] and transcriptomes (methylome) in human cells and mouse tissues and revealed m^6A enrichments within long exons and around stop codons [84, 85], further suggesting fundamental regulatory roles of m^6A. 5-methylcytidine (m^5C) is another mRNA modification that plays a role in gene regulation [86, 87]. Anti-m^5C RIP is based on the use of monoclonal antibodies that specifically bind 5-methylcytosine [88] and that have been broadly used in pull-down experiments of modified DNA molecules. These antibodies were raised against 5-methylcytosine nucleotide conjugated to ovalbumin without the ribose or deoxyribose sugar and are blind to the DNA/RNA context. Both m^6A Me-RIP and m^5C RIP methods appear to be better suited for establishing global mRNA modification patterns rather than for the specific recovery of targeted mRNAs, and they are usually associated to other epitranscriptome characterization technologies (for a review see [89]).

2.4.2 Trimethylguanosine-capped mRNA immunoprecipitation (TMG-IP)

We have designed immunoprecipitation methods for the detection of endogenous $m_3^{2,2,7}G$-capped selenoprotein mRNAs from total cell extracts or after polysome fractionation of cytoplasmic extracts [90]. The detection of the $m_3^{2,2,7}G$-capped RNAs can be performed by immunoprecipitation experiments using a highly specific anti-$m_3^{2,2,7}G$ cap R1131 serum (or anti-TMG serum) that was demonstrated not to recognize monomethylated caps [91, 92]. On the contrary, antibodies developed against m^7G-cap recognize both m^7G and $m_3^{2,2,7}G$ modifications. Experiments based on the use of anti-TMG serum are called TMG-IP and have been used successfully to characterize the $m_3^{2,2,7}G$-capped small non-coding RNAs such as snRNAs and snoRNAs. We have adapted this assay coupled to real time quantitative PCR for the detection of endogenous $m_3^{2,2,7}G$-capped selenoprotein mRNAs *in vivo* from cultured cells (Figure 2.4B) [31, 90]. The fraction of the $m_3^{2,2,7}G$-capped selenoprotein mRNAs recovered by TMG-IP does not exceed 5–15%, whereas this level can reach close to 100% for $m_3^{2,2,7}G$-capped snoRNAs. Determinants such as the larger size and lower abundance as well as stability and shorter half-life of mRNAs compared to non-coding RNAs may contribute to their

lower recovery in the TMG-IP. The contribution of potential mRNA folded structures in the 5'UTR must also be taken into consideration in the recognition processes of individual mRNAs by the antibody. The TMG-IP method can also be applied to evaluate the ability of $m_{3\,2,2,7}G$-capped mRNAs to associate with actively translating ribosomes and to recover mRNAs after polysome fractionation of cytoplasmic mRNAs (Figure 2.4C). To this end cytoplasmic extracts must be recovered from cycloheximide-treated cells in order to block the translation elongation step and fractionated on linear 7–47% sucrose gradients. TMG-IP experiments can then be performed on pooled fractions that contain either free or non-polysome associated RNAs. This allows to estimate the ratio of ribosome bound compared to free $m_{3\,2,2,7}G$-capped mRNAs [31].

2.5 Conclusion

Protein-RNA interactions and their modularity play a central role in the mRNA fate including their assembly, modification, transport, translation and degradation. While many methods have been developed to examine mRNP complexes, there are still significant challenges that need to be addressed. The precise nature of the protein complexes that interact with most RNAs in the cell is still poorly understood and exploring low abundance transcripts remains a difficult task. The advent of high-throughput methods as well as new computational approaches to address protein–protein and protein-RNA networks is transforming our understanding of mRNP architecture as well as biology. In this chapter we have presented protein-centric techniques designed to reveal and validate targeted mRNA–protein interactions.

Acknowledgments The authors are supported by institutional funds from the French Centre National de la Recherche Scientifique (CNRS), the French Agence Nationale de la Recherche [grant ANR H4translation to GE, FM and CA], Doctoral grants of the University of Strasbourg (to LG and FA), and a Lebanese Doctoral grant to HH.

References

1. G. Singh, G. Pratt, G.W. Yeo, et al. (2015) The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion, Annual Review of Biochemistry. 84, 325–354.
2. M.J. Moore (2005) From Birth to Death: The Complex Lives of Eukaryotic mRNAs, Science. 309, 1514–1518.
3. S.F. Mitchell and R. Parker (2014) Principles and Properties of Eukaryotic mRNPs, Molecular Cell. 54, 547–558.
4. P. Björk and L. Wieslander (2017) Integration of mRNP formation and export, Cellular and Molecular Life Sciences. 74, 2875–2897.
5. M. Müller-McNicoll and K.M. Neugebauer (2013) How cells get the message: dynamic assembly and function of mRNA–protein complexes, Nature Reviews Genetics. 14, 275.
6. J.D. Keene and S.A. Tenenbaum (2002) Eukaryotic mRNPs May Represent Posttranscriptional Operons, Molecular Cell. 9, 1161–1167.
7. R.N. Singh, M.D. Howell, E.W. Ottesen, et al. (2017) Diverse role of survival motor neuron protein, Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 1860, 299–315.
8. S. Massenet, E. Bertrand, and C. Verheggen (2017) Assembly and trafficking of box C/D and H/ACA snoRNPs, RNA Biology. 14, 680–692.
9. C.-Y.A. Chen and A.-B. Shyu (2014) Emerging mechanisms of mRNP remodeling regulation, Wiley interdisciplinary reviews. RNA. 5, 713–722.
10. C.A. McHugh, P. Russell, and M. Guttman (2014), Methods for comprehensive experimental identification of RNA-protein interactions, Genome Biol. 15, 203.
11. D. Marchese, N.S. de Groot, N. Lorenzo Gotor, et al. (2016) Advances in the characterization of RNA-binding proteins, Wiley Interdisciplinary Reviews: RNA. 7, 793–810.
12. X. Li, J. Song, and C. Yi (2014), Genome-wide Mapping of Cellular Protein-RNA Interactions Enabled by Chemical Crosslinking, Genomics Proteomics Bioinformatics. 12, 72-78.
13. S.A. Tenenbaum, C.C. Carson, P.J. Lager, et al. (2000) Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays, Proceedings of the National

Academy of Sciences of the United States of America. 97, 14085–14090.
14. J. König, K. Zarnack, G. Rot, et al. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution, Nature structural & molecular biology. 17, 909–915.
15. M. Hafner, M. Landthaler, L. Burger, et al. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP, Cell. 141, 129–141.
16. N. Kastelic and M. Landthaler (2017) mRNA interactome capture in mammalian cells, Methods. 126, 38–43.
17. A. Castello, R. Horos, C. Strein, et al. (2013) System-wide identification of RNA-binding proteins by interactome capture, Nature Protocols. 8, 491.
18. A. Castello, C.K. Frese, B. Fischer, et al. (2017) Identification of RNA-binding domains of RNA-binding proteins in cultured cells on a system-wide scale with RBDmap, Nature Protocols. 12, 2447.
19. S.P. Ryder (2016) Protein-mRNA interactome capture: cartography of the mRNP landscape, F1000Research. 5, 2627.
20. T. Köster, C. Marondedze, K. Meyer, et al. (2017) RNA-Binding Proteins Revisited – The Emerging Arabidopsis mRNA Interactome, Trends in Plant Science. 22, 512–526.
21. S. Mili and J. Steitz (2004) Evidence for reassociation of RNA-binding proteins after cell lysis: Implications for the interpretation of immunoprecipitation analyses, RNA. 10, 1692–1694.
22. J.D. Keene, J.M. Komisarow, and M.B. Friedersdorf (2006) RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts, Nat. Protocols. 1, 302–307.
23. A. Mayeda and A.R. Krainer (1999) Preparation of Hela Cell Nuclear and Cytosolic S100 Extracts for In Vitro Splicing BT - RNA-Protein Interaction Protocols, Presented at the .
24. N. Kataoka and G. Dreyfuss (2008) Preparation of efficient splicing extracts from whole cells, nuclei, and cytoplasmic fractions, Methods in Molecular Biology. 488, 357–365.
25. S.M. Abmayr, T. Yao, T. Parmely, et al. (2006) Preparation of nuclear and cytoplasmic extracts from mammalian cells, Current Protocols in Pharmacology. 12, 3.
26. J.D. Dignam, R.M. Lebovitz, and R.G. Roeder (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei, Nucleic Acids Res. 11, 1475–1489.
27. C.-H.T. Webb and K.J. Hertel (2014) Preparation of Splicing Competent Nuclear Extracts, Methods in molecular biology (Clifton, N.J.). 1126, 117–121.
28. S. Boulon, N. Marmier-Gourrier, B. Pradet-Balade, et al. (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery, J Cell Biol. 180, 579–595.
29. A.-S. Gribling-Burrer, M. Leichter, L. Wurth, et al. (2017) SECIS-binding protein 2 interacts with the SMN complex and the methylosome for selenoprotein mRNP assembly and translation, Nucleic Acids Research. 45, 5399-5413.
30. S.A. Bustin, V. Benes, J.A. Garson, et al. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments, Clinical Chemistry. 55, 611–622.
31. L. Wurth, A.-S. Gribling-Burrer, C. Verheggen, et al. (2014) Hypermethylated-capped selenoprotein mRNAs in mammals, Nucleic Acids Research. 42, 8663–8677.
32. E.C. Wheeler, E.L. Van Nostrand, and G.W. Yeo (2018) Advances and challenges in the detection of transcriptome-wide protein-RNA interactions, Wiley Interdisciplinary Reviews: RNA. 9, e1436.
33. F. Zambelli and G. Pavesi (2015) RIP-Seq Data Analysis to Determine RNA–Protein Associations, Methods Mol Biol. 1269, 293-303.
34. M. Selbach and M. Mann (2006) Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK), Nature Methods. 3, 981.
35. C. Meyer, A. Garzia, and T. Tuschl (2017) Simultaneous detection of the subcellular localization of RNAs and proteins in cultured cells by combined multicolor RNA-FISH and IF, Methods. 118–119, 101–110.
36. W. Van Criekinge and R. Beyaert (1999) Yeast Two-Hybrid: State of the Art, Biological Procedures Online. 2, 1–38.
37. D.R. Setzer (1999) Measuring Equilibrium and Kinetic Constants Using Gel Retardation Assays BT - RNA-Protein Interaction Protocols. 115-128,.
38. L.M. Hellman and M.G. Fried (2007) Electrophoretic Mobility Shift Assay (EMSA) for Detecting Protein-Nucleic Acid Interactions, Nature protocols. 2, 1849–1861.
39. L. Trinkle-Mulcahy (2012), Resolving protein interactions and complexes by affinity purification followed by label-based quantitative mass spectrometry, Proteomics. 10, 1623-1638.
40. E. Gerace and D. Moazed (2015) Affinity Pull-Down of Proteins Using Anti-FLAG M2 Agarose Beads, Methods in Enzymology. 559, 99–110.
41. L. Trinkle-Mulcahy, S. Boulon, Y.W. Lam, et al. (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes, Journal of Cell Biology. 183, 223–239.
42. E.A. Hoffman, B.L. Frey, L.M. Smith, et al. (2015) Formaldehyde Crosslinking: A Tool for the Study of Chromatin Complexes, Journal of Biological Chemistry. 290, 26404–26411.
43. V. Orlando (2000) Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation, Trends in Biochemical Sciences. 25, 99–104.
44. B.W. Sutherland, J. Toews, and J. Kast (2008), Utility of formaldehyde cross-linking and

mass spectrometry in the study of protein-protein interactions, Journal of Mass Spectrometry. 43, 699-715.
45. P.C.K. Au, C. Helliwell, and M.-B. Wang (2014) Characterizing RNA–protein interaction using cross-linking and metabolite supplemented nuclear RNA-immunoprecipitation, Molecular Biology Reports. 41, 2971–2977.
46. S. Kaneko and J.L. Manley (2005) The Mammalian RNA Polymerase II C-Terminal Domain Interacts with RNA to Suppress Transcription-Coupled 3′ End Formation, Molecular Cell. 20, 91–103.
47. C. Gilbert, A. Kristjuhan, G.S. Winkler, et al. (2004) Elongator Interactions with Nascent mRNA Revealed by RNA Immunoprecipitation, Molecular Cell. 14, 457–464.
48. A. V Jivotovskaya, L. Valášek, A.G. Hinnebusch, et al. (2006) Eukaryotic Translation Initiation Factor 3 (eIF3) and eIF2 Can Promote mRNA Binding to 40S Subunits Independently of eIF4G in Yeast, Molecular and Cellular Biology. 26, 1355–1372.
49. S. Niranjanakumari, E. Lasda, R. Brazas, et al. (2002) Reversible cross-linking combined with immunoprecipitation to study RNA–protein interactions in vivo, Methods. 26, 182–190.
50. I.G. Pashev, S.I. Dimitrov, and D. Angelov (1991) Crosslinking proteins to nucleic acids by ultraviolet laser irradiation, Trends Biochem Sci. 16, 323–326.
51. K.B. Jensen and R.B. Darnell (2008) CLIP: Crosslinking and immunoprecipitation of in vivo RNA targets of RNA-binding proteins, Methods in Molecular Biology. 488, 85–98.
52. S. Kishore, L. Jaskiewicz, L. Burger, et al. (2011) A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins, Nature Methods. 8, 559–567.
53. S. Hussain, A.A. Sajini, S. Blanco, et al. (2013) NSun2-Mediated Cytosine-5 Methylation of Vault Noncoding RNA Determines Its Processing into Regulatory Small RNAs, Cell Reports. 4, 255–261.
54. M.T. Bohnsack, R. Martin, S. Granneman, et al. (2009) Prp43 Bound at Different Sites on the Pre-rRNA Performs Distinct Functions in Ribosome Synthesis, Molecular Cell. 36, 583–592.
55. S. Granneman, G. Kudla, E. Petfalski, et al. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs, Proceedings of the National Academy of Sciences. 106, 9613–9618.
56. C. Strein, A.M. Alleaume, U. Rothbauer, et al. (2014) A versatile assay for RNA-binding proteins in living cells, RNA. 20, 721–731.
57. Y.H. Choi and C.H. Hagedorn (2003) Purifying mRNAs with a high-affinity eIF4E mutant identifies the short 3' poly(A) end phenotype, Proc Natl Acad Sci U S A. 100, 7033–7038.
58. A.S.Y. Lee, P.J. Kranzusch, and J.H.D. Cate (2015) eIF3 targets cell proliferation mRNAs for translational activation or repression, Nature. 522, 111–114.
59. K. Stade, J. Rinke-Appel, and R. Brimacombe (1989) Site-directed cross-linking of mRNA analogues to the Escherichia coli ribosome; identification of 30S ribosomal components that can be cross-linked to the mRNA at various points 5' with respect to the decoding site., Nucleic Acids Research. 17, 9889–9908.
60. F. Bélanger, H. Baigude, and T.M. Rana (2009) U30 of 7SK RNA forms a specific photo-crosslink with Hexim1 in the context of both a minimal RNA-binding site and a full reconstituted 7SK/Hexim1/P-TEFb ribonucleoprotein complex, Journal of molecular biology. 386, 1094–1107.
61. S. López de Quinto, E. Lafuente, and E. Martínez-Salas (2001) IRES interaction with translation initiation factors: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII, RNA. 7, 1213–1226.
62. A. V Pisarev, V.G. Kolupaeva, M.M. Yusupov, et al. (2008) Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes, The EMBO Journal. 27, 1609–1621.
63. A.S.Y. Lee, P.J. Kranzusch, J.A. Doudna, et al. (2016) eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation, Nature. 536, 96.
64. A. Pacheco, S.L. de Quinto, J. Ramajo, et al. (2009) A novel role for Gemin5 in mRNA translation, Nucleic Acids Research. 37, 582–590.
65. L. Gross, L. Schaeffer, F. Alghoul, et al. (2018) Tracking the m7G-cap during translation initiation by crosslinking methods, Methods. 137, 3-10.
66. N. Sonenberg and A.J. Shatkin (1977) Reovirus mRNA can be covalently crosslinked via the 5' cap to proteins in initiation complexes, Proceedings of the National Academy of Sciences of the United States of America. 74, 4288–92.
67. F. Martin, S. Barends, S. Jaeger, et al. (2011) Cap-assisted internal initiation of translation of histone H4, Mol Cell. 41, 197–209.
68. P.J. Batista (2017) The RNA Modification N6-methyladenosine and Its Implications in Human Disease, Genomics, Proteomics & Bioinformatics. 15, 154–163.
69. J.E. Squires, H.R. Patel, M. Nousch, et al. (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA, Nucleic Acids Research. 40, 5023–5033.
70. R. Desrosiers, K. Friderici, and F. Rottman (1974) Identification of Methylated Nucleosides in Messenger RNA from Novikoff Hepatoma Cells, Proceedings of the National Academy of Sciences . 71, 3971–3975.
71. A.J. Shatkin (1976) Capping of eucaryotic mRNAs, Cell. 9, 645–653.
72. I. Topisirovic, Y. V Svitkin, N. Sonenberg, et al. (2011) Cap and cap-binding proteins in the control of gene expression, Wiley interdisciplinary reviews. RNA. 2, 277–298.
73. J. Houseley and D. Tollervey (2009) The many pathways of RNA degradation, Cell. 136, 763–776.
74. A. Kohler and E. Hurt (2007) Exporting RNA from the nucleus to the cytoplasm, Nat Rev

Mol Cell Biol. 8, 761–773.
75. P. Fortes, T. Inada, T. Preiss, et al. (2000) The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation, Mol Cell. 6, 191–196.
76. C.J. Wilusz, M. Wormington, and S.W. Peltz (2001) The cap-to-tail guide to mRNA turnover, Nat Rev Mol Cell Biol. 2, 237–246.
77. E. Izaurralde, J. Lewis, C. Gamberi, et al. (1995) A cap-binding protein complex mediating U snRNA export, Nature. 376, 709–712.
78. M.T. Tuck (1992) The formation of internal 6-methyladenine residues in eucaryotic messenger rna, International Journal of Biochemistry. 24, 379–386.
79. X. Wang, Z. Lu, A. Gomez, et al. (2013) N6-methyladenosine-dependent regulation of messenger RNA stability, Nature. 505, 117–120.
80. G. Jia, Y. Fu, and C. He (2013) Reversible RNA adenosine methylation in biological regulation, Trends in Genetics. 29, 108–115.
81. A.S. Warda, J. Kretschmer, P. Hackert, et al. (2017) Human METTL16 is a N6-methyladenosine (m6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs, EMBO reports. 18, 2004-2014.
82. D. Dominissini, S. Moshitch-Moshkovitz, S. Schwartz, et al. (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq, Nature. 485, 201.
83. N. Liu, K.I. Zhou, M. Parisien, et al. (2017) N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein, Nucleic Acids Research. 45, 6051–6063.
84. D. Dominissini, S. Moshitch-Moshkovitz, M. Salmon-Divon, et al. (2013) Transcriptome-wide mapping of N6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing, Nature Protocols. 8, 176.
85. K.D. Meyer, Y. Saletore, P. Zumbo, et al. (2012) Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3′ UTRs and near Stop Codons, Cell. 149, 1635–1646.
86. B. Delatte, F. Wang, L.V. Ngoc, et al. (2016) Transcriptome-wide distribution and function of RNA hydroxymethylcytosine, Science. 351, 282 -285.
87. X. Yang, Y. Yang, B.-F. Sun, et al. (2017) 5-methylcytosine promotes mRNA export — NSUN2 as the methyltransferase and ALYREF as an m5C reader, Cell Research. 27, 606.
88. S. Edelheit, S. Schwartz, M.R. Mumbach, et al. (2013) Transcriptome-Wide Mapping of 5-methylcytidine RNA Modifications in Bacteria, Archaea, and Yeast Reveals m5C within Archaeal mRNAs, PLOS Genetics. 9, e1003602.
89. X. Li, X. Xiong, and C. Yi (2016) Epitranscriptome sequencing technologies: decoding RNA modifications, Nature Methods. 14, 23.
90. A.-S. Gribling-Burrer, G. Eriani, and C. Allmang (2018) Modification of selenoprotein mRNAs by cap tri-methylation, Methods in Molecular Biology. 1661, 125–141.
91. R. Luhrmann, B. Appel, P. Bringmann, et al. (1982) Isolation and characterization of rabbit anti-m3 2,2,7G antibodies, Nucleic Acids Res. 10, 7103–7113.
92. K.T. Tycowski, A. Aab, and J.A. Steitz (2004) Guide RNAs with 5’ caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa, Curr Biol. 14, 1985–1995.

Table 2.1: List of protein-centric methods for the purification of mRNP complexes

	Method of mRNP purification	**Applications**	**Limitations**
In vivo RIP	Native RIP	Endogenous RNP complexes purification	Possible RNP rearrangements Possible nonspecific interactions Requires highly specific anti-RBP antibodies
	Epitope tagging and RIP	Confirmation of native RIP Analysis of low abundance target proteins and their target mRNA Efficient RNP affinity purification	Possible impact of the epitope tag on the target protein: misfolding, loss of function, mislocalization Potential disruption of functional protein complexes Nonspecific interactions with the epitope tag
	Formaldehyde cross-linking and RIP	Capture of RNA-protein interactions Stabilization of transitory interactions and indirect RNA-binding factors	False positives due to excess of cross-link False negatives due to low cross-linking efficiency
	UV cross-linking and RIP	Identification of direct mRNA-protein interactions Trapping of functional interactions Genome-wide analysis: PAR-CLIP HITS-CLIP; iCLIP; iCLAP CRAC	False negatives due to low cross-linking efficiency
Cell free mRNP purification	GST pull-down	Reconstitution and screening for mRNA-protein interactions	Possible artefacts due to protein misfolding, lack of post-translational modifications Nonspecific interactions with the GST-tag
	Cross-linking	Identification RBPs using *in vitro* transcribed mRNAs: 4-thioU internally labelled mRNAs 6-thioG capped mRNAs, tracking of cap-binding proteins	Optimization of *in vitro* 4-thioU incorporation and 6-thioG cap labelling
Direct mRNA IP	m^6A Me-RIP m^5C RIP	Global identification of mRNA transcriptional modifications patterns	Specificity of the antibodies
	TMG-IP	Detection of endogenous $m_3^{2,2,7}G$-capped mRNAs	Low recovery levels of mRNAs possibly due to their abundance, stability, short-half life

RIP, RNA Immunoprecipitation, PAR-CLIP, Photoactivatable ribonucleoside-enhanced Cross-Linking and ImmunoPrecipitation, HITS-CLIP, High-throughput sequencing CLIP, iCLIP, individual nucleotide resolution CLIP, iCLAP, individual-nucleotide-resolution UV-cross-linking and affinity purification, CRAC UV cross-linking and analysis of cDNA, GST (Glutathione-S-transferase) pull-down, m^6A Me-RIP, anti N^6-methyladenosine RIP, m^5C RIP, anti 5-methylcytidine RIP, TMG-IP anti $m_3^{2,2,7}G$ IP.

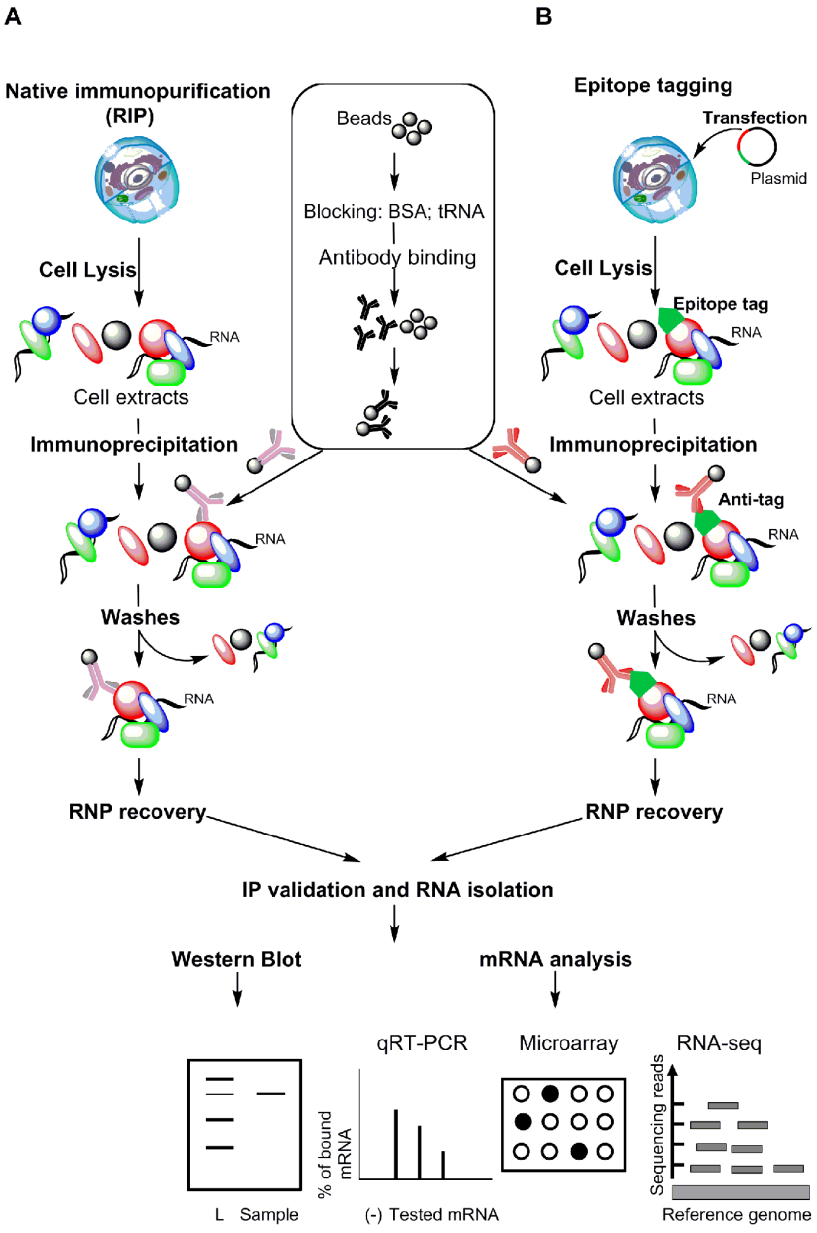


Fig. 2.1. Schematic representation of RNA immunoprecipitation (RIP) methods for the identification of endogenous protein-mRNA interactions. Native mRNPs can be recovered directly from HEK293 or HeLa cell lysates (**A**) or after transfection and expression of the epitope tagged protein (**B**). In both cases RNAs bound to specific proteins of interest are immunoprecipitated from cell lysates in native conditions by using either a protein-specific antibody (**A**) or anti-tag antibody (**B**). After several washing steps the RNA-protein complexes (RNPs) are recovered by elution in either denaturing or native conditions. The validity of the immunoprecipitation experiment is verified by western blot. The RNA is extracted and interacting mRNAs are identified by targeted RNA detection methods (qRT-PCR, microarray) or global high throughput sequencing analysis (RNA-seq). The insert represents the preparation of the affinity matrix before immunoprecipitation. This includes pre-coating of the beads with purified BSA and yeast total tRNA to reduce nonspecific background before binding the antibodies.

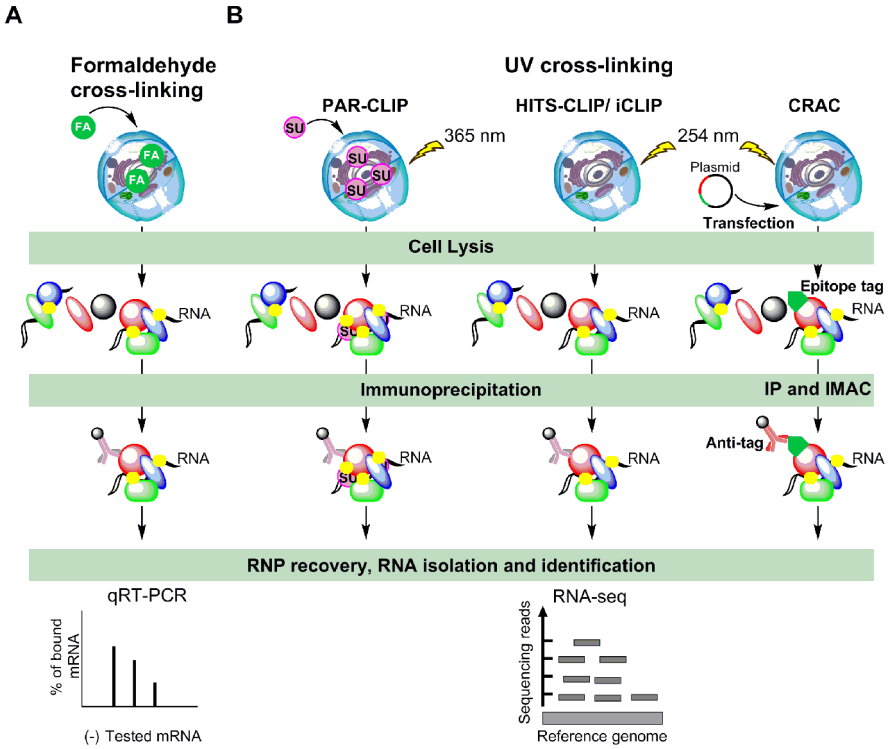


Fig. 2.2. Outline of chemical and UV cross-linking immunoprecipitation methods. (**A**) Formaldehyde (FA) cross-linking. (**B**) UV cross-linking methods coupled to high-throughput analysis are represented from left to right: photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP), High-throughput sequencing CLIP (HITS-CLIP) and individual-nucleotide resolution CLIP (iCLIP), individual nucleotide crosslinking and cDNA analysis (CRAC). PAR-CLIP uses 4-thioU ribonucleosides; UV at 365 nm induces zero length cross-links between the photoactivatable 4-thioU moiety of the RNA and RNA-binding proteins (RBPs), while the other three methods utilize UV cross-linking at 254 nm. Isolation of RNP complexes is achieved by immunoprecipitation (IP) or double affinity purification by IP and immobilized metal ion affinity chromatography (IMAC) in the case of CRAC. FA represents formaldehyde and SU and represents 4-thioU nucleotides respectively. Yellow dots symbolize cross-links between RNA and proteins.

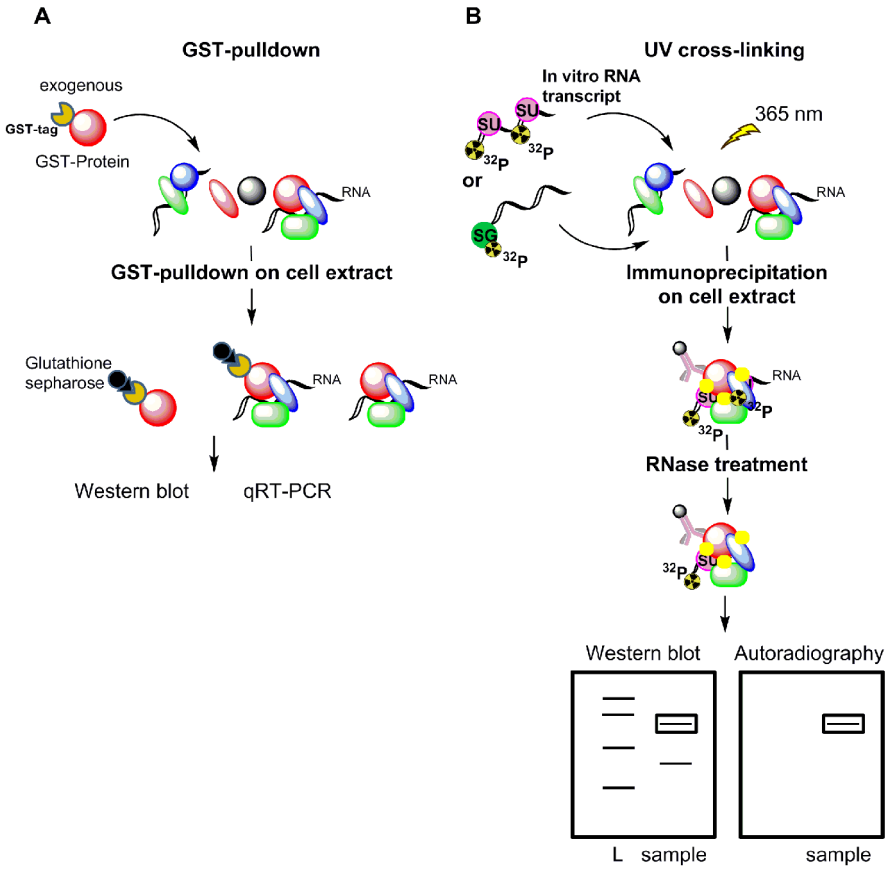


Fig. 2.3. Cell-free extract mRNP immunoprecipitation methods. (A) GST pull-down. An exogenous GST-tagged recombinant protein produced in *E. coli* is incubated with cell extracts to reconstitute RNA-protein interactions. After affinity purification using Glutathione sepharose beads, the co-precipitated mRNA can be identified by target mRNA identification methods such as qRT-PCR. **(B)** Illustration of cross-linking methods using exogenous 4-thioU or 6-thioG labelled mRNAs. 4-thioU and ^{32}P labelled transcripts or 6-thioG/^{32}P cap modified mRNA transcripts are incubated with cell free extracts and cross-linked to RBPs at 365 nm. Partial RNase digestion allows transferring the fragments of radiolabelled mRNA to the interacting protein. After immunoprecipitation comparison of Western blot and Phosphorimager analysis of an SDS-gels allows to identify the 5'-^{32}P-labeled protein. SU and SG represent 4-thioU and 6-thioG nucleotides respectively. Cross-links are illustrated as in Fig. 2.2.

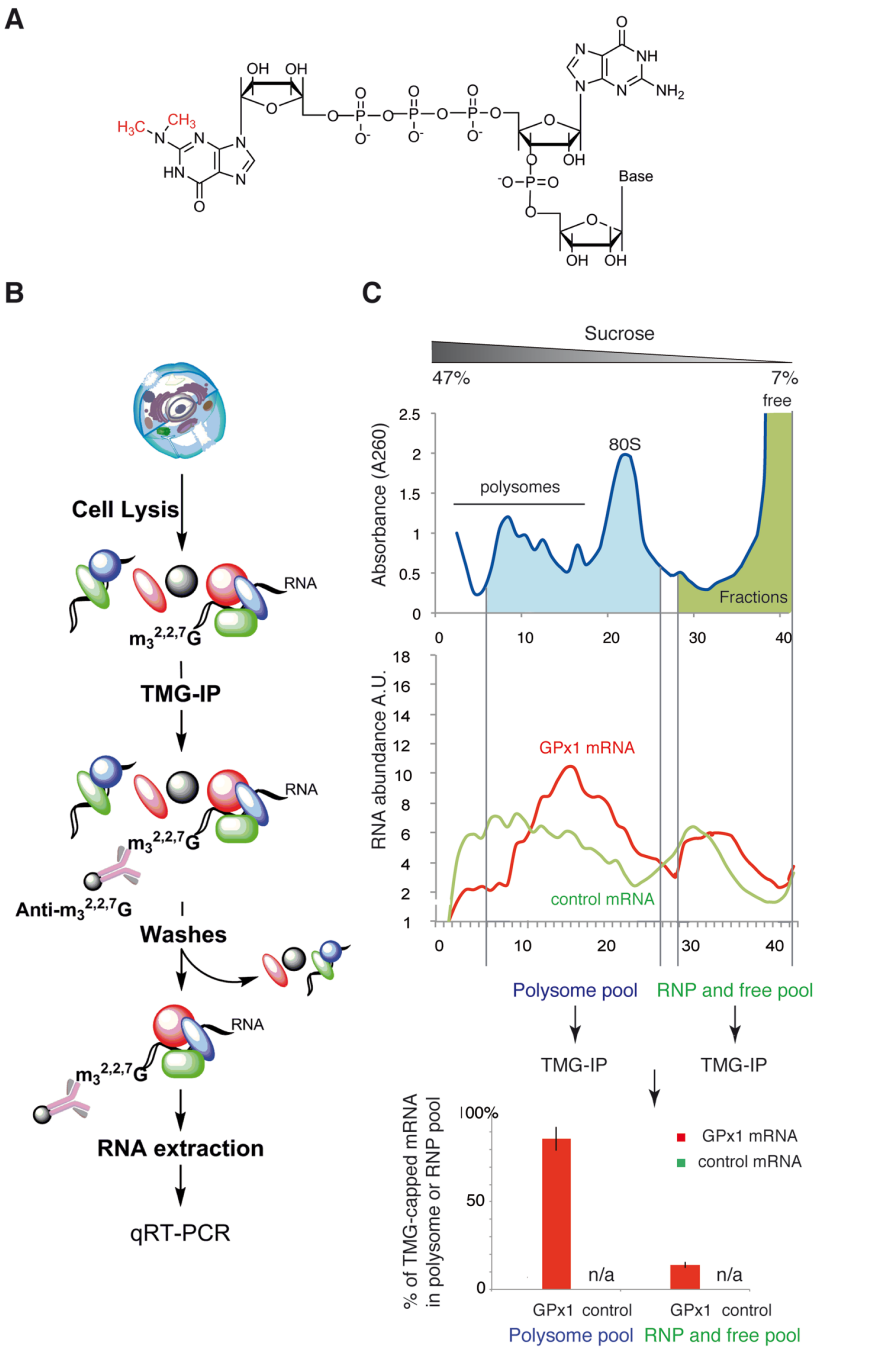


Fig. 2.4. Anti-$m_3^{2,2,7}G$ cap immunoprecipitation (TMG-IP) experiment. (**A**) Structure of the $m_3^{2,2,7}G$ cap. Compared to the canonical m^7G cap, two additional methyl groups are present on the exocyclic N2; they are represented in red. (**B**) TMG-IP workflow. RNA extracted from cellular extracts is immunoprecipitated with anti-TMG serum (anti-$m_3^{2,2,7}G$ Ab, R1131). The bound RNA is analysed by qRT-PCR. (**C**) Polysome analysis of the distribution of TMG-capped mRNAs in ribosome bound and free polysome fractions. Cytoplasmic extracts are fractionated onto 7–47% (w/v) linear sucrose gradient after cycloheximide treatment in order to block translation elongation. A typical absorbance profile is represented in the top panel and the positions of the polysomes, 80S ribosome as well as free RNAs are indicated. The middle panel represents relative mRNA abundance in each fraction measured by qRT-PCR in the case of the GPx1 selenoprotein mRNA and a control mRNA. Vertical bars mark the position of the polysome as well as RNP and free fractions that are pooled and can be analysed by TMG-IP. The amount of RNA immunoprecipitated from the polysome and RNP pools are measured separately by qRT-PCR and normalized to 100%. Error bars represent standard deviations of an average of two independent experiments

ANNEXE 3 : PUBLICATION 3

Purification of in vivo or in vitro-assembled RNA-protein complexes by RNA centric methods

Janvier, A., Hayek, H., Alghoul, F., Gross, L., Allmang, C., Martin, F., and Eriani, G.

***Book chapter** : In Advanced Technologies for Protein Complex Production and Characterization. Editors: Vega, M. Cristina Springer International Publishing Switzerland (**in press**)*

Chapter 3. Purification of *in vivo* or *in vitro*-assembled RNA-protein complexes by RNA centric methods

Aurélie Janvier, Hassan Hayek, Fatima Alghoul, Lauriane Gross, Christine Allmang, Franck Martin, and Gilbert Eriani*

Architecture et Réactivité de l'ARN, Université de Strasbourg, Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 67084 Strasbourg, France

*Correspondence to g.eriani@unistra.fr

Running title: purification of RNA-protein complexes

Abstract Throughout their entire life cycle, RNAs are associated with RNA-binding proteins (RBPs), forming ribonucleoprotein (RNP) complexes with highly dynamic compositions and very diverse functions in RNA metabolism, including splicing, translational regulation, ribosome assembly. Many RNPs remain poorly characterized due to the challenges inherent in their purification and subsequent biochemical characterization. Therefore, developing methods to isolate specific RNA-protein complexes is an important initial step toward understanding their function. Many elegant methodologies have been developed to isolate RNPs. This chapter describes different approaches and methods devised for RNA-specific purification of a target RNP. We focused on general methods for selecting RNPs that target a given RNA under conditions favourable for the copurification of associated factors including RNAs and protein components of the RNP.

Keywords RNA Protein complexes, RNP ribonucleoprotein, Ribosomes, Affinity purification, Mammalian cell-free extracts, Complex assembly, RNA aptamer, RNA tag

3.1 Introduction: issues and challenges of RNP purifications

The use of small, genetically introduced affinity tags for the production and purification of recombinant proteins and their complexes with other proteins has greatly improved our knowledge on their function in a variety of research fields. The usual protein affinity tags include polyhistidine (poly-His), the hemagglutinin epitope (HA-tag), myc epitope, glutathione S-transferase (GST-tag), Maltose-Binding Protein (MBP-tag), Strep-tag, FLAG epitope (Flag-tag), protein A, etc. These affinity tags all bind with high affinity to a ligand that can be immobilized on a chromatography resin for further purification. In most cases, the bound complexes can either be released from the resin by competitive elution or cleaved off by a protease with a recognition site incorporated in the fusion protein. The widespread use and success of protein affinity tags has led to the development of comparable tags for nucleic acids leading to new RNA-centric methods for affinity-purification of RNA-binding proteins (RBDs) or ribonucleo-particles (RNPs). RNA-centric methods target RBPs or RNPs binding to a single RNA of interest. The majority of the existing methods use tagged RNAs as baits to capture and study complexes bound to it. During these last years, particular emphasis has been placed on developing new affinity tags for RNA molecules. Between the early studies in the 90s and today's methods, a large variety of diverse RNA tags have been described with various applications including functional analyses, in

addition to mass-spectrometry, cell imaging, or three-dimensional structure determination.

Different approaches have been used to tag RNA molecules. (i) RNAs can be chemically tagged during *in vitro* transcription through the incorporation of modified ribonucleotides that contain biotin, fluorescent dyes, or other compounds. (ii) Well-characterized protein-binding RNA sequences can be incorporated during *in vitro* or *in vivo* transcription. Such natural sequences might be further optimized in size and affinity. (iii) Similarly, artificially selected RNA aptamers can be incorporated during transcription. Because of their small sizes, such aptamer tags are easy to insert into the RNA of interest. The tagged RNAs and assembled RNPs are subsequently purified by affinity-purification on an adapted resin. The limiting step often consists in the elution of the assembled complexes in native conditions since the binding affinities are typically high. (iv) Hybridization of biotinylated oligonucleotides that are complementary to the RNA target is an efficient purification strategy. It requires accessible single-stranded regions and elution can be achieved under denaturing conditions, or displacement by a competitor oligonucleotide, or by targeted RNase degradation. (v) Alternatively, the RNA sequence of interest may be ligated to a biotinylated-DNA oligonucleotide in order to yield chimeric RNA-DNA molecules that can efficiently bind streptavidin beads. Benefits of this procedure include the absence of constraints resulting from the use of a foreign structured tag. In addition, the elution is performed in native conditions by DNase or targeted RNase degradation that releases the RNA-protein complexes without any extra sequence. Each of these methods has particular advantages and disadvantages for the tagging or affinity isolation or elution of RNPs. They will be discussed hereafter.

3.2 First affinity purification methods

Two early studies performed in the 90s established the foundations of affinity purification based on natural RNA regulatory sequence as bait [1, 2]. The studies focused on the Iron Responsive Element (IRE) that binds the Iron Responsive Element Binding-Protein (IRE-BP) involved in iron homeostasis. In the first study, the RNA transcript containing the regulatory sequences upstream of a biotinylated unstructured tail was added in solution to unfractionated lysates containing the protein target. Then, the high affinity interaction between biotin and avidin was used to attach the IRE/IRE-BP to a solid matrix. For that, the assembled mRNA-protein complexes were bound to biotin-agarose beads through a succinylated-avidin intermediate (Fig. 1A). IRE-BP was subsequently eluted from the RNA with 2 M KCl [1]. In the second study, *in vitro* transcribed polyadenylated RNA was hybridized to a poly(U)-Sepharose resin and incubated with pre-purified cellular extracts. The RNA-protein complexes were assembled and the adsorbed proteins were recovered by elution with 1 M KCl [2] (Fig. 1B). Although these two seminal studies reached the mRNP assembly on a specific RNA target, the complex was dissociated during the high-salt elution. Nevertheless, they successfully tested two RNA-tagging procedures. The first was using biotinylated nucleotides inserted in the RNA for binding to the beads, and the second procedure took advantage of the RNA-RNA hybridization to immobilize RNA-protein complexes on a solid matrix.

Over the same period, a more general affinity purification method that allowed the isolation of specific RNAs and RNA-protein complexes was described [3]. The approach, illustrated in Figure 1C, hinges on the specific interaction between the bacteriophage R17 coat protein and a short hairpin found in its genomic RNA. In this strategy, the R17 coat protein (CP) was covalently attached to beads in order to retain RNA sequences containing the short hairpin sequence. To improve the interaction between the CP and RNA hairpin, a high affinity variant of the RNA hairpin was used (Ka = 3.5 □ 10^{9}M [4]). A hybrid RNA containing two R17 recognition sites fused to the RNA sequence of interest was prepared either *in vitro*, by transcription with a phage polymerase, or *in vivo*, by cellular transcription of a transfected or injected DNA template. After incubating the

chimeric RNA with appropriate factors in the cell or lysate, the resulting RNA-factor complexes were retained on the resin coupled to the R17 coat protein. The binding of RNAs to the beads turned out to be rapid, efficient and highly selective. After eluting the column with an excess of R17 recognition sites, the authors could obtain biologically active factors and complexes of interest [4].

3.3 Affinity purification based on artificially-selected or natural RNA motifs

Over the years, a large variety of aptamer-based affinity purification approaches have been developed for the isolation of *in vitro*-assembled RNA–protein complexes. These include the use of StreptoTag aptamers [5], streptavidin aptamer [6], Sephadex aptamer [7], tobramycin aptamers [8], and Mango aptamer [9]. Other natural RNA elements have been used such as the bacteriophage R17 MS2 sequence (binding site of the coat protein CP), or the equivalent PP7 RNA from *Pseudomonas aeruginosa* bacteriophage PP7 [10–13].

3.3.1 Antibiotic-binding RNA aptamers

One of the first methods for affinity-purification was based on the RNA sequence named StreptoTag aptamer [5]. This 48-nt long RNA element binds with high binding specificity to the antibiotic streptomycin. It was isolated via SELEX and binds to the antibiotic with a dissociation constant (Kd) of around 1 μM (Fig. 2A). The StreptoTag was inserted by *in vitro* transcription in a hybrid RNA and was incubated with crude extract. The assembly reaction was then applied to a dihydrostreptomycin-Sepharose column. The resulting bound complexes were washed and specifically eluted upon addition of free streptomycin. Using this purification scheme, properly assembled spliceosomal U1A protein and the bacteriophage MS2 coat protein could be isolated via their appropriate RNA motif [5] as well as 48S ribosomal complexes assembled on an IRES element in milligram quantities [14]. The method turned out to be more specific than the procedures that use poly-U Sepharose [2] or biotin/streptavidin beads [1] that often copurify unspecific proteins [5]. One of the drawbacks of the method consists in the preparation of the dihydrostreptomycin-Sepharose column. Although the protocol for coupling streptomycin to Sepharose 6B is well established, it is a time-consuming process and the coupled matrix can only be stored for a maximum of 4 weeks. In addition, the quality of the matrix might be subjected to batch-to-batch variation.

A similar affinity-purification procedure was developed using an RNA aptamer selected to bind with 5 nM affinity the aminoglycoside antibiotic tobramycin [15, 16] (Fig. 2A). The tobramycin-binding aptamer was used to purify preparative amounts of human pre-spliceosome complex. The method turned out to be ideally suited to isolate high molecular weight complexes in their native form and for functional and structural studies [8, 16]. However, the drawbacks of the method are similar to those of the StreptoTag and mainly concern the preparation and preservation of the tobramycin-coupled Sepharose support.

3.3.2 Streptavidin and Sephadex aptamers

The S1 Streptavidin-RNA binding aptamer was selected on the basis of a tight binding to a commonly available target molecule in such a way that the ligand–RNP complex can be selectively and gently dissociated afterwards. The S1 Streptavidin tag binds specifically to streptavidin and can be eluted by competition with biotin. Streptavidin is a protein produced by *Streptomyces*

avidinii, which binds to biotin with extraordinary high affinity. The Kd for the biotin interaction is one of the strongest non-covalent biological interactions found in nature (10^{-14} M). Streptavidin is commercially available in many forms, either as a purified protein or conjugated with enzymes, dyes, or many supporting matrices. The Streptavidin aptamer, as an affinity tag (Fig. 2A), inserted into the RNA component of the large RNA subunit of RNase P, was successfully used to purify the active form of the ribonucleoprotein RNase P [6]. The S1 aptamer was further used for the isolation of an mRNP involved in translation activation [17] and purification of mRNA-interacting proteins from human cells and Drosophila embryo extracts using *in vitro* transcribed RNAs attached to streptavidin via the S1 tag [18, 19]. It was also shown that the addition of a tRNA scaffold to the Streptavidin aptamer increased binding efficiency of about 10-fold. Similarly, by optimizing the RNA aptamer S1 in structure and repeat conformation, the affinity for Streptavidin was increased and found to be optimal with a 4-fold repeat of S1m (4×S1m) and be more efficient than the established MS2 and PP7 systems from bacteriophages [20].

The advantages of the Streptavidin tag include an elution step with biotin, which is done under mild conditions. Mild conditions preserve the particle integrity and increase purity since non-specific contaminants that bind to the column support would not elute (on the contrary to stringent elution with high salt, pH changes, or with denaturing agents). In addition, Streptavidin-matrices are widely distributed as agarose beads or magnetic particles. This eliminates the necessity of preparing a specific affinity resin.

Sephadex-binding RNA aptamers were selected against Sephadex, a commonly used matrix in gel filtration (Fig. 2A) [7]. Sephadex is widely used in many laboratories and readily available. It is relatively inexpensive, making large-scale purification more affordable. Sephadex is stable and can be regenerated several times. Its basic structure, formed of glucose repeats connected via α-1,6 glucosidic bonds, provides a high binding capacity. Using Sephadex G-100, the Sephadex aptamer could be purified from a complex mixture of cellular RNA, with an enrichment of at least 60 000-fold. Similarly, yeast nuclear RNase P containing one RNA molecule tightly associated with 9 protein subunits could be purified and eluted by competition with soluble dextran B512 [21].

3.4 Affinity purification with RNA tags derived from natural sequences

A variety of studies have focused on natural RNA tags that are small and form structured RNA motifs that can fold independently of the tagged RNA (Fig. 2B). They bind specific ligands with high affinity. They include the MS2 RNA hairpin [10–13, 22, 23], the *Pseudomonas* phage 7 (PP7) RNA hairpin for coat protein [24–28], the 19-nts of the B-box of lambda bacteriophage that bind the antiterminator protein N (22 amino acid RNA-binding domain) [29–31], or the U1 snRNA hairpin II that binds U1 A protein [32].

The most widely used affinity purification with RNA tag is that of the MS2 coat protein (CP) and its cognate RNA. MS2 is a member of a family of closely related enterobacterial viruses that includes bacteriophage f2, bacteriophage Qβ, and R17. The coat proteins of these single-stranded RNA bacteriophages are translational repressors of their viral replicase. They achieve this function by specifically binding an RNA hairpin that encompasses the replicase start codon. This interaction is generally conserved among the RNA bacteriophages with variations in the sequences of the coat proteins and RNA hairpins during evolution.

Initially developed by Bardwell & Wickens in 1990 (see above) with the R17 recognition sequence originating from the phage R17, the MS2 tag of bacteriophage MS2 has similar sequence and binds the coat protein (CP) with similar high affinity. Three copies of the MS2 tag are usually placed in the RNA of interest in order to improve its binding efficiency. Since the affinity of the repeated MS2 sequence for MS2 CP is too high to consider an efficient elution by

direct affinity purification, a hybrid protein is used as an intermediate to affinity-bind the MS2 sequences on one side (CP moiety) in fusion with the Maltose-Binding Peptide (MBP moiety) in order to interact with the resin on the other side. Therefore, there is no requirement for the MS2 CP to be immobilized, and the isolation can be performed using the MBP and not the MS2 coat protein. The most used fusion protein involves the N-terminal domain of the MBP combined with the MS2 CP part with a double mutation (V75Q and A81G) that prevents the oligomerization [33]. The MS2 CP-MBP fusion protein is expressed in *Escherichia coli* and after affinity purification the complex can be released from the affinity resin by elution with maltose (Fig. 2B, 2C). Alternatively, the elution can be performed using a protease cleavage site located between the MS2 coat protein site and the protein that binds directly to the affinity resin. In this case, purified complexes are released from the resin on by cleavage with a specific protease. Regardless of whether the elution proceeds via the maltose or by protease cleavage, the MS2 CP-MBP or MS2 CP remain bound to the mRNA-protein complex, which may represent a disadvantage for structural or functional studies. This approach has been successfully used to purify many large molecular weight complexes such as U2 snRNP with the ATP-independent spliceosomal complex [10], ribosomes on IRES elements [34] or native spliceosomes [11, 13, 22, 35].

Variations of the fusion MS2 CP-MBP include the MS2 CP-GST which binds to Glutathione-beads with an elution step by glutathione [36]. An immunoglobulin variant was proposed with the MS2 CP domain fused to the immunoglobulin-binding domain of the *Staphylococcus aureus* protein A (ZZ domain) (Fig. 2C). The complex of interest could be purified from a cell lysate using immunoglobulin-conjugated beads [37]. The MS2-BioTRAP method is another RNA tagging system designed for purification of *in vivo* assembled RNPs. In this method, the RNA of interest harbouring multiple MS2 RNA elements is co-expressed with a MS2 CP fused to the HisBio tag (HB tag). The HB tag adds to the MS2 CP two hexa-His tags, a TEV cleavage site, and a signal sequence for *in vivo* biotinylation. Here, the high affinity of biotin for streptavidin is used to isolate the endogenously biotinylated HB-tagged protein associated with the MS2-tagged RNA within the RNPs [38]. In addition, the MS2 aptamer inspired numerous derived methods for targeting RNPs or miRNAs such as RaPID [39], MS2-TRAP [36], and Ribotrap [37]. The MS2-CP was also fused with Green Fluorescent Protein (GFP) to visualize mRNAs bearing the MS2 aptamer *in vivo* [40–42]. MS2-CP-GFP could also be fused with the Streptavidin-Binding Protein (SBP) in order to give a MS2-CP-GFP-SBP, which could be visualized by fluorescence and purified using streptavidin-conjugated beads [39], however these methods go far beyond the scope of this article.

Generally, all these methods are relatively flexible and widely applicable, and in some cases, their *in vivo* applicability allows the study of protein–RNA interactions in physiological conditions. Tagged RNAs can be stably expressed in appropriate mammalian cell lines and RNPs assembled on the tagged RNAs may be purified from extracts using the appropriate affinity purification method. The primary use of such techniques has involved identification of mRNP components by mass spectrometry, but purified RNPs may also be used for downstream functional assays. However, in some cases, the incorporation of a tag into the RNA bait may alter its secondary structures and possibly the formation of ribonucleoprotein complexes. On the other hand, the presence of the exogenous tag may also induce binding of additional nonspecific RNA binding proteins.

3.4 RNP capture based on antisense biotinylated oligonucleotides

These methods are using antisense oligoribonucleotides that are complementary to the RNA of interest (mRNA, or RNP). These complexes are subsequently enriched and purified from cell extracts by hybridization. Usually, one or several biotin residues are included at the 5' end or internally of the oligonucleotide to immobilize it to a Streptavidin column. Following hybridization, binding and

extensive washing steps, the RNA of interest is eluted by a displacement oligonucleotide, or by targeted RNase degradation (Fig. 3A). In these methods, the RNA of interest does not need a special preparation; the RNA oligonucleotide is chosen to be specific of the RNA sequence. Usually, 2'-O-alkyl (methyl or allyl) oligoribonucleotides are preferred since they are resistant to nuclease degradation in crude extracts containing high levels of endogenous RNase or DNase activity [43]. Elution of bound RNP from the antisense oligonucleotide can be achieved under non-denaturing conditions, using a displacement oligonucleotide, which can form a thermodynamically more stable duplex with the antisense oligonucleotide [44–46]. Alternatively, a DNA oligonucleotide complementary to another part of the RNA of interest is added together with RNase H which will specifically cleave the DNA/RNA duplex and induce the elution of the RNP in native conditions [47, 48].

To successfully apply this technique, an unstructured and accessible region of RNA sequence must be available to hybridize the antisense oligoribonucleotide. Therefore, this method can be difficult for highly structured RNAs and RNPs. The need for duplex formation of a previously single-stranded region of the RNA may tend to destabilize the RNP architecture. On the other hand, compared to methods that incorporate a structured RNA tag, this method avoids folding problems of the RNA due to the foreign sequence. In addition, several different hybridization sites in the RNA target can be quickly tested for accessibility to synthetic oligonucleotide probes.

3.5 RNP purification based on ligation of a DNA-oligonucleotide tag

This method brings together several advantages from all the other methods. The RNA sequence of interest is ligated to an unstructured DNA oligonucleotide also called adapter, forming a chimeric RNA-DNA molecule. In this way, standard chemically synthesized DNA oligonucleotides carrying a variety of modifications, including biotin, can be ligated to RNA. The resulting chimera can efficiently bind streptavidin resins or beads where the RNP complexes are subsequently assembled starting from crude cell preparations. Here the absence of a foreign structured tag reduces constraints on RNP formation and interaction with non-specific molecules.

3.5.1. Tagging RNA by splint-dependent ligation with T4 DNA ligase

In this version of the method, a standard biotinylated-DNA adapter is ligated by T4 DNA ligase, which is able to ligate hybrids of ribo- and deoxyribonucleotide homo-polymers in double stranded regions [49–52]. Therefore, a splint DNA oligonucleotide hybridizing both mRNA and biotinylated DNA oligonucleotide is added to form the double-strand and allow ligation (Fig. 3B). Efficient ligation of RNA/DNA duplexes requires stoichiometric or greater concentrations of T4 DNA ligase because this enzyme cannot turnover effectively on RNA-containing duplexes [51]. In addition, since T7 RNA polymerase produces significant amount of N+1 RNA products by run-off transcription, a second splint oligonucleotide is added with an extra nucleotide (mix of four nucleotides at this position) at the junction site (Fig. 4B). Prior to incubation with crude cell extracts, the splint oligonucleotide is removed from the duplex to prevent cleavage by endogenous RNase H (which cleaves RNA/DNA duplexes). To do so, the immobilized chimeric mRNAs is incubated 2 min at 95 °C to unwind mRNA-splint duplexes and then 10 min at room temperature in the presence of a 10-fold excess of an antisense splint DNA oligonucleotide containing the complementary sequence to trap the splint and avoid re-annealing on the mRNA. After these washing steps, the immobilized mRNA–DNA chimera is ready for RNP complex capture. It is

possible to accommodate 3□-sequence variety of the RNAs of interest by using a combination of splints, whereas the biotinylated oligonucleotide remains constant. The elution of RNP is performed by DNase I, which degrades the DNA moiety of the chimera (the adapter), releasing in native conditions the RNP assembled on an intact RNA without extra tag.

It is also possible to extend the 3'-end of the RNAs by a poly (CAA)$_n$ unstructured extension that favors hybridizing the splint. For the assembly of translation initiation complexes, the chimeric molecule is obtained by ligating the DNA oligonucleotide to the 3' end of the mRNA in order to enable ribosome binding to occur freely on the 5' end. However, the method can be adapted to isolate RNP complexes assembled on the 3' end of the mRNA, since nothing prevents ligation of the biotinylated DNA oligonucleotide to the 5' end of the mRNA.

The method has been used to assemble translation initiation complexes on histone H4 mRNA, hepatitis C virus (HCV) internal ribosome entry site (IRES), β-globin mRNA, and cricket paralysis virus IRES, and these complexes were suitable for cryo-EM and mass spectrometry studies [53–56].

3.5.2 Direct ligation of 5'-adenylated deoxyoligonucleotides tag by T4 RNA ligase 2

The advent of deep sequencing has led to the development of new methods to generate very large libraries of DNA molecules. To improve ligation efficiency and limit concatemer formation, the ligation reaction is split in two consecutive steps. ATP is first used to form a 5'-adenylated DNA oligonucleotide (5'-AppDNA) that is the biochemical intermediate of enzymes as DNA- or RNA-ligase. This activated intermediate is a good substrate for various enzymatic reactions, forming a building block to assemble nucleic acids under specific conditions. Then, the 5'-AppDNA molecule acts as substrate during the second step catalyzed by robust T4 ligases that work in the absence of ATP and do not concatenate products or form circles.

The first step of 5'-adenylation of the oligonucleotide is catalysed by a thermostable RNA ligase of *Methanobacterium thermoautotrophicum* (MthRnl) [57, 58]. The enzyme usually catalyzes the intramolecular ligation of single-strand RNA through ligase-adenylate and 5'-AppRNA intermediates. When using 3'-terminal blocked DNA oligonucleotides, and suitable ATP concentration, MthRnl accumulates 5'-AppDNA that can serve for efficient ligation of RNA or DNA molecules.

The second enzyme that ligates the RNA of interest to the activated 5'-AppDNA adapter is a mutated form of truncated T4 RNA ligase 2 (T4 Rn12tr). T4 Rnl2tr is defective in self-adenylation and readily accepts pre-adenylated substrates for ligation [59]. Adding the R55K and K227Q mutations to T4 Rnl2tr leads to very low background in concatenation and circularization of the RNA molecules [59]. Because the optimized enzyme does not use ATP for ligation but pre-adenylated adapters, up to 75% of efficiency is reached during ligation [59]. As stated above, this procedure is commonly used to optimize adapter ligation for high-throughput sequencing technology. Here, we recently adapted the protocol for the formation of the mRNA-DNA chimera for RNP purification, and more precisely for purification of translation initiation complexes assembled on mRNA. With this method, we are routinely reaching 50% efficiency when ligating a biotinylated DNA adapter to the 3´ OH end of the mRNA (Fig. 3C), which is much higher than the result obtained by the splint-dependent ligation with T4 DNA ligase (about 10%). The following steps of RNP assembly and washes basically follow the conventional protocols. At the end, elution is performed with RNase H after adding a complementary DNA oligonucleotide that forms the RNase H cleavage site in the 3' part of the mRNA. We found that RNase H was less sensitive than DNase I to steric clashes with the matrix beads and larger amounts of RNP are recovered when compared with the DNase elution.

3.6 Outlook

In the past few years, many approaches targeting the assembly of RNPs from cell lysates have been described. The capture of RNPs assembled *in vitro* remains valuable for the structural or biochemical analysis of a large variety of complexes and systems, including the spliceosome, ribosome, and other smaller RNPs such as the telomerase and RNase P. A critical step of these purification processes remains the capture of the RNA sequence of interest to a solid support that will allow separation of the total cell extract. This step relies on the presence, in the RNA target, of an external element with affinity for the solid matrix. Various types or elements have been selected and tested. Synthetic RNA aptamers have been artificially selected to target a specific molecule. Once bound to a solid support this molecule can be used for affinity purification of the RNP. Other elements are natural RNA sequences exhibiting specific binding sites for well-characterized RNA-binding proteins (phage elements). These methods are attractive because RNA aptamers are generally small, directly bind bead matrixes, and are eluted with small molecules. However to successfully tag a given RNA, the aptamer should be placed in a solvent-accessible region. It is important to consider the steric effect of using such aptamers and the potential interference with the folding, structure, assembly and normal function or interactions of the RNP under study. Conversely, the RNP assembly may disrupt aptamer folding and therefore its ability to bind the solid support. Therefore after designing the hybrid RNA it may be useful to predict the folding of the sequence. In case attempts to purify RNP complexes all fail, it might be necessary to reiterate the design of the construct by trying different tags, changing their location in the RNA or adding flexible spacers that would prevent steric hindrance.

The use of biotinylated oligonucleotides, which are complementary to accessible single-stranded regions or ligated to the RNA of interest, is an efficient alternative approach when the RNA structure is too constrained and incompatible with the aptamer approach. Compared to aptamers, biotin is a discrete molecule that only slightly modifies the RNP complex. The biotinylated RNA, in the form of a DNA-RNA hybrid or ligated chimera, exhibits high affinity to streptavidin affinity resins. Efficient assembly and stringent washing steps allow the decrease of the background of contaminating proteins.

In addition, elution of the RNP complexes can be achieved under native conditions through the use of competing oligonucleotides or DNase/RNase which releases the RNP without extra RNA sequence.

In conclusion, we have presented here several general strategies for selecting RNPs *in vitro* assembled in crude cell extracts and exhibiting *in vivo* potential application. Although these methods have been developed for the preparation of specific complexes, most of them have been successfully applied to the preparation of RNP involved in fundamental cellular processes.

Acknowledgments The authors are supported by institutional funds from the Centre National de la Recherche Scientifique (CNRS), the Université de Strasbourg, and grant ANR-17-CE11-0024-01 and ANR-17-CE12-0025-01.

References

1. Rouault TA, Hentze MW, Haile DJ, Harford JB, Klausner RD (1989) The iron-responsive element binding protein: a method for the affinity purification of a regulatory RNA-binding protein. Proc Natl Acad Sci U S A 86:5768–5772
2. Neupert B, Thompson NA, Meyer C, Kuhn LC (1990) A high yield affinity purification method for specific RNA-binding proteins: isolation of the iron regulatory factor from human placenta. Nucleic Acids Res 18:51–55
3. Bardwell VJ, Wickens M (1990) Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method. Nucleic Acids Res 18:6587–6594
4. Lowary PT, Uhlenbeck OC (1987) An RNA mutation that increases the affinity of an RNA–protein interaction. Nucleic Acids Res 15:10483–10493
5. Bachler M, Schroeder R, von Ahsen U (1999) StreptoTag: a novel method for the isolation of RNA-binding proteins. RNA 5:1509–1516

6. Srisawat C, Engelke DR (2001) Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. RNA 7:632–641
7. Srisawat C, Goldstein IJ, Engelke DR (2001) Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures. Nucleic Acids Res 29:E4
8. Hartmuth K, et al. (2002) Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. Proc Natl Acad Sci U S A 99:16719–16724
9. Panchapakesan SSS, et al. (2017) Ribonucleoprotein purification and characterization using RNA Mango. RNA 23:1592–1599
10. Das R, Zhou Z, Reed R (2000) Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. Mol Cell 5:779–787
11. Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ (2002) Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. RNA 8:426–439
12. Ji H, Fraser CS, Yu Y, Leary J, Doudna JA (2004) Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. Proc Natl Acad Sci U S A 101:16990–16995
13. Zhou Z, Sim J, Griffith J, Reed R (2002) Purification and electron microscopic visualization of functional human spliceosomes. Proc Natl Acad Sci U S A 99:12203–12207
14. Locker N, Easton LE, Lukavsky PJ (2006) Affinity purification of eukaryotic 48S initiation complexes. RNA 12:683–690
15. Wang Y, Rando RR (1995) Specific binding of aminoglycoside antibiotics to RNA. Chem Biol 2:281–290
16. Hartmuth K, Vornlocher H-P, Lührmann R (2004) Tobramycin affinity tag purification of spliceosomes. Methods Mol Biol 257:47–64
17. Vasudevan S, Steitz JA (2007) AU-Rich-Element-Mediated Upregulation of Translation by FXR1 and Argonaute 2. Cell 128:1105–1118
18. Iioka H, Loiselle D, Haystead TA, Macara IG (2011) Efficient detection of RNA-protein interactions using tethered RNAs. Nucleic Acids Res 39:e53
19. Dix CI, et al. (2013) Lissencephaly-1 promotes the recruitment of dynein and dynactin to transported mRNAs. J Cell Biol 202:479–494
20. Leppek K, Stoecklin G (2014) An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. Nucleic Acids Res 42:e13
21. Srisawat C, Engelke DR (2002) RNA affinity tags for purification of RNAs and ribonucleoprotein complexes. Methods 26:156–161
22. Deckert J, et al. (2006) Protein Composition and Electron Microscopy Structure of Affinity-Purified Human Spliceosomal B Complexes Isolated under Physiological Conditions. Mol Cell Biol 26(14):5528–5543
23. Kastner B, et al. (2008) GraFix: sample preparation for single-particle electron cryomicroscopy. Nat Methods 5:53–55
24. Lim F, Downey TP, Peabody DS (2001) Translational repression and specific RNA binding by the coat protein of the Pseudomonas phage PP7. J Biol Chem 276:22507–22513
25. Lim F, Peabody DS (2002) RNA recognition site of PP7 coat protein. Nucleic Acids Res 30:4138–4144
26. Hogg JR, Collins K (2007) RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. RNA 13:868–880
27. Hogg JR, Goff SP (2010) Upf1 Senses 3□UTR Length to Potentiate mRNA Decay. Cell 143:379–389
28. Fritz SE, Haque N, Hogg JR (2018) Highly efficient in vitro translation of authentic affinity-purified messenger ribonucleoprotein complexes. RNA 24:982–989
29. Baron-Benhamou J, Gehring NH, Kulozik AE, Hentze MW (2004) Using the λN Peptide to Tether Proteins to RNAs. MRNA Processing and Metabolism (Humana Press, New Jersey), pp 135–154
30. Gehring NH, Hentze MW, Kulozik AE (2008) Chapter 23 Tethering Assays to Investigate Nonsense□Mediated mRNA Decay Activating Proteins. Methods in Enzymology, pp 467–482
31. Ramanathan M, et al. (2018) RNA–protein interaction detection in living cells. Nat Methods 15:207–212
32. Piekna-Przybylska D, Liu B, Fournier MJ (2007) The U1 snRNA hairpin II as a RNA affinity tag for selecting snoRNP complexes. Methods Enzymol 425:317–353
33. LeCuyer KA, Behlen LS, Uhlenbeck OC (1995) Mutants of the bacteriophage MS2 coat protein that alter its cooperative binding to RNA. Biochemistry 34:10600–10606
34. Ji H, Fraser CS, Yu Y, Leary J, Doudna JA (2004) Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. Proc Natl Acad Sci U S A 101:16990–16995
35. Bertram K, et al. (2017) Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. Cell 170:701–713.e11
36. Yoon J-H, Srikantan S, Gorospe M (2012) MS2-TRAP (MS2-tagged RNA affinity purification): tagging RNA to identify associated miRNAs. Methods 58:81–87
37. Beach DL, Keene JD (2008) Ribotrap□: targeted purification of RNA-specific RNPs from cell lysates through immunoaffinity precipitation to identify regulatory proteins and RNAs. Methods Mol Biol 419:69–91
38. Tsai BP, Wang X, Huang L, Waterman ML (2011) Quantitative profiling of in vivo-

assembled RNA-protein complexes using a novel integrated proteomic approach. Mol Cell Proteomics 10(4):M110.007385
39. Slobodin B, Gerst JE (2010) A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. RNA 16:2277–2290
40. Bertrand E, et al. (1998) Localization of ASH1 mRNA particles in living yeast. Mol Cell 2:437–445
41. Beach DL, Salmon ED, Bloom K (1999) Localization and anchoring of mRNA in budding yeast. Curr Biol 9:569–578
42. Haim L, Zipor G, Aronov S, Gerst JE (2007) A genomic integration method to visualize localization of endogenous mRNAs in living yeast. Nat Methods 4:409–412
43. Iribarren AM, et al. (1990) 2'-O-alkyl oligoribonucleotides as antisense probes. Proc Natl Acad Sci U S A 87:7747–7751
44. Lingner J, Cech TR (1996) Purification of telomerase from *Euplotes aediculatus*: requirement of a primer 3' overhang. Proc Natl Acad Sci U S A 93:10712–10717
45. Schnapp G, Rodi HP, Rettig WJ, Schnapp A, Damm K (1998) One-step affinity purification protocol for human telomerase. Nucleic Acids Res 26:3311–3313
46. Kurth I, Cristofari G, Lingner J (2008) An affinity oligonucleotide displacement strategy to purify ribonucleoprotein complexes applied to human telomerase. Methods Mol Biol 488:9–22
47. Namy O, Moran SJ, Stuart DI, Gilbert RJ, Brierley I (2006) A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. Nature 441:244–247
48. Eliseev B, et al. (2018) Structure of a human cap-dependent 48S translation pre-initiation complex. Nucleic Acids Res 46:2678-2689
49. Kleppe K, Van de Sande JH, Khorana HG (1970) Polynucleotide ligase-catalyzed joining of deoxyribo-oligonucleotides on ribopolynucleotide templates and of ribo-oligonucleotides on deoxyribopolynucleotide templates. Proc Natl Acad Sci U S A 67:68–73
50. Moore MJ, Sharp PA (1992) Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. Science 256:992–997
51. Moore MJ, Query CC (2000) Joining of RNAs by splinted ligation. Methods Enzym 317:109–123
52. Fareed GC, Wilt EM, Richardson CC (1971) Enzymatic breakage and joining of deoxyribonucleic acid. 8. Hybrids of ribo- and deoxyribonucleotide homopolymers as substrates for polynucleotide ligase of bacteriophage T4. J Biol Chem 246:925–932
53. Prongidi-Fix L, et al. (2013) Rapid purification of ribosomal particles assembled on histone H4 mRNA: a new method based on mRNA-DNA chimaeras. Biochem J 449:719–28
54. Chicher J, et al. (2015) Purification of mRNA-programmed translation initiation complexes suitable for mass spectrometry analysis. Proteomics 15:2417–2425
55. Martin F, et al. (2016) Ribosomal 18S rRNA base pairs with mRNA during eukaryotic translation initiation. Nat Commun 7:12622
56. Gross L, et al. (2017) The IRES 5□UTR of the dicistrovirus cricket paralysis virus is a type III IRES containing an essential pseudoknot structure. Nucleic Acids Res 45:8993–9004
57. Torchia C, Takagi Y, Ho CK (2008) Archaeal RNA ligase is a homodimeric protein that catalyzes intramolecular ligation of single-stranded RNA and DNA. Nucleic Acids Res 36:6218–6227
58. Zhelkovsky AM, McReynolds LA (2012) Structure-function analysis of *Methanobacterium thermoautotrophicum* RNA ligase – engineering a thermostable ATP independent enzyme. BMC Mol Biol 13:24
59. Viollet S, Fuchs RT, Munafo DB, Zhuang F, Robb GB (2011) T4 RNA ligase 2 truncated active site mutants: improved tools for RNA analysis. BMC Biotechnol 11:72

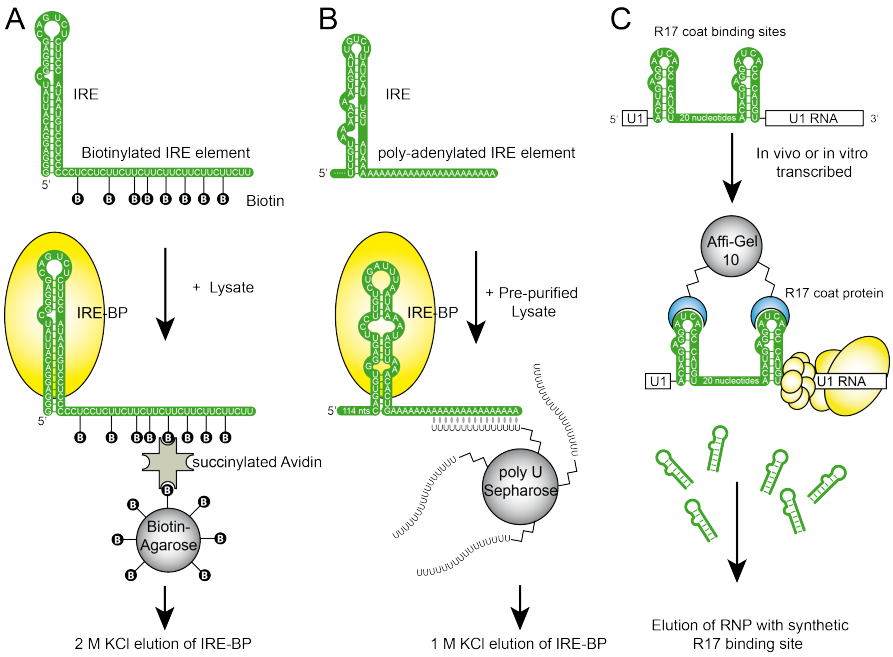


Fig. 3.1 Outline of the three methods of purification of a specific RNA protein complex from a mixture of proteins. (**A**) Schematic representation of the affinity purification of IRE-BP based on Biotin-Avidin affinity. The allylamine uridine triphosphate was randomly incorporated throughout the entire transcript before coupling with succinimidyl-biotin. Bound IRE-BP was subsequently eluted with 2 M KCI. IRE, Iron Responsive Element; IRE-BP, IRE-Binding protein. (**B**) Representation of the affinity purification of IRE-BP based on polyU resin. In vitro transcribed polyadenylated RNA was bound to poly(U)-Sepharose. Prepurified IRE-BP was specifically adsorbed to the IRE and subsequently eluted with 1 M KCI. The IRE here shown is one of the transferrin IRE elements. (**C**) Affinity purification of snRNP U1 based on the affinity of R17 coat protein for its RNA-binding element. A chimeric RNA containing two R17 recognition sites inserted in U1 RNA was prepared either *in vitro*, by transcription with a phage polymerase, or *in vivo*, by cellular transcription of a transfected or injected DNA template. The chimeric RNA binds to appropriate factors in the cell extract. The resulting complexes are selectively retained on a support to which R17 coat protein has been covalently coupled. Specific RNA molecules and any associated factors are eluted with an excess of R17 recognition elements

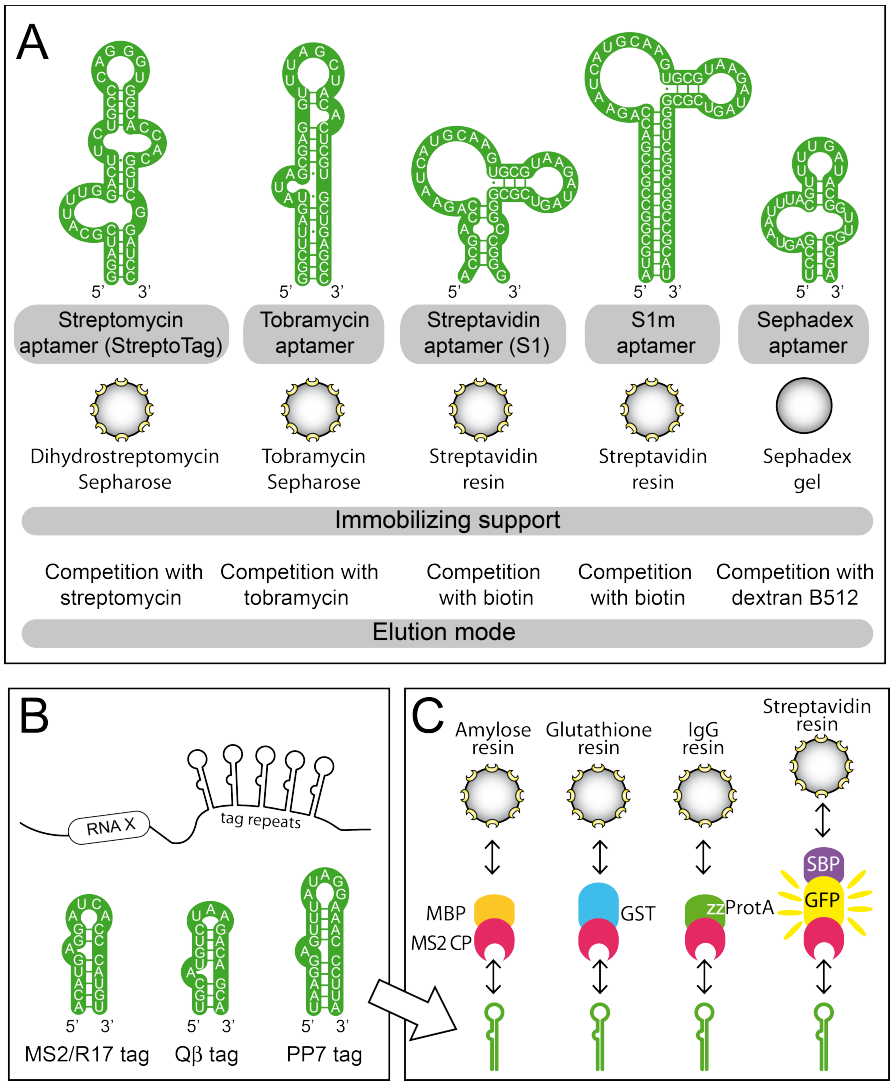


Fig. 3.2 Strategies of purification of RNPs based on artificial RNA aptamers or natural tags. (**A**) Sequences and predicted secondary structures of different RNA aptamers that have been used to develop purification strategies. The streptomycin RNA aptamer sequence (StreptoTag) binds with high affinity to dihydrostreptomycin coupled to a sepharose column matrix. The tobramycin aptamer binds efficiently to tobramycin-derivatized sepharose beads. The streptavidin aptamer (S1) specifically binds to streptavidin conjugated to many commercial supports. S1 was optimized in S1m that can be repeated (4xS1m). Elution is obtained under mild conditions by adding biotin. The Sephadex aptamer is specific to the Sephadex matrix. Elution is performed by dextran B512, which is the base material of Sephadex. (**B**) Several natural sequences from bacteriophages have been used as tags to affinity-purify RNPs. The MS2 hairpin is the most frequently used, with a variable number of copies in the RNA target. (**C**) Since the affinity of the repeated MS2 sequence for the coat protein (CP) is too high for an efficient elution, several fusion proteins have been engineered in order to affinity-bind the MS2 tag on one side (CP) and additional peptides with different binding properties. The Maltose-Binding Peptide (MBP) binds to amylose columns and release of the complexes is obtained by adding maltose or by protease cleavage at specific sites inserted between the fused domains. Other domains have been fused with the CP domain. The GST domain binds to glutathione resin, the zz domain of protein A (zzProtA) binds to IgG coated beads and the fusion with the Streptavidin-Binding Peptide (SBP) allows binding to streptavidin coated resins whereas the GFP domain adds fluorescent properties. In all that cases, co-expression or co-purification of the fusion proteins is needed

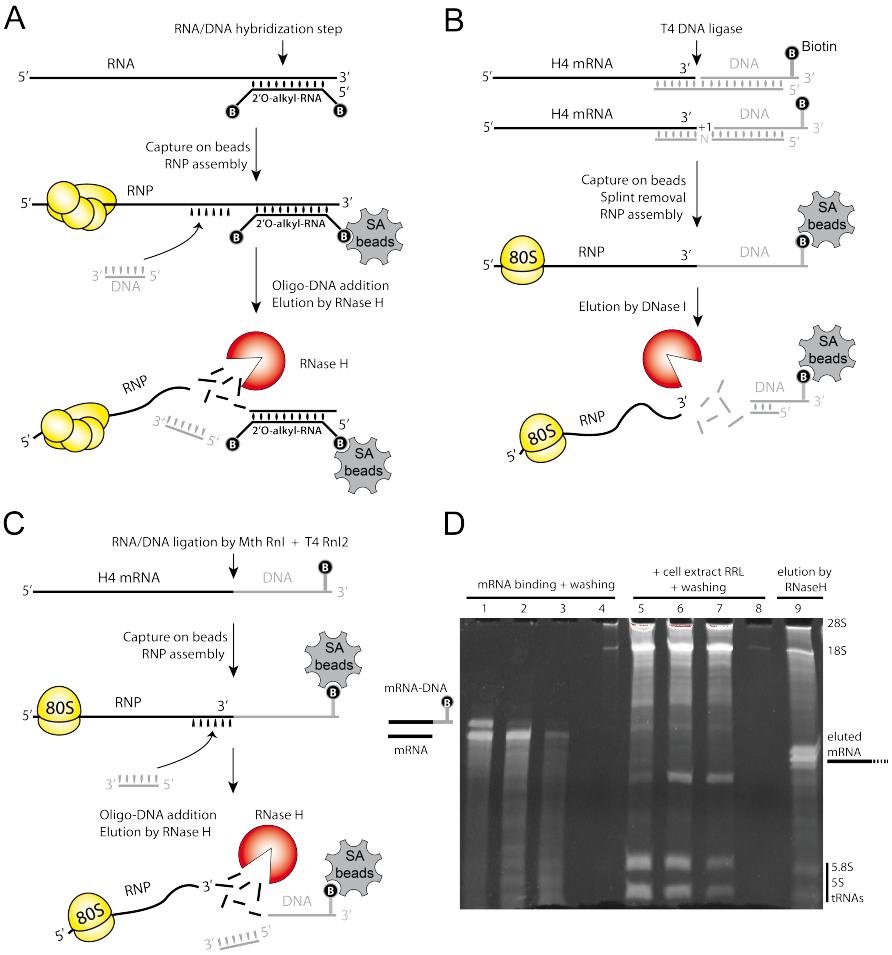


Fig. 3.3 Strategies of purification of RNPs based on hybridization and ligation of biotinylated oligonucleotides. (**A**) Affinity purification by oligonucleotide antisense and elution with RNase H (Lower). The RNP of interest is assembled to the RNA immobilized via an antisense biotinylated RNA oligonucleotide (grey color) to streptavidin beads. The antisense oligonucleotide contains 2'-0-alkyl sugars in order to resist to degradation in cell extracts. After washing, the elution of the RNP is obtained by adding of a DNA oligonucleotide, which forms the cleavage site for RNase H. (**B**) This method is based on the formation of RNA–DNA chimeric molecules that resist to the RNase H activity present in cell extracts. T4 DNA ligase catalyses the ligation of RNA-DNA hybrids if these are joined by base pairing to a "splint oligonucleotide". To circumvent the problem of imperfect pairing due to heterogeneous 3□ ends due to the tendency of RNA polymerases to incorporate *in vitro* non-encoded nucleotides, a degenerated splint carrying a mixture of all four nts is used. After removal of the splint oligo-mixture, RNPs are assembled in rabbit reticulocytes lysate (RRL). Release of the complexes is performed with DNase I or RQ1 DNase, which specifically cleave the DNA part of the chimera. (**C**) In this approach, histone H4 mRNA is ligated to a biotinylated DNA oligonucleotide (in grey) in two successive steps using *M. thermoautotrophicum* RNA ligase (MthRnl) and truncated T4 RNA ligase, forming a chimeric RNA/DNA molecule. Assembly of translation initiation complexes on the mRNA is performed in RRL. Elution is performed with RNase H as in panel A. (**D**) Representative denaturing urea PAGE gel showing the purification of 80S particles assembled on H4 mRNA by the method described in panel C. The chimeric mRNA–biotinylated DNA (lane 1) was loaded on the beads (7 pmoles are shown, 50% of the mRNA molecules were ligated to the biotinylated DNA oligonucleotide). Unbound mRNAs are shown in the flow-through (lane 2) and washing fractions (lanes 3-4). Translation-initiation complexes were assembled in RRL. Unbound rRNAs and tRNAs are shown in the flow-through (lane 5) and washing fractions (lanes 6-8). The bound 80S particles were eluted with RNase H (lane 9) after adding the DNA oligonucleotide forming the RNase H cleavage site. The positions of the ribosomal RNAs, tRNAs and histone H4 mRNA are indicated

BIBLIOGRAPHIE

BIBLIOGRAPHIE

Ahlemann, M., Zeidler, R., Lang, S., Mack, B., Münz, M., and Gires, O. (2006). Carcinoma-associated eIF3i overexpression facilitates mTOR-dependent growth transformation. Mol. Carcinog. *45*, 957–967.

Aitken, C.E., Beznosková, P., Vlčkova, V., Chiu, W.-L., Zhou, F., Valášek, L.S., Hinnebusch, A.G., and Lorsch, J.R. (2016). Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. Elife *5*.

Alard, A., Marboeuf, C., Fabre, B., Jean, C., Martineau, Y., Lopez, F., Vende, P., Poncet, D., Schneider, R.J., Bousquet, C., et al. (2019). Differential Regulation of the Three Eukaryotic mRNA Translation Initiation Factor (eIF) 4Gs by the Proteasome. Front. Genet. *10*, 254.

Arake de Tacca, L.M., Pulos-Holmes, M.C., Floor, S.N., and Cate, J.H.D. (2019). PTBP1 mRNA isoforms and regulation of their translation. RNA rna.070193.118.

Arents, G., and Moudrianakis, E.N. (1995). The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. Proc. Natl. Acad. Sci. U. S. A. *92*, 11170–11174.

Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E., and Moudrianakis, E.N. (1991). The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc. Natl. Acad. Sci. U. S. A. *88*, 10148–10152.

Asano, K., Clayton, J., Shalev, A., and Hinnebusch, A.G. (2000). A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. Genes Dev. *14*, 2534–2546.

Au, H.H., and Jan, E. (2014). Novel viral translation strategies. Wiley Interdiscip. Rev. RNA *5*, 779–801.

Aylett, C.H.S., Boehringer, D., Erzberger, J.P., Schaefer, T., and Ban, N. (2015). Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. Nat. Struct. Mol. Biol. *22*, 269–271.

Banaszynski, L.A., Allis, C.D., and Lewis, P.W. (2010). Histone Variants in Metazoan Development. Dev. Cell *19*, 662–674.

Bandyopadhyay, A., Lakshmanan, V., Matsumoto, T., Chang, E.C., and Maitra, U. (2002). Moe1 and spInt6, the fission yeast homologues of mammalian translation initiation factor 3 subunits p66 (eIF3d) and p48 (eIF3e), respectively, are required for stable association of eIF3 subunits. J. Biol. Chem. *277*, 2360–2367.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Barth-Baus, D., Bhasker, C.R., Zoll, W., and Merrick, W.C. (2013). Influence of translation factor activities on start site selection in six different mRNAs. Transl. (Austin, Tex.) *1*, e24419.

Beckmann, K., Grskovic, M., Gebauer, F., and Hentze, M.W. (2005). A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in Drosophila. Cell *122*, 529–540.

Ben-Asouli, Y., Banai, Y., Pel-Or, Y., Shir, A., and Kaempfer, R. (2002). Human Interferon-γ mRNA Autoregulates Its Translation through a Pseudoknot that Activates the Interferon-Inducible Protein Kinase PKR. Cell *108*, 221–232.

Beznosková, P., Wagner, S., Jansen, M.E., von der Haar, T., and Valášek, L.S. (2015). Translation initiation factor eIF3 promotes programmed stop codon readthrough. Nucleic Acids Res. *43*, 5099–5111.

Bhardwaj, U., Powell, P., and Goss, D.J. (2019). Eukaryotic initiation factor (eIF) 3 mediates Barley Yellow Dwarf Viral mRNA 3'-5' UTR interactions and 40S ribosomal subunit binding to facilitate cap-independent translation. Nucleic Acids Res. *47*, 6225–6235.

Bhat, M., Robichaud, N., Hulea, L., Sonenberg, N., Pelletier, J., and Topisirovic, I. (2015). Targeting the translation machinery in cancer. Nat. Rev. Drug Discov. *14*, 261–278.

Bostock, C.J., Prescott, D.M., and Kirkpatrick, J.B. (1971). An evaluation of the double thymidine block for synchronizing mammalian cells at the G1-S border. Exp. Cell Res. *68*, 163–168.

Browning, K.S., Gallie, D.R., Hershey, J.W., Hinnebusch, A.G., Maitra, U., Merrick, W.C., and Norbury, C. (2001). Unified nomenclature for the subunits of eukaryotic initiation factor 3. Trends Biochem. Sci. *26*, 284.

Bukhari, S.I.A., Truesdell, S.S., Lee, S., Kollu, S., Classon, A., Boukhali, M., Jain, E., Mortensen, R.D., Yanagiya, A., Sadreyev, R.I., et al. (2016). A Specialized Mechanism of Translation Mediated by FXR1a-Associated MicroRNP in Cellular Quiescence. Mol. Cell *61*, 760–773.

Bushell, M., Wood, W., Carpenter, G., Pain, V.M., Morley, S.J., and Clemens, M.J. (2001). Disruption of the interaction of mammalian protein synthesis eukaryotic initiation factor 4B with the poly(A)-binding protein by caspase- and viral protease-mediated cleavages. J. Biol. Chem. *276*, 23922–23928.

Cai, Q., Todorovic, A., Andaya, A., Gao, J., Leary, J.A., and Cate, J.H.D. (2010). Distinct regions of human eIF3 are sufficient for binding to the HCV IRES and the 40S ribosomal subunit. J. Mol. Biol. *403*, 185–196.

Cakmakci, N.G., Lerner, R.S., Wagner, E.J., Zheng, L., and Marzluff, W.F. (2008). SLIP1, a Factor Required for Activation of Histone mRNA Translation by the Stem-Loop Binding

Protein. Mol. Cell. Biol. *28*, 1182–1194.

Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc. Natl. Acad. Sci. U. S. A. *106*, 7507–7512.

Cammas, A., and Millevoi, S. (2016). RNA G-quadruplexes: emerging mechanisms in disease. Nucleic Acids Res. *45*, gkw1280.

Chakravarthy, S., Park, Y.-J., Chodaparambil, J., Edayathumangalam, R.S., and Luger, K. (2005). Structure and dynamic properties of nucleosome core particles. FEBS Lett. *579*, 895–898.

Chapat, C., Jafarnejad, S.M., Matta-Camacho, E., Hesketh, G.G., Gelbart, I.A., Attig, J., Gkogkas, C.G., Alain, T., Stern-Ginossar, N., Fabian, M.R., et al. (2017). Cap-binding protein 4EHP effects translation silencing by microRNAs. Proc. Natl. Acad. Sci. *114*, 5425–5430.

Chaudhuri, J., Chowdhury, D., and Maitra, U. (1999). Distinct functions of eukaryotic translation initiation factors eIF1A and eIF3 in the formation of the 40 S ribosomal preinitiation complex. J. Biol. Chem. *274*, 17975–17980.

Cho, P.F., Poulin, F., Cho-Park, Y.A., Cho-Park, I.B., Chicoine, J.D., Lasko, P., and Sonenberg, N. (2005). A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell *121*, 411–423.

Chu, J., Cargnello, M., Topisirovic, I., and Pelletier, J. (2016). Translation Initiation Factors: Reprogramming Protein Synthesis in Cancer. Trends Cell Biol. *26*, 918–933.

Cohen-Chalamish, S., Hasson, A., Weinberg, D., Namer, L.S., Banai, Y., Osman, F., and Kaempfer, R. (2009). Dynamic refolding of IFN-gamma mRNA enables it to function as PKR activator and translation template. Nat. Chem. Biol. *5*, 896–903.

Collart, M.A. (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip. Rev. RNA *7*, 438–454.

Colón-Ramos, D.A., Shenvi, C.L., Weitzel, D.H., Gan, E.C., Matts, R., Cate, J., and Kornbluth, S. (2006). Direct ribosomal binding by a cellular inhibitor of translation. Nat. Struct. Mol. Biol. *13*, 103–111.

Conroy, S.C., Dever, T.E., Owens, C.L., and Merrick, W.C. (1990). Characterization of the 46,000-Dalton subunit of eIF-4F. Arch. Biochem. Biophys. *282*, 363–371.

Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M., and Borden, K.L.B. (2006). eIF4E is a central node of an RNA regulon that governs cellular proliferation. J. Cell Biol. *175*, 415–426.

D'Orchymont, A. (2013). Etudes structurales de l'ARN messager de l'histone H4.

Dabrowski, M., Bukowy-Bieryllo, Z., and Zietkiewicz, E. (2015). Translational readthrough

potential of natural termination codons in eucaryotes--The impact of RNA sequence. RNA Biol. *12*, 950–958.

Damoc, E., Fraser, C.S., Zhou, M., Videler, H., Mayeur, G.L., Hershey, J.W.B., Doudna, J.A., Robinson, C. V, and Leary, J.A. (2007). Structural characterization of the human eukaryotic initiation factor 3 protein complex by mass spectrometry. Mol. Cell. Proteomics *6*, 1135–1146.

Das, S., Ghosh, R., and Maitra, U. (2001). Eukaryotic translation initiation factor 5 functions as a GTPase-activating protein. J. Biol. Chem. *276*, 6720–6726.

Davuluri, R. V. (2000). CART Classification of Human 5' UTR Sequences. Genome Res. *10*, 1807–1816.

Deniz, N., Lenarcic, E.M., Landry, D.M., and Thompson, S.R. (2009). Translation initiation factors are not required for Dicistroviridae IRES function in vivo. RNA *15*, 932–946.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb. Perspect. Biol. *4*, a013706.

Diallo, L.H., Tatin, F., David, F., Godet, A.-C., Zamora, A., Prats, A.-C., Garmy-Susini, B., and Lacazette, E. (2019). How are circRNAs translated by non-canonical initiation mechanisms? Biochimie.

Doldan, A., Chandramouli, A., Shanas, R., Bhattacharyya, A., Cunningham, J.T., Nelson, M.A., and Shi, J. (2008). Loss of the eukaryotic initiation factor 3f in pancreatic cancer. Mol. Carcinog. *47*, 235–244.

Dominski, Z., and Marzluff, W.F. (1999). Formation of the 3' end of histone mRNA. Gene *239*, 1–14.

Dominski, Z., Yang, X., and Marzluff, W.F. (2005). The Polyadenylation Factor CPSF-73 Is Involved in Histone-Pre-mRNA Processing. Cell *123*, 37–48.

Dong, Z., and Zhang, J.-T. (2006). Initiation factor eIF3 and regulation of mRNA translation, cell growth, and cancer. Crit. Rev. Oncol. Hematol. *59*, 169–180.

ElAntak, L., Tzakos, A.G., Locker, N., and Lukavsky, P.J. (2007). Structure of eIF3b RNA recognition motif and its interaction with eIF3j: structural insights into the recruitment of eIF3b to the 40 S ribosomal subunit. J. Biol. Chem. *282*, 8165–8174.

Elfakess, R., Sinvani, H., Haimov, O., Svitkin, Y., Sonenberg, N., and Dikstein, R. (2011). Unique translation initiation of mRNAs-containing TISU element. Nucleic Acids Res. *39*, 7598–7609.

Ellisdon, A.M., and Stewart, M. (2012). Structural biology of the PCI-protein fold. Bioarchitecture *2*, 118–123.

Enchev, R.I., Schreiber, A., Beuron, F., and Morris, E.P. (2010). Structural insights into the COP9 signalosome and its common architecture with the 26S proteasome lid and eIF3.

Structure *18*, 518–527.

Eriani, G., and Martin, F. (2018). START: STructure-Assisted RNA Translation. RNA Biol. *15*, 1250–1253.

Erzberger, J.P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C.H.S., Cimermančič, P., Boehringer, D., Sali, A., Aebersold, R., et al. (2014). Molecular Architecture of the 40S·eIF1·eIF3 Translation Initiation Complex. Cell *159*, 1227–1228.

Fan, Y., and Guo, Y. (2015). Knockdown of eIF3D inhibits breast cancer cell proliferation and invasion through suppressing the Wnt/β-catenin signaling pathway. Int. J. Clin. Exp. Pathol. *8*, 10420–10427.

Fernández-Miranda, G., and Méndez, R. (2012). The CPEB-family of proteins, translational control in senescence and cancer. Ageing Res. Rev. *11*, 460–472.

Fernández, I.S., Bai, X.-C., Murshudov, G., Scheres, S.H.W., and Ramakrishnan, V. (2014). Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. Cell *157*, 823–831.

Fraser, C.S., and Doudna, J.A. (2007). Structural and mechanistic insights into hepatitis C viral translation initiation. Nat. Rev. Microbiol. *5*, 29–38.

Fraser, C.S., Lee, J.Y., Mayeur, G.L., Bushell, M., Doudna, J.A., and Hershey, J.W.B. (2004). The j-Subunit of Human Translation Initiation Factor eIF3 Is Required for the Stable Binding of eIF3 and Its Subcomplexes to 40 S Ribosomal Subunits *in Vitro*. J. Biol. Chem. *279*, 8946–8956.

Fraser, C.S., Berry, K.E., Hershey, J.W.B., and Doudna, J.A. (2007). eIF3j Is Located in the Decoding Center of the Human 40S Ribosomal Subunit. Mol. Cell *26*, 811–819.

Gandin, V., Masvidal, L., Hulea, L., Gravel, S.-P., Cargnello, M., McLaughlan, S., Cai, Y., Balanathan, P., Morita, M., Rajakumar, A., et al. (2016). nanoCAGE reveals 5' UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs. Genome Res. *26*, 636–648.

Gebauer, F., and Hentze, M.W. (2004). Molecular mechanisms of translational control. Nat. Rev. Mol. Cell Biol. *5*, 827–835.

des Georges, A., Dhote, V., Kuhn, L., Hellen, C.U.T., Pestova, T. V, Frank, J., and Hashem, Y. (2015). Structure of mammalian eIF3 in the context of the 43S preinitiation complex. Nature *525*, 491–495.

Gillis, L.D., and Lewis, S.M. (2013). Decreased eIF3e/Int6 expression causes epithelial-to-mesenchymal transition in breast epithelial cells. Oncogene *32*, 3598–3605.

Gingras, A.-C., Raught, B., and Sonenberg, N. (1999). eIF4 Initiation Factors: Effectors of mRNA Recruitment to Ribosomes and Regulators of Translation. Annu. Rev. Biochem.

68, 913–963.

Goh, S.-H., Hong, S.-H., Hong, S.-H., Lee, B.-C., Ju, M.-H., Jeong, J.-S., Cho, Y.-R., Kim, I.-H., and Lee, Y.-S. (2011). eIF3m expression influences the regulation of tumorigenesis-related genes in human colon cancer. Oncogene *30*, 398–409.

Gomes-Duarte, A., Lacerda, R., Menezes, J., and Romão, L. (2018). eIF3: a factor for human health and disease. RNA Biol. *15*, 26–34.

Gradi, A., Svitkin, Y. V, Imataka, H., and Sonenberg, N. (1998). Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. Proc. Natl. Acad. Sci. U. S. A. *95*, 11089–11094.

Greenberg, J.R., Phan, L., Gu, Z., deSilva, A., Apolito, C., Sherman, F., Hinnebusch, A.G., and Goldfarb, D.S. (1998). Nip1p Associates with 40 S Ribosomes and the Prt1p Subunit of Eukaryotic Initiation Factor 3 and Is Required for Efficient Translation Initiation. J. Biol. Chem. *273*, 23485–23494.

Gribling-Burrer, A.-S., Leichter, M., Wurth, L., Huttin, A., Schlotter, F., Troffer-Charlier, N., Cura, V., Barkats, M., Cavarelli, J., Massenet, S., et al. (2017). SECIS-binding protein 2 interacts with the SMN complex and the methylosome for selenoprotein mRNP assembly and translation. Nucleic Acids Res. *45*, 5399–5413.

Gross, L. (2018). Le piratage de l'appareil de traduction de la cellule hôte par le virus de la paralysie du cricket (CrPV). Strasbourg.

Gross, L., Schaeffer, L., Alghoul, F., Hayek, H., Allmang, C., Eriani, G., and Martin, F. (2018). Tracking the m^7G-cap during translation initiation by crosslinking methods. Methods *137*.

Grzmil, M., Rzymski, T., Milani, M., Harris, A.L., Capper, R.G., Saunders, N.J., Salhan, A., Ragoussis, J., and Norbury, C.J. (2010). An oncogenic role of eIF3e/INT6 in human breast cancer. Oncogene *29*, 4080–4089.

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 *559*, 130–134.

H. Hayek; L. Gross; F. Alghoul; F. Martin; G. Eriani; and C. Allmang Immunoprecipitation methods to isolate messenger ribonucleoprotein complexes (mRNP). In Advanced Technologies for Protein Complex Production and Characterization, M.C. Vega, ed. (In press), p.

Haghighat, A., Svitkin, Y., Novoa, I., Kuechler, E., Skern, T., and Sonenberg, N. (1996). The eIF4G-eIF4E complex is the target for direct cleavage by the rhinovirus 2A proteinase. J. Virol. *70*, 8444–8450.

Haimov, O., Sinvani, H., Martin, F., Ulitsky, I., Emmanuel, R., Tamarkin-Ben-Harush, A., Vardy, A., and Dikstein, R. (2017). Efficient and Accurate Translation Initiation Directed by TISU Involves RPS3 and RPS10e Binding and Differential Eukaryotic Initiation Factor 1A Regulation. Mol. Cell. Biol. *37,* 00150-17

Halder, K., Wieland, M., and Hartig, J.S. (2009). Predictable suppression of gene expression by 5'-UTR-based RNA quadruplexes. Nucleic Acids Res. *37*, 6811–6817.

Hanachi, P., Hershey, J.W.B., and Vornlocher, H.-P. (1999). Characterization of the p33 Subunit of Eukaryotic Translation Initiation Factor-3 from *Saccharomyces cerevisiae*. J. Biol. Chem. *274*, 8546–8553.

Harcourt, E.M., Kietrys, A.M., and Kool, E.T. (2017). Chemical and structural effects of base modifications in messenger RNA. Nature *541*, 339–346.

Harel-Sharvit, L., Eldad, N., Haimovich, G., Barkai, O., Duek, L., and Choder, M. (2010). RNA Polymerase II Subunits Link Transcription and mRNA Decay to Translation. Cell *143*, 552–563.

Harris, M.E., Böhni, R., Schneiderman, M.H., Ramamurthy, L., Schümperli, D., and Marzluff, W.F. (1991). Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. Mol. Cell. Biol. *11*, 2416–2424.

Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., Hellen, C.U.T., Pestova, T. V, and Frank, J. (2013a). Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. Cell *153*, 1108–1119.

Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., Pestova, T. V, Hellen, C.U.T., and Frank, J. (2013b). Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. Nature *503*, 539–543.

Herrmannová, A., Daujotyte, D., Yang, J.-C., Cuchalová, L., Gorrec, F., Wagner, S., Dányi, I., Lukavsky, P.J., and Valásek, L.S. (2012). Structural analysis of an eIF3 subcomplex reveals conserved interactions required for a stable and proper translation pre-initiation complex assembly. Nucleic Acids Res. *40*, 2294–2311.

Hershey, J.W.B. (2010). Regulation of protein synthesis and the role of eIF3 in cancer. Brazilian J. Med. Biol. Res. = Rev. Bras. Pesqui. Medicas e Biol. *43*, 920–930.

Hinnebusch, A.G. (2005). TRANSLATIONAL REGULATION OF *GCN4* AND THE GENERAL AMINO ACID CONTROL OF YEAST. Annu. Rev. Microbiol. *59*, 407–450.

Hinnebusch, A.G. (2006). eIF3: a versatile scaffold for translation initiation complexes. Trends Biochem. Sci. *31*, 553–562.

Hinnebusch, A.G. (2014). The Scanning Mechanism of Eukaryotic Translation Initiation. Annu. Rev. Biochem. *83*, 779–812.

Hinnebusch, A.G. (2017). Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. Trends Biochem. Sci. *42*, 589–

611.

Hinnebusch, A.G., and Lorsch, J.R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harb. Perspect. Biol. *4*, a011544.

Hinnebusch, A.G., Asano, K., Olsen, D.S., Phan, L., Nielsen, K.H., and Valásek, L. (2004). Study of translational control of eukaryotic gene expression using yeast. Ann. N. Y. Acad. Sci. *1038*, 60–74.

Hronová, V., Mohammad, M.P., Wagner, S., Pánek, J., Gunišová, S., Zeman, J., Poncová, K., and Valášek, L.S. (2017). Does eIF3 promote reinitiation after translation of short upstream ORFs also in mammalian cells? RNA Biol. *14*, 1660–1667.

Hsieh, A.C., Liu, Y., Edlind, M.P., Ingolia, N.T., Janes, M.R., Sher, A., Shi, E.Y., Stumpf, C.R., Christensen, C., Bonham, M.J., et al. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature *485*, 55–61.

Huang, J.-S., Chao, C.-C., Su, T.-L., Yeh, S.-H., Chen, D.-S., Chen, C.-T., Chen, P.-J., and Jou, Y.-S. (2004). Diverse cellular transformation capability of overexpressed genes in human hepatocellular carcinoma. Biochem. Biophys. Res. Commun. *315*, 950–958.

Hundsdoerfer, P., Thoma, C., and Hentze, M.W. (2005). Eukaryotic translation initiation factor 4GI and p97 promote cellular internal ribosome entry sequence-driven translation. Proc. Natl. Acad. Sci. U. S. A. *102*, 13421–13426.

Huynen, M.A., Konings, D.A., and Hogeweg, P. (1992). Equal G and C contents in histone genes indicate selection pressures on mRNA secondary structure. J. Mol. Evol. *34*, 280–291.

Hwang, J., Sato, H., Tang, Y., Matsuda, D., and Maquat, L.E. (2010). UPF1 Association with the Cap-Binding Protein, CBP80, Promotes Nonsense-Mediated mRNA Decay at Two Distinct Steps. Mol. Cell *39*, 396–409.

Jackson, R.J., and Kaminski, A. (1995). Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. RNA *1*, 985–1000.

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.

Jackson, R.J., Hellen, C.U.T., and Pestova, T. V. (2012). Termination and post-termination events in eukaryotic translation. In Advances in Protein Chemistry and Structural Biology, pp. 45–93.

Jaeger, S., Martin, F., Rudinger-Thirion, J., Giegé, R., and Eriani, G. (2006). Binding of human SLBP on the 3'-UTR of histone precursor H4-12 mRNA induces structural rearrangements that enable U7 snRNA anchoring. Nucleic Acids Res. *34*, 4987–4995.

Jäger, S., Cimermancic, P., Gulbahce, N., Johnson, J.R., McGovern, K.E., Clarke, S.C.,

Shales, M., Mercenne, G., Pache, L., Li, K., et al. (2011). Global landscape of HIV-human protein complexes. Nature *481*, 365–370.

Jang, S.K., Kräusslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., and Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J. Virol. *62*, 2636–2643.

Jennings, M.D., and Pavitt, G.D. (2010). eIF5 has GDI activity necessary for translational control by eIF2 phosphorylation. Nature *465*, 378–381.

Ji, H., Fraser, C.S., Yu, Y., Leary, J., and Doudna, J.A. (2004). Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. Proc. Natl. Acad. Sci. *101*, 16990–16995.

Jivotovskaya, A. V, Valásek, L., Hinnebusch, A.G., and Nielsen, K.H. (2006). Eukaryotic translation initiation factor 3 (eIF3) and eIF2 can promote mRNA binding to 40S subunits independently of eIF4G in yeast. Mol. Cell. Biol. *26*, 1355–1372.

Johnstone, O., and Lasko, P. (2001). Translational Regulation and RNA Localization in *Drosophila* Oocytes and Embryos. Annu. Rev. Genet. *35*, 365–406.

Johnstone, T.G., Bazzini, A.A., and Giraldez, A.J. (2016). Upstream ORFs are prevalent translational repressors in vertebrates. EMBO J. *35*, 706–723.

Jonas, S., and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. Nat. Rev. Genet. *16*, 421–433.

Kahvejian, A., Roy, G., and Sonenberg, N. (2001). The mRNA closed-loop model: the function of PABP and PABP-interacting proteins in mRNA translation. Cold Spring Harb. Symp. Quant. Biol. *66*, 293–300.

Kamakaka, R.T., and Biggins, S. (2005). Histone variants: deviants? Genes Dev. *19*, 295–310.

Kamenska, A., Simpson, C., and Standart, N. (2014a). eIF4E-binding proteins: new factors, new locations, new roles. Biochem. Soc. Trans. *42*, 1238–1245.

Kamenska, A., Lu, W.-T., Kubacka, D., Broomhead, H., Minshall, N., Bushell, M., and Standart, N. (2014b). Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. Nucleic Acids Res. *42*, 3298–3313.

Kamenska, A., Simpson, C., Vindry, C., Broomhead, H., Bénard, M., Ernoult-Lange, M., Lee, B.P., Harries, L.W., Weil, D., and Standart, N. (2016). The DDX6-4E-T interaction mediates translational repression and P-body assembly. Nucleic Acids Res. *44*, 6318–6334.

Kapp, L.D., and Lorsch, J.R. (2004). GTP-dependent Recognition of the Methionine Moiety on Initiator tRNA by Translation Factor eIF2. J. Mol. Biol. *335*, 923–936.

Karásková, M., Gunišová, S., Herrmannová, A., Wagner, S., Munzarová, V., and Valášek,

L.S. (2012). Functional characterization of the role of the N-terminal domain of the c/Nip1 subunit of eukaryotic initiation factor 3 (eIF3) in AUG recognition. J. Biol. Chem. *287*, 28420–28434.

Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. Cell *116*, 259–272.

Khoshnevis, S., Gunišová, S., Vlčková, V., Kouba, T., Neumann, P., Beznosková, P., Ficner, R., and Valášek, L.S. (2014). Structural integrity of the PCI domain of eIF3a/TIF32 is required for mRNA recruitment to the 43S pre-initiation complexes. Nucleic Acids Res. *42*, 4123–4139.

Kieft, J.S., Zhou, K., Jubin, R., and Doudna, J.A. (2001). Mechanism of ribosome recruitment by hepatitis C IRES RNA. RNA *7*, 194–206.

Kim, B.-H., Cai, X., Vaughn, J.N., and von Arnim, A.G. (2007). On the functions of the h subunit of eukaryotic initiation factor 3 in late stages of translation initiation. Genome Biol. *8*, R60.

Kolupaeva, V.G., Unbehaun, A., Lomakin, I.B., Hellen, C.U.T., and Pestova, T. V (2005). Binding of eukaryotic initiation factor 3 to ribosomal 40S subunits and its role in ribosomal dissociation and anti-association. RNA *11*, 470–486.

Komar, A.A., and Hatzoglou, M. (2011). Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell Cycle *10*, 229–240.

Kondrashov, N., Pusic, A., Stumpf, C.R., Shimizu, K., Hsieh, A.C., Xue, S., Ishijima, J., Shiroishi, T., and Barna, M. (2011). Ribosome-Mediated Specificity in Hox mRNA Translation and Vertebrate Tissue Patterning. Cell *145*, 383–397.

Kornberg, R.D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA. Science (80-.). *184*, 868–871.

Kornberg, R.D., and Thonmas, J.O. (1974). Chromatin Structure: Oligomers of the Histones. Science (80-.). *184*, 865–868.

Korneeva, N.L., Lamphear, B.J., Hennigan, F.L., Merrick, W.C., and Rhoads, R.E. (2001). Characterization of the two eIF4A-binding sites on human eIF4G-1. J. Biol. Chem. *276*, 2872–2879.

Koyama, M., and Kurumizaka, H. (2018). Structural diversity of the nucleosome. J. Biochem. *163*, 85–95.

Kozak, M. (1978). How do eucaryotic ribosomes select initiation regions in messenger RNA? Cell *15*, 1109–1123.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. *12*, 857–872.

Kozak, M. (1990). Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. Proc. Natl. Acad. Sci. U. S. A. *87*, 8301–8305.

Kumari, S., Bugaut, A., and Balasubramanian, S. (2008). Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5' UTR of the NRAS proto-oncogene. Biochemistry *47*, 12664–12669.

de la Parra, C., Ernlund, A., Alard, A., Ruggles, K., Ueberheide, B., and Schneider, R.J. (2018). A widespread alternate form of cap-dependent mRNA translation initiation. Nat. Commun. *9*, 3068.

Lacerda, R., Menezes, J., and Romão, L. (2017). More than just scanning: the importance of cap-independent mRNA translation initiation for cellular stress response and cancer. Cell. Mol. Life Sci. *74*, 1659–1680.

Lacerda, R., Menezes, J., and Candeias, M.M. (2019). Alternative Mechanisms of mRNA Translation Initiation in Cellular Stress Response and Cancer. In Advances in Experimental Medicine and Biology, pp. 117–132.

Lamphear, B.J., Kirchweger, R., Skern, T., and Rhoads, R.E. (1995). Mapping of Functional Domains in Eukaryotic Protein Synthesis Initiation Factor 4G (eIF4G) with Picornaviral Proteases. J. Biol. Chem. *270*, 21975–21983.

Lee, S.H., and McCormick, F. (2006). p97/DAP5 is a ribosome-associated factor that facilitates protein synthesis and cell proliferation by modulating the synthesis of cell cycle proteins. EMBO J. *25*, 4008–4019.

Lee, A.S.Y., Kranzusch, P.J., and Cate, J.H.D. (2015). eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. Nature *522*, 111–114.

Lee, A.S.Y., Kranzusch, P.J., Doudna, J.A., and Cate, J.H.D. (2016). eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. Nature *536*, 96–99.

Lee, K.-H., Kim, S.-H., Kim, H.-J., Kim, W., Lee, H.-R., Jung, Y., Choi, J.-H., Hong, K.Y., Jang, S.K., and Kim, K.-T. (2014). AUF1 contributes to Cryptochrome1 mRNA degradation and rhythmic translation. Nucleic Acids Res. *42*, 3590–3606.

Lee, K.-M., Chen, C.-J., and Shih, S.-R. (2017). Regulation Mechanisms of Viral IRES-Driven Translation. Trends Microbiol. *25*, 546–561.

Leppek, K., Das, R., and Barna, M. (2018). Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. Nat. Rev. Mol. Cell Biol. *19*, 158–174.

Lewis, S.M., and Holcik, M. (2008). For IRES trans-acting factors, it is all about location. Oncogene *27*, 1033–1035.

Li, T., Li, S., Chen, D., Chen, B., Yu, T., Zhao, F., Wang, Q., Yao, M., Huang, S., Chen, Z., et al. (2017). Transcriptomic analyses of RNA-binding proteins reveal *eIF3c* promotes cell proliferation in hepatocellular carcinoma. Cancer Sci. *108*, 877–885.

Lin, L., Holbro, T., Alonso, G., Gerosa, D., and Burger, M.M. (2001). Molecular interaction

between human tumor marker protein p150, the largest subunit of eIF3, and intermediate filament protein K7. J. Cell. Biochem. *80*, 483–490.

Lin, Y., Zhang, R., and Zhang, P. (2016). Eukaryotic translation initiation factor 3 subunit D overexpression is associated with the occurrence and development of ovarian cancer. FEBS Open Bio *6*, 1201–1210.

Locker, N., Easton, L.E., and Lukavsky, P.J. (2007). HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. EMBO J. *26*, 795–805.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251–260.

Lyons, S.M., Cunningham, C.H., Welch, J.D., Groh, B., Guo, A.Y., Wei, B., Whitfield, M.L., Xiong, Y., and Marzluff, W.F. (2016). A subset of replication-dependent histone mRNAs are expressed as polyadenylated RNAs in terminally differentiated tissues. Nucleic Acids Res. *44*, 9190–9205.

MacDonald, J.I., Verdi, J.M., and Meakin, S.O. (1999). Activity-dependent interaction of the intracellular domain of rat trkA with intermediate filament proteins, the beta-6 proteasomal subunit, Ras-GRF1, and the p162 subunit of eIF3. J. Mol. Neurosci. *13*, 141–158.

Mader, S., Lee, H., Pause, A., and Sonenberg, N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. Mol. Cell. Biol. *15*, 4990–4997.

Mailliot, J., and Martin, F. (2018). Viral internal ribosomal entry sites: four classes for one goal. Wiley Interdiscip. Rev. RNA *9*, e1458.

Majzoub, K., Hafirassou, M.L., Meignin, C., Goto, A., Marzi, S., Fedorova, A., Verdier, Y., Vinh, J., Hoffmann, J.A., Martin, F., et al. (2014). RACK1 controls IRES-mediated translation of viruses. Cell *159*, 1086–1095.

Marash, L., Liberman, N., Henis-Korenblit, S., Sivan, G., Reem, E., Elroy-Stein, O., and Kimchi, A. (2008). DAP5 promotes cap-independent translation of Bcl-2 and CDK1 to facilitate cell survival during mitosis. Mol. Cell *30*, 447–459.

Marchetti, A., Buttitta, F., Pellegrini, S., Bertacca, G., and Callahan, R. (2001). Reduced expression of INT-6/eIF3-p48 in human tumors. Int. J. Oncol. *18*, 175–179.

Martin, F., Schaller, A., Eglite, S., Schümperli, D., and Müller, B. (1997). The gene for histone RNA hairpin binding protein is located on human chromosome 4 and encodes a novel type of RNA binding protein. EMBO J. *16*, 769–778.

Martin, F., Barends, S., Jaeger, S., Schaeffer, L., Prongidi-Fix, L., and Eriani, G. (2011). Cap-Assisted Internal Initiation of Translation of Histone H4. Mol. Cell *41*, 197–209.

Martin, F., Ménétret, J.-F., Simonetti, A., Myasnikov, A.G., Vicens, Q., Prongidi-Fix, L.,

Natchiar, S.K., Klaholz, B.P., and Eriani, G. (2016). Ribosomal 18S rRNA base pairs with mRNA during eukaryotic translation initiation. Nat. Commun. *7*, 12622.

Martin, R., Chen, C., and Sarnow, P. (2003). Internal Ribosome Entry Sites Tests with Circular mRNAs. In Protein Synthesis, (New Jersey: Humana Press), pp. 355–364.

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

Martineau, Y., Wang, X., Alain, T., Petroulakis, E., Shahbazian, D., Fabre, B., Bousquet-Dubouch, M.-P., Monsarrat, B., Pyronnet, S., and Sonenberg, N. (2014). Control of Paip1-eukayrotic translation initiation factor 3 interaction by amino acids through S6 kinase. Mol. Cell. Biol. *34*, 1046–1053.

Martínez-Salas, E., Lozano, G., Fernandez-Chamorro, J., Francisco-Velilla, R., Galan, A., and Diaz, R. (2013). RNA-Binding Proteins Impacting on Internal Initiation of Translation. Int. J. Mol. Sci. *14*, 21705–21726.

Marzluff, W.F., and Duronio, R.J. (2002). Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. Curr. Opin. Cell Biol. *14*, 692–699.

Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat. Rev. Genet. *9*, 843–854.

Matsuda, S., Rouault, J.-P., Magaud, J.-P., and Berthet, C. (2001). In search of a function for the TIS21/PC3/BTG1/TOB family. FEBS Lett. *497*, 67–72.

Maze, I., Noh, K.-M., Soshnev, A.A., and Allis, C.D. (2014). Every amino acid matters: essential contributions of histone variants to mammalian development and disease. Nat. Rev. Genet. *15*, 259–271.

McCormack, J.C., Yuan, X., Yingling, Y.G., Kasprzak, W., Zamora, R.E., Shapiro, B.A., and Simon, A.E. (2008). Structural domains within the 3' untranslated region of Turnip crinkle virus. J. Virol. *82*, 8706–8720.

Meeks-Wagner, D., and Hartwell, L.H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell *44*, 43–52.

Meier, V.S., Böhni, R., Schumperli, D., Bohni, R., and Schumperli, D. (1989). Nucleotide sequence of two mouse histone H4 genes. Nucleic Acids Res. *17*, 795–795.

van der Meijden, C.M.J., Lapointe, D.S., Luong, M.X., Peric-Hupkes, D., Cho, B., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (2002). Gene profiling of cell cycle progression through S-phase reveals sequential expression of genes required for DNA replication and nucleosome assembly. Cancer Res. *62*, 3233–3243.

Mejlvang, J., Feng, Y., Alabert, C., Neelsen, K.J., Jasencakova, Z., Zhao, X., Lees, M.,

Sandelin, A., Pasero, P., Lopes, M., et al. (2014). New histone supply regulates replication fork speed and PCNA unloading. J. Cell Biol. *204*, 29–43.

Méthot, N., Song, M.S., and Sonenberg, N. (1996). A region rich in aspartic acid, arginine, tyrosine, and glycine (DRYG) mediates eukaryotic initiation factor 4B (eIF4B) self-association and interaction with eIF3. Mol. Cell. Biol. *16*, 5328–5334.

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.

Min, K.-W., Davila, S., Zealy, R.W., Lloyd, L.T., Lee, I.Y., Lee, R., Roh, K.H., Jung, A., Jemielity, J., Choi, E.-J., et al. (2017). eIF4E phosphorylation by MST1 reduces translation of a subset of mRNAs, but increases lncRNA translation. Biochim. Biophys. Acta. Gene Regul. Mech. *1860*, 761–772.

Minshall, N., Reiter, M.H., Weil, D., and Standart, N. (2007). CPEB Interacts with an Ovary-specific eIF4E and 4E-T in Early *Xenopus* Oocytes. J. Biol. Chem. *282*, 37389–37401.

Mocquet, V., Neusiedler, J., Rende, F., Cluet, D., Robin, J.-P., Terme, J.-M., Duc Dodon, M., Wittmann, J., Morris, C., Le Hir, H., et al. (2012). The Human T-Lymphotropic Virus Type 1 Tax Protein Inhibits Nonsense-Mediated mRNA Decay by Interacting with INT6/EIF3E and UPF1. J. Virol. *86*, 7530–7543.

von Moeller, H., Lerner, R., Ricciardi, A., Basquin, C., Marzluff, W.F., and Conti, E. (2013). Structural and biochemical studies of SLIP1–SLBP identify DBP5 and eIF3g as SLIP1-binding proteins. Nucleic Acids Res. *41*, 7960–7971.

Mohammad, M.P., Munzarová Pondělíčková, V., Zeman, J., Gunišová, S., and Valášek, L.S. (2017). *In vivo* evidence that eIF3 stays bound to ribosomes elongating and terminating on short upstream ORFs to promote reinitiation. Nucleic Acids Res. *45*, gkx049.

Mokrejs, M., Masek, T., Vopálensky, V., Hlubucek, P., Delbos, P., and Pospísek, M. (2010). IRESite--a tool for the examination of viral and cellular internal ribosome entry sites. Nucleic Acids Res. *38*, D131-6.

Morris, C., Wittmann, J., Jäck, H.-M., and Jalinot, P. (2007). Human INT6/eIF3e is required for nonsense-mediated mRNA decay. EMBO Rep. *8*, 596–602.

Mowry, K.L., and Steitz, J.A. (1987). Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNA's. Science *238*, 1682–1687.

Muhs, M., Hilal, T., Mielke, T., Skabkin, M.A., Sanbonmatsu, K.Y., Pestova, T. V., and Spahn, C.M.T. (2015). Cryo-EM of ribosomal 80S complexes with termination factors reveals the translocated cricket paralysis virus IRES. Mol. Cell *57*, 422–432.

Murray, J., Savva, C.G., Shin, B.-S., Dever, T.E., Ramakrishnan, V., and Fernández, I.S. (2016). Structural characterization of ribosome recruitment and translocation by type IV

IRES. Elife *5*.

Nakamura, A., Sato, K., and Hanyu-Nakamura, K. (2004). Drosophila cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. Dev. Cell *6*, 69–78.

Naranda, T., Kainuma, M., MacMillan, S.E., and Hershey, J.W. (1997). The 39-kilodalton subunit of eukaryotic translation initiation factor 3 is essential for the complex's integrity and for cell viability in Saccharomyces cerevisiae. Mol. Cell. Biol. *17*, 145–153.

Neusiedler, J., Mocquet, V., Limousin, T., Ohlmann, T., Morris, C., and Jalinot, P. (2012). INT6 interacts with MIF4GD/SLIP1 and is necessary for efficient histone mRNA translation. RNA *18*, 1163–1177.

Nevins, T.A., Harder, Z.M., Korneluk, R.G., and Holcík, M. (2003). Distinct regulation of internal ribosome entry site-mediated translation following cellular stress is mediated by apoptotic fragments of eIF4G translation initiation factor family members eIF4GI and p97/DAP5/NAT1. J. Biol. Chem. *278*, 3572–3579.

Newburn, L.R., and White, K.A. (2015). Cis-acting RNA elements in positive-strand RNA plant virus genomes. Virology *479–480*, 434–443.

Nicholson, B.L., Wu, B., Chevtchenko, I., and White, K.A. (2010). Tombusvirus recruitment of host translational machinery via the 3' UTR. RNA *16*, 1402–1419.

Niepmann, M. (2013). Hepatitis C Virus RNA Translation. pp. 143–166.

Obayashi, E., Luna, R.E., Nagata, T., Martin-Marcos, P., Hiraishi, H., Singh, C.R., Erzberger, J.P., Zhang, F., Arthanari, H., Morris, J., et al. (2017). Molecular Landscape of the Ribosome Pre-initiation Complex during mRNA Scanning: Structural Role for eIF3c and Its Control by eIF5. Cell Rep. *18*, 2651–2663.

Ohlmann, T., Rau, M., Pain, V.M., and Morley, S.J. (1996). The C-terminal domain of eukaryotic protein synthesis initiation factor (eIF) 4G is sufficient to support cap-independent translation in the absence of eIF4E. EMBO J. *15*, 1371–1382.

Osley, M.A. (1991). The Regulation of Histone Synthesis in the Cell Cycle. Annu. Rev. Biochem. *60*, 827–861.

Otto, G.A., and Puglisi, J.D. (2004). The Pathway of HCV IRES-Mediated Translation Initiation. Cell *119*, 369–380.

Ozgur, S., Basquin, J., Kamenska, A., Filipowicz, W., Standart, N., and Conti, E. (2015). Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex. Cell Rep. *13*, 703–711.

Pause, A., Belsham, G.J., Gingras, A.-C., Donzé, O., Lin, T.-A., Lawrence, J.C., and Sonenberg, N. (1994). Insulin-dependent stimulation of protein synthesis by

phosphorylation of a regulator of 5'-cap function. Nature *371*, 762–767.

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

Pena, V., Liu, S., Bujnicki, J.M., Lührmann, R., and Wahl, M.C. (2007). Structure of a multipartite protein-protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa. Mol. Cell *25*, 615–624.

Pestova*, T. V., and Hellen, C.U.T. (2000). The structure and function of initiation factors in eukaryotic protein synthesis. Cell. Mol. Life Sci. *57*, 651–674.

Pestova, T. V, and Kolupaeva, V.G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. Genes Dev. *16*, 2906–2922.

Pestova, T. V, Shatsky, I.N., Fletcher, S.P., Jackson, R.J., and Hellen, C.U. (1998). A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes Dev. *12*, 67–83.

Phan, L., Zhang, X., Asano, K., Anderson, J., Vornlocher, H.-P., Greenberg, J.R., Qin, J., and Hinnebusch, A.G. (1998). Identification of a Translation Initiation Factor 3 (eIF3) Core Complex, Conserved in Yeast and Mammals, That Interacts with eIF5. Mol. Cell. Biol. *18*, 4935–4946.

Pick, E., Hofmann, K., and Glickman, M.H. (2009). PCI complexes: Beyond the proteasome, CSN, and eIF3 Troika. Mol. Cell *35*, 260–264.

Pincheira, R., Chen, Q., and Zhang, J.T. (2001). Identification of a 170-kDa protein over-expressed in lung cancers. Br. J. Cancer *84*, 1520–1527.

Piontkivska, H., Rooney, A.P., and Nei, M. (2002). Purifying Selection and Birth-and-death Evolution in the Histone H4 Gene Family. Mol. Biol. Evol. *19*, 689–697.

Pisarev, A. V., Shirokikh, N.E., and Hellen, C.U.T. (2005). Translation initiation by factor-independent binding of eukaryotic ribosomes to internal ribosomal entry sites. C. R. Biol. *328*, 589–605.

Pisareva, V.P., Pisarev, A. V, Komar, A.A., Hellen, C.U.T., and Pestova, T. V (2008). Translation initiation on mammalian mRNAs with structured 5'UTRs requires DExH-box protein DHX29. Cell *135*, 1237–1250.

Prévôt, D., Darlix, J.-L., and Ohlmann, T. (2003). Conducting the initiation of protein synthesis: the role of eIF4G. Biol. Cell *95*, 141–156.

Pyronnet, S., Dostie, J., and Sonenberg, N. (2001). Suppression of cap-dependent translation in mitosis. Genes Dev. *15*, 2083–2093.

Pyronnet S., Imataka H., Gingras A.C., Fukunaga R., Hunter T., and Sonenberg N. Human (1999). eukaryotic translation initiation factor 4G (eIF4G) recruits Mnk1 to phosphorylate eIF4E. EMBO J., *18*, 270-279.

Qu, J., Ero, R., Feng, C., Ong, L.-T., Tan, H.-F., Lee, H.-S., Ismail, M.H., Bu, W.-T., Nama, S., Sampath, P., et al. (2016). Kindlin-3 interacts with the ribosome and regulates c-Myc expression required for proliferation of chronic myeloid leukemia cells. Sci. Rep. *5*, 18491.

Querol-Audi, J., Sun, C., Vogan, J.M., Smith, M.D., Gu, Y., Cate, J.H.D., and Nogales, E. (2013). Architecture of Human Translation Initiation Factor 3. Structure *21*, 920–928.

Rajasekhar, V.K., and Holland, E.C. (2004). Postgenomic global analysis of translational control induced by oncogenic signaling. Oncogene *23*, 3248–3264.

Ramanathan, A., Robb, G.B., and Chan, S.-H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res. *44*, 7511.

Reynolds, J.E., Kaminski, A., Kettinen, H.J., Grace, K., Clarke, B.E., Carroll, A.R., Rowlands, D.J., and Jackson, R.J. (1995). Unique features of internal initiation of hepatitis C virus RNA translation. EMBO J. *14*, 6010–6020.

Rhoads, R.E. (2009). eIF4E: New Family Members, New Binding Partners, New Roles. J. Biol. Chem. *284*, 16711–16715.

Ricci, E.P., Limousin, T., Soto-Rifo, R., Rubilar, P.S., Decimo, D., and Ohlmann, T. (2013). miRNA repression of translation in vitro takes place during 43S ribosomal scanning. Nucleic Acids Res. *41*, 586–598.

Riis, B., Rattan, S.I.S., Clark, B.F.C., and Merrick, W.C. (1990). Eukaryotic protein elongation factors. Trends Biochem. Sci. *15*, 420–424.

Robichaud, N., and Sonenberg, N. (2017). Translational control and the cancer cell response to stress. Curr. Opin. Cell Biol. *45*, 102–109.

Rogers, G.W., Richter, N.J., Lima, W.F., and Merrick, W.C. (2001). Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. J. Biol. Chem. *276*, 30914–30922.

Roux, P.P., and Topisirovic, I. (2018). Signaling Pathways Involved in the Regulation of mRNA Translation. Mol. Cell. Biol. *38*, 00070-18

Rouya, C., Siddiqui, N., Morita, M., Duchaine, T.F., Fabian, M.R., and Sonenberg, N. (2014). Human DDX6 effects miRNA-mediated gene silencing via direct binding to CNOT1. RNA *20*, 1398–1409.

Salzman, J. (2016). Circular RNA Expression: Its Potential Regulation and Function. Trends

Genet. *32*, 309–316.

Sànchez, R., and Marzluff, W.F. (2002). The stem-loop binding protein is required for efficient translation of histone mRNA in vivo and in vitro. Mol. Cell. Biol. *22*, 7093–7104.

Sasaki, J., and Nakashima, N. (1999). Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus in vitro. J. Virol. *73*, 1219–1226.

Schueren, F., and Thoms, S. (2016). Functional Translational Readthrough: A Systems Biology Perspective. PLOS Genet. *12*, e1006196.

Scoles, D.R., Yong, W.H., Qin, Y., Wawrowsky, K., and Pulst, S.M. (2006). Schwannomin inhibits tumorigenesis through direct interaction with the eukaryotic initiation factor subunit c (eIF3c). Hum. Mol. Genet. *15*, 1059–1070.

Sesen, J., Cammas, A., Scotland, S.J., Elefterion, B., Lemarié, A., Millevoi, S., Mathew, L.K., Seva, C., Toulas, C., Moyal, E.C.-J., et al. (2014). Int6/eIF3e is essential for proliferation and survival of human glioblastoma cells. Int. J. Mol. Sci. *15*, 2172–2190.

Sesen, J., Casaos, J., Scotland, S.J., Seva, C., Eisinger-Mathason, T.S.K., and Skuli, N. (2017). The Bad, the Good and eIF3e/INT6. Front. Biosci. (Landmark Ed. *22*, 1–20.

Shah, M., Su, D., Scheliga, J.S., Pluskal, T., Boronat, S., Motamedchaboki, K., Campos, A.R., Qi, F., Hidalgo, E., Yanagida, M., et al. (2016). A Transcript-Specific eIF3 Complex Mediates Global Translational Control of Energy Metabolism. Cell Rep. *16*, 1891–1902.

Shanmugam, M.K., Arfuso, F., Arumugam, S., Chinnathambi, A., Jinsong, B., Warrier, S., Wang, L.Z., Kumar, A.P., Ahn, K.S., Sethi, G., et al. (2018). Role of novel histone modifications in cancer. Oncotarget *9*, 11414.

Shatsky, I.N., Dmitriev, S.E., Terenin, I.M., and Andreev, D.E. (2010). Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs. Mol. Cells *30*, 285–293.

Shatsky, I.N., Terenin, I.M., Smirnova, V. V., and Andreev, D.E. (2018). Cap-Independent Translation: What's in a Name? Trends Biochem. Sci. *43*, 882–895.

Shi, J., Kahle, A., Hershey, J.W.B., Honchak, B.M., Warneke, J.A., Leong, S.P.L., and Nelson, M.A. (2006). Decreased expression of eukaryotic initiation factor 3f deregulates translation and apoptosis in tumor cells. Oncogene *25*, 4923–4936.

Simon, A.E., and Miller, W.A. (2013). 3′ Cap-Independent Translation Enhancers of Plant Viruses. Annu. Rev. Microbiol. *67*, 21–42.

Sizova, D. V, Kolupaeva, V.G., Pestova, T. V, Shatsky, I.N., and Hellen, C.U. (1998). Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. J. Virol. *72*, 4775–4782.

Smith, M.D., Arake-Tacca, L., Nitido, A., Montabana, E., Park, A., and Cate, J.H. (2016).

Assembly of eIF3 Mediated by Mutually Dependent Subunit Insertion. Structure *24*, 886–896.

Sokabe, M., Fraser, C.S., and Hershey, J.W.B. (2012). The human translation initiation multi-factor complex promotes methionyl-tRNAi binding to the 40S ribosomal subunit. Nucleic Acids Res. *40*, 905–913.

Sonenberg, N., and Dever, T.E. (2003). Eukaryotic translation initiation factors and regulators. Curr. Opin. Struct. Biol. *13*, 56–63.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell *136*, 731–745.

Sonenberg, N., and Meerovitch, K. (1990). Translation of poliovirus mRNA. Enzyme *44*, 278–291.

Sonenberg, N., Rupprecht, K.M., Hecht, S.M., and Shatkin, A.J. (1979). Eukaryotic mRNA cap binding protein: purification by affinity chromatography on sepharose-coupled m7GDP. Proc. Natl. Acad. Sci. *76*, 4345–4349.

Song, N., Wang, Y., Gu, X., Chen, Z., and Shi, L. (2013). Effect of siRNA-mediated knockdown of eIF3c gene on survival of colon cancer cells. J. Zhejiang Univ. Sci. B *14*, 451–459.

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.

Spriggs, K.A., Stoneley, M., Bushell, M., and Willis, A.E. (2008). Re-programming of translation following cell stress allows IRES-mediated translation to predominate. Biol. Cell *100*, 27–38.

Spriggs, K.A., Bushell, M., and Willis, A.E. (2010). Translational Regulation of Gene Expression during Conditions of Cell Stress. Mol. Cell *40*, 228–237.

Spycher, C., Streit, A., Stefanovic, B., Albrecht, D., Koning, T.H., and Schümperli, D. (1994). 3' end processing of mouse histone pre-mRNA: evidence for additional base-pairing between U7 snRNA and pre-mRNA. Nucleic Acids Res. *22*, 4023–4030.

Stoneley, M., Paulin, F.E., Quesne, J.P. Le, Chappell, S.A., and Willis, A.E. (1998). C-Myc 5′ untranslated region contains an internal ribosome entry segment. Oncogene *16*, 423–428.

Stupina, V.A., Meskauskas, A., McCormack, J.C., Yingling, Y.G., Shapiro, B.A., Dinman, J.D., and Simon, A.E. (2008). The 3' proximal translational enhancer of Turnip crinkle virus binds to 60S ribosomal subunits. RNA *14*, 2379–2393.

Sugiyama, H., Takahashi, K., Yamamoto, T., Iwasaki, M., Narita, M., Nakamura, M., Rand, T.A., Nakagawa, M., Watanabe, A., and Yamanaka, S. (2017). Nat1 promotes translation of specific proteins that induce differentiation of mouse embryonic stem cells.

Proc. Natl. Acad. Sci. U. S. A. *114*, 340–345.

Sun, C., Todorovic, A., Querol-Audí, J., Bai, Y., Villa, N., Snyder, M., Ashchyan, J., Lewis, C.S., Hartland, A., Gradia, S., et al. (2011). Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). Proc. Natl. Acad. Sci. U. S. A. *108*, 20473–20478.

Sun, C., Querol-Audí, J., Mortimer, S.A., Arias-Palomo, E., Doudna, J.A., Nogales, E., and Cate, J.H.D. (2013). Two RNA-binding motifs in eIF3 direct HCV IRES-dependent translation. Nucleic Acids Res. *41*, 7512–7521.

Suo, J., Medina, D., Herrera, S., Zheng, Z.-Y., Jin, L., Chamness, G.C., Contreras, A., Gutierrez, C., Hilsenbeck, S., Umar, A., et al. (2015). Int6 reduction activates stromal fibroblasts to enhance transforming activity in breast epithelial cells. Cell Biosci. *5*, 10.

Szostak, E., García-Beyaert, M., Guitart, T., Graindorge, A., Coll, O., and Gebauer, F. (2018). Hrp48 and eIF3d contribute to msl-2 mRNA translational repression. Nucleic Acids Res. *46*, 4099–4113.

Talbert, P.B., Ahmad, K., Almouzni, G., Ausió, J., Berger, F., Bhalla, P.L., Bonner, W.M., Cande, W.Z., Chadwick, B.P., Chan, S.W.L., et al. (2012). A unified phylogeny-based nomenclature for histone variants. Epigenetics Chromatin *5*, 7.

Terenin, I.M., Smirnova, V. V, Andreev, D.E., Dmitriev, S.E., and Shatsky, I.N. (2017). A researcher's guide to the galaxy of IRESs. Cell. Mol. Life Sci. *74*, 1431–1455.

Thoreen, C.C., Chantranupong, L., Keys, H.R., Wang, T., Gray, N.S., and Sabatini, D.M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. Nature *485*, 109–113.

Tong, J., Flavell, R.A., and Li, H.-B. (2018). RNA m6A modification and its function in diseases. Front. Med. *12*, 481–489.

Treder, K., Kneller, E.L.P., Allen, E.M., Wang, Z., Browning, K.S., and Miller, W.A. (2008). The 3' cap-independent translation element of Barley yellow dwarf virus binds eIF4F via the eIF4G subunit to initiate translation. RNA *14*, 134–147.

Ujino, S., Nishitsuji, H., Sugiyama, R., Suzuki, H., Hishiki, T., Sugiyama, K., Shimotohno, K., and Takaku, H. (2012). The interaction between human initiation factor eIF3 subunit c and heat-shock protein 90: A necessary factor for translation mediated by the hepatitis C virus internal ribosome entry site. Virus Res. *163*, 390–395.

Valásek, L.S. (2012). 'Ribozoomin'--translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). Curr. Protein Pept. Sci. *13*, 305–330.

Valásek, L., Trachsel, H., Hasek, J., and Ruis, H. (1998). Rpg1, the *Saccharomyces cerevisiae* Homologue of the Largest Subunit of Mammalian Translation Initiation Factor

3, Is Required for Translational Activity. J. Biol. Chem. *273*, 21253–21260.

Valásek, L., Hasek, J., Nielsen, K.H., and Hinnebusch, A.G. (2001). Dual function of eIF3j/Hcr1p in processing 20 S pre-rRNA and translation initiation. J. Biol. Chem. *276*, 43351–43360.

Valásek, L., Nielsen, K.H., and Hinnebusch, A.G. (2002). Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation in vivo. EMBO J. *21*, 5886–5898.

Valásek, L., Nielsen, K.H., Zhang, F., Fekete, C.A., and Hinnebusch, A.G. (2004). Interactions of eukaryotic translation initiation factor 3 (eIF3) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. Mol. Cell. Biol. *24*, 9437–9455.

Valášek, L.S., Zeman, J., Wagner, S., Beznosková, P., Pavlíková, Z., Mohammad, M.P., Hronová, V., Herrmannová, A., Hashem, Y., and Gunišová, S. (2017). Embraced by eIF3: structural and functional insights into the roles of eIF3 across the translation cycle. Nucleic Acids Res. *45*, 10948–10968.

Villa, N., Do, A., Hershey, J.W.B., and Fraser, C.S. (2013). Human Eukaryotic Initiation Factor 4G (eIF4G) Protein Binds to eIF3c, -d, and -e to Promote mRNA Recruitment to the Ribosome. J. Biol. Chem. *288*, 32932–32940.

Vleugel, M.M., Greijer, A.E., Bos, R., van der Wall, E., and van Diest, P.J. (2006). c-Jun activation is associated with proliferation and angiogenesis in invasive breast cancer. Hum. Pathol. *37*, 668–674.

Wagner, S., Herrmannová, A., Malík, R., Peclinovská, L., and Valášek, L.S. (2014). Functional and biochemical characterization of human eukaryotic translation initiation factor 3 in living cells. Mol. Cell. Biol. *34*, 3041–3052.

Wagner, S., Herrmannová, A., Šikrová, D., and Valášek, L.S. (2016). Human eIF3b and eIF3a serve as the nucleation core for the assembly of eIF3 into two interconnected modules: the yeast-like core and the octamer. Nucleic Acids Res. *44*, 10772–10788.

Walsh, D., and Mohr, I. (2014). Coupling 40S ribosome recruitment to modification of a cap-binding initiation factor by eIF3 subunit e. Genes Dev. *28*, 835–840.

Wang, J., and Pantopoulos, K. (2011). Regulation of cellular iron metabolism. Biochem. J. *434*, 365–381.

Wang, L., and Ouyang, L. (2019). Effects of EIF3B gene downregulation on apoptosis and proliferation of human ovarian cancer SKOV3 and HO-8910 cells. Biomed. Pharmacother. *109*, 831–837.

Wang, H., Ru, Y., Sanchez-Carbayo, M., Wang, X., Kieft, J.S., and Theodorescu, D. (2013a). Translation initiation factor eIF3b expression in human cancer and its role in tumor growth and lung colonization. Clin. Cancer Res. *19*, 2850–2860.

Wang, Y.-W., Lin, K.-T., Chen, S.-C., Gu, D.-L., Chen, C.-F., Tu, P.-H., and Jou, Y.-S. (2013b). Overexpressed-eIF3I interacted and activated oncogenic Akt1 is a theranostic target in human hepatocellular carcinoma. Hepatology *58*, 239–250.

Wang, Z., Treder, K., and Miller, W.A. (2009). Structure of a viral cap-independent translation element that functions via high affinity binding to the eIF4E subunit of eIF4F. J. Biol. Chem. *284*, 14189–14202.

Wei, Z., Zhang, P., Zhou, Z., Cheng, Z., Wan, M., and Gong, W. (2004). Crystal structure of human eIF3k, the first structure of eIF3 subunits. J. Biol. Chem. *279*, 34983–34990.

Weingarten-Gabbay, S., Khan, D., Liberman, N., Yoffe, Y., Bialik, S., Das, S., Oren, M., and Kimchi, A. (2014). The translation initiation factor DAP5 promotes IRES-driven translation of p53 mRNA. Oncogene *33*, 611–618.

Wen, F., Zhou, R., Shen, A., Choi, A., Uribe, D., and Shi, J. (2012). The tumor suppressive role of eIF3f and its function in translation inhibition and rRNA degradation. PLoS One *7*, e34194.

Wen, F., Wu, Z.-Y., Nie, L., Zhang, Q.-Z., Qin, Y.-K., Zhou, Z., Wu, J.-J., Zhao, X., Tan, J., Sawmiller, D., et al. (2019). Eukaryotic initiation factor 3, subunit C silencing inhibits cell proliferation and promotes apoptosis in human ovarian cancer cells. Biosci. Rep. 39

Wethmar, K. (2014). The regulatory potential of upstream open reading frames in eukaryotic gene expression. Wiley Interdiscip. Rev. RNA *5*, 765–768.

Wethmar, K., Barbosa-Silva, A., Andrade-Navarro, M.A., and Leutz, A. (2014). uORFdb--a comprehensive literature database on eukaryotic uORF biology. Nucleic Acids Res. *42*, D60-7.

Whitfield, M.L., Zheng, L.-X., Baldwin, A., Ohta, T., Hurt, M.M., and Marzluff, W.F. (2000). Stem-Loop Binding Protein, the Protein That Binds the 3' End of Histone mRNA, Is Cell Cycle Regulated by Both Translational and Posttranslational Mechanisms. Mol. Cell. Biol. *20*, 4188–4198.

Wilhelm, J.E., and Smibert, C.A. (2005). Mechanisms of translational regulation in Drosophila. Biol. Cell *97*, 235–252.

Wilson, J.E., Pestova, T. V, Hellen, C.U., and Sarnow, P. (2000). Initiation of protein synthesis from the A site of the ribosome. Cell *102*, 511–520.

Wurth, L., Gribling-Burrer, A.S., Verheggen, C., Leichter, M., Takeuchi, A., Baudrey, S., Martin, F., Krol, A., Bertrand, E., and Allmang, C. (2014). Hypermethylated capped selenoprotein mRNAs in mammals. Nucl. Acids Res. *42*, 8663–8677.

Xue, S., Tian, S., Fujii, K., Kladwang, W., Das, R., and Barna, M. (2015). RNA regulons in Hox 5′ UTRs confer ribosome specificity to gene regulation. Nature *517*, 33–38.

Yang, C., Zhang, Y., Du, W., Cheng, H., and Li, C. (2019). Eukaryotic translation initiation

factor 3 subunit G promotes human colorectal cancer. Am. J. Transl. Res. *11*, 612–623.

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

Yoffe, Y., David, M., Kalaora, R., Povodovski, L., Friedlander, G., Feldmesser, E., Ainbinder, E., Saada, A., Bialik, S., and Kimchi, A. (2016). Cap-independent translation by DAP5 controls cell fate decisions in human embryonic stem cells. Genes Dev. *30*, 1991–2004.

Young, D.J., and Guydosh, N.R. (2019). Hcr1/eIF3j Is a 60S Ribosomal Subunit Recycling Accessory Factor In Vivo. Cell Rep. *28*, 39-50.e4.

Zeman, J., Itoh, Y., Kukačka, Z., Rosůlek, M., Kavan, D., Kouba, T., Jansen, M.E., Mohammad, M.P., Novák, P., and Valášek, L.S. (2019). Binding of eIF3 in complex with eIF5 and eIF1 to the 40S ribosomal subunit is accompanied by dramatic structural changes. Nucleic Acids Res. *47*, 8282-8300

Zeng, L., Wan, Y., Li, D., Wu, J., Shao, M., Chen, J., Hui, L., Ji, H., and Zhu, X. (2013). The m subunit of murine translation initiation factor eIF3 maintains the integrity of the eIF3 complex and is required for embryonic development, homeostasis, and organ size control. J. Biol. Chem. *288*, 30087–30093.

Zhang, L., Pan, X., and Hershey, J.W.B. (2007). Individual Overexpression of Five Subunits of Human Translation Initiation Factor eIF3 Promotes Malignant Transformation of Immortal Fibroblast Cells. J. Biol. Chem. *282*, 5790–5800.

Zhang, L., Smit-McBride, Z., Pan, X., Rheinhardt, J., and Hershey, J.W.B. (2008). An oncogenic role for the phosphorylated h-subunit of human translation initiation factor eIF3. J. Biol. Chem. *283*, 24047–24060.

Zhao, B.S., Roundtree, I.A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. Nat. Rev. Mol. Cell Biol. *18*, 31–42.

Zheng, L., Dominski, Z., Yang, X.-C., Elms, P., Raska, C.S., Borchers, C.H., and Marzluff, W.F. (2003). Phosphorylation of Stem-Loop Binding Protein (SLBP) on Two Threonines Triggers Degradation of SLBP, the Sole Cell Cycle-Regulated Factor Required for Regulation of Histone mRNA Processing, at the End of S Phase. Mol. Cell. Biol. *23*, 1590–1601.

Zheng, Q., Liu, H., Ye, J., Zhang, H., Jia, Z., and Cao, J. (2016). Nuclear distribution of eIF3g and its interacting nuclear proteins in breast cancer cells. Mol. Med. Rep. *13*, 2973–2980.

Zhou, C., Arslan, F., Wee, S., Krishnan, S., Ivanov, A.R., Oliva, A., Leatherwood, J., and Wolf, D.A. (2005). PCI proteins eIF3e and eIF3m define distinct translation initiation factor 3 complexes. BMC Biol. *3*, 14.

Zhou, J., Rode, K.A., and Qian, S.-B. (2016). m ^{6} A: A novel hallmark of translation. Cell Cycle *15*, 309–310.

Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L., and Philippe, M. (1995). Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. EMBO J. *14*, 4065–4072.

UNIVERSITÉ DE STRASBOURG

Hassan Hayek

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

Laboratoire: Evolution des systèmes d'initiation de la traduction chez les eucaryotes, UPR9002-IBMC

Rôle du facteur d'initiation eIF3 dans la traduction de l'ARNm de l'histone H4

Chez la plupart des ARNm eucaryotes, l'initiation de la traduction est coiffe-dépendante et fait appel au scanning de l'extrémité 5' UTR par les ribosomes. Ce modèle standard ne s'applique pas à tous les ARNm cellulaires. C'est le cas de l'ARNm de l'histone H4 dont la traduction est initiée par un mécanisme original reprenant à la fois certaines caractéristiques cellulaires (fixation à la coiffe) et virales (absence de scanning et entrée interne des ribosomes). Les ribosomes sont recrutés par des facteurs d'initiation fixés sur des séquences de l'ARNm situées dans la phase codante de la protéine puis déposés directement sur l'AUG initiateur grâce à une structure repliée en jonction à 3 hélices qui participe à leur positionnement. Il s'en suit une initiation de la traduction extrêmement efficace qui permet la production massive des histones requises durant la phase S pour empaqueter l'ADN néosynthétisé. Récemment, il a été montré que le facteur d'initiation eIF3 assurait le recrutement sélectif d'ARNm cellulaires, dont celui de l'histone H4, afin de contrôler leur expression. Chez l'homme, eIF3 est un complexe multi-protéique qui comprend 13 sous-unités non identiques (eIF3a-m) d'un poids moléculaire combiné d'environ 800 kDa.

Mon projet de thèse a consisté à démontrer et caractériser l'interaction d'eIF3 avec l'ARNm de l'histone H4 et à analyser sa fonction. Dans un premier temps, j'ai reproduit cette interaction *in vitro* à partir de facteur eIF3 purifié et d'ARNm H4. Par immunoprécipitation de complexes ribonucléoprotéiques (RNP) *in vivo* j'ai montré que les ARNm de H4 et des histones H1, H2A, H2B et H3 interagissaient avec eIF3 dans les cellules HEK293FT. Grâce à des expériences de pontage du complexe ARNm H4-thioU/eIF3 puis de fractionnement des produits de pontage sur gel bidimensionnel et d'immunodétection, j'ai identifié quatre sous-unités, eIF3c, d, e et g, en interaction directe avec l'ARNm H4. La capacité de ces 4 sous-unités exprimées sous forme recombinante dans *E. coli* à interagir avec les ARNm d'histones a ensuite été testée par des expériences de GST pull-down couplées à la détection des ARNm par RT-PCR. Nous avons montré que les trois sous-unités eIF3 c, d et g sont capables d'interagir avec les ARNm d'histones H1, H2A, H2B, H3 et H4 *in vitro* indépendamment du complexe eIF3. eIF3e en revanche interagit uniquement avec l'ARNm de H1 et non avec H4 comme le suggéraient nos expériences de pontage.

Une seconde partie de mon travail de thèse a consisté à rechercher la fonction cellulaire de l'interaction d'eIF3 avec les ARNm d'histones. Dans une première approche, j'ai testé l'effet de la tige-boucle de l'ARNm H4 interagissant avec eIF3 sur la traduction d'un gène rapporteur R-Luc, *in vitro* à l'aide d'ARNm monocistroniques purifiés, ou *in vivo* dans un système de bicistrons. Des effets opposés sur la traduction ont été obtenus selon la localisation la tige-boucle en 5' ou 3' de la phase codante de la R-Luc.

Enfin, j'ai analysé le rôle fonctionnel d'eIF3 sur la synthèse des histones *in vivo* en étudiant l'impact de l'inhibition des sous-unités eIF3 c, d, e et g par siARN sur la traduction des histones. L'inhibition simple ou multiple des sous-unités pendant 48 heures stimule préférentiellement la traduction des ARNm d'histones alors que l'inactivation double et quadruple pendant 72 heures cause une diminution de leur synthèse en même temps qu'un effondrement de l'activité métabolique des cellules. L'ensemble de ces résultats permet de conclure à la présence d'une interaction privilégiée d'eIF3 avec les ARNm des histones. Cette interaction, répressive en temps normal, pourrait favoriser la synthèse des histones dans certaines situations métaboliques lorsque le nombre de molécules d'eIF3 vient à être limitant.

Mots clés : ARNm d'histone H4, eIF3, Initiation de la traduction

Université de Strasbourg

Hassan Alaskari HAYEK

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

Rôle du facteur d'initiation eIF3 dans la traduction de l'ARNm de l'histone H4

Résumé en français

La traduction de l'ARNm de l'histone H4 est initiée par un mécanisme original combinant à la fois certaines caractéristiques canoniques (fixation à la coiffe) et virales (absence de scanning et entrée interne des ribosomes). Le facteur d'initiation eIF3, un complexe de 13 sous-unités (a-m), assure le recrutement sélectif d'ARNm cellulaires, dont celui de l'histone H4, afin de contrôler leur expression. Dans cette thèse nous avons démontré l'interaction entre le facteur eIF3 et les ARNm de H4 et des histones H1, H2A, H2B et H3 *in vivo* et identifié 4 sous-unités eIF3c, d, e et g en interaction avec l'ARNm H4. Celles-ci se lient aux ARNm d'histones *in vitro* indépendamment du complexe eIF3. Nous avons analysé le rôle fonctionnel d'eIF3 sur la synthèse des histones *in vivo* et montré que l'inhibition des 4 sous-unités par siARN augmente la traduction des histones au début de la phase S du cycle cellulaire. eIF3 semble réprimer la traduction des ARNm d'histones quand il n'est pas limitant.

Mots clés : ARNm d'histone H4, eIF3, Initiation de la traduction

Résumé en anglais

The translation of the histone H4 mRNA is initiated by an original mechanism that combines both canonical (binding to the cap) and viral (absence of scanning and internal entry of ribosomes) characteristics. The initiation factor eIF3, a complex of 13 subunits (a-m), selectively recruits cellular mRNAs, including histone H4, to control their expression. In this thesis we demonstrated the interaction between eIF3 and the mRNAs of H4 and histones H1, H2A, H2B and H3 *in vivo* and identified 4 subunits eIF3c, d, e and g in interaction with H4 mRNA. These subunits bind to histone mRNAs *in vitro* independently of the eIF3 complex. We analyzed the functional role of eIF3 on histone synthesis *in vivo* and showed that inhibition of the 4 subunits by siRNA increased histone translation at the beginning of the S phase of the cell cycle. eIF3 appears to suppress translation of histone mRNAs when it is not limiting.

Keywords : Histone H4 mRNA, eIF3, Translation initiation