



HAL
open science

Etudes structurales de l'ARN messenger de l'histone H4

Arnaud d'Orchymont

► **To cite this version:**

Arnaud d'Orchymont. Etudes structurales de l'ARN messenger de l'histone H4. Biologie cellulaire. Université de Strasbourg, 2013. Français. NNT : 2013STRAJ039 . tel-01037935

HAL Id: tel-01037935

<https://theses.hal.science/tel-01037935>

Submitted on 23 Jul 2014

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

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

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ
UPR 9002 & UMR 7104

THÈSE présentée par : **Arnaud D'ORCHYMONT**

soutenue le : **27 novembre 2013**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**
Discipline/ Spécialité : **Aspects Moléculaires et Cellulaires de la Biologie**

Etudes structurales de l'ARN messenger de l'histone H4

THÈSE dirigée par :

M. Eriani Gilbert
M. Cavarelli Jean

Dr, IBMC
Professeur, Université de Strasbourg

RAPPORTEURS :

Mme Dock-Bregeon Anne-Catherine
M. Mechulam Yves

Dr, IBENS
Dr, Ecole polytechnique

AUTRES MEMBRES DU JURY :

Mme Romby Pascale
Mme Favier Frédérique

Dr, IBMC
MCF, Université de Lorraine

Abréviations	5
Avant propos de la thèse.....	7
Introduction.....	9
A. Les histones, constituants de la chromatine	9
a. Structure et fonction des histones.....	9
b. Synthèse et structure des ARNm d'histones.....	10
B. Traduction	10
a. Mécanisme d'initiation de la traduction coiffe dépendante	11
b. Mécanisme d'initiation coiffe indépendant induit par des IRES	14
i. IRES de type I.....	15
ii. IRES de type II.....	16
iii. IRES de type III.....	17
iv. IRES de type IV.....	17
c. Modèle d'initiation de la traduction de l'histone H4.....	18
Chapitre 1 : Purification et repliement de l'ARNm H4.....	21
1.1 Introduction.....	21
1.2 Synthèse et purification de l'ARNm H4.....	21
1.3 Contrôle de l'ARN purifié par séquençage	23
1.4 Synthèse de l'ARN H4 présentant une extrémité 3' homogène	23
1.5 Repliement de l'ARNm H4	24
1.5.1 Introduction	24
1.5.2 Recherche d'un protocole de repliement de l'ARNm	25
1.5.2.1 Approche utilisée et élaboration d'un premier protocole de renaturation	25
1.5.2.2 Caractérisation du protocole retenu et optimisation	26
1.5.2.3 Conclusion	28
1.5.3 Recherche d'un second protocole de repliement de l'ARNm	28
1.5.3.1 Elaboration d'un deuxième protocole de renaturation.....	28
1.5.3.2 Caractérisation du protocole par paliers et optimisation.....	29
1.5.3.3 Conclusion	31
1.5.4 Test d'un troisième protocole de renaturation.....	31
1.5.5 Test d'un quatrième protocole de criblage	32
1.5.6 Pureté des conformères H et B et dégradation au cours des processus de renaturation	33
1.6 Effet de la coiffe sur le repliement de l'ARNm H4	34
1.7 Conclusion	35
Chapitre 2 : Caractérisation structurale préliminaire des conformations H et B	37
2.1 Stabilité des conformations H et B	37
2.1.1 Test de stabilité en fonction du temps.....	37
2.1.2 Test de stabilité en fonction de la température.....	38
2.2 Détermination du rayon hydrodynamique par diffusion dynamique de la lumière (DLS)..	38
2.3 Détermination du rayon de giration par diffusion des rayons X aux petits angles (SAXS)..	40
2.4 Détermination de la masse des conformations H et B par équilibre de sédimentation.....	41
2.5 Conclusion.....	42

Chapitre 3 : Caractérisation fonctionnelle des conformations H et B de l'ARNm H4	43
3.1 Tests de retard sur gel en présence d'eIF4E	43
3.2 Tests de traduction <i>in vitro</i>	44
3.3 Expériences de « toe-print »	45
3.4 Test de pontage de la coiffe aux UV	47
3.5 Conclusion	49
Chapitre 4 : Etude de la structure secondaire des conformations H et B.....	51
4.1 Sondage en solution des conformations H et B.....	52
4.2 Réalisation d'un modèle de structure secondaire	58
Chapitre 5 : Essais de cristallisation sur l'ARNm H4 complet ou tronqué.....	61
5.1 Essais de cristallisation sur l'ARNm H4 complet	61
5.1.1 Introduction	61
5.1.2 Inventaire et résultats des essais de cristallisation réalisés	61
5.2 Essais de cristallisation sur l'ARNm H4 tronqué	63
5.2.1 Sélection d'ARNm tronqués en 5' ou 3' ou 3' et 5'	63
5.2.2 Étude du repliement des ARN tronqués.....	64
5.2.3 Inventaire et résultats des essais de cristallisation réalisés	67
5.3 Conclusion	70
Chapitre 6 : Réexamen du modèle de structure secondaire de l'ARNm H4	71
6.1 Ré analyse des modèles de structure secondaire	71
6.1.1 Région 1-135 présentant la jonction à trois hélices	71
6.1.2 Région 136-188	73
6.1.3 Région 189-249 présentant l'élément 4E-SE.....	74
6.1.4 Région 250-340	75
6.2 Modèle actuel de structure secondaire de l'ARNm H4 complet	77
Conclusion générale et perspectives	79
ANNEXE : Matériel et méthodes	83
A. Méthodes couramment utilisées lors de la manipulation des acides nucléiques	83
a. Solutions couramment utilisées.....	83
b. Extraction des protéines au phénol	83
c. Précipitation à l'éthanol.....	84
d. Détermination de la concentration de l'ARN en solution	84
B. Synthèse et purification des ARN étudiés	84
a. Synthèse par amplification PCR d'une matrice à transcrire.....	84
b. Transcription <i>in vitro</i> par la T7 ARN polymérase	86
c. Purification des produits de transcription sur gel de polyacrylamide dénaturant	87
d. Coiffage des transcrits purifiés.....	87
e. Ultrafiltration des produits de coiffage.....	88
f. Elimination des agrégats par ultracentrifugation	88

C. Marquages au [³²P] des ARN étudiés et des amorces utilisées pour réaliser des transcriptions inverses	89
a. Marquage d'ARN au [³² P] en 5'	89
b. Marquage d'ARN au [³² P] en 3'	89
c. Marquage d'amorce au [³² P] en 5'	90
D. Etude fonctionnelle de l'ARNm H4	90
a. Expériences de retard sur gel en présence d'eIF4E	90
b. Tests de traduction <i>in vitro</i>	90
i. Traduction dans des lysats de reticulocytes de lapin	91
ii. Traduction dans des extraits de germe de blé	91
E. Etude structurale préliminaire de l'ARNm H4	92
a. Diffusion dynamique de la lumière	92
b. Ultracentrifugation analytique	93
c. Expériences de pontage aux UV	94
i. Préparation de l'ARNm H4 modifié	94
ii. Pontage aux UV et détection des modifications par transcription inverse	95
iii. Séquençage de Sanger par transcription inverse	95
iv. Analyse des résultats sur gel de polyacrylamide dénaturant	96
F. Etude du repliement de l'ARN en solution	96
a. Sondes utilisées et conditions d'utilisation	96
i. Sondes enzymatiques	96
ii. Sondage aux ions Pb ²⁺	98
iii. « In-line probing »	98
b. Réalisation d'échelles de coupure	99
G. Essais de cristallisation	99
a. Principe	100
b. Techniques utilisées	100
c. Matrices et approches utilisées	101
Bibliographie	103
Résumé	111

Abréviations

ABA	<u>A</u> crylamide- <u>b</u> is- <u>a</u> crylamide
ADN	<u>A</u> cide <u>d</u> ésoxyribo <u>n</u> ucléique
ARNm	<u>A</u> cide <u>r</u> ibo <u>n</u> ucléique <u>m</u> essenger
ARNt	<u>A</u> cide <u>r</u> ibo <u>n</u> ucléique de <u>t</u> ransfert
ARNm H4	<u>A</u> cide <u>r</u> ibo <u>n</u> ucléique <u>m</u> essenger de l'histone H4
ATP	<u>A</u> dénosine-5' - <u>t</u> riphosphate
DNase	<u>D</u> ésoxyribo <u>n</u> ucléase
DLS	<u>D</u> ynamic <u>l</u> ight <u>s</u> cattering
DMSO	<u>D</u> iméthyl <u>s</u> ulfoxyde
dNTP	<u>D</u> ésoxyribo <u>n</u> ucléotide <u>t</u> ri- <u>p</u> hosphate
DTT	<u>D</u> ithio <u>t</u> hréitol
EDTA	<u>E</u> thylene <u>d</u> iamine <u>t</u> etra <u>a</u> cetic acid
eIF	<u>E</u> karyotic <u>i</u> nitiation <u>f</u> actor
GTP	<u>G</u> uanosine-5' - <u>t</u> riphosphate
HEPES	4-(2- <u>h</u> ydroxy <u>e</u> thyl)-1-piperazine <u>e</u> thane <u>s</u> ulfonic acid
IRES	<u>I</u> nternal <u>r</u> ibosome <u>e</u> ntry <u>s</u> ite
MPD	2- <u>M</u> ethyl-2,4-pentanediol
nt(s)	nucleotide(s)
ORF	<u>O</u> pen <u>r</u> eading <u>f</u> rame
PEG	<u>P</u> oly <u>e</u> thylène <u>g</u> lycol
PCR	<u>P</u> olymerase <u>c</u> hain <u>r</u> eaction
p/v	<u>P</u> oids à <u>v</u> olume
RNase	<u>R</u> ibo <u>n</u> ucléase
rNTP	<u>R</u> ibo <u>n</u> ucleotide <u>t</u> ri- <u>p</u> hosphate
rpm	<u>R</u> otation par <u>m</u> inute
SAXS	<u>S</u> mall- <u>a</u> nge <u>X</u> -ray <u>s</u> cattering
SDS	<u>D</u> odecyl <u>S</u> ulfate de <u>S</u> odium
SE	<u>S</u> edimentation <u>e</u> quilibrium
Tris	<u>T</u> ris-(hydroxyméthyl) aminométhane
UTR	<u>U</u> n <u>t</u> ranslated <u>r</u> egion
UV	<u>U</u> ltraviolet
v/v	<u>V</u> olume à <u>v</u> olume
WT	<u>W</u> ild type

Avant propos de la thèse

Mon travail de thèse s'est placé dans le contexte de l'étude de l'initiation de la traduction de l'histone H4 menée dans l'équipe « Evolution des systèmes d'initiation de la traduction chez les eucaryotes » de l'UPR9002 dans l'Institut de Biologie Moléculaire et Cellulaire du CNRS. Les travaux antérieurs menés au laboratoire avaient conduit à décrire un nouveau mécanisme d'initiation de la traduction conduisant à une expression massive de l'histone H4. Ce mécanisme distinct du processus de « scanning », baptisé « tethering », repose sur plusieurs structures essentielles présentes dans la phase codante de l'ARNm. Ainsi, la phase codante de l'ARNm joue à la fois le rôle classique de messenger porteur de l'information génétique et d'élément fonctionnel participant au recrutement des facteurs d'initiation et des ribosomes. Cette double propriété a conduit au développement d'un projet d'étude structurale avec l'équipe « Biologie structurale de cibles épigénétiques » de l'UMR7104 de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire dans laquelle j'avais effectué mon stage de recherche de master 2. J'avais alors réalisé des expériences préliminaires de purification de l'ARNm H4 et de cristallisation.

À l'issue de mon master, j'ai décidé de poursuivre l'étude structurale sur l'ARNm de l'histone H4 en commençant une thèse en co-direction entre les deux laboratoires afin d'optimiser l'apprentissage des techniques à la fois biochimiques et biophysiques et des connaissances fonctionnelles et structurales. Les résultats obtenus durant les 3 années de ma thèse sont développés dans les chapitres qui suivent ; les grandes lignes sont brièvement présentées ci-dessous.

Dans une première étude, j'ai optimisé les protocoles de synthèse et de purification de l'ARNm H4 afin d'obtenir de grandes quantités d'ARN pur et homogène du côté 3'. Puis des conditions de renaturation ont été mises au point afin d'isoler deux formes d'ARNm dont le repliement était stable en conditions natives. Ces deux formes ont été caractérisées d'un point de vue biophysique (effet du temps, température, études DLS, SAXS et équilibre de sédimentation) et fonctionnel (aptitude à être traduites, à fixer eIF4E, la coiffe et les ribosomes).

Pour progresser dans l'étude cristallographique, j'ai entrepris de diversifier les ARNm à cristalliser en recherchant des fragments de tailles plus réduites et plus adaptées aux études structurales. Pour cela, j'ai réalisé des études de sondage en solution sur l'ARNm afin d'établir un modèle de repliement permettant de délimiter des domaines minimaux susceptibles d'être plus facilement cristallisables. Des sondes enzymatiques et chimiques ont été utilisées. Les résultats ont montré que les deux formes d'ARNm partagent une grande majorité de signaux de coupure et de réactivité mais qu'elles présentent aussi des différences notables dans certaines zones de l'ARN. Malgré l'abondance des données de sondage, il n'a pas été possible de déduire deux modèles de structure secondaire distincts et exclusifs. Il semblerait que les deux formes d'ARNm ne différencieraient que par des interactions à distance entre des régions éloignées de l'ARNm. Ce modèle de repliement secondaire a permis d'effectuer un choix de fragments d'ARNm, dans le but d'isoler des domaines structurellement indépendants et homogènes d'un point de vue du repliement. Au total, 21 fragments de l'ARNm H4 ont été synthétisés et purifiés. Nous avons réussi à obtenir une population structurellement homogène en condition native d'un grand nombre de ces délétants mais malgré de nombreux tests de cristallisation, aucun cristal d'ARN n'a été obtenu à ce jour.

Introduction

A. Les histones, constituants de la chromatine

a. Structure et fonction des histones

Le nucléosome, unité de base de la chromatine, est formé par un octamère d'histone constitué de deux hétérodimères H2A-H2B et d'un tétramère H3-H4 autour desquels l'ADN est enroulé en formant 1,7 tour d'une super hélice gauche (soit 145 à 147 paire de base) (Arents et al., 1991; Kornberg, 1974; Kornberg and Thomas, 1974; Luger et al., 1997). Sur cet octamère d'histone se fixe l'histone H1, dont la fonction est d'assurer la compaction du nucléosome et de rigidifier la structure de l'ADN. Chaque histone possède une extension amino-terminale dépassant du nucléosome. Ces queues sont sujettes à un ensemble de modifications post-traductionnelles, comme l'acétylation des lysines, la méthylation des lysines et des arginines, la phosphorylation des sérines et des thréonines, l'ubiquitination des lysines, la sumoylation des lysines et la désamination des arginines (Allfrey and Mirsky, 1964; Khorasanizadeh, 2004). Ces modifications peuvent avoir lieu séquentiellement ou de manière combinée pour former un code histone permettant de définir l'état transcriptionnel de la chromatine (Strahl and Allis, 2000). Ces modifications covalentes régulent les fonctions génomiques telles que la transcription, la réparation et la réplication (Bedford and Richard, 2005; Pahlich et al., 2006). Les histones, particulièrement H3 et H4, figurent parmi les protéines les mieux conservées chez les eucaryotes (Piontkivska et al., 2002). Les alignements multiples des gènes d'histone H4 indiquent que la conservation de la région codante est inhabituellement élevée au niveau nucléotidique (supérieur à 90%) et excède le besoin pour maintenir la séquence protéique, ce qui suggère que cette conservation au niveau nucléotidique reflète une pression de sélection sur la structure secondaire de ces ARNm qui s'est exercé au cours de l'évolution (Huynen et al., 1992).

b. Synthèse et structure des ARNm d'histones

Chez les métazoaires, les histones sont codées par une famille particulière de gènes, les gènes d'histones réplication-dépendants. Ces gènes ont pour particularité d'être dépourvus d'introns et leur expression est strictement restreinte à la phase S du cycle cellulaire afin de compacter l'ADN néo synthétisé (Osley, 1991; Pyronnet and Sonenberg, 2001). L'expression des gènes d'histones doit être stœchiométrique sous peine d'induire une instabilité génétique et une perte de chromosomes (Meeks-Wagner and Hartwell, 1986). En effet, la présence d'histones libres dans la cellule est extrêmement délétère pour le développement de la drosophile et de *C. elegans* (Kodama et al., 2002; Pettitt et al., 2002; Salzler et al., 2009; Sullivan et al., 2001). Des analyses réalisées à partir de puces à ADN sur des cellules HeLa synchronisées ont montré que les ARNm codant pour les histones sont détectés en milieu de phase S, soit 3 h après le début de la phase S (van der Meijden et al., 2002). En 1 h, 30 millions de nucléosomes doivent être synthétisés, ce qui suggère qu'un mécanisme particulièrement efficace est utilisé pour synthétiser ces histones (Gunjan et al., 2005; Osley, 1991).

Au niveau structural, les ARNm d'histone possèdent des extrémités non traduites (UTR) très courtes (entre 20 et 60 nucléotides). Les ARNm des histones sont également les seuls ARNm eucaryotiques dépourvus de queue poly(A) à leur extrémité 3' (Piontkivska et al., 2002). À la place, l'extrémité 3' se termine par une séquence de 26 nucléotides fortement conservée et contenant une structure en tige boucle de 16 nucléotides. Cette tige boucle participe à l'étape de « processing » des pré-ARNm d'histones (Jaeger et al., 2005; Martin et al., 1997; Marzluff et al., 2008). Cette organisation structurale particulière est un autre indice qui suggère que ces ARNm sont traduits à l'aide d'un mécanisme non conventionnel.

B. Traduction

La traduction des ARNm est l'étape finale de l'expression des gènes. C'est un processus cellulaire complexe, indispensable à la vie cellulaire. Ce processus nécessaire mais non suffisant à l'expression génétique assure la constitution du protéome à partir de l'information génétique contenue dans l'ARNm. Ce mécanisme a été hautement conservé au cours de l'évolution, si bien que les cellules procaryotes et eucaryotes partagent les mêmes fondamentaux. La traduction se divise en 4 étapes : initiation, élongation, terminaison et recyclage. La traduction est un processus cyclique : les particules ribosomiques impliquées dans l'initiation de la traduction sont issues d'un processus de recyclage ciblant les complexes post traductionnels.

L'initiation de la traduction est le processus dont la finalité est d'assembler un ribosome avec un ARNt initiateur dont l'anticodon est apparié avec l'AUG initiateur dans le site P du ribosome. La différence principale entre eucaryotes et procaryotes se situe au niveau du mécanisme utilisé pour la reconnaissance du codon d'initiation.

a. Mécanisme d'initiation de la traduction coiffe dépendante

Le mécanisme d'initiation de la traduction dépendant de la coiffe (« coiffe dépendante ») ou modèle consensus d'initiation de la traduction chez les eucaryotes est utilisé par environ 97% des ARNm eucaryotes connus. Ce processus nécessite l'intervention de 9 à 13 facteurs protéiques (**Figure 1, Tableau 1**) appelés eIF (eukaryotic Initiation Factor) ainsi que la formation de nombreux complexes intermédiaires, ce qui fait d'elle l'étape limitante de la traduction (Sonenberg and Hinnebusch, 2009; Trachsel et al., 1977). Durant l'étape de transcription par l'ARN polymérase II, les ARNm sont coiffés à leur extrémité 5' et sont polyadénylés au niveau de leur extrémité 3'. Ces éléments structuraux sont ensuite spécifiquement reconnu par des facteurs agissant en *trans*, afin d'activer l'ARNm.

Le processus d'initiation débute par la formation du complexe eIF4F sur la coiffe de l'ARNm. Ce complexe est constitué des facteurs eIF4A, eIF4E et eIF4G dont la fonction est de respectivement déplier la partie 5' UTR de l'ARNm (activité ARN hélicase), de fixer la coiffe de l'ARNm, et d'assurer la cohésion du complexe (Asano et al., 2001; Edery and Sonenberg, 1985; Etchison et al., 1982; He et al., 2003; Hentze, 1997; Lamphear et al., 1995; Pause et al., 1994; Sonenberg et al., 1978). L'interaction entre la coiffe et la surface concave d'eIF4E repose essentiellement sur des interactions de type « π stacking » entre la purine méthylée et 2 résidus tryptophane conservés (W102 et 156) et sur des interactions hydrogène entre le méthyl de la purine et un troisième résidu tryptophane conservé (W166) (Marcotrigiano et al., 1997). Une fois formé, ce complexe va déplier les structures secondaires présentes sur la partie 5' UTR de l'ARNm grâce à l'activité hélicase d'eIF4A (Merrick and Hershey, 1996). Suite à la formation du complexe eIF4F, eIF4B, un facteur auxiliaire d'eIF4A, se joint au complexe (Rozen et al., 1990). En parallèle, la queue poly(A) de l'ARNm se fixe à une protéine spécifique : la PABP (PolyA Binding protein). La PABP interagit à son tour avec eIF4G, afin d'induire la circularisation de l'ARNm en rapprochant l'extrémité 3' de l'extrémité 5' (Banerjee, 1980; Gingras et al., 1999; Tarun and Sachs, 1996). L'ARNm est alors dit activé (Wells et al., 1998).

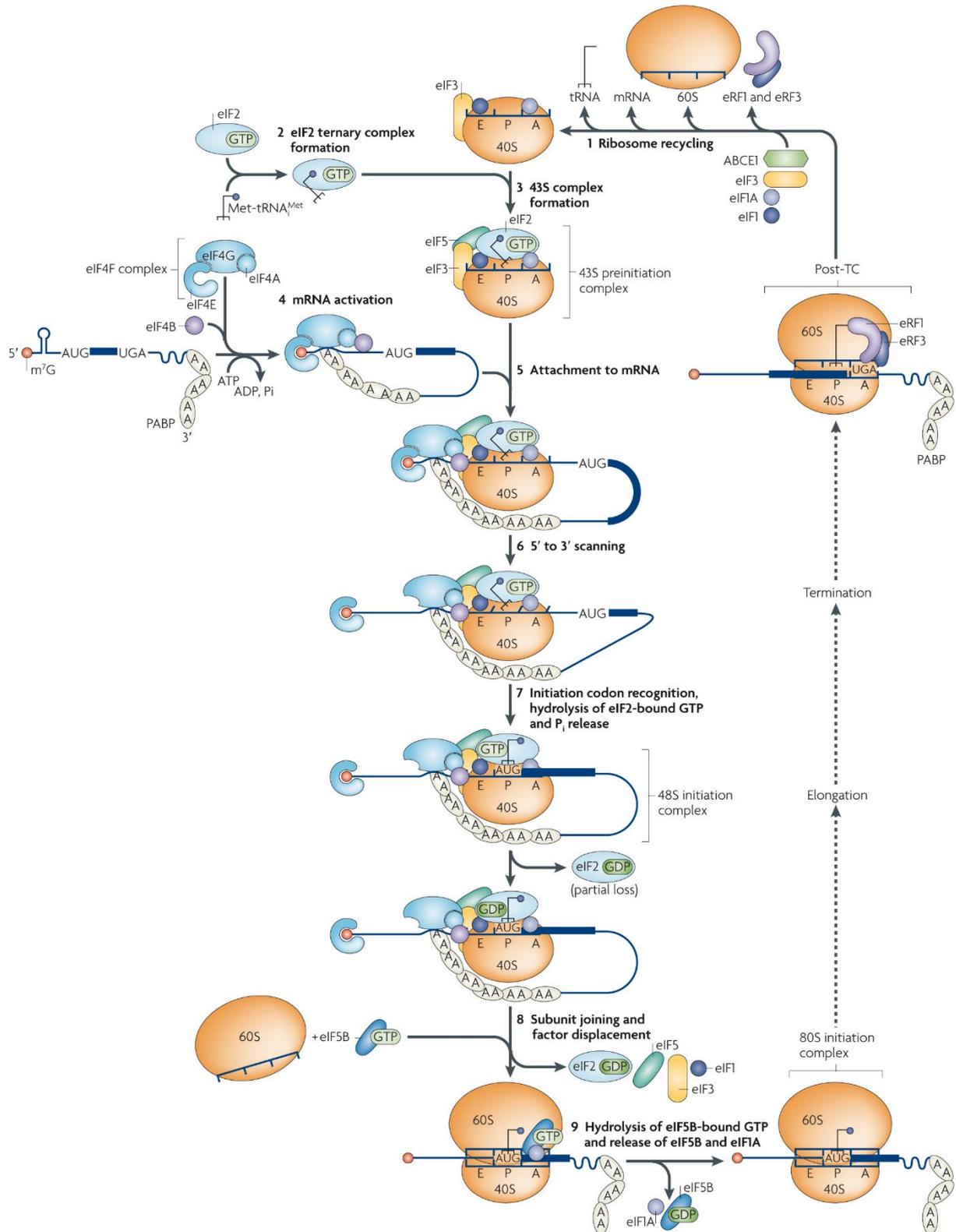


Figure 1. Modèle du mécanisme d'initiation de la traduction coiffe dépendante. Ce processus est constitué de 8 étapes (2 à 9). La traduction est un processus cyclique : les particules ribosomiques impliquées dans l'initiation de la traduction sont issues d'un processus de recyclage ciblant les ribosomes ayant achevé de traduire (étape 1) (figure reprise de (Jackson et al., 2010)).

Parallèlement à l'activation de l'ARNm, se forme le complexe de pré-initiation 43S, constitué de la sous unité 40S, du complexe ternaire eIF2-GTP-Met-ARNt^{met}, d'eIF3, d'eIF1, d'eIF1A et probablement d'eIF5 (Dever, 2002; Majumdar et al., 2003; Trachsel and Staehelin, 1978, 1979). Ce complexe est ensuite recruté au niveau de la partie 5' UTR de l'ARNm grâce à l'interaction entre eIF4G et eIF3. Une fois le complexe de pré-initiation 43S recruté, celui-ci va glisser de 5' en 3' sur l'ARNm (mécanisme dit de « scanning ») jusqu'au codon d'initiation AUG (Kozak, 1989, 1991; Mitchell and Lorsch, 2008; Pestova and Kolupaeva, 2002). Ce processus de « scanning » est assuré par le facteur eIF1 dont la fonction est de stabiliser la sous unité 40S dans une conformation dite « ouverte », c'est-à-dire compétente pour glisser sur l'ARNm (Passmore et al., 2007; Pestova et al., 1998a; Pestova and Hellen, 2000). Au cours de ce processus, la reconnaissance du codon d'initiation est assurée par le facteur eIF1 qui a la capacité de discriminer les codons trop proches de l'extrémité 5' de l'ARNm (Pestova and Kolupaeva, 2002). La reconnaissance et le positionnement du codon d'initiation au niveau du site P du ribosome induit un changement de conformation du complexe de scanning : celui-ci adopte alors une conformation dite fermée. Ce changement de conformation provoque le déplacement d'eIF1, afin de permettre à eIF5 de catalyser l'hydrolyse du GTP fixé sur eIF2 et par conséquent de permettre à l'anticodon du Met-ARNt^{Met} initiateur de s'ancrer dans le site P du ribosome et d'interagir avec le codon d'initiation AUG (Algire et al., 2005; Huang and Carmichael, 1997; Lee et al., 2002; Lorsch and Herschlag, 1999; Maiti et al., 2000; Passmore et al., 2007). Le complexe 48S est alors formé.

Le facteur eIF5B-GTP va ensuite induire dans un premier temps la dissociation d'eIF1, d'eIF3 et d'eIF2-GDP, ainsi que l'association de la sous unité 60S avec la sous unité 40S (Pestova and Hellen, 2000). L'hydrolyse du GTP fixé sur eIF5B n'est pas nécessaire pour cette étape, mais une fois le ribosome 80S assemblé, elle est requise pour que ce facteur puisse libérer eIF1A et se détacher de la sous unité 40S (Acker and Lorsch, 2008; Merrick, 1979). Le ribosome est alors prêt à entrer en phase d'élongation et le processus d'initiation est alors terminé. Bien que la plupart des ARNm utilise ce mécanisme dit de « scanning » pour former le complexe 48S, sur d'autres ARNm, la formation du complexe 48S est induite par un mécanisme de type « IRES ».

Facteurs d'initiation principaux	Nombre de sous unités et masses moléculaires totales (kDa)	Fonction
eIF2	3 (125,6)	Impliqué dans la formation d'un complexe ternaire avec un ARNt ^{met} initiateur et du GTP. GTPase permettant la fixation du Met-tRNA _i ^{met} au niveau du site P de la sous-unité 40S.
eIF3	13 (800)	Catalyse la fixation du complexe ternaire sur la sous unité 40S du ribosome. Impliqué dans la formation du complexe de pré-initiation 48S, en interagissant simultanément avec la sous unité 40S du ribosome, eIF1, eIF4G et eIF5. Ce qui fait de ce facteur la clef de voûte du complexe de pré-initiation 48S. Prévient l'association spontanée des sous unités 40S et 60S et possède la capacité de dissocier ces 2 sous unités.
eIF1	1 (12,7)	Assure l'étape de « scanning » et favorise la fixation du complexe ternaire sur la sous unité 40S du ribosome. Impliqué dans la reconnaissance du codon AUG initiateur au cours de l'étape de « scanning ». Inhibe l'activité GTPase d'eIF2, afin de prévenir une hydrolyse prématurée induite par eIF5.
eIF1A	1 (16,5)	Favorise la fixation du complexe ternaire sur la sous unité 40S du ribosome. Impliqué dans la reconnaissance du codon AUG initiateur au cours de l'étape de « scanning ».
eIF4E	1 (24,5)	Fixe la coiffe de l'ARNm
eIF4A	1 (46,1)	ARN hélicase ATP dépendante
eIF4G	1 (175,5)	Plate-forme multifactorielle impliquée dans la formation du complexe eIF4F et du complexe 48S en interagissant simultanément avec eIF4E, eIF4A, eIF3, PABP et l'ARNm. Stimule l'activité hélicase d'eIF4A.
eIF4F	3 (246,1)	Complexe recruté par la coiffe de l'ARNm et constitué des facteurs eIF4E, eIF4A et eIF4G
eIF4B	1 (69,3)	Assure la processivité d'eIF4A.
eIF4H	1 (27,4)	Assure la processivité d'eIF4A.
eIF5	1 (49,2)	Stimule l'activité GTPase d'eIF2 suite à l'appariement de l'ARNt ^{met} initiateur sur le codon d'initiation AUG.
eIF5B	1 (138,9)	GTPase permettant la dissociation des facteurs eIF1, eIF1A, eIF2-GDP et eIF3. Impliqué dans l'assemblage des 2 particules ribosomiques.
eIF2B	5 (262,9)	Facteur impliqué dans l'étape de recyclage, remplace le GDP d'eIF2 par du GTP.
Facteurs d'initiation auxiliaires		
eIF6	1 (26,6)	Empêche l'association spontanée des 2 sous unités du ribosome et possède la capacité de dissocier ces 2 sous unités. Pourrait être davantage un facteur impliqué dans la biogenèse des ribosomes.
PABP	1 (70,7)	Fixe la queue poly(A) de l'ARNm, eIF4G et eIF3 Améliore l'affinité d'eIF4F pour la coiffe

Tableau 1. Tableau récapitulant les facteurs impliqués directement ou indirectement dans l'initiation de la traduction.

b. Mécanisme d'initiation coiffe indépendant induit par des IRES

Le mécanisme de type IRES (« Internal Ribosome Entry Site ») est un mécanisme alternatif d'initiation de la traduction utilisé notamment par les ARNm viraux (Jackson et al., 1990; Jackson and Kaminski, 1995; Pelletier and Sonenberg, 1988). Au cours de l'évolution, les ARNm viraux ont mis en place des stratégies pour détourner la machinerie traductionnelle de la cellule hôte afin de favoriser leur traduction au dépend des ARNm cellulaires, ce qui leur permet de se multiplier et de se propager efficacement. La machinerie traductionnelle est détournée à l'aide d'éléments structuraux dont la fonction est de recruter les particules ribosomiques et/ou les facteurs d'initiation. Ces éléments structuraux sont appelés IRES.

Jusqu'à présent, 4 types d'IRES différents ont été caractérisés (**Figure 2** ; (Jackson et al., 2010)). Ces différents types d'IRES diffèrent par leurs architectures et par le nombre d'eIF requis. On estime qu'environ 3% des ARNm cellulaires utilisent également un mécanisme de type IRES (Johannes et al., 1999), parmi ces ARNm, on peut citer les ARNm codant pour Antennapedia (Antp), BiP (Binding immunoglobulin Protein) et FGF2 (Fibroblast Growth Factors) (Macejak and Sarnow, 1991; Oh et al., 1992; Vagner et al., 1995).

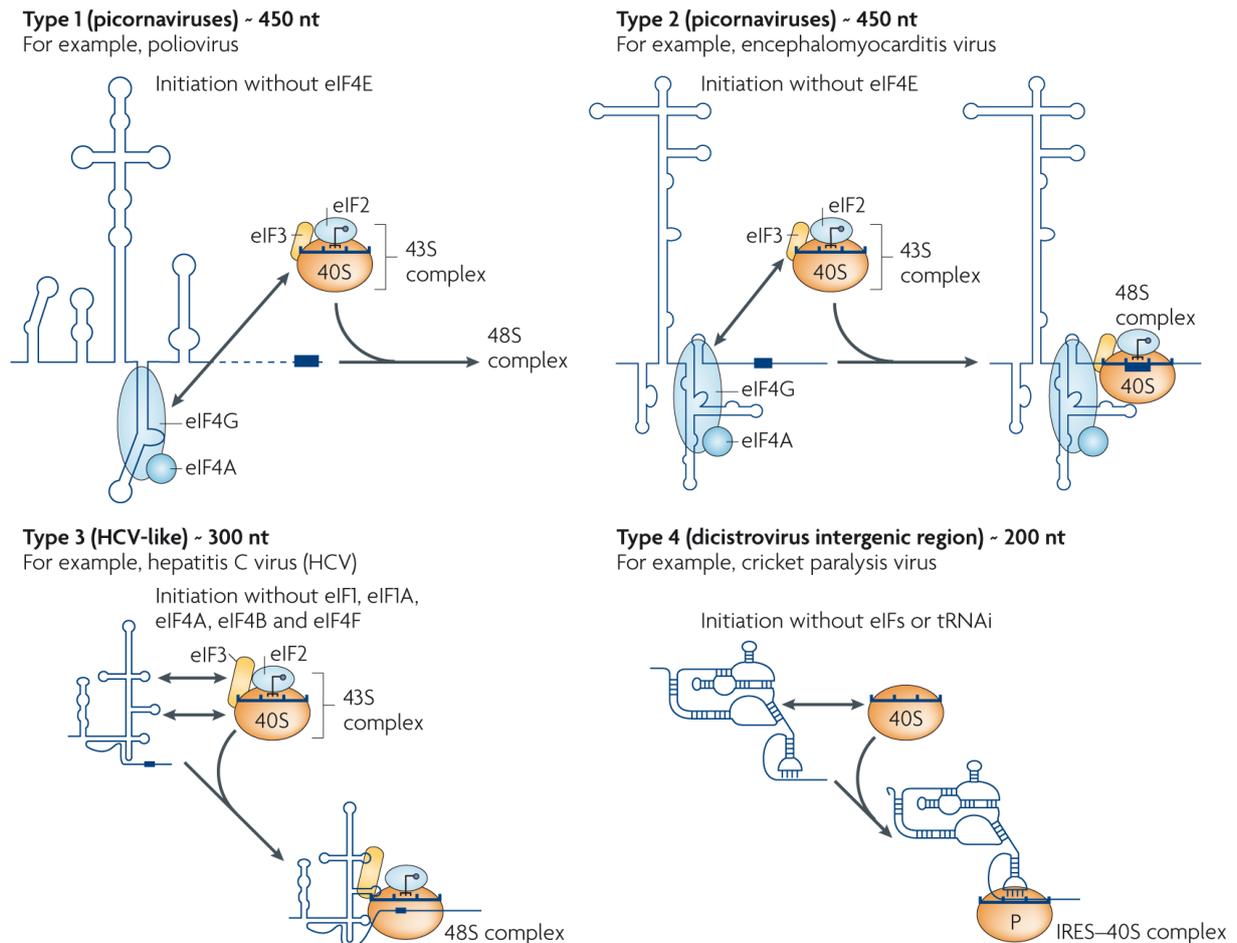


Figure 2. Modèles du mécanisme d'initiation de la traduction coiffe indépendante induit par des IRES (figure reprise de (Jackson et al., 2010)).

i. IRES de type I

Les IRES de type I sont retrouvés dans le génome des *Picornaviridae*, une famille de virus à ARN de polarité positive, c'est-à-dire directement traduits par la machinerie cellulaire. Parmi les virus qui utilisent ce type d'IRES, on peut citer le virus de la polio (Pelletier and Sonenberg, 1988). Afin d'assembler un ribosome fonctionnel au niveau de leur AUG initiateur, les IRES de type I nécessitent tous les eIFs impliqués dans le modèle consensus d'initiation de la traduction chez les eucaryotes, sauf eIF4E et son domaine d'interaction sur eIF4F (la partie N-terminale) (Sonenberg and Meerovitch, 1990). Au début de l'infection virale, le poliovirus

synthétise une protéase (2Apro) dont la fonction est de cliver eIF4F en deux parties (domaine N-terminal et P100) afin d'inactiver l'initiation coiffe dépendante utilisée par la cellule hôte, ce qui permet au virus de s'approprier la machinerie de traduction cellulaire (Davies et al., 1991; Gradi et al., 1998; Svitkin et al., 1999). Le recrutement du complexe de pré-initiation 43S est réalisé en amont du codon d'initiation AUG, sur un AUG silencieux, par l'intermédiaire d'eIF4G tronqué en N-terminal (Sonenberg and Meerovitch, 1990). Une fois le complexe de pré-initiation 43S recruté, celui-ci va glisser sur l'ARNm de 5' en 3' (mécanisme dit de « scanning ») jusqu'au codon d'initiation AUG (Belsham, 1992). Le complexe 80S est ensuite assemblé de la même manière que dans le cas du modèle d'initiation dépendant de la coiffe (Sonenberg and Meerovitch, 1990). En plus des facteurs d'initiation de la traduction, ce mécanisme utilise d'autres protéines cellulaires : les ITAF (« IRES Trans-Acting Factor »). Le poliovirus utilise ces ITAF pour moduler et stabiliser l'architecture de son IRES dans le but de recruter efficacement le complexe de pré-initiation 43S. Dans le cas du poliovirus, on dénombre au moins 6 ITAF impliqués dans la régulation de l'initiation de la traduction, il s'agit des protéines PCBP, PTB, La, SRp20, DRBP76 et unr (Gamarnik and Andino, 1996; Hellen et al., 1993; Hunt and Jackson, 1999; Meerovitch et al., 1989; Meerovitch et al., 1993; Pestova et al., 1991; Svitkin et al., 1994a; Svitkin et al., 1994b). La fonction précise de ces ITAF dans le cas d'une infection par le poliovirus est encore mal comprise et continue à être étudiée (Niepmann, 2009).

ii. IRES de type II

Les IRES de type II sont également retrouvés dans le génome des *Picornaviridae*. Contrairement aux IRES de type I, ces IRES sont traduits avec une bonne efficacité dans les réticulocytes de lapin. Parmi les virus qui utilisent ce mécanisme, on peut citer l'exemple du virus de l'encéphalomyocardite virale (EMCV) (Jang et al., 1988). Pour assurer l'étape d'initiation de la traduction, les IRES de types II utilisent tous les facteurs d'initiation conventionnels à l'exception d'eIF4E et de la partie N-terminale d'eIF4F (Kolupaeva et al., 1998; Pestova et al., 1996a; Pestova et al., 1996b). Les facteurs requis sont donc les mêmes que pour les IRES de type I. L'une des différences majeures entre ces 2 types d'IRES se situe au niveau de la stratégie utilisée pour s'approprier de manière exclusive la machinerie traductionnelle de la cellule hôte. Dans le cas des IRES de type II, cette stratégie est basée sur l'activation d'un répresseur cellulaire de la traduction : 4E-BP1 (Pause et al., 1994). Ce répresseur inhibe spécifiquement la traduction coiffe dépendante en se fixant sur le domaine d'interaction avec eIF4G d'eIF4E, ce qui inhibe la formation du complexe eIF4F (Gingras et

al., 1996). L'autre différence majeure avec les IRES de type I est que le complexe de pré-initiation 43S est recruté directement sur le codon d'initiation AUG, il n'y a donc pas d'étape de « scanning » (Kaminski et al., 1990).

iii. IRES de type III

Ces IRES sont retrouvés dans le génome des *Flaviviridae*, une famille de virus à ARN de polarité positive directement traduits par la machinerie cellulaire. Parmi les virus qui utilisent ce type d'IRES, on peut citer le virus de l'hépatite C (HCV) (Reynolds et al., 1995; Tsukiyama-Kohara et al., 1992; Wang et al., 1993). Dans le cas de HCV, la formation du complexe 48S est réalisée sans étape de « scanning ». Seul les facteurs eIF3 et eIF2 sont requis pour assurer la formation du complexe 48S (Pestova et al., 1998b). Ceci s'explique par la haute affinité de certains éléments structuraux de l'IRES pour la sous unité 40S (1,9 nM) et pour eIF3 (35 nM) (Ji et al., 2004; Kieft et al., 2001a; Kieft et al., 2001b; Kolupaeva et al., 2000; Otto and Puglisi, 2004). Cette haute affinité permet à l'IRES de former un complexe stable avec la sous unité 40S et de recruter le complexe ternaire eIF2-GTP-Met-ARNt^{Met} dont le Met-ARNt^{Met} va interagir directement avec l'AUG dans le site P du ribosome (Pestova et al., 1998b). Contrairement au mécanisme d'initiation de la traduction coiffe dépendant, eIF3 se fixe directement sur un élément structural de l'ARNm (Kieft et al., 2001b). Le complexe 80S est ensuite assemblé à l'aide des facteurs eIF5B-GTP et eIF5 de la même manière que dans le cas du modèle d'initiation de la traduction coiffe dépendant (Locker et al., 2007; Pestova et al., 1998b).

iv. IRES de type IV

Ces IRES sont retrouvés dans le génome des *Dicistroviridae*, une famille de virus à ARN de polarité positive. Parmi les virus utilisant ce type d'IRES on peut citer le virus de la paralysie du criquet (CrPV) (Wilson et al., 2000b). Ces IRES se différencient des autres types d'IRES par leur faculté à assembler un ribosome fonctionnel sans l'aide de facteurs d'initiation et de Met-ARNt^{Met}. La totalité du processus d'initiation est assuré par des éléments structuraux de l'IRES (Wilson et al., 2000a). Une autre particularité de ce mécanisme est que la traduction est initiée dans le site A du ribosome sur un codon GCU (Wilson et al., 2000a).

c. Modèle d'initiation de la traduction de l'histone H4

Messenger atypique, l'ARNm de l'histone H4 possède une séquence codante courte (309 nt) possédant un biais en G/C (61,4%). L'extrémité 5' UTR de l'ARNm de l'histone H4-12 de la souris est particulièrement courte. Avec seulement 9 nucléotides, elle est l'extrémité la plus courte connue parmi les ARNm d'histones H4 (Piontkivska et al., 2002). Les études réalisées dans mon laboratoire d'accueil à l'IBMC ont montré que l'ORF de l'ARNm H4 contenait au moins 2 éléments structuraux impliqués dans le processus d'initiation de la traduction. Le premier élément, localisé à proximité de l'extrémité 5', est replié sous forme d'une jonction à 3 hélices permettant un positionnement direct du ribosome sur l'AUG sans avoir recours à l'étape de scanning (Martin et al., 2011). Le deuxième élément structural est replié sous la forme d'une double tige boucle similaire à la structure de l'élément 4E-SE (« eIF4E Sensitive Element ») trouvé dans la région 3' UTR de l'ARNm de la cycline D1 (Culjkovic et al., 2007; Culjkovic et al., 2006). La fonction de ce domaine est de fixer le facteur d'initiation eIF4E du complexe eIF4F et de là, recruter les particules ribosomiques 43S pour les déposer à proximité de l'AUG initiateur (Martin et al., 2011). Contrairement au modèle consensus d'initiation de la traduction, le recrutement d'eIF4E se fait donc indépendamment de la coiffe. Pour assurer cette interaction, l'extension N-terminale d'eIF4E est essentielle mais non suffisante. Curieusement, cette extrémité N-terminale n'est pas phylogénétiquement conservée et jusqu'à présent aucune autre fonction ne lui est attribuée. Bien que la coiffe ne soit pas nécessaire pour recruter le complexe eIF4F, celle-ci est indispensable pour que le ribosome puisse entrer en phase d'élongation (Martin et al., 2011).

Dans le modèle du mécanisme d'initiation proposé dans la **figure 3** (Martin et al., 2011), la coiffe est séquestrée dans un premier temps à l'intérieur de la structure repliée en jonction à 3 hélices afin que la coiffe n'entre pas en compétition avec l'élément 4E-SE pour recruter eIF4E. La fixation du complexe eIF4F sur l'ORF de l'ARNm H4 entraîne un changement de conformation de l'ARNm H4 qui conduit à la libération de la coiffe. L'interaction entre eIF4E et la coiffe pourrait entraîner le rapprochement d'eIF4A de la jonction à 3 hélices de l'ARNm H4. Une fois positionné, eIF4A peut déplier la jonction en 3 hélices libérant ainsi le ribosome pour l'étape d'élongation. Ce modèle explique la dépendance vis-à-vis de la coiffe et l'assemblage direct du ribosome au niveau du codon AUG. Le mécanisme proposé est un mécanisme hybride combinant des caractéristiques propres au modèle consensus d'initiation de la traduction chez les eucaryotes (utilisation de la coiffe et d'eIF4E) et des caractéristiques retrouvées chez les IRES (présence d'éléments structuraux impliqués dans le recrutement interne de facteur d'initiation de la traduction). Ce mécanisme d'initiation par « tethering » de

la machinerie d'initiation n'a pas encore été observé chez d'autres ARNm cellulaires et il est vraisemblablement à l'origine de la très grande efficacité de traduction qui caractérise les ARNm d'histones lors de la phase S du cycle cellulaire.

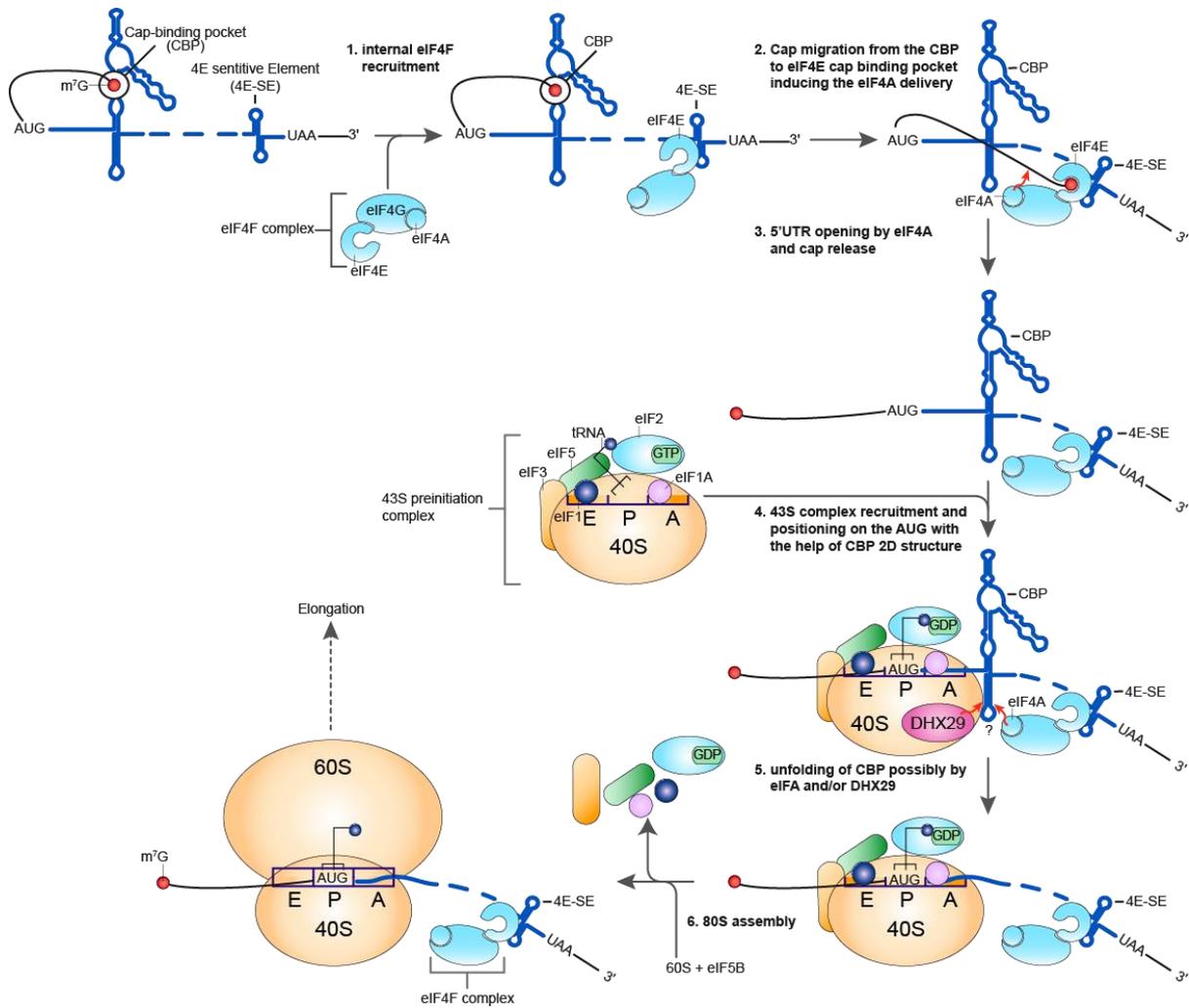


Figure 3. Modèle du mécanisme d'initiation de la traduction de l'ARNm de l'histone H4. La région codante contient 2 éléments structuraux impliqués dans le processus d'initiation de la traduction : une jonction à 3 hélices (CBP) et une double tige boucle (4E-SE) (figure reprise de (Martin, 2012)).

Chapitre 1 : Purification et repliement de l'ARNm H4

1.1 Introduction

À mon arrivée au laboratoire, l'ARNm de l'histone H4 (ARNm H4 ci-après) était couramment synthétisé et purifié pour les études fonctionnelles et les expériences de sondage en solution destinées à élucider le repliement de l'ARNm. Or, ces expériences avaient révélé la présence de réactivités contradictoires dans certaines régions de l'ARNm où des coupures double et simple brin co-existaient (Jaeger, 2005). Il avait aussi été démontré qu'après un protocole de repliement classique, l'ARNm H4 se repliait en deux formes majeures bien observables par électrophorèse sur gel natif. Ces deux conformations de l'ARNm pouvaient expliquer les différences de réactivité et suggéraient que les deux formes différaient par leur contenu en régions appariées.

Pour élucider le repliement de ces deux formes d'ARNm, pour tester la fonctionnalité de ces deux formes mais aussi pour isoler des ARNm les plus homogènes possible pour les études structurales, j'ai entrepris une première étude visant à établir les conditions de repliement permettant d'obtenir isolément chacune des deux formes majeures de l'ARNm H4.

1.2 Synthèse et purification de l'ARNm H4

L'ARNm H4 a été synthétisé par transcription *in vitro* par l'ARN polymérase du phage T7 à partir d'une matrice d'ADN constituée d'un fragment du gène de l'histone H4 porteur du promoteur T7 (amplifié par PCR à partir d'un plasmide pUC19-H4). La transcription du gène H4 est généralement très efficace comme le montre le gel présenté dans la **figure 4A**. Une bande unique est visible, suggérant que le produit de transcription est largement majoritaire. L'ARNm synthétisé est ensuite purifié en conditions dénaturantes sur gel ABA-8 M urée pour éliminer les transcrits abortifs, les NTP résiduels et les produits de dégradation de la matrice de départ (digérée par la DNase en fin de transcription). Après séparation électrophorétique, la bande correspondant à l'ARNm H4 est découpée, éluee par électrophorèse puis analysée sur un gel ABA-8 M urée. La **figure 4B** montre que l'ARN final ne présente aucun produit de dégradation.

Ce protocole nous a permis d'obtenir de manière reproductible un ARNm H4 pur, avec un rendement de 2 mg d'ARN par millilitre de transcription. Durant mon stage de master, j'avais testé l'utilisation d'une colonne MonoQ comme outil de purification principal à la place de la séparation électrophorétique et l'électroélution, mais la colonne ne permettait pas d'éliminer les fragments de dégradation de taille voisine du fragment entier (données non présentées). De même, l'utilisation de colonne de type « gel filtration » n'ajoutait pas à la qualité de l'ARNm purifié (données non présentées).

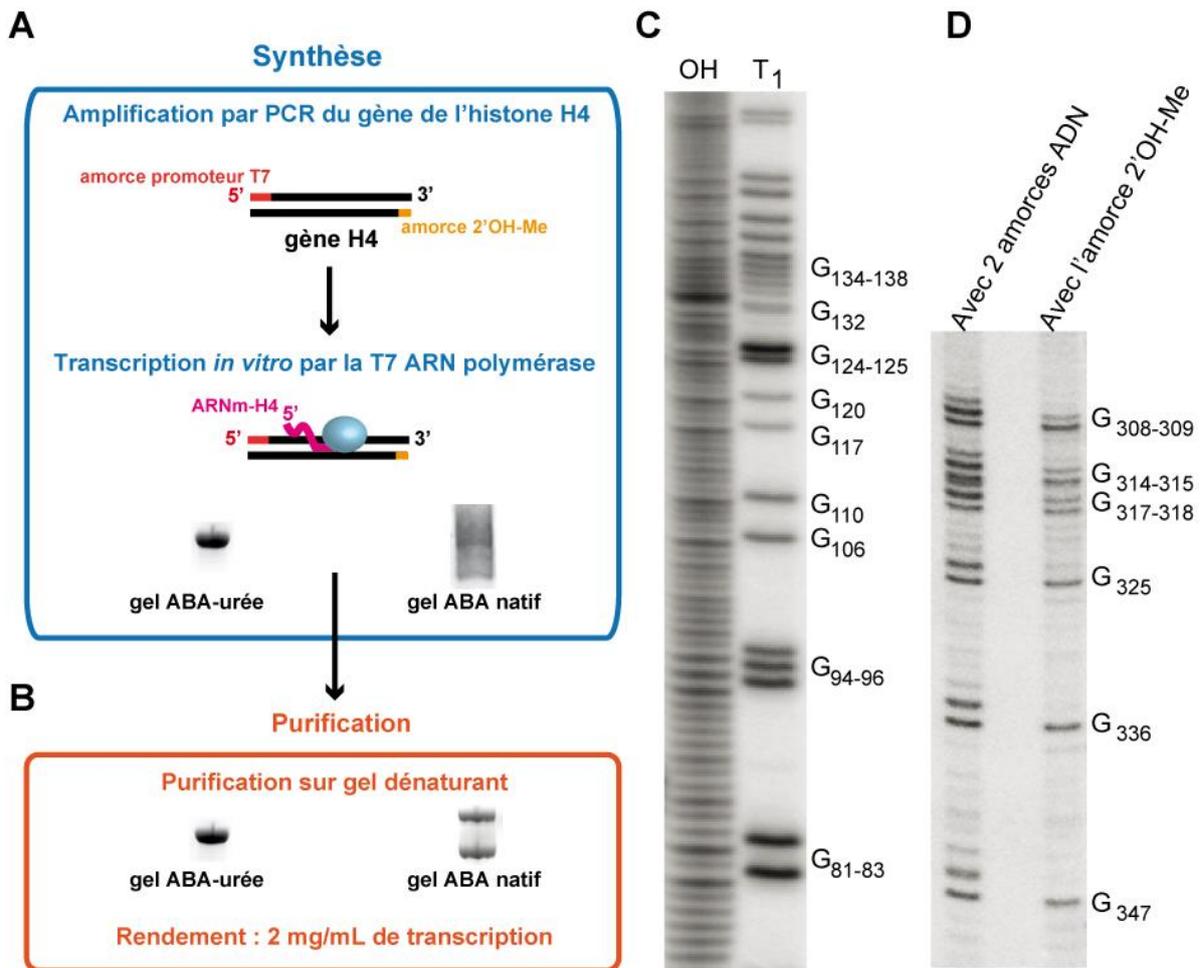


Figure 4. Protocole de synthèse et de purification de l'ARNm H4, hydrolyse T1 de contrôle et effet de l'utilisation d'une amorce 2'OH-Me. (A,B) Le schéma résume les étapes de purification et la pureté de l'ARNm. (C) Séquençage de l'ARNm par digestion à la RNase T1 (piste T1) en conditions dénaturantes. Les produits de clivage sont identifiés par comparaison à la piste d'hydrolyse alcaline partielle (OH). Dans cette expérience, l'ARNm a été marqué au $[^{32}\text{P}]$ en 5' sur la coiffe. (D) Contrôle de la fidélité de terminaison de l'ARNm par la T7 ARN polymérase. Les ARNm ont été transcrits soit à partir d'un fragment de PCR obtenu avec deux amorces ADN ordinaires, soit en utilisant une amorce 2'OH méthylée en 3' du gène. Dans ce cas, l'ARNm a été marqué par ligation en 3' de $[^{32}\text{P}]$ pCp. En absence de 2'OH-Me sur l'amorce 3', on note un doublement des bandes au niveau des sites de coupure ce qui révèle la présence d'un nucléotide additionnel sur environ 50% des molécules synthétisés. En présence de 2'OH-Me, l'ARNm obtenu est homogène et se termine bien au niveau d'un nucléotide unique.

1.3 Contrôle de l'ARN purifié par séquençage

Afin de vérifier la qualité de l'ARNm purifié, une hydrolyse alcaline partielle ainsi qu'une digestion partielle à la RNase T1 en conditions dénaturantes ont été réalisées sur cet ARNm préalablement marqué au [³²P] en 5' au niveau de la coiffe. L'hydrolyse alcaline permet d'obtenir une « échelle » complète de coupure permettant de visualiser tous les nucléotides de l'ARNm, tandis que la digestion à la RNase T1 permet d'identifier tous les fragments d'ARN se terminant par un G. Les fragments d'ARNm ainsi obtenus ont été fractionnés sur gel ABA-8M urée et révélés par autoradiographie, ce qui permet d'en déduire la séquence (**Figure 4C**). L'analyse des 2 échelles nous a permis de séquencer partiellement l'ARNm purifié et d'en vérifier l'identité, mais aussi de nous assurer de l'homogénéité de l'extrémité 5' grâce au marquage à cette extrémité.

1.4 Synthèse de l'ARN H4 présentant une extrémité 3' homogène

Un problème récurrent de la transcription *in vitro* par la T7 ARN polymérase est l'incorporation supplémentaire et aléatoire de 1 à 3 nucléotides à l'extrémité 3' des transcrits. Cette micro hétérogénéité en 3' peut réduire la capacité des ARN à cristalliser ; en conséquence, nous avons recherché des conditions d'obtention des transcrits se terminant par une extrémité 3' homogène. Plutôt que d'ajouter en 3' du gène H4 un ribozyme autocatalytique permettant de générer un clivage précis, nous avons préféré synthétiser par PCR des matrices se terminant (à la fin du gène H4) par des nucléotides porteurs de groupements méthyles sur les derniers riboses (Kao et al., 1999). Ces nucléotides présentent la particularité d'entraîner un arrêt net de la T7 ARN polymérase sur le dernier nucléotide. L'homogénéité de l'extrémité 3' de l'ARNm H4 synthétisé par cette méthode a été analysée à partir d'un ARNm H4 marqué au [³²P] en 3' partiellement digéré par la RNase T1. Le profil de digestion obtenu à partir de cet ARNm a été comparé à celui obtenu à partir d'un ARNm synthétisé de manière classique (**Figure 4D**). Il en résulte que l'ARNm synthétisé à partir d'une matrice ADN 2'OH méthylée sur les derniers riboses présente un profil de coupure conforme à la séquence alors que l'ARNm synthétisé de manière classique contient un nombre de fragments anormalement élevé résultant d'un dédoublement, voir triplement de certaines coupures. Ce phénomène reflète la micro hétérogénéité obtenue lorsque des fragments de PCR ordinaires sont utilisés pour la transcription. Cette expérience prouve également qu'il est relativement aisé d'améliorer la qualité de l'extrémité 3' des transcrits.

1.5 Repliement de l'ARNm H4

1.5.1 Introduction

Habituellement, après purification en conditions dénaturantes, les ARN subissent une étape de renaturation afin qu'ils adoptent une conformation la plus homogène possible et que les expériences qui suivent reflètent cette homogénéité et non pas la somme de différentes molécules présentant des conformations et des activités différentes. En général, l'ARN dissout dans l'eau est chauffé brièvement à 95°C (2 min) puis refroidi brutalement en plongeant le tube dans la glace (5 min). Ce protocole rapide et l'absence d'ajout d'ions divalents permettent de limiter la dégradation de l'ARNm (il est supposé que le minimum d'ions divalents indispensables sont co-purifiés avec l'ARNm). Cependant, pour de grands ARN comme l'ARNm H4, ce protocole ne permet pas d'obtenir un ARN replié en une conformation unique (Jaeger, 2005). Deux formes majeures de repliement de l'ARNm H4 sont visibles sur les gels analytiques réalisés en conditions natives sur la **figure 4B** ; parfois une troisième bande et des variations dans la distribution des bandes sont également observées. Ces résultats préliminaires montrent que l'obtention de molécules d'ARNm repliées en une conformation unique allait devenir un point critique de mon travail de thèse.

Dans un premier temps, pour éviter l'étape de dénaturation de l'ARNm inhérente à la purification sur gel ABA-urée, j'ai examiné sur gel ABA natif la conformation de l'ARNm H4 à l'issue de la transcription par la T7 ARN polymérase. En effet, il était envisageable de purifier l'ARNm H4 en conditions natives. Malheureusement, l'ARNm présent dans le milieu de transcription contient déjà de multiples conformations difficilement séparables (**Figure 4A**). De ce fait, l'étape de purification à partir de gel natif a été écartée, car moins résolutive qu'un gel dénaturant et contenant davantage de conformations. Cela m'a conduit à développer des protocoles de repliement spécifiques permettant d'isoler les 2 formes majeures de l'ARNm H4.

Le repliement d'ARN est potentiellement l'un des problèmes les plus difficiles que l'on puisse rencontrer. En effet, replier en une seule conformation toutes les molécules d'ARN présentes dans un échantillon est un problème multiparamétrique n'obéissant pas à des règles standard et généralisables à tous les ARN. Contrairement à la cristallisation, il n'y a pas de méthodologie et peu de littérature sur le sujet, ce qui renforce la difficulté du problème. J'ai donc dû déterminer les conditions optimales de repliement de l'ARNm H4 en développant de nouveaux protocoles.

1.5.2 Recherche d'un protocole de repliement de l'ARNm

1.5.2.1 Approche utilisée et élaboration d'un premier protocole de renaturation

L'objectif initial était de trouver un protocole permettant de replier toutes les molécules d'ARNm H4 en une seule et unique conformation pour pouvoir la caractériser et la cristalliser. Dans un premier essai, j'ai tenté d'améliorer le protocole standard utilisé au laboratoire qui consistait à chauffer l'ARN durant 2 min à 95°C puis à le refroidir 5 min dans la glace. Mon premier essai a consisté à ajouter différents cations, à des concentrations variables avant l'étape de dénaturation à 95°C. Ce choix était motivé par le fait que la structure d'un certain nombre d'ARN est stabilisée par des cations, or le protocole initial n'en utilise pas et ne contient que les traces d'ions co-purifiés avec l'ARN.

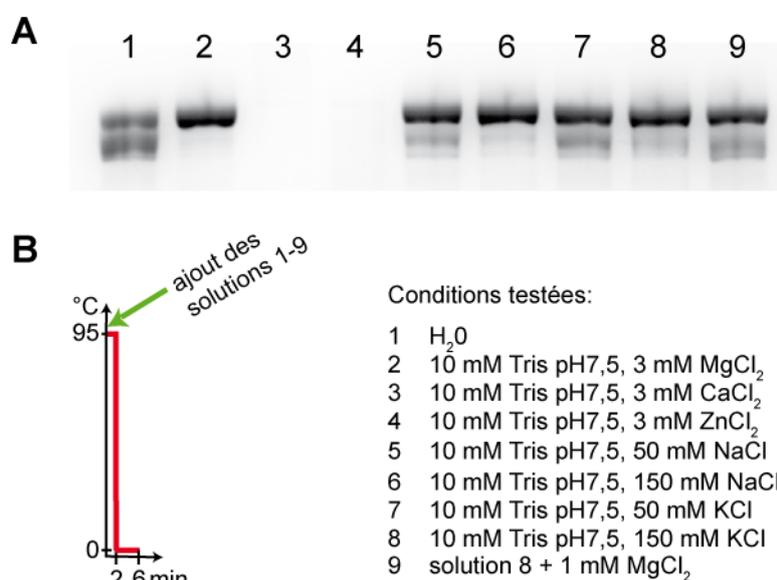


Figure 5. (A) Analyse sur gel ABA-natif des résultats du premier protocole de criblage. (B) Les différentes conditions utilisées ainsi que la cinétique de retour à 4°C sont indiquées ci-joint. Cinq µg d'ARNm ont été utilisés dans chacune des pistes (12,4 µM final); les gels ont été colorés au bleu de toluidine ; ces conditions ont été reprises pour la réalisation des expériences qui suivent sauf mention particulière.

Différents cations essentiels à la cellule comme les ions potassium, sodium, magnésium, calcium et zinc ont été testés (**Figure 5**). Pour les cations monovalents, j'ai utilisé des concentrations physiologiques, c'est-à-dire entre 50 et 150 mM. Pour les cations divalents, afin de limiter la dégradation de l'ARN, j'ai utilisé des concentrations variant entre 1 et 3 mM. J'ai particulièrement insisté sur les ions K⁺ et Mg²⁺, car ces cations sont souvent impliqués dans la stabilisation des structures d'ARN. Ces expériences ont été réalisées à une concentration élevée en ARNm H4 coiffé (12,4 µM, soit 1,5 mg/mL) afin que les résultats obtenus puissent avoir

une application en cristallographie. Les ARN étaient ensuite analysés par électrophorèse sur gel ABA natif (Tris-glycine, sans ions Mg^{2+}) et révélés au bleu de toluidine. Les électrophorèses ont été réalisées à 1 W, à 20°C, pendant 2h30 et dans du tampon Tris-glycine.

Les résultats observés sur le gel présenté dans la **figure 5** montrent que ce protocole de repliement de l'ARNm H4 entraîne l'apparition de plusieurs bandes, dont une qui migre plus lentement trouvée en haut du gel. Une concentration de 3 mM d'ions Mg^{2+} permet l'obtention de cette forme haute à l'état quasiment pur (piste 2). Les autres ions divalents (Ca^{2+} , Zn^{2+}) entraînent la disparition de l'ARNm, tandis que les ions monovalents (Na^+ , K^+) induisent l'apparition d'autres bandes présentant des mobilités électrophorétiques augmentées. Seule la solution contenant 150 mM de NaCl (piste 6) entraîne un repliement similaire à celui observé dans la piste 2 présentant 3 mM de $MgCl_2$. Cette conformation de l'ARNm qui migre le plus lentement a été nommée indifféremment conformère haut ou forme haute (H) dans la suite du manuscrit.

1.5.2.2 Caractérisation du protocole retenu et optimisation

Lors des tests fonctionnels et des essais de cristallisation, la concentration en ARNm peut être amenée à être considérablement augmentée ou diminuée. En conséquence, j'ai vérifié si le protocole de repliement permettant d'isoler la conformation haute était reproductible quel que soit la concentration en ARNm H4. J'ai retenu le protocole utilisant 3 mM de $MgCl_2$ (**Figure 5**, piste 2), car c'est le plus adéquat pour réaliser les essais de cristallisation et les tests fonctionnels. De plus, on peut espérer remplacer les ions Mg^{2+} par des ions Mn^{2+} ce qui peut permettre de résoudre le problème des phases en cas d'obtention de cristaux d'ARN. D'autre part, l'ajout d'ions Zn^{2+} et Ca^{2+} aux systèmes de traduction *in vitro* comme les extraits de réticulocytes de lapin ou les extraits de germes de blé risquait d'inhiber la traduction, ce qui excluait les protocoles 3 et 4. Restaient les protocoles de repliement utilisant soit les ions Na^+ , K^+ ou $K^+ + Mg^{2+}$ (protocoles 5 à 9) ; bien que permettant d'obtenir le conformère H, ces protocoles n'ont pas été retenus car contrairement au protocole 2 ils nécessitent l'ajout de concentrations plus élevées d'ions ce qui peut être un désavantage dans la suite de l'étude.

Pour optimiser le protocole 2 et mieux comprendre ses limites, j'ai réalisé un gradient de concentration en ARNm H4 en gardant la concentration d'ions Mg^{2+} constante. La concentration initiale en ARNm H4 (12,4 μM) a été divisée et multipliée par un facteur 2 (6,2 et 24,8 μM respectivement). Le gel natif présenté sur la **figure 6** montre l'apparition d'un mélange de 2 conformations lorsque la concentration d'ARN est doublée (24,8 μM). Aucun changement n'est observé pour une concentration en ARN de 6,2 μM . Ce résultat suggère que

ce protocole de repliement dépend dans une certaine mesure de la concentration en ARN c'est-à-dire du rapport ARN/Mg²⁺.

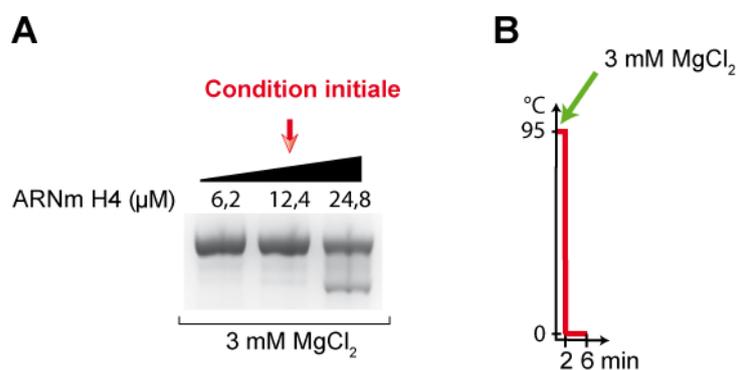


Figure 6. Analyse sur gel ABA-natif de l'effet de la concentration de l'ARNm H4 (A) sur le repliement en suivant le protocole de repliement représenté (B). La concentration en ions divalents Mg²⁺ est maintenue constante tandis que la concentration en ARNm varie.

Suite à cette première expérience, l'étude du rapport ARN/Mg²⁺ a été approfondie en réalisant un gradient d'ions Mg²⁺ au voisinage de la concentration de référence (3 mM MgCl₂). Sur le gel ABA natif présenté dans la **figure 7**, un mélange de 2 conformations est observable aux concentrations de 1 et 2 mM MgCl₂. Aux concentrations de 3 et 4 mM MgCl₂, le conformère H est obtenu pur, tandis qu'à 5 mM, l'ARNm H4 n'est plus visible suggérant

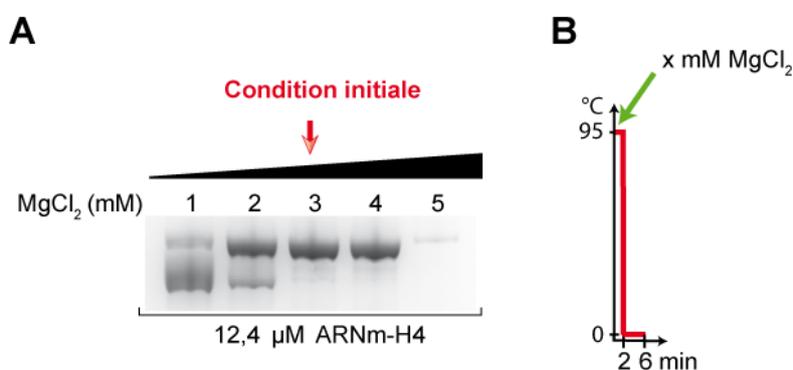


Figure 7. Analyse sur gel ABA-natif de l'effet de la concentration en MgCl₂ (A) sur le repliement de l'ARNm H4 en suivant sur le protocole de repliement représenté (B). La concentration en ARNm H4 est maintenue constante tandis que la concentration en ions divalents Mg²⁺ varie.

qu'il est agrégé, précipité ou dégradé. Ce résultat confirme l'importance du rapport ARN/Mg²⁺ pour l'obtention du conformère H. Aux basses concentrations d'ions Mg²⁺, plusieurs repliements sont obtenus (conformément au résultat du protocole de repliement dans l'eau) ; un excès d'ions Mg²⁺ est fortement préjudiciable au repliement et/ou à la solubilité de l'ARNm.

Les ions Mg^{2+} favorisant la dégradation des ARN à haute température, j'ai optimisé le temps de dénaturation de l'ARNm à $95^{\circ}C$ afin de limiter la dégradation de l'ARNm H4. En pratique, dans un bain sec et dans un volume de $50 \mu L$, le temps nécessaire pour qu'un échantillon placé dans un microtube à parois normales passe de 24° à $95^{\circ}C$ est de 50 sec environ. En conséquence, le temps d'incubation à $95^{\circ}C$ a été réduit à 1 min. Les résultats obtenus sont présentés sur le gel de la **figure 8**. Dans ces nouvelles conditions, le protocole permet toujours d'isoler la conformation H.

Cette optimisation du temps de dénaturation sera employée par la suite lors des expériences de routine pour la préparation des ARNm H4.

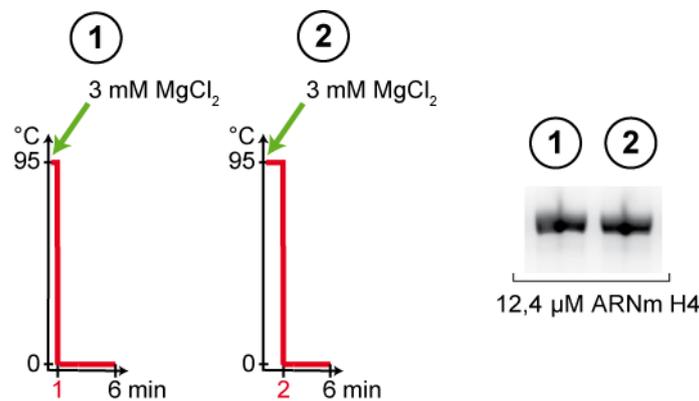


Figure 8. Analyse sur gel natif de l'effet du temps d'incubation à $95^{\circ}C$. Les 2 protocoles testés diffèrent par le temps de dénaturation à $95^{\circ}C$. Aucune différence significative n'est observée sur le repliement de l'ARNm H4.

1.5.2.3 Conclusion

A l'aide de ce protocole de dénaturation/renaturation en présence d'ions Mg^{2+} , il est possible d'isoler la conformation H de l'ARNm H4. Le critère important est un rapport molaire ARNm H4/ Mg^{2+} situé entre 1/241 et 1/376 ; au-delà, à partir d'une certaine concentration en ARN, l'ARNm H4 précipite ou s'agrége au cours du processus de repliement.

1.5.3 Recherche d'un second protocole de repliement de l'ARNm

1.5.3.1 Elaboration d'un deuxième protocole de renaturation

Après avoir trouvé un protocole permettant d'isoler la forme haute de l'ARNm H4, une nouvelle étude a été entreprise afin d'isoler la forme basse de l'ARNm. Cette forme migre plus rapidement sur gel natif ce qui suggère que son repliement est plus compact et plus achevé que la forme haute. Un deuxième protocole de repliement a été élaboré en testant à nouveau différentes conditions ioniques et en explorant des cinétiques de retour à $0^{\circ}C$ utilisant 2 paliers

intermédiaires pour favoriser l'établissement des interactions tertiaires lors du processus de repliement (**Figure 9**). Le palier à 60°C devait favoriser la formation de structures secondaires très stables, comme par exemple les tiges-boucles avec un motif GNRA, UNCG ou CUUG, car la plupart de ces tiges-boucles ont un point de fusion supérieur à cette température. Le palier à 30°C avait pour but de favoriser la formation de structures tertiaires stables suite à un ajout de cations couplé à un brusque changement de température. La température de 30°C a été choisie car proche de la température physiologique et suffisamment éloignée de 60°C pour permettre un refroidissement rapide de l'échantillon.

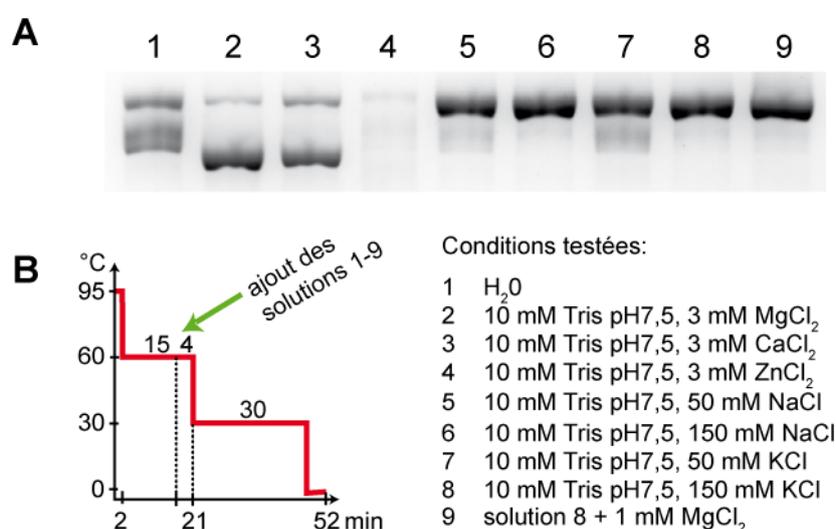


Figure 9. Analyse sur gel natif des résultats du deuxième protocole de criblage. Les différentes conditions utilisées ainsi que la cinétique de retour à 4°C sont indiquées.

Pour ce deuxième protocole, les solutions de cations utilisées lors du premier criblage ont été réemployées. Le gel ABA natif de la **figure 9**, montre que la cinétique de refroidissement induit des repliements sensiblement différents en présence d'ions divalents. En présence de 3 mM d'ions Mg²⁺ ou Ca²⁺ (pistes 2 et 3), une forme qui migre rapidement a été obtenue; elle a été nommée conformère bas ou forme basse (B). Comme dans le premier protocole, l'ARNm disparaît en présence d'ions Zn²⁺ (piste 4) alors que les ions monovalents (pistes 5-9) permettent l'isolement de la forme haute avec un état de pureté légèrement amélioré.

1.5.3.2 Caractérisation du protocole par paliers et optimisation

Pour les raisons exposées précédemment et pour obtenir la conformation B pure, une optimisation du protocole a été entreprise. En partant du protocole utilisant 3 mM MgCl₂, 3 concentrations en ARNm H4 ont été soumises au protocole de renaturation et le produit a été

analysé sur gel ABA natif (**Figure 10**). À concentration élevée d'ARN (24,8 μM) le protocole conduit à la formation du conformère H. À la plus basse concentration d'ARN (4,1 μM) on retrouve la conformation B, mais l'intensité de la bande est diminuée par rapport à la concentration de 12,4 μM ce qui suggère qu'une partie a été dégradée ou s'est agrégée dans ces conditions.

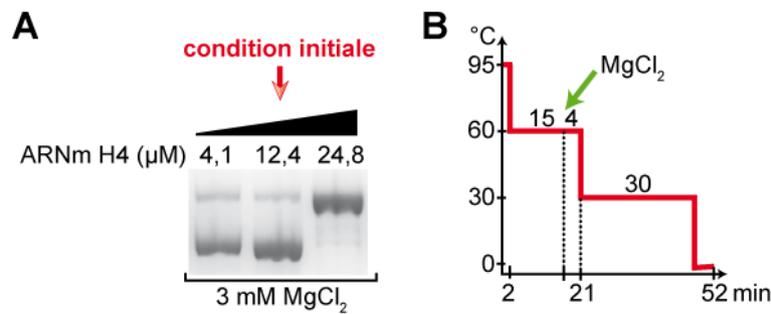


Figure 10. Analyse sur gel ABA natif (**A**) de l'effet de la concentration en ARNm H4 sur le protocole de repliement présenté en (**B**). Des quantités équivalentes d'ARNm (3 μg) ont été déposées sur gel.

Dans une seconde étape, l'impact du rapport $\text{ARN}/\text{Mg}^{2+}$ a été testé sur le nouveau protocole de renaturation. Un gradient de MgCl_2 centré sur la condition initiale de 3 mM MgCl_2 a été utilisé et les produits de renaturation ont été analysés sur gel. La **figure 11** permet de constater qu'à 0,5 mM de MgCl_2 , on obtient plusieurs conformations semblables à celles retrouvées lors d'un repliement dans l'eau ou lors du protocole précédent en une étape (**Figure 7**). À 1 mM de MgCl_2 , c'est la conformation haute (H) qui est retrouvée, tandis qu'à 2, 3 ou 5 mM MgCl_2 on retrouve majoritairement la conformation basse (B). A 10 mM MgCl_2 une partie de l'ARN replié dans la conformation B est manquante, ce qui suggère une dégradation et/ou une agrégation. Bien que ces tests n'aient pas permis d'obtenir le conformère B pur, ils ont néanmoins permis de confirmer que le repliement de l'ARN est également gouverné par le rapport $\text{ARNm}/\text{Mg}^{2+}$ employé.

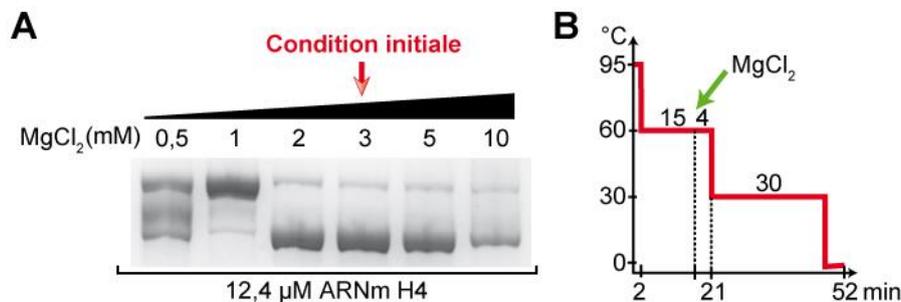


Figure 11. Analyse sur gel ABA natif (**A**) de l'effet de la concentration en ions Mg^{2+} sur le protocole de repliement de l'ARNm H4 présenté en (**B**).

Un dernier test a consisté à optimiser le point d'introduction des ions Mg^{2+} dans le protocole par étape. Les ions Mg^{2+} ont été introduits à 95°C et à 30°C en plus du point initial à 60°C (**Figure 12**). La cinétique de renaturation est restée constante, seul le point d'ajout de la solution de $MgCl_2$ a été modifié. Le gel de la **figure 12**, montre que l'on retrouve la conformation H lors de l'ajout à 95°C ; par contre lorsque les ions sont ajoutés après 15 min d'incubation à 30°C, on obtient un mélange constitué principalement de conformère B et d'un conformère intermédiaire, donnant un aspect diffus et élargi à la bande. Il y a également davantage de conformère H. Cette expérience confirme l'hypothèse de départ et montre qu'il est critique d'ajouter les ions Mg^{2+} à 60°C pour obtenir majoritairement la conformation B de l'ARNm H4.

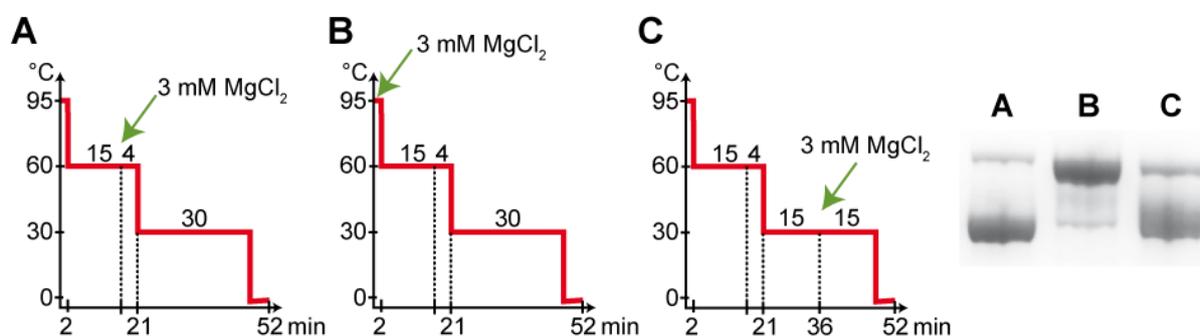


Figure 12. Analyse sur gel natif d'ARNm H4 replié à l'aide de 3 protocoles différents (A, B et C).

1.5.3.3 Conclusion

A partir de ce second protocole de renaturation, il est possible d'isoler la conformation H ou la conformation B de l'ARNm H4. Pour d'isoler la conformation H, il faut un rapport molaire ARNm H4/ Mg^{2+} d'environ 1/80. Pour obtenir majoritairement la conformation B, il faut un rapport molaire ARNm H4/ Mg^{2+} situé entre 1/161 et 1/376. Au-delà, l'ARNm H4 commence à se dégrader de manière significative ou à s'agréger. Il est également essentiel de rajouter le $MgCl_2$ durant l'incubation à 60°C. Il serait intéressant de faire varier cette température d'incubation afin peut-être d'obtenir 100% d'ARNm adoptant la conformation B.

1.5.4 Test d'un troisième protocole de renaturation

Un troisième protocole de renaturation de l'ARNm a été testé en se basant sur des protocoles publiés (Henriet et al., 2007). L'idée était d'améliorer la pureté du conformère B et éventuellement d'obtenir le conformère intermédiaire (conformère I ci-après). Pour ce troisième criblage, en plus des cations précédemment testés lors des 2 premiers criblages, une polyamine (spermidine) a été introduite car il a été montré que dans certaines conditions, les

structures d'ARN peuvent être stabilisées par les polyamines. Dans ce protocole de renaturation, après l'étape de dénaturation à 95°C, l'ARN est refroidi brutalement à 0°C puis les tampons contenant les cations et polyamines sont ajoutés avant une incubation à 30°C durant 30 min. Les résultats de cette expérience sont visibles sur le gel de la **figure 13**. Aucune des conditions testées ne permet d'obtenir majoritairement la forme intermédiaire I. Dans la plupart des cas, c'est un mélange des 2 conformations (intermédiaire et B) qui est obtenu. La présence de spermidine a conduit à la précipitation ou agrégation de l'ARNm H4 (pistes 3 et 10).

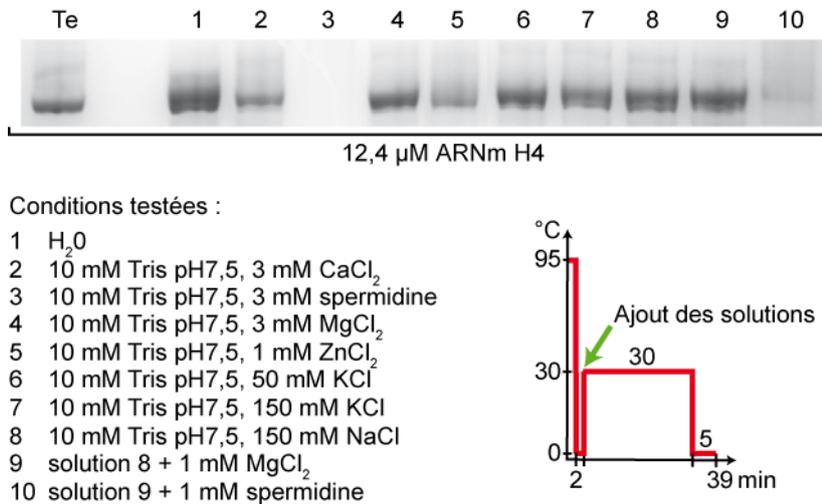


Figure 13. Analyse sur gel natif des résultats du troisième protocole de renaturation. Les différentes conditions utilisées ainsi que la cinétique de retour à 4°C sont indiquées. Afin d'identifier les conformations obtenus à l'aide de protocole, un échantillon contenant la conformation B a également été déposé sur gel (piste Te). Trois µg d'ARNm (8,2 µM final) ont été déposés dans chacune des pistes.

1.5.5 Test d'un quatrième protocole de criblage

Un quatrième protocole qui avait été décrit dans la littérature a été utilisé (Doudna et al., 1993). Dans ce protocole, l'ARNm n'est pas dénaturé à 95°C mais à 60°C. Il subit ensuite un refroidissement lent pendant une heure jusqu'à la température de 20°C. Le gel ABA natif présenté dans la **figure 14**, montre que ce protocole tend à favoriser la forme haute de l'ARNm, sans obtention de forme intermédiaire et peu de forme basse. Là aussi, la spermidine induit la précipitation ou l'agrégation de l'ARNm H4 (pistes 3 et 10).

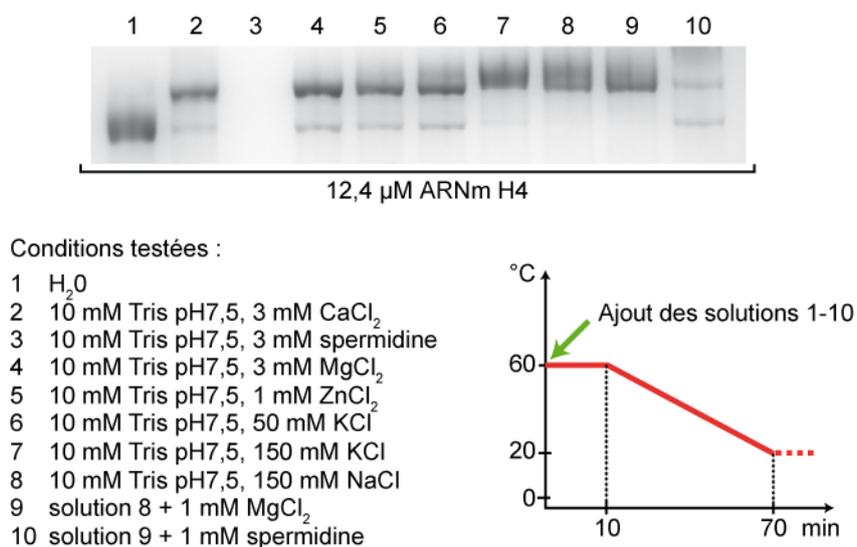


Figure 14. Analyse sur gel natif des résultats du quatrième protocole de renaturation. Les différentes conditions utilisées ainsi que la cinétique de retour à 20°C sont représentées. Trois μg d'ARNm H4 (8,2 μM final) ont été déposés dans chacune des pistes.

1.5.6 Pureté des conformères H et B et dégradation au cours des processus de renaturation

En utilisant le premier protocole de renaturation, le conformère H est obtenu pur et exempt de contamination en conformère B. Par contre, le conformère B obtenu par la seconde méthode de renaturation par paliers, reste contaminé par une certaine quantité de conformère H. Pour quantifier cette contamination, « quelques dizaines de milliers de cpm » d'ARNm H4 marqué au [³²P] en 5' ont été mélangés à de l'ARNm H4 non marqué. Puis le mélange a subi le second protocole de renaturation et les ARN ont été séparés sur gel ABA natif avant d'être quantifiés après exposition sous Imaging plate. Les résultats des quantifications sont reportés dans le **tableau 2**. L'analyse des valeurs montre que 86% de l'ARNm H4 adopte la conformation B lorsqu'il est replié à l'aide de ce protocole.

Echantillon	[³² P] ARNm H4 (cpm)	Répartition des 2 espèces moléculaires
Conformère H	3965	14%
Conformère B	24120	86%

Tableau 2. Quantification de la contamination du conformère B par le conformère H à l'issue du protocole de renaturation par paliers. Le mélange d'ARNm renaturé a été séparé sur gel ABA natif puis le gel a été séché et exposé sous Imaging plate.

Afin de déterminer si les échantillons repliés en présence de MgCl₂ n'avaient pas subi de clivage excessif (préjudiciable pour les essais de cristallisation et autres tests fonctionnels), j'ai déterminé le pourcentage d'ARN dégradé à l'issue des deux protocoles de repliement

retenus. Comme précédemment, une certaine quantité d'ARNm H4 marqué au [³²P] en 5' et mélangé avec de l'ARNm H4 non marqué a été renaturé en suivant les deux protocoles. Puis les échantillons ont été fractionnés sur gel ABA natif en présence d'un témoin ARNm H4 n'ayant pas subi le traitement. Après séchage du gel, la radioactivité des bandes des ARN entiers a été mesurée, de même que la radioactivité correspondant aux fragments de dégradation qui migrent plus rapidement sur gel. Les chiffres présentés dans le **tableau 3** montrent qu'environ 15% de l'ARNm se dégrade durant les traitements ou lors de la migration sur le gel (alternativement, l'ARNm pourrait déjà être dégradé avant son introduction dans le tube). Le conformère H issu du premier protocole de renaturation en une étape est légèrement plus dégradé que l'ARN issu du protocole de repliement par paliers (2,5% davantage dégradé). Cette valeur n'est certainement pas significative puisque durant le premier protocole, l'ARNm H4 est exposé moins longtemps à la température que dans le protocole par paliers. Globalement, les chiffres montrent que les deux protocoles n'induisent pas de dégradation significative de l'ARNm par rapport à un ARNm témoin n'ayant subi aucun traitement.

Echantillon	Bandes analysées	[³² P] ARNm H4 (cpm)	Dégradation (%)
ARNm H4 non renaturé	ARN entier	34484	15
	Fragments de dégradation	6240	
ARNm H4 conformation haute H	ARN entier	32552	17,5
	Fragments de dégradation	7067	
ARNm H4 conformation basse B	ARN entier	28086	15
	Fragments de dégradation	4090	

Tableau 3. Résultat des quantifications réalisées sur gel afin de déterminer le pourcentage d'ARN dégradé après une étape de repliement en présence de 3 mM MgCl₂.

1.6 Effet de la coiffe sur le repliement de l'ARNm H4

Les précédents tests de repliement ayant été réalisés sur des molécules d'ARNm H4 coiffées, nous avons entrepris de tester si la coiffe participait au processus de repliement des deux conformations H ou B. Des tests de repliement ont donc été réalisés sur de l'ARNm H4 coiffé ou non coiffé à l'extrémité 5'. Les résultats de cette expérience sont visibles sur le gel de la **figure 15**. Les résultats obtenus montrent que l'ARNm H4 coiffé en 5' se comporte de façon identique à l'ARNm H4 non coiffé. La coiffe ne participe donc pas au processus de repliement des deux conformations H et B.

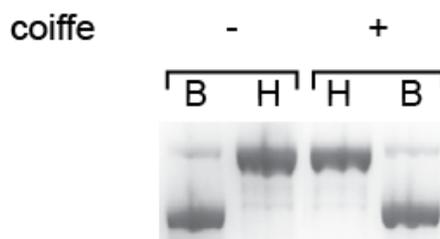


Figure 15. Analyse sur gel ABA natif de l'effet de la coiffe sur les protocoles de repliement de l'ARNm H4 permettant d'isoler les conformations H et B.

1.7 Conclusion

Le protocole utilisé pour synthétiser et purifier l'ARNm H4 permet d'obtenir de manière reproductible un ARNm H4 pur et avec un rendement de 2 mg d'ARN par millilitre de transcription. Cependant, cet ARNm possède la propriété intrinsèque de se replier sous plusieurs formes qui diffèrent par leur mobilité électrophorétique en conditions natives.

La recherche de conditions de renaturation de l'ARNm H4 a conduit à la mise au point de 2 protocoles qui permettent d'isoler 2 des 3 conformations majeures observées. L'optimisation des protocoles a conduit aux 2 schémas expérimentaux rappelés dans la **figure 16**.

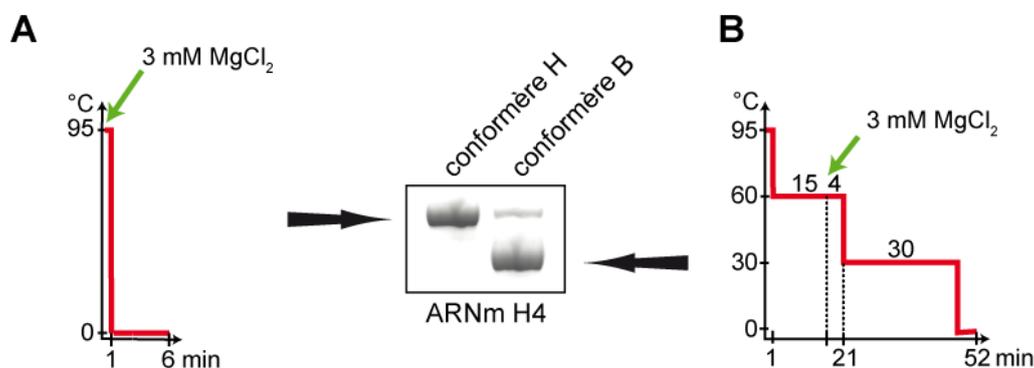


Figure 16. Rappel des schémas de dénaturation/renaturation utilisés pour l'ARNm H4. Le protocole **A** permet d'obtenir la forme haute de l'ARNm H4 (conformère H). Le protocole **B** utilisant des paliers de renaturation permet d'obtenir la forme basse de l'ARNm H4 (conformère B).

Le premier protocole en une étape permet d'isoler la conformation haute (H) de l'ARNm. Le second protocole de renaturation par paliers permet d'affiner le repliement de l'ARNm qui adopte alors une conformation migrant plus rapidement (conformation basse, B). Bien que ce second protocole permette également d'obtenir la forme haute (en modifiant la concentration en ions Mg^{2+}), sa mise en œuvre est plus délicate et le premier protocole reste préféré pour

l'obtention de la conformation haute. Il n'a malheureusement pas été possible de trouver un protocole permettant d'isoler la forme intermédiaire de l'ARNm H4, migrant entre la conformation haute et basse.

Chapitre 2 : Caractérisation structurale préliminaire des conformations H et B de l'ARNm H4

2.1 Stabilité des conformations H et B

En prélude à des tests fonctionnels et des études structurales, nous avons contrôlé la stabilité des deux conformations de l'ARNm H4 obtenues avec les deux protocoles de repliement. Nous avons testé leur stabilité en fonction du temps et de la température.

2.1.1 Test de stabilité en fonction du temps

La vitesse de croissance de cristaux de macromolécule biologiques peut être très lente et s'étaler sur plusieurs jours, voire plusieurs mois. Dans ces conditions, une bonne stabilité des macromolécules est un pré-requis indispensable pour aborder une étude structurale *de novo*. La stabilité des deux formes d'ARNm H4 a été observée pendant 5 jours à 20°C pour vérifier ce point. Puis les échantillons ont été analysés sur gel ABA natif à 20°C et comparés à des ARN fraîchement repliés. L'examen du gel sur la **figure 17** montre qu'il n'y a pas de différence significative entre un ARN fraîchement replié et un ARN replié 5 jours avant, et ceci pour les deux formes H et B. Ce résultat suggère que la barrière énergétique séparant les deux conformations est suffisamment élevée pour que les conformations H et B soient structurellement stables à 20°C pendant au moins 5 jours.

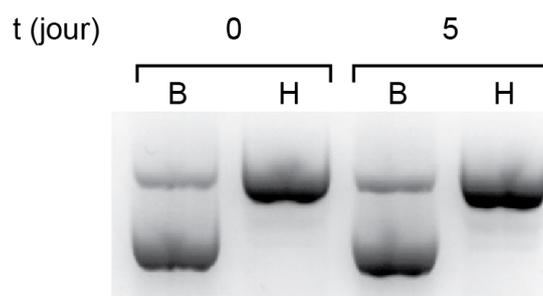


Figure 17. Mesure de la stabilité dans le temps des deux formes d'ARNm H4. Les deux formes d'ARNm préalablement incubées durant 5 jours à 20°C sont comparées aux mêmes ARNm repliés extemporanément. Les ARN (5 µg) ont été fractionnés sur gel ABA natif puis colorés au bleu de toluidine.

2.1.2 Test de stabilité en fonction de la température

Les études fonctionnelles et structurales réalisées ci-après ont été réalisées dans une gamme de température s'étendant entre 4°C et 37°C. Par conséquent, la stabilité des conformations H et B a été étudiée à ces températures. Après repliement par les deux protocoles, les échantillons d'ARNm H4 ont été incubés pendant 60 min à 4°C ou à 37°C, puis les échantillons ont été analysés par électrophorèse sur gel ABA natif. Pour être plus pertinente, l'électrophorèse a été réalisée à la même température, c'est-à-dire 4°C et 37°C. Les gels présentés dans la **figure 18** montrent qu'il n'y a pas de différence significative de stabilité des conformères à 4°C ou 37°C, les résultats étant très semblables à ceux observés à 20°C.

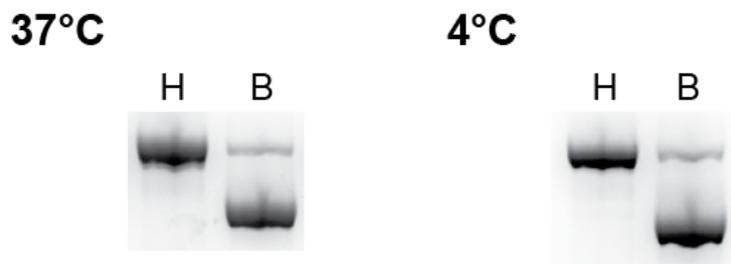


Figure 18. Analyse sur gel ABA natif des conformations H et B après incubation à 4°C ou 37° et électrophorèse à la même température.

2.2 Détermination du rayon hydrodynamique des conformations H et B par diffusion dynamique de la lumière (DLS)

Une étude par diffusion dynamique de la lumière a été réalisée sur les deux conformations afin de déterminer le rayon hydrodynamique et la polydispersité des formes H et B. Ces paramètres permettent d'identifier la conformation la plus compacte, la plus homogène et la plus prédisposée à cristalliser. Après avoir réalisé l'étape de renaturation, une étape additionnelle de purification par ultracentrifugation a été effectuée afin d'éliminer les agrégats présents en solution (100 000 × g, 40 min à 20°C). Dans une étude préliminaire, j'avais pu montrer la présence d'un petit nombre d'agrégats (représentant uniquement 1% de la quantité totale d'ARNm) qui gêne l'interprétation des données obtenus par DLS, par SAXS et par équilibre de sédimentation.

Les résultats de DLS présentés dans le **tableau 4** et la **figure 19** montrent que quelle que soit la conformation étudiée, une seule population est observée. La conformation H présente un rayon hydrodynamique moyen de 6,5 nm, tandis que la conformation B a un rayon hydrodynamique moyen de 7,1 nm, soit une différence de 10% entre les deux conformations. L'écart entre les deux conformations étant faible, les expériences ont été répétées 5 fois de

manière indépendante et pour chaque expérience, 10 mesures ont été réalisées et moyennées. Ces résultats suggèrent que l'ARNm de la forme H est plus compact que l'ARNm de la forme B alors que celui-ci migre plus rapidement sur gel ABA natif. D'autre part, les deux conformations H et B ont une polydispersité moyenne supérieure à 20%, elles sont donc considérées comme étant polydisperses.

Echantillon	Nombre de coups (Kcps)	Coefficient de diffusion ($\mu\text{m}^2/\text{s}$)	R_H (nm)	Polydispersité (%)
ARN H4 conformation B	256,4	49,3	7,2	35,64%
ARN H4 conformation B	291,8	45,4	7,81	46,15%
ARN H4 conformation B	261,3	50,5	7,02	29,50%
ARN H4 conformation B	259,5	51,3	6,91	30,50%
ARN H4 conformation B	256	49,6	7,16	22,36%
ARN H4 conformation H	576,7	55,3	6,415	26,27%
ARN H4 conformation H	588,7	54,4	6,525	30,66%
ARN H4 conformation H	576,8	53,9	6,58	31,78%
ARN H4 conformation H	591,6	55,4	6,405	24,70%
ARN H4 conformation H	562,8	56,4	6,285	22,80%

Tableau 4. Résultats des expériences réalisées par DLS sur les conformations H et B. Chaque ligne correspond à une expérience réalisée de manière indépendante et correspond à une moyenne de 10 mesures.

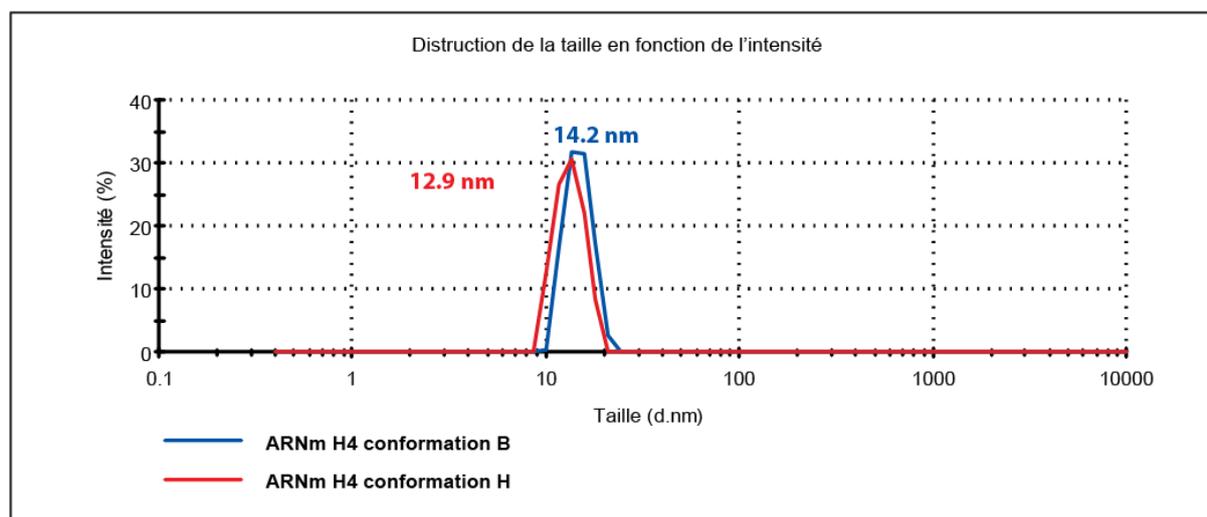


Figure 19. Graphique représentant la distribution de la taille de chaque conformation en fonction de leur intensité. La polydispersité de chaque conformation est calculée à partir de cette représentation.

2.3 Détermination du rayon de giration des conformations H et B par diffusion des rayons X aux petits angles (SAXS)

Afin de confirmer les résultats obtenus par DLS, une étude par diffusion des rayons X aux petits angles a été réalisée sur les conformations H et B. Après repliement et élimination des agrégats par centrifugation, les études par SAXS ont été réalisées et traitées à l'IGBMC par Vincent Cura à l'aide d'un appareil BioSAXS-1000 (Rigaku). Les données ont été analysées à l'aide du logiciel GNOM afin d'obtenir la longueur maximale (D_{\max}), le rayon de giration (R_g) et la fonction de répartition des distances $p(r)$ de chaque conformation. La fonction de répartition des distances est une représentation dans l'espace réel des données de diffusion et permet de représenter graphiquement la forme de la particule étudiée. Concrètement, elle indique le nombre de paires d'atomes séparées par une distance r . D'après les résultats obtenus par SAXS (**Tableau 5 et Figure 20**), la conformation H présente un rayon de giration moyen de 6,3 nm et une longueur maximale moyenne de 19,15 nm, tandis que la conformation B a un rayon hydrodynamique moyen de 7,8 nm et une longueur maximale moyenne de 27,6 nm. Ces valeurs ont été moyennées à partir d'expériences réalisées à différentes concentrations. Entre les 2 conformations, il y a donc une différence de 16% au niveau du rayon hydrodynamique R_g et une différence de 30% au niveau de la distance maximale D_{\max} . D'après ces résultats, la conformation B a une forme plus allongée que la conformation H, ce qui est en accord avec les résultats obtenus précédemment par DLS.

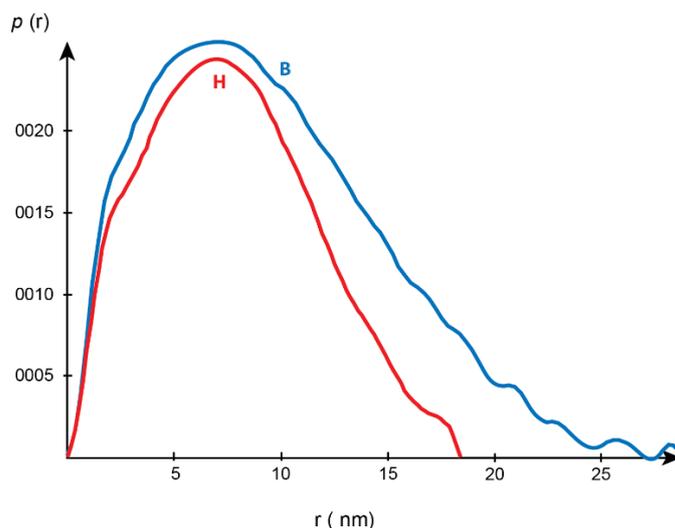


Figure 20. Graphiques des fonctions de répartition des distances obtenus à partir des conformations H et B.

Echantillon	Température (°C)	Concentration (mg/mL)	Dmax (nm)	Rg (nm)
ARN H4 Conformation H	4	1,5	18,3	6,1 ±0.02
ARN H4 Conformation H	4	1	18	6,02 ±0.02
ARN H4 Conformation H	4	0,75	20	6,45 ±0.03
ARN H4 Conformation B	4	1,5	28,7	7,7 ±0,14
ARN H4 Conformation B	4	1	25	7,6 ±0,14
ARN H4 Conformation B	4	0,75	29,3	8,1 ±0,12

Tableau 5. Résultats des expériences réalisées par SAXS sur les conformations H et B.

2.4 Détermination de la masse des conformations H et B par équilibre de sédimentation

Une étude par équilibre de sédimentation a été réalisée sur les deux formes de l'ARN_m H4 pour déterminer leur masse en solution. L'objectif était de caractériser l'état oligomérique des conformations H et B. Les expériences d'équilibre de sédimentation ont été réalisées à 3 vitesses différentes : 7 000, 10 000 et 13 000 rpm. D'après les différents gradients de concentration obtenus (**Figure 21**), les conformations H et B ont atteint l'équilibre de sédimentation à ces 3 vitesses. Dans le cas de la conformation H, la masse obtenue à partir de la distribution des 3 gradients dans la cellule est de 123 007 Da, tandis que pour la conformation B, la masse obtenue est de 118 216 Da. La masse théorique d'un ARN_m H4 monomérique étant de 121 408 Da, les masses mesurées présentent donc un écart de 1,3% (conformation H) et 2,7% (conformation B) avec la masse théorique. D'après ces résultats, les conformations H et B sont donc des monomères.

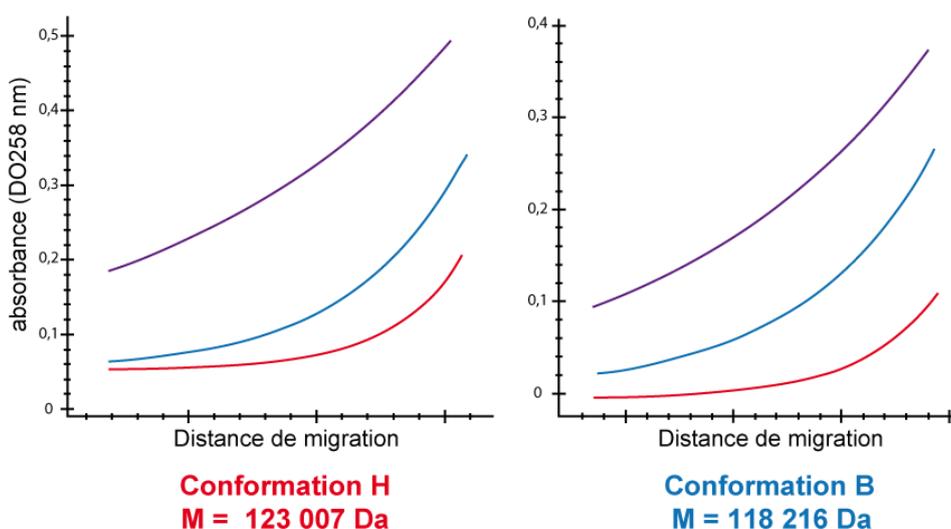


Figure 21. Graphiques représentant pour chaque expérience la DO à 258 nm mesurée le long de la cellule. Les données obtenues à 7 000, 10 000 et 13 000 rpm sont indiquées respectivement en rouge, bleu et violet.

2.5 Conclusion

Les tests de stabilité en fonction du temps et de la température ont permis de démontrer que les conformations H et B sont stables durant plusieurs jours à 20°C et pendant au moins 1 h à 4°C et 37°C. Ces résultats permettent de réaliser et d'interpréter un large éventail d'études structurales et fonctionnelles sur ces conformations.

Les études structurales préliminaires réalisées par DLS et SAXS ont permis de montrer que bien que la conformation H migre plus lentement que la conformation B sur gel ABA natif, la conformation H semble être plus compacte. La vitesse de migration d'une macromolécule sur gel ABA natif dépendant de sa forme et de ses charges accessibles, cela signifie que la conformation B présente davantage de charges accessibles que la conformation H.

Les études par équilibre de sédimentation ont permis de déterminer la masse des conformations H et B. Les résultats obtenus confirment que les deux conformations H et B sont bien des monomères et que les différences de migration sur gel résultent bien de différences de repliement de leur structure et non de la formation de dimères.

Chapitre 3 : Caractérisation fonctionnelle des conformations H et B de l'ARNm H4

Les études présentées dans le chapitre 2 ont montré que l'ARNm H4 replié dans les formes H et B est suffisamment stable pour permettre la réalisation de tests fonctionnels. Par conséquent, nous avons effectué quatre expériences afin de déceler d'éventuelles différences de comportement ou de réactivité entre les deux formes d'ARNm.

3.1 Tests de retard sur gel en présence d'eIF4E

La première expérience a consisté en un test de fixation d'eIF4E visualisé par retard sur gel de l'ARNm H4, aussi appelé EMSA pour Electro Mobility Shift Assay. Le facteur eIF4E est une protéine de 25 kDa qui existe sous deux formes dans la cellule : libre ou en complexe avec eIF4G (220 kDa) et eIF4A (50 kDa) pour former le complexe trimérique eIF4F. La fonction principale d'eIF4E dans le complexe eIF4F est de se fixer sur la coiffe située à l'extrémité 5' des ARNm et de déclencher le processus d'initiation de la traduction par le biais du recrutement de la particule ribosomique 43S (voir introduction). Dans le modèle d'initiation de la traduction de l'histone H4, eIF4E peut se fixer sur des structures secondaires de l'ARNm H4 situées dans la phase codante de l'ARNm H4 (Martin et al., 2011). Cette interaction est indépendante de la coiffe et fait intervenir l'extension N-terminale de la protéine eIF4E. Sur l'ARNm H4, le fragment permettant de conserver intacte la liaison à eIF4E est constitué d'environ 150 nt (bien que des fragments de tailles plus réduites fixent encore eIF4E mais moins efficacement). La nécessité d'un grand domaine de liaison suggère que ce domaine pourrait subir des variations de repliement dans les différents conformères de l'ARNm H4. Nous avons donc entrepris de comparer la fixation d'eIF4E sur les deux formes de l'ARNm H4.

Des tests d'interaction d'eIF4E « par retard sur gel » ont été réalisés sur les deux formes H et B de l'ARNm H4. L'ARNm H4 préalablement purifié a été marqué au [³²P] par coiffage catalysé par la « Vaccina Capping Enzyme ». L'ARNm H4 a ensuite subi les protocoles de repliement permettant d'obtenir les formes H et B et les deux échantillons d'ARNm ont été incubés dans un tampon de fixation (10 mM Tris-HCl pH 7,5, 50 mM KCl, 1 mM DTT, glycérol 10% (v/v)) en présence d'eIF4E (1 μM final) et d'un extrait total d'ARNt de levure (4 μg/μL

final). Après une étape d'incubation, les échantillons ont été déposés sur gel ABA natif. La **figure 22** montre que la migration des deux formes H et B est retardée en présence d'eIF4E (pistes 3 et 6). Pour les deux ARNm, un contrôle permettant de vérifier que le tampon ne modifie pas le repliement de l'ARNm a été réalisé. Ces contrôles (pistes 2 et 5), qui ne contenaient pas d'eIF4E, ont confirmé la stabilité des conformères d'ARN dans le tampon. L'analyse plus détaillée des bandes retardées dans les pistes 3 et 6 montre que les complexes ARNm-eIF4E ont conservé la différence de migration propre aux formes H et B, cependant le complexe ARNm-eIF4E de la piste 3 présente deux bandes minoritaires dont l'une est déjà présente dans l'ARN libre (pistes 1 et 2) ainsi qu'une bande plus basse dont la mobilité est très voisine de celle du complexe ARNm-eIF4E de la forme B (piste 6). Ce résultat suggère qu'eIF4E est capable de remodeler partiellement la structure de la forme H.

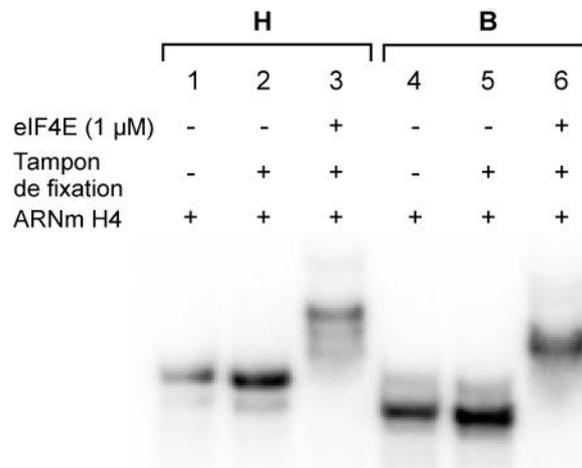


Figure 22. Analyse sur gel ABA natif du profil de migration des conformations H et B en présence ou en l'absence de eIF4E. L'ARNm H4 est marqué au [³²P] en 5'.

3.2 Tests de traduction *in vitro*

Des tests de traduction *in vitro* dans des lysats de réticulocytes de lapin (RRL) et des extraits de germe de blé (WGE) ont été effectués afin de comparer l'efficacité de traduction des deux formes d'ARN.

Les ARNm coiffés et ayant subi les deux protocoles de repliement ont été introduits dans les extraits de traduction et les réactions de traduction ont été effectuées en suivant le protocole décrit en annexe. Les produits de traduction ont ensuite été analysés par électrophorèse sur gel ABA SDS. L'efficacité de traduction de chaque forme d'ARNm a été déterminée en quantifiant la [³⁵S]-méthionine incorporée dans l'histone H4 synthétisée. Les résultats obtenus sur gel (**Figure 23**) montrent que les ARNm dans les conformations H et B sont traduits avec quasiment la même efficacité, que se soit dans les WGE ou dans les RRL.

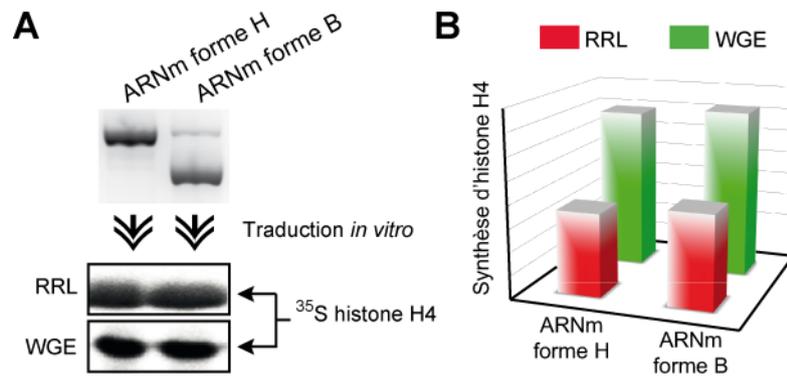


Figure 23. Analyse sur gel ABA dénaturant des produits de traduction obtenus à partir des conformations H ou B et de lysats de réticulocytes de lapin (RRL) ou d'extraits de germe de blé (WGE).

3.3 Expériences de « toe-print »

Les tests de traduction réalisés ci-dessus n'ont pas permis de mettre en évidence de différence majeure dans l'efficacité de traduction des deux formes H et B de l'ARNm H4. Lors des expériences de traduction, les molécules d'ARNm sont assemblées aux ribosomes puis dépliées dans leur totalité lors du passage des ribosomes. Il s'en suit un retour à une conformation inconnue qui tient compte de paramètres de repliement différents de ceux utilisés dans nos protocoles de repliement expérimentaux. Ainsi, les deux conformères H et B pourraient se retrouver dans le même état de repliement après le premier cycle de traduction (« pioneer round of translation »). Par conséquent, le test de traduction ne permet pas de visualiser des différences dans l'efficacité de traduction des deux formes, ces différences pouvant disparaître aux cycles suivants, occasionnant une « dilution » de l'effet initial.

Nous avons alors décidé d'utiliser un autre test qui ne prend en compte que l'étape de fixation initiale des ribosomes et qui reflète donc l'effet de la structure sur l'étape d'initiation. Des expériences de « ribosome toe-print » (extension d'amorce en présence de ribosomes assemblés sur le codon AUG initiateur) ont été réalisées avec Angelita Simonetti au laboratoire. Des ARNm H4 coiffés et repliés dans les conformations H et B ont été soumis au test.

Après repliement, les échantillons d'ARNm H4 ont été incubés en présence d'extraits de RRL et d'un inhibiteur de la traduction. Trois inhibiteurs différents ont été utilisés pour bloquer l'initiation à différents stades :

- l'édéine, un inhibiteur universel de l'initiation de la traduction, qui inhibe de manière spécifique la fixation d'aminoacyl-ARNt dans le site P de la sous unité 30S ou 40S du ribosome (Kozak and Shatkin, 1978b).
- La cycloheximide, un inhibiteur de l'étape d'élongation de la traduction chez les eucaryotes, elle inhibe le mouvement de translocation du ribosome (Merrick and Hershey, 1996).

- Le GMP-PNP ou guanosine 5'-[β,γ -imido]triphosphate est un analogue non hydrolysable du GTP. Il bloque les facteurs utilisant le GTP comme eIF2 ou le facteur eIF5B responsable de l'assemblage avec la grande sous-unité 60S (Gray and Hentze, 1994).

Après incubation des extraits de traduction avec les ARNm et les inhibiteurs, une étape d'ultracentrifugation a été réalisée afin de séparer les complexes ARNm-ribosomes des complexes ARNm-« RNA binding proteins » qui pourraient également donner des signaux d'arrêt de la reverse transcriptase. Des fractions P (« pellet ») et S (« surnageant ») ont été ainsi obtenues et les ARNm présents dans ces fractions ont subi une étape d'extension d'amorce catalysée par une transcriptase inverse (AMV reverse transcriptase). L'amorce étant marquée au [32 P] en 5' et étant complémentaire à la région 110-130 de l'ARNm H4, il est possible de visualiser les fragments d'ADNc synthétisés et les arrêts de synthèse dus aux ribosomes, protéines ou structures secondaires de l'ARN. L'utilisation de pistes contrôles sans ribosomes (ctl) et le fractionnement en surnageants (S) et culots d'ultracentrifugation (P), permet ensuite d'identifier les arrêts dus aux ribosomes.

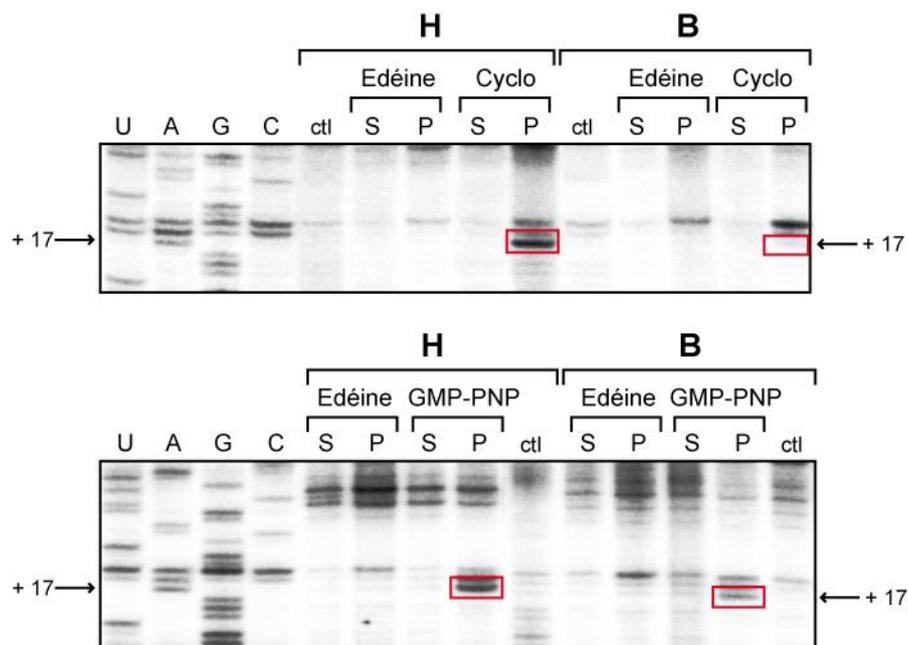


Figure 24. Expériences de toe-print sur les formes H et B de l'ARNm H4. Gels ABA dénaturants des produits de transcription inverse réalisées sur les fractions P (pellet) ou S (surnageants) ou ARNm seuls (pistes ctl). Une réaction de séquençage de l'ARNm H4 a été réalisée par transcription inverse selon la technique de Sanger. La position +17 du toe-print est indiquée. Cyclo : cycloheximide.

La **figure 24** montre la séparation, par électrophorèse sur gel ABA dénaturant, des produits de transcription inverse. En présence d'édéine, aucune pause due aux ribosomes n'est visible, que se soit pour la conformation H ou B. Cette absence d'assemblage des ribosomes est conforme à l'effet attendu de l'édéine et il confirme nos précédents résultats (Prongidi-Fix et al., 2013) tout en constituant un contrôle négatif supplémentaire. En présence de cycloheximide,

une forte bande est visible dans la piste P (contenant les ribosomes) à la position +17 dans la forme H de l'ARNm. Cette bande est d'intensité nettement moins forte dans la forme B. En présence de GMP-PNP, un arrêt similaire est trouvé à +17 mais la différence d'intensité entre les formes H et B est cependant moins importante qu'en présence de cycloheximide.

La cycloheximide bloque les ribosomes au stade 80S par inhibition de l'étape de translocation. Les complexes d'initiation sont alors incapables de s'engager dans l'étape d'élongation. Le GMP-PNP bloque l'assemblage au stade 48S, c'est-à-dire au stade de petite sous-unité fixée sur l'ARNm avec l'ARNt initiateur apparié à l'AUG initiateur dans le site P. L'expérience de toe-print a montré qu'en présence de ces deux inhibiteurs, une différence était visible dans l'assemblage des formes H et B de l'ARNm H4. La forme H semble plus propice à assembler les particules ribosomiques et cet effet est davantage marqué au stade 80S.

3.4 Test de pontage de la coiffe aux UV

Des expériences de pontage aux UV ont ensuite été réalisées sur les deux formes H et B de l'ARNm H4 grâce à un groupement photo-activable placé sur la coiffe de l'ARNm. Au laboratoire, des expériences de ce type avaient déjà été réalisées précédemment sur un mélange de conformères de l'ARNm H4 replié dans l'eau. Ces expériences avaient montré que la coiffe pontait les nucléotides C₄₁ et U₄₂, suggérant que ces résidus sont localisés à distance d'interaction (moins de 2 Å de la coiffe). L'interaction de la coiffe avec cette région de l'ARNm contraindrait eIF4E à interagir avec l'élément 4E-SE localisé dans la partie centrale de l'ARNm plutôt qu'à l'extrémité 5'. Nous avons donc voulu vérifier le pontage de la coiffe de l'ARNm H4 lorsque celui-ci est replié dans la forme H ou B.

Après purification de l'ARNm H4, le groupement photo-activable a été incorporé lors de l'étape de coiffage catalysé par la « Vaccina Capping Enzyme », en remplaçant le GTP (substrat initial de l'enzyme) par du 6-Thio-GTP. L'ARNm H4 a ensuite subi les protocoles de repliement permettant d'obtenir les formes H et B et les deux échantillons d'ARNm ont été exposés à un rayonnement UV (365 nm) afin de déclencher la réaction de pontage. Les ARNm ont ensuite subi une étape d'extension d'amorce catalysée par une transcriptase inverse (AMV reverse transcriptase). Différentes amorces marquées au [³²P] en 5' et complémentaires à différentes régions de l'ARNm H4 ont été utilisées afin d'analyser la quasi-totalité de l'ARNm H4. L'utilisation de pistes témoin contenant des ARNm H4 non exposés aux UV a permis ensuite d'identifier les arrêts dus à la formation d'une liaison covalente entre la coiffe et l'ARNm H4.

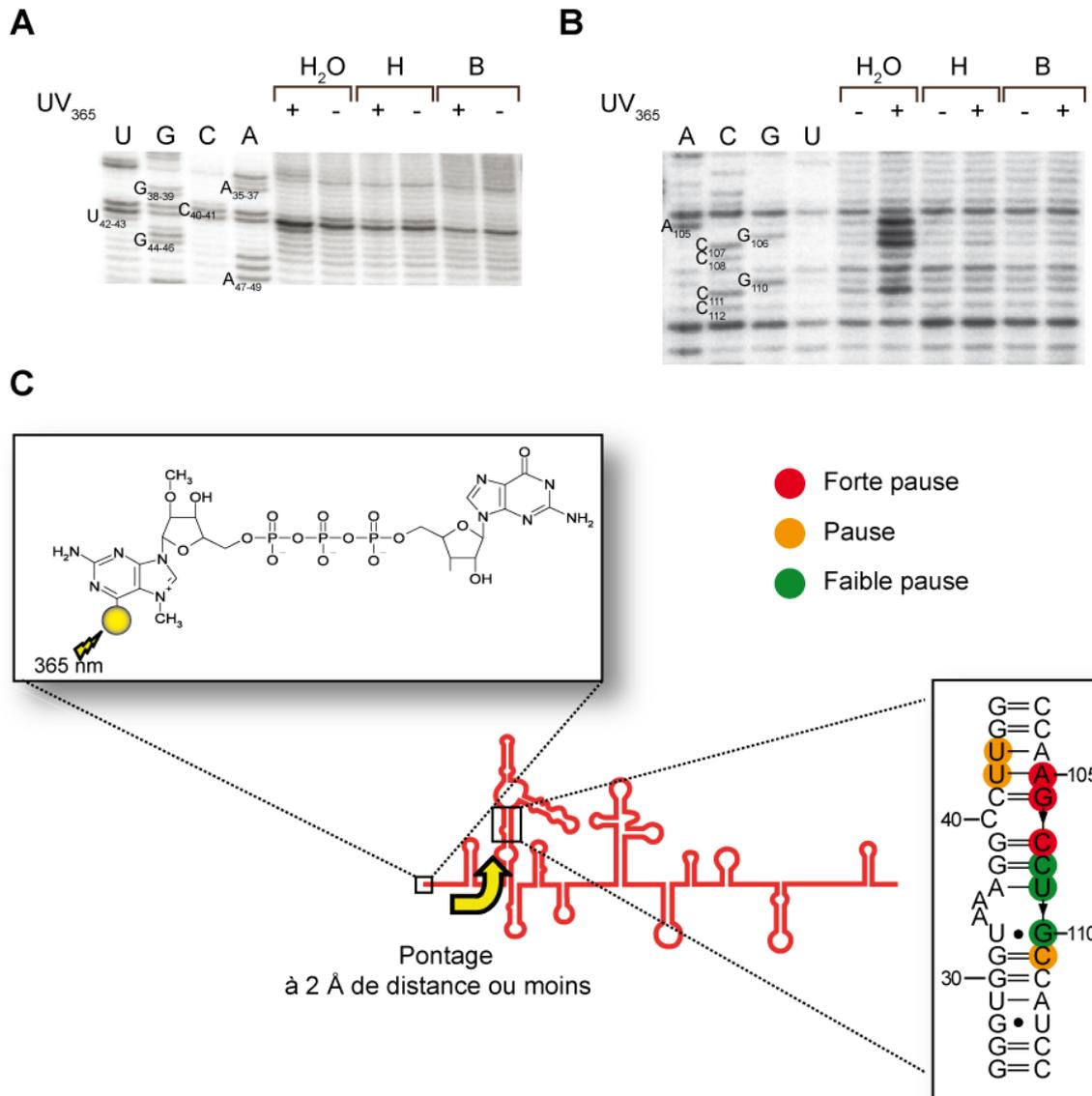


Figure 25. Expériences de pontage aux UV sur l'ARNm H4 préalablement replié dans l'eau (H₂O) ou suivant les protocoles permettant d'obtenir la forme haute (H) et basse (B). (**A,B**) Gels ABA dénaturants des produits de transcription inverse réalisées sur des ARNm exposés ou non aux UV (+) et (-). Une réaction de séquençage de l'ARNm H4 a été réalisée par transcription inverse selon la technique de Sanger. Les pauses présentées ont été détectées à l'aide d'une amorce complémentaire à la région 160-180. Les pauses détectées ont ensuite été reportées sur un modèle de structure secondaire (**C**).

Les résultats obtenus sur gel (**Figure 25**) montrent que l'ARNm H4 replié dans la forme H ou B n'est pas capable de ponter la coiffe placée à l'extrémité 5'. Absolument aucun signal de pontage n'est observé, même en examinant l'ensemble de la molécule. L'ARNm H4 témoin qui a été replié dans l'eau présente les pontages attendus aux positions U₄₂-U₄₃, et A₁₀₅-C₁₁₁, c'est-à-dire dans une des hélices de la structure repliée en jonction triple. Ces nucléotides se font face dans l'hélice ce qui confirme la position de la coiffe à cet endroit. Par contre l'utilisation de nouvelles lampes UV irradiant à 265 nm a entraîné un léger décalage du pontage C₄₁-U₄₂ aux positions U₄₂-U₄₃.

L'absence de pontage de la coiffe des formes H et B de l'ARNm H4 soulève certaines interrogations quant à leur repliement. L'ARNm H4 replié dans l'eau présente plusieurs conformations visibles sur gel natif, dont les conformations H et B ainsi qu'une bande intermédiaire qui pourrait être responsable du pontage de la coiffe. Alternativement, il est possible que l'ajout des ions Mg^{2+} dans les protocoles de repliement conduisant aux formes H et B soit responsable de l'inhibition de la réaction de pontage en agissant comme un contre ion d'un groupement chimique critique pour la réaction de pontage.

3.5 Conclusion

Nous avons réalisé 4 tests destinés à observer d'éventuelles différences de fonction des deux repliements de l'ARNm H4.

Le test de fixation d'eIF4E n'a pas révélé de différence majeure entre les deux formes d'ARNm. On note cependant qu'en présence d'eIF4E, la forme H semble partiellement évoluer vers un repliement de type B. Le facteur eIF4E présenterait donc une certaine capacité à remodeler la structure de l'ARNm. En fait, *in vivo*, bien qu'eIF4E existe à l'état libre, l'ARNm fixe le facteur eIF4F (eIF4E+eIF4G+eIF4A) lors de l'étape d'initiation. Ne disposant pas d'un tel facteur nous avons utilisé eIF4E isolé, mais nous ne pouvons exclure qu'eIF4F, qui est bien plus encombrant qu'eIF4E (300 kDa contre 25 kDa), ait une préférence pour l'une des deux formes de l'ARNm H4.

Les tests de traduction réalisés *in vitro* dans les extraits de germes de blé et de réticulocytes de lapin indiquent que les conformations H et B sont traduites de manière équivalente. Ces résultats suggèrent donc que ces deux conformations sont reconnues de manière équivalente par la machinerie de traduction cellulaire. Cependant, ce test d'activité ne reflète peut-être pas la vitesse du premier cycle de traduction de l'ARNm qui pourrait dépendre de la structure de l'ARNm. Il serait intéressant d'étudier ce premier cycle de traduction en réduisant le temps d'incubation. La vitesse de traduction étant de 5,6 acides aminés par sec (Bostrom et al., 1986; Ingolia et al., 2011), une vingtaine de sec est théoriquement suffisante pour traduire la totalité de l'ARNm. À cela se rajoute le temps d'assemblage des ribosomes sur l'ARNm qui est d'environ 1 min pour l'ARNm H4 (observations récentes réalisées au laboratoire). Additionnées, ces deux valeurs permettent d'envisager d'effectuer des cinétiques de traduction sur les deux formes d'ARNm afin de déterminer si elles sont réellement traduites avec la même efficacité. De même, il pourrait être intéressant de mesurer la vitesse d'assemblage des particules 80S sur les deux formes car l'étape d'initiation pourrait être affectée en priorité par les différences de conformation des ARNm.

Les expériences de « toe-print » ont montré que les ribosomes s'assemblent plus efficacement sur la forme H de l'ARNm. Il est donc probable qu'après l'étape d'assemblage des ribosomes, la structure de l'ARNm est totalement dépliée par le passage des ribosomes en cours de traduction. Les ribosomes pourraient alors « gommer » les différences structurales existant entre les conformères d'ARN ce qui expliquerait qu'après plusieurs cycles de traduction, la même synthèse globale d'histone H4 est observée.

Les expériences de pontage de la coiffe aux UV n'ont pas apporté de données interprétables. Les deux formes de l'ARNm H4 ne pontent pas la coiffe contrairement à l'ARNm H4 replié dans l'eau et nous ignorons si les structures des formes H et B sont en cause ou s'il s'agit d'un effet ionique lié à l'ajout d'ions Mg^{2+} .

À l'issue de ces tests se pose la question de la fonction, voire de l'existence de ces conformations d'ARNm dans la cellule. Il serait envisageable de sonder la structure de l'ARNm H4 *in vivo* afin de détecter la présence de l'une ou l'autre des conformations. Malheureusement, les ARNm d'histone ne sont présents que pendant la phase S du cycle cellulaire durant laquelle ils sont massivement traduits pour synthétiser les histones nécessaires à la compaction de l'ADN fraîchement dupliqué. Il y a donc peu de chance de les trouver à l'état libre, dans un état permettant leur sondage *in vivo*. Par contre, il est envisageable de bloquer leur traduction par un inhibiteur comme la cycloheximide et de sonder leur repliement dans ces conditions.

Chapitre 4 : Etude de la structure secondaire des conformations H et B

Dans les chapitres précédents, nous avons montré qu'il était possible d'obtenir avec une pureté satisfaisante deux des conformations principales de l'ARNm de l'histone H4 observées dans l'eau. Nous avons vérifié que les deux formes d'ARN conduisaient à la synthèse d'histone H4 et pu observer que des différences sensibles existaient lors du processus d'assemblage des particules ribosomiques, la forme H étant sensiblement plus efficace. Par ailleurs, la différence de migration entre les deux conformations de l'ARNm suggère que des différences notables pourraient exister dans leurs structures.

En conséquence, nous avons repris un projet déjà ancien visant à élucider la structure secondaire (2D) des deux formes de l'ARNm H4. La structure secondaire peut apporter des informations capitales pour la compréhension du mécanisme d'initiation par « tethering » observé pour l'ARNm de l'histone H4. Par ailleurs, dans le cadre de mon projet d'étude structurale de l'ARNm H4, l'information 2D est utile pour identifier les domaines flexibles ou impliqués dans le repliement de l'ARN. Elle peut permettre d'identifier des domaines structurellement indépendants et réduire la taille de l'ARNm à des tailles plus adaptées à l'étude cristallographique.

Au laboratoire, une étude cristallographique avait été réalisée précédemment sur le mélange de conformères de l'ARNm H4 replié dans l'eau. Une structure 2D expliquant bon nombre des réactivités avait été proposée, mais il subsistait un certain nombre d'ambiguïtés dans certaines régions de l'ARNm où coupures simple et double brin coexistaient (**Figure 26**). J'ai donc repris ce projet en mettant à profit mon expérience dans la purification des deux conformations principales de l'ARNm H4 pour résoudre les structures secondaires en solution de ces deux formes à l'aide de sondes enzymatiques (RNases) et chimiques (ions Pb^{2+} , hydrolyse spontanée). Grâce aux nouveaux procédés de marquage disponibles en 5' (Scriptcapping) et 3' (ligation de pCp sur extrémité homogène), il a été possible de révéler les coupures « en direct » sans étape de transcription inverse. Cette amélioration des méthodologies a beaucoup contribué à la réalisation rapide du sondage dont les résultats sont présentés ci-dessous.

phosphate adopte une certaine géométrie (dite en ligne) permettant le clivage avec libération d'une extrémité 3' se terminant par un phosphate 2'-3' cyclique.

Pour réaliser les réactions de sondage, l'ARNm H4 pur a été marqué radioactivement à l'extrémité 5' par l'enzyme virale VCE (Vaccinia virus Capping Enzyme) du kit « ScriptCap™ m7G Capping System » ou à l'extrémité 3' par ligation de [³²P]-pCP (cytidine-3',5'-bisphosphate) par la T4 RNA ligase. Dans les deux cas, l'ARNm marqué a été re-purifié sur gel ABA dénaturant puis élué du gel passivement. Les ARN marqués ont ensuite subi les protocoles de repliement permettant d'obtenir les deux formes d'ARN. Puis les ARNm radioactifs ont été incubés en présence d'une sonde ou du tampon « In Line ». Différentes concentrations de sondes ont été testées de façon à identifier des conditions permettant d'obtenir statistiquement une seule coupure par molécule d'ARN. Pour identifier les fragments d'ARN clivés par les différentes sondes, des échelles de coupure par la RNase T1 et par hydrolyse alcaline partielle ont été réalisées. Les produits de sondage, les échelles de coupures et les témoins ont ensuite été séparés par électrophorèse sur gel ABA dénaturant durant différents temps de migration de façon à pouvoir identifier jusqu'à 200 résidus sur les gels.

Les **figures 27, 28 et 29** présentent des gels caractéristiques des réactions de sondage. Les profils de coupure des deux formes de l'ARNm H4 présentent de grandes similitudes, mais des différences notables sont également observables dans les régions 75-85, 123-125, 190-195, 200-205, 210-220, 235-245 et 295-305. Cependant, moins de 20% des nucléotides de l'ARNm H4 présentent des différences de réactivité significatives.

Les profils de coupure ont été interprétés et classés en trois catégories correspondant à des réactivités fortes, faibles ou nulles. L'ensemble des coupures a ensuite été répertorié dans des histogrammes afin de faciliter la comparaison des résultats (**Figure 30**). Ces données ont ensuite servi de base pour la construction d'un modèle de repliement en deux dimensions.

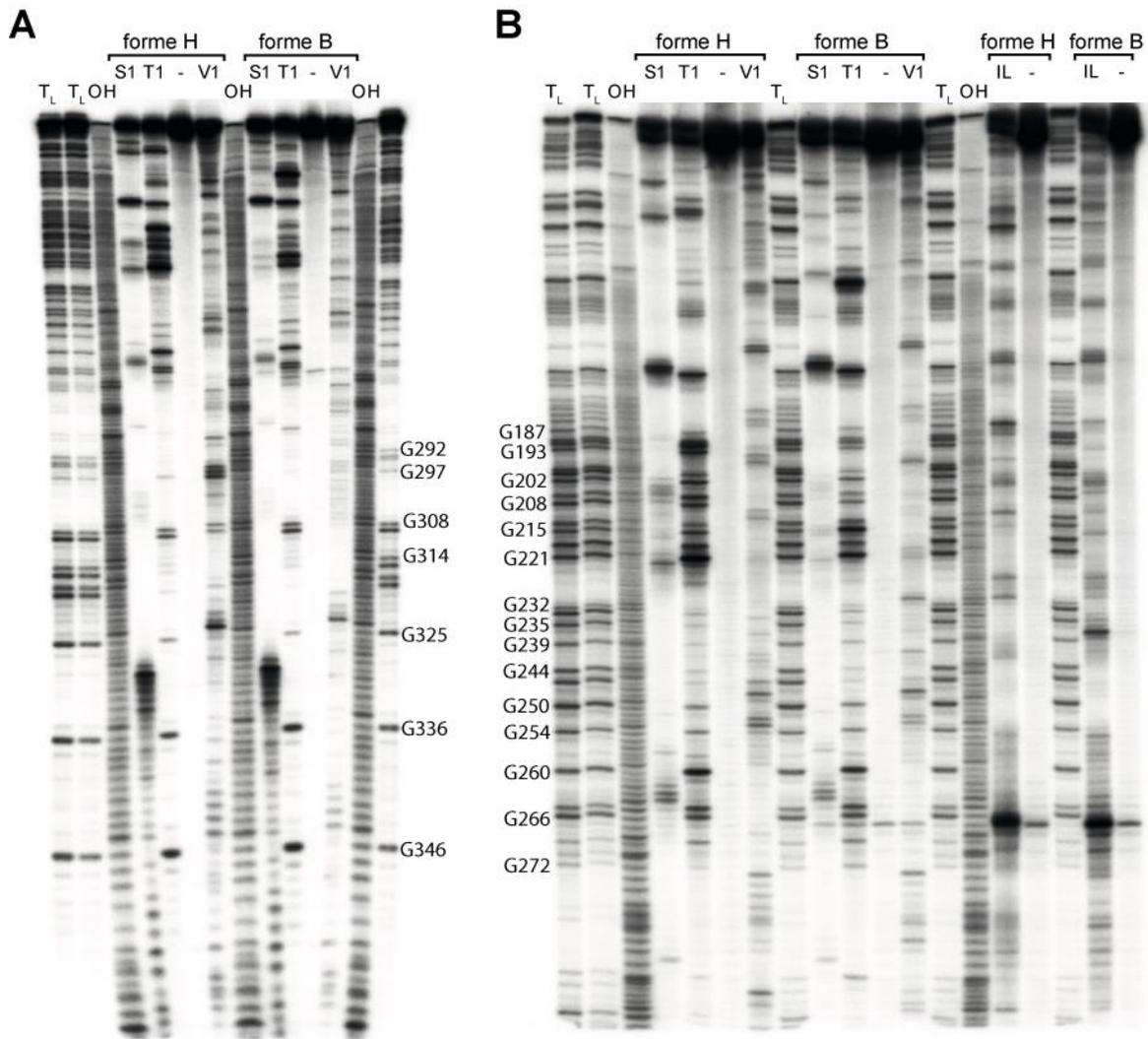


Figure 27. Fractionnement sur gel ABA 8% dénaturant des produits de sondage obtenus à partir d'un ARNm H4 marqué au $[^{32}\text{P}]$ à l'extrémité 3'. Les sondes enzymatiques utilisées sont les RNases S1, T1 et V1. L'hydrolyse spontanée des ARN a été observée dans le tampon « IL » (« In Line probing ») favorisant les coupures lorsque les résidus ribose et phosphate sont dans une géométrie « in line ». Les pistes (-) contiennent des ARNm témoins incubés sans sonde. Une échelle de coupure par la RNase T1 (pistes T_L) et par hydrolyse alcaline partielle (pistes OH) ont été réalisées. (A) Migration courte (bleu de bromophénol au bas du gel) ; (B) migration longue (xylène cyanol au bas du gel).

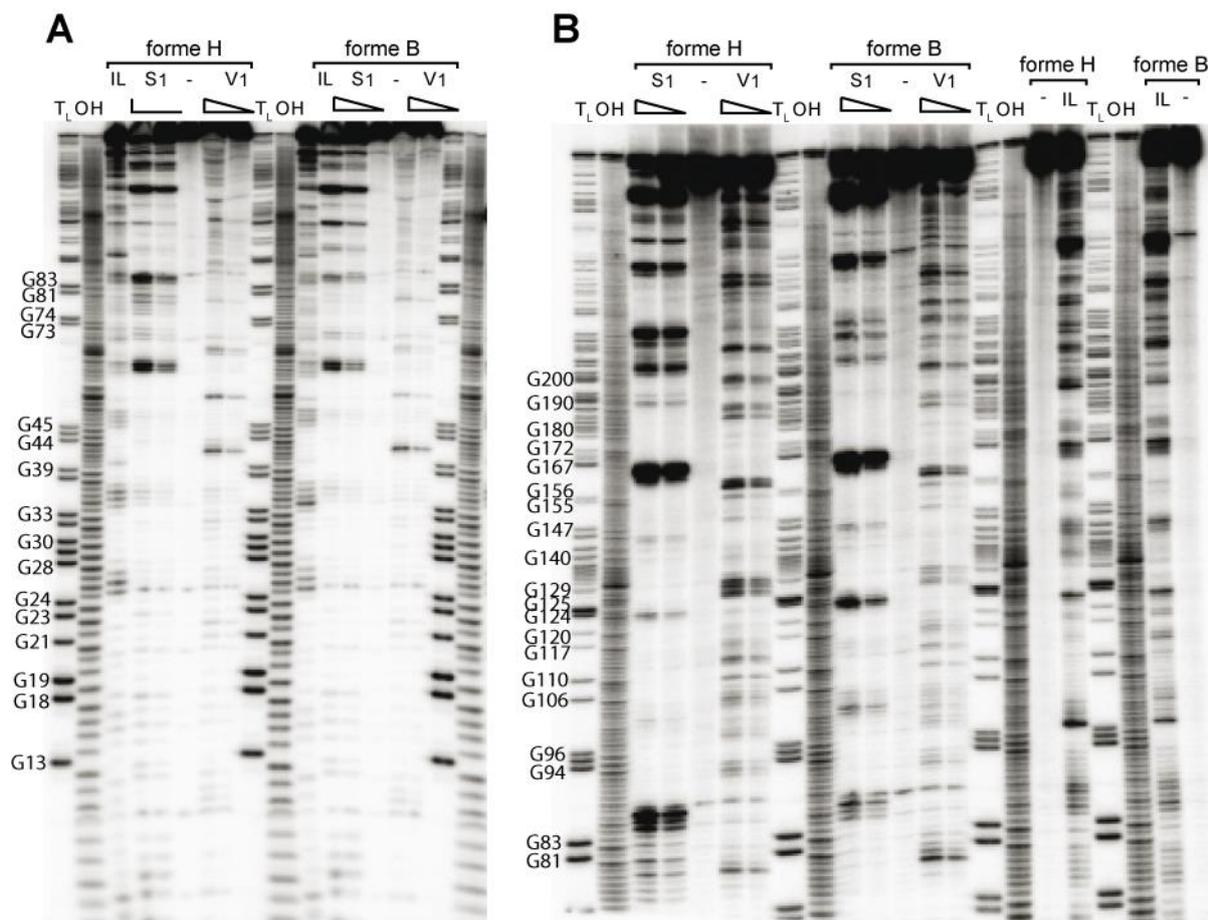


Figure 28. Fractionnement sur gel ABA dénaturant de produits de sondage obtenu à partir d'un ARNm H4 marqué au $[^{32}\text{P}]$ à l'extrémité 5'. Les sondes enzymatiques utilisées sont les RNases S1 et V1. L'hydrolyse spontanée des ARN a été observée dans le tampon « IL » (« In Line probing ») favorisant les coupures lorsque les résidus ribose et phosphate sont dans une géométrie « in line ». Les pistes (-) correspondent à l'ARNm incubés en l'absence de sonde. (A) Un gel ABA 12% a été utilisé (temps court : bleu de bromophénol au bas du gel). (B) Gel ABA 6% (temps long : xylène cyanol au bas du gel). Des échelles de coupure par la RNase T1 (pistes T_L) et d'hydrolyse alcaline partielle (OH) ont été réalisées.

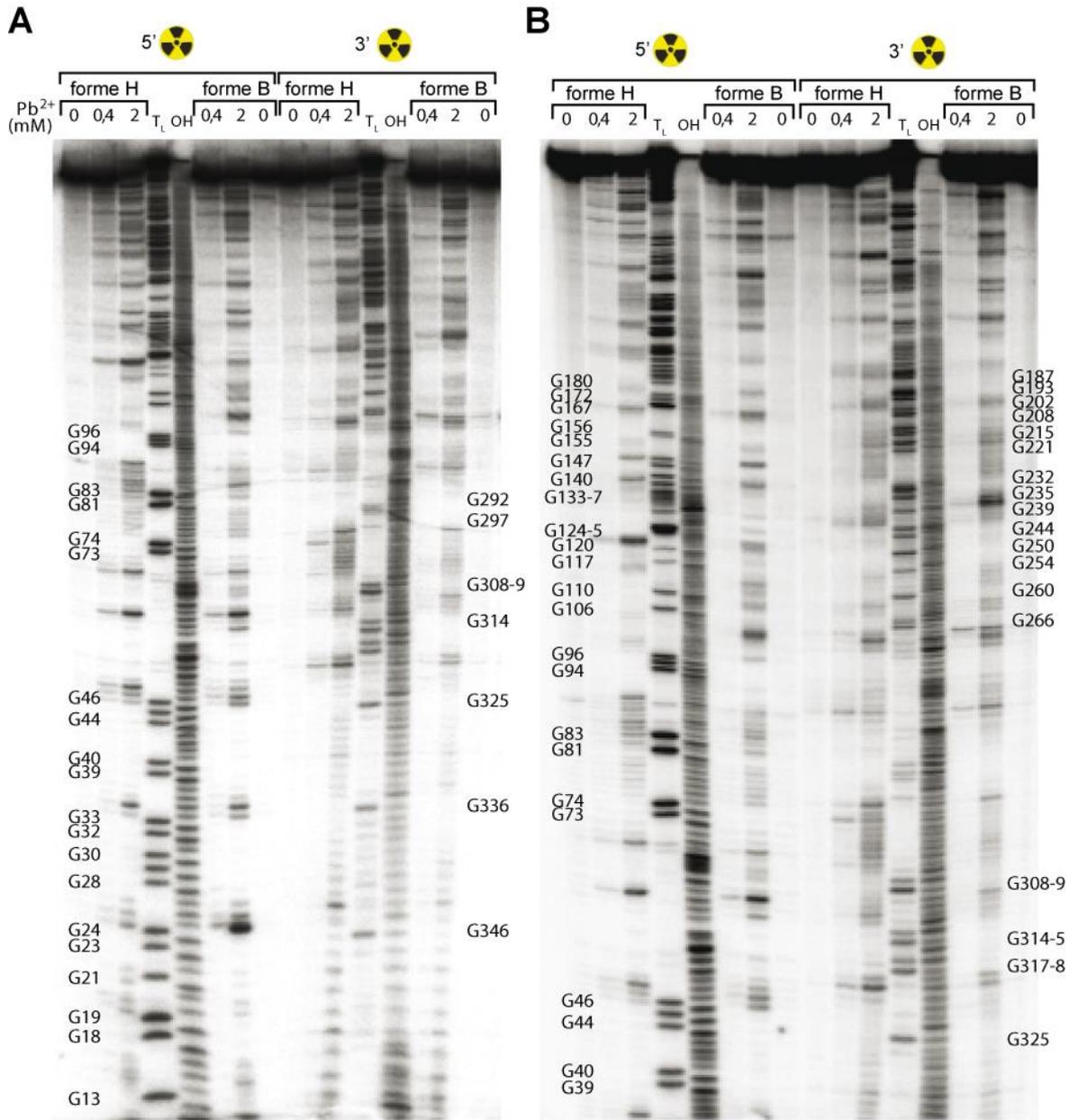


Figure 29. Fractionnement sur gel ABA dénaturant des produits de clivage de l'ARNm H4 par les ions Pb^{2+} . L'ARNm a été marqué au $[^{32}P]$ en 3' ou en 5'. Les profils de coupure sont comparés à un profil sans ions Pb^{2+} (pistes 0). Les produits de sondage ont été séparés et analysé sur un gel ABA 8% jusqu'à ce que bleu de bromophénol atteigne le bas du gel (**A**) et sur un gel ABA 6% jusqu'à ce que le xylène cyanol arrive au bas du gel (**B**). Des échelles de coupure par la RNase T1 (pistes T_L) et d'hydrolyse alcaline partielle (pistes OH) complètent la migration afin de permettre l'identification des fragments.



Figure 30. Histogrammes présentant les réactivités observées lors du sondage en solution des deux conformations de l'ARNm H4. Pour chaque conformation (H et B), les réactivités observées en présence de sondes chimiques (IL et Pb^{2+}) et enzymatiques (S1, T1 et V1) ont été reportées.

4.2 Réalisation d'un modèle de structure secondaire

Nous avons ensuite entrepris de construire un modèle de structure secondaire de l'ARNm H4 à partir des données de sondage. Dans un premier temps nous avons testé si des logiciels pouvaient nous aider dans l'établissement du modèle. Les logiciels RNAfold, CONTRAfold, KineFold, Pknots, Mfold, pknotsRG, Centroid Fold, et RNA structure ont été utilisés et se sont tous révélés inefficaces pour obtenir un modèle 2D en accord avec les données de sondage en solution.

Les logiciels INFERENCE of RNA ALIGNMENT utilisant les co-variations de séquences dans une population de séquences homologues n'ont pas permis non plus d'obtenir un modèle satisfaisant. Rappelons que les séquences nucléotidiques correspondant à la phase codante de l'histone H4 sont très conservées (**Figure 31**), avec des conservations de nucléotides de plus de 90% entre organismes différents. Cette forte conservation au niveau nucléotidique avait conduit des auteurs à proposer qu'elles pourraient refléter l'existence de structures secondaires dont la fonction serait essentielle pour justifier leur maintien au cours de l'évolution (Huynen et al., 1992).

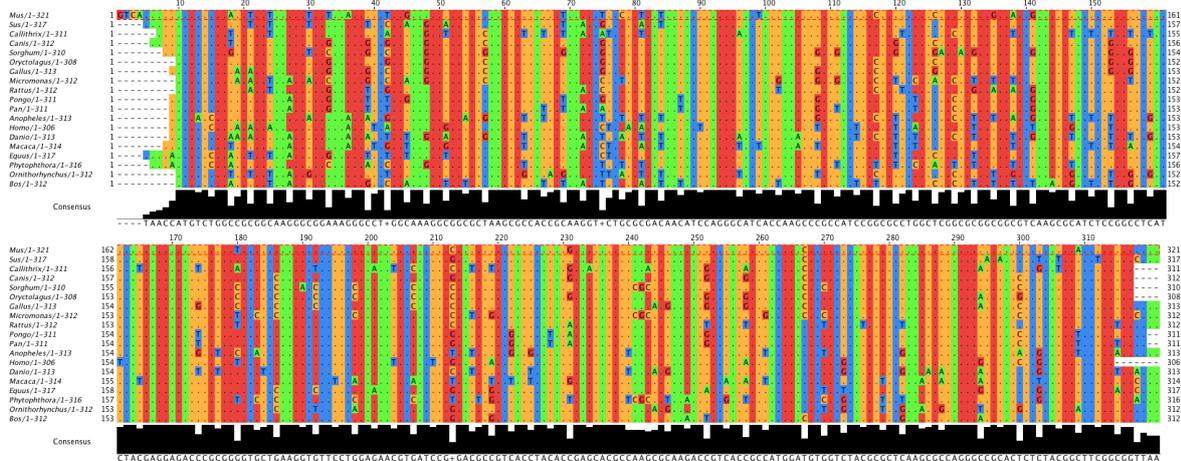


Figure 31. Alignement multiple réalisé sur des séquences cDNA de l'histone H4 issus de différents organismes eucaryotiques. Les nucléotides G, A, T et C sont respectivement coloriés en rouge, vert, jaune et orange. Les nucléotides sont affichés sont ceux qui diffèrent de la séquence consensus.

La construction du modèle de structure secondaire a finalement été réalisée manuellement. Nous sommes partis du modèle existant (**Figure 26**) en y insérant les nouvelles données. Un certain nombre de réactivités ambiguës ont disparu lors du sondage des deux conformations séparées. Par ailleurs, le sondage « direct » de l'ARNm marqué en 5' ou 3' a considérablement accru la qualité des données par rapport au sondage « indirect » et comblé les lacunes des zones « blanches » masquées par les amorces utilisées par la reverse transcriptase.

Par contre, les signaux ont parfois été difficiles à attribuer au centre de la molécule, lorsque la taille des fragments dépassait 175 nt. Deux éléments structuraux du premier modèle ont été conservés : il s'agit de la structure en jonction triple dont la fonction est d'aider la particule 43S à identifier l'AUG initiateur (**Figure 32**, tige boucles II et III) et de l'élément 4E-SE dont la fonction est de fixer eIF4E par son domaine NH₂-terminal (**Figure 32**, tiges boucles VII et VIII) (Martin et al., 2011). Les tiges boucles IV et V ont également été conservées, ainsi que la XIV qui est la tige boucle terminale impliquée dans le processus de maturation de l'extrémité 3' (Dominski and Marzluff, 2007).

Le modèle initial (**Figure 26**) a été considérablement modifié par l'ouverture de l'appariement à distance qui refermait l'ARNm et l'ajout d'un certain nombre de tiges boucles supplémentaires. Bien que présentant peu de réactivités, le début de la molécule a été replié pour former la tige boucle I. L'élément 4E-SE a été englobé au sein d'une structure plus étendue à l'allure de tRNA (tiges boucles VII, VIII et IX). Cette structure repose sur un appariement de 8 paires de bases et sur la conservation de la structure 4E-SE. Pour connecter l'ensemble, 4 nouvelles tiges boucles (VI, X, XI et XII) ont été rajoutées. Globalement, le modèle présenté dans la **figure 32** est en bon accord avec les données de sondage, mais il ne permet pas d'expliquer les différences de réactivité observées entre les conformations H et B. Il se pourrait que les conformations ne diffèrent que par des interactions à distance qui n'ont pas encore été élucidées.

Bien qu'encore inachevé et n'expliquant pas toutes les réactivités, ce modèle a servi de base de réflexion pour l'étude cristallographique qui a suivi.

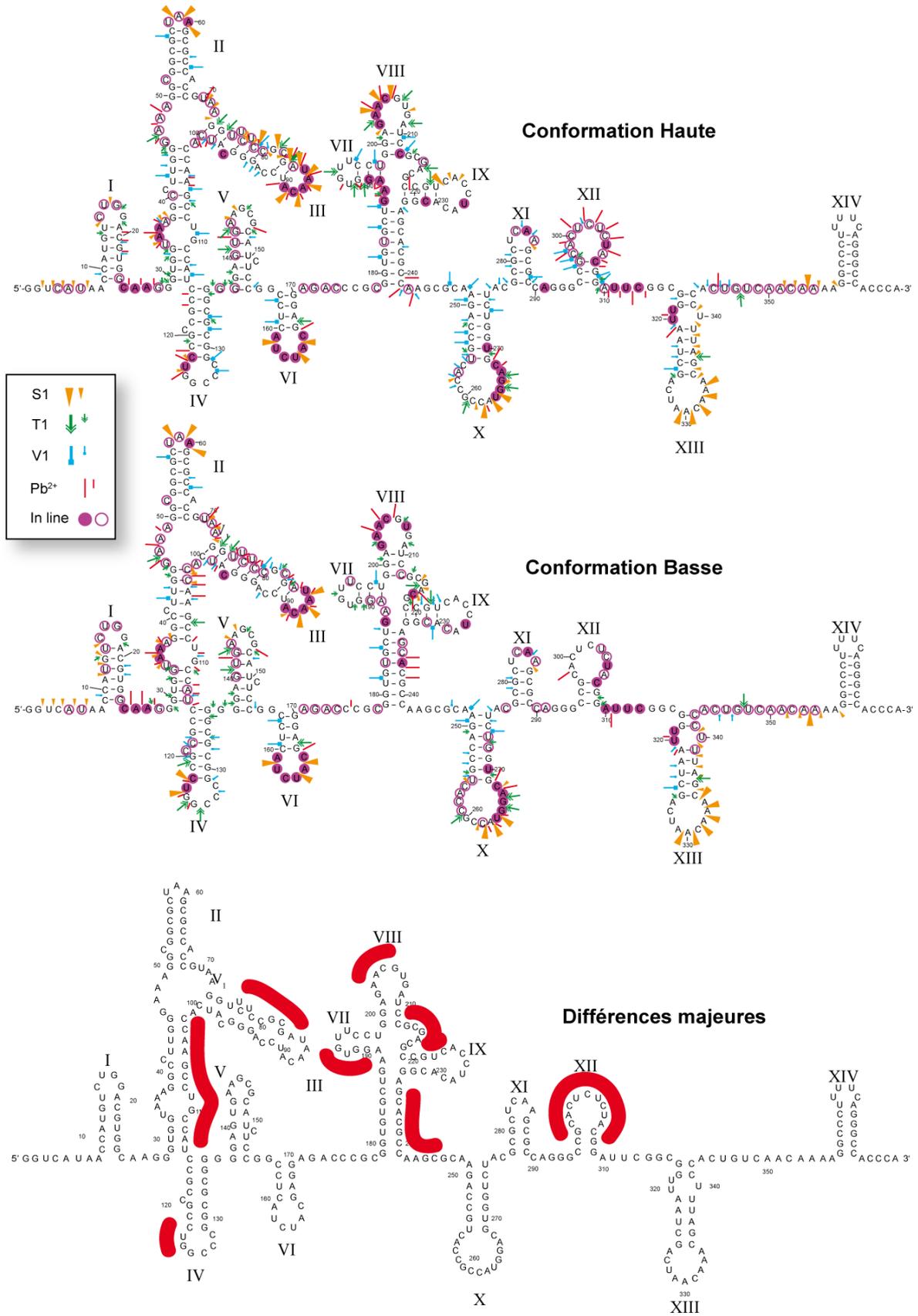


Figure 32. Modèle de structure secondaire de l'ARNm H4 réalisé avec les données de sondage en solution. Les réactivités obtenues avec la forme haute sont reportées sur le modèle du haut tandis que les réactivités obtenues avec la forme basse ont été reportées sur le modèle du milieu. Les différences de réactivité les plus fortes sont indiquées sur le modèle du bas.

Chapitre 5 : Essais de cristallisation sur l'ARNm H4 complet ou tronqué

Afin de résoudre la structure de l'ARNm H4 par la méthode de diffraction des rayons X sur des monocristaux, des essais de cristallisation ont été réalisés dans un premier temps sur les formes H et B de l'ARNm H4 puis, dans un second temps, sur des formes délétées de l'ARNm.

5.1 Essais de cristallisation sur l'ARNm H4 complet

5.1.1 Introduction

Durant cette étape, l'ARN est exposé à différentes solutions, issues de matrices commerciales, dont les compositions ont été choisies de façon à couvrir une large gamme de tampons, de pH, de sels et d'agents précipitants. Ces matrices commerciales contenant chacune 96 conditions différentes ont été utilisées à différentes températures (17, 20, 24 et 37°C). Les ARNm étant sensibles aux pH alcalins ainsi qu'aux solutions contenant de fortes concentrations en cations divalents, des matrices contenant peu de conditions dénaturantes ($\text{pH} \leq 7,5$; $[\text{MgCl}_2] \leq 50 \text{ mM}$) pour les ARN ont été privilégiées. Ainsi, les matrices « The PEGs », « σ RNA » « The Classics », « The MPD », « Index » et « Nucleix » ont été utilisées en utilisant la technique par diffusion de vapeur (gouttes en position assise). Toutes les gouttes réalisées ont été observées après 1, 2,4 et 8 semaines d'incubation. Cette analyse a été systématiquement réalisée sur différents plans focaux (minimum 3).

5.1.2 Inventaire et résultats des essais de cristallisation réalisés

Afin de déterminer la concentration optimale en ARNm H4 pour les essais de cristallisation, des premiers tests ont été réalisés à partir de matrices systématiques contenant différents gradients d'agents précipitants (PEG, MPD et AmSO_4). Suite aux résultats obtenus (données non présentées), nous avons décidé de retenir une concentration en ARNm H4 de 1,5 mg/mL. La forme H étant la plus compacte et la plus pure à l'issue du processus de repliement, nous l'avons testée en priorité. De plus, cette forme peut être obtenue en présence de NaCl ou de MgCl_2 (**Figure 5**), ce qui nous a permis de tester l'ARN replié avec l'un ou l'autre de ces sels. Au total, plus de 3 500 essais de cristallisation ont été réalisés sur l'ARNm H4, (**Tableau**

6), mais aucun cristal d'ARN n'a été obtenu après deux mois d'incubation, seuls quelques cristaux de sels ont été obtenus (**Figure 33**). Bien que décevante, cette absence de cristaux d'ARN n'est pas surprenante puisqu'à ce jour, aucune structure d'ARNm complet

Construction	Matrice utilisée	Information additionnelle	Température (°C)	Concentration en ARN (mg/mL)	Nombre de précipités (%)
ARN H4 (H)	The PEGs	Repliement avec 150 mM NaCl	37	1,5	47
ARN H4 (H)	The PEGs	Repliement avec 150 mM NaCl	37	3	54
ARN H4 (H)	The PEGs	Repliement avec 150 mM NaCl	17	1,5	72
ARN H4 (H)	The PEGs	Repliement avec 150 mM NaCl	17	3	75
ARN H4 (H)	The PEGs	Repliement avec 3 mM MgCl ₂	20	3	38
ARN H4 (H)	The Classics	Repliement avec 150 mM NaCl	37	1,5	56
ARN H4 (H)	The Classics	Repliement avec 150 mM NaCl	37	3	60
ARN H4 (H)	The Classics	Repliement avec 150 mM NaCl	17	1,5	60
ARN H4 (H)	The Classics	Repliement avec 150 mM NaCl	17	3	73
ARN H4 (H)	Index	Repliement avec 150 mM NaCl	37	1,5	57
ARN H4 (H)	Index	Repliement avec 150 mM NaCl	37	3	75
ARN H4 (H)	Index	Repliement avec 150 mM NaCl	17	1,5	82
ARN H4 (H)	Index	Repliement avec 150 mM NaCl	17	3	86
ARN H4 (H)	σ RNA	Repliement avec 150 mM NaCl	37	1,5	37,5
ARN H4 (H)	σ RNA	Repliement avec 150 mM NaCl	37	3	48
ARN H4 (H)	σ RNA	Repliement avec 150 mM NaCl	17	1,5	65
ARN H4 (H)	σ RNA	Repliement avec 150 mM NaCl	17	3	67
ARN H4 (H)	Nucleix	Repliement avec 150 mM NaCl	24	1,5	80
ARN H4 (H)	Nucleix	Repliement avec 3 mM MgCl ₂	20	3	73
ARN H4 (H)	The MPD	Repliement avec 3 mM MgCl ₂	20	3	58
ARN H4 (H)	Additif 10% pH 8 + AmSO ₄	Repliement avec 150 mM NaCl	24	1,5	44
ARN H4 (H)	Additif 10% pH 7 + AmSO ₄	Repliement avec 150 mM NaCl	24	1,5	28
ARN H4 (H)	Additif 10% pH 6 + AmSO ₄	Repliement avec 150 mM NaCl	24	1,5	20
ARN H4 (H)	Additif 10% pH 8 + PEG 3350	Repliement avec 150 mM NaCl	24	1,5	18
ARN H4 (H)	Additif 10% pH 7 + PEG 3350	Repliement avec 150 mM NaCl	24	1,5	16
ARN H4 (H)	Additif 10% pH 6 + PEG 3350	Repliement avec 150 mM NaCl	24	1,5	15
ARN H4 (H) 5' m7G	σ RNA	Repliement avec 3 mM MgCl ₂	24	1,5	34
ARN H4 (H) 5' m7G	The Classics	Repliement avec 3 mM MgCl ₂	24	1,5	46
ARN H4 (H) 5' m7G	Nucleix	Repliement avec 3 mM MgCl ₂	24	1,5	67
ARN H4 (H) 5' m7G	Index	Repliement avec 3 mM MgCl ₂	24	1,5	66
ARN H4 (B)	The PEGs	Repliement avec 3 mM MgCl ₂	20	3	55
ARN H4 (B)	The MPD	Repliement avec 3 mM MgCl ₂	20	3	66
ARN H4 (B)	Nucleix	Repliement avec 3 mM MgCl ₂	20	3	87
ARN H4 (B) 5' m7G	σ RNA	Repliement avec 3 mM MgCl ₂	20	1,5	32
ARN H4 (B) 5' m7G	The Classics	Repliement avec 3 mM MgCl ₂	20	1,5	67
ARN H4 (B) 5' m7G	Nucleix	Repliement avec 3 mM MgCl ₂	20	1,5	80
ARN H4 (B) 5' m7G	Index	Repliement avec 3 mM MgCl ₂	20	1,5	80

Tableau 6. Résumé des essais de cristallisation réalisés à partir de matrices commerciales sur les conformations Haute et Basse de l'ARN H4. Le type et la quantité de cations ajoutés lors du processus de repliement sont indiqués. La matrice « Additif » a été utilisée suite à une dilution d'un facteur 10 dans un milieu contenant soit 13% (v/v) de PEG ou bien 900 mM d'AmSO₄ (concentration finale). Le pH de la matrice « Additif » a été ajusté soit à l'aide de 100 mM Tris-HCl (pH 8) ou à partir de 100 mM Bis-Tris-HCl (pH 7 et 6). Pour chaque matrice utilisée, le pourcentage de conditions ayant conduit à une précipitation de l'ARN a été indiqué.

n'a été résolue par diffraction des rayons X. Bien que de taille raisonnable (376 bases), l'ARNm H4 est probablement trop grand et sa structure est trop flexible pour cristalliser dans les conditions testées. En conséquence, nous avons décidé de réduire la taille de l'ARNm H4 en utilisant le modèle de structure secondaire proposé dans le chapitre 4 (**Figure 32**) comme base de réflexion.



100 mM Bis-Tris-HCl pH 7, 150 mM NaCl, 900 mM AmSO_4 , 10 mM SrCl_2 , 24°C



100 mM Hepes pH 7, 150 mM NaCl, 30% PEG 3350, 37°C

Figure 33. Exemples de cristaux de sels obtenus lors des essais de cristallisation sur l'ARNm H4 complet.

5.2 Essais de cristallisation sur l'ARNm H4 tronqué

5.2.1 Sélection d'ARNm tronqués en 5' ou 3' ou 3' et 5'

Suite aux premiers essais de cristallisation de l'ARNm entier, des ARNm H4 tronqués en 5', ou en 3', ou en 3' et 5' ont été choisis en se basant sur le modèle de structure secondaire établi à partir des données de sondage en solution et proposé dans le chapitre 4 (**Figure 32**). L'objectif principal de ces délétions était de disposer de fragments moins flexibles et plus aptes à cristalliser. La plupart des sites de coupure ont été choisis de façon à générer des ARN s'achevant par des hélices complètes (**Figure 34**), car les extrémités 5' et 3' repliées en double brin sont souvent impliquées dans les contacts cristallins. L'absence de zones simple brin permet également de faciliter l'étape de repliement.

Parmi les fragments sélectionnés, on retrouve le domaine minimal de fixation d'eIF4E (179-342) ainsi que le domaine de fixation optimal de la coiffe (1-135), capable d'assembler des particules 80S (résultats non publiés).

Les différents fragments délétés ont été testés dans les deux procédés de repliement dans le but d'identifier les séquences impliquées dans le repliement dans l'une ou l'autre des conformations.

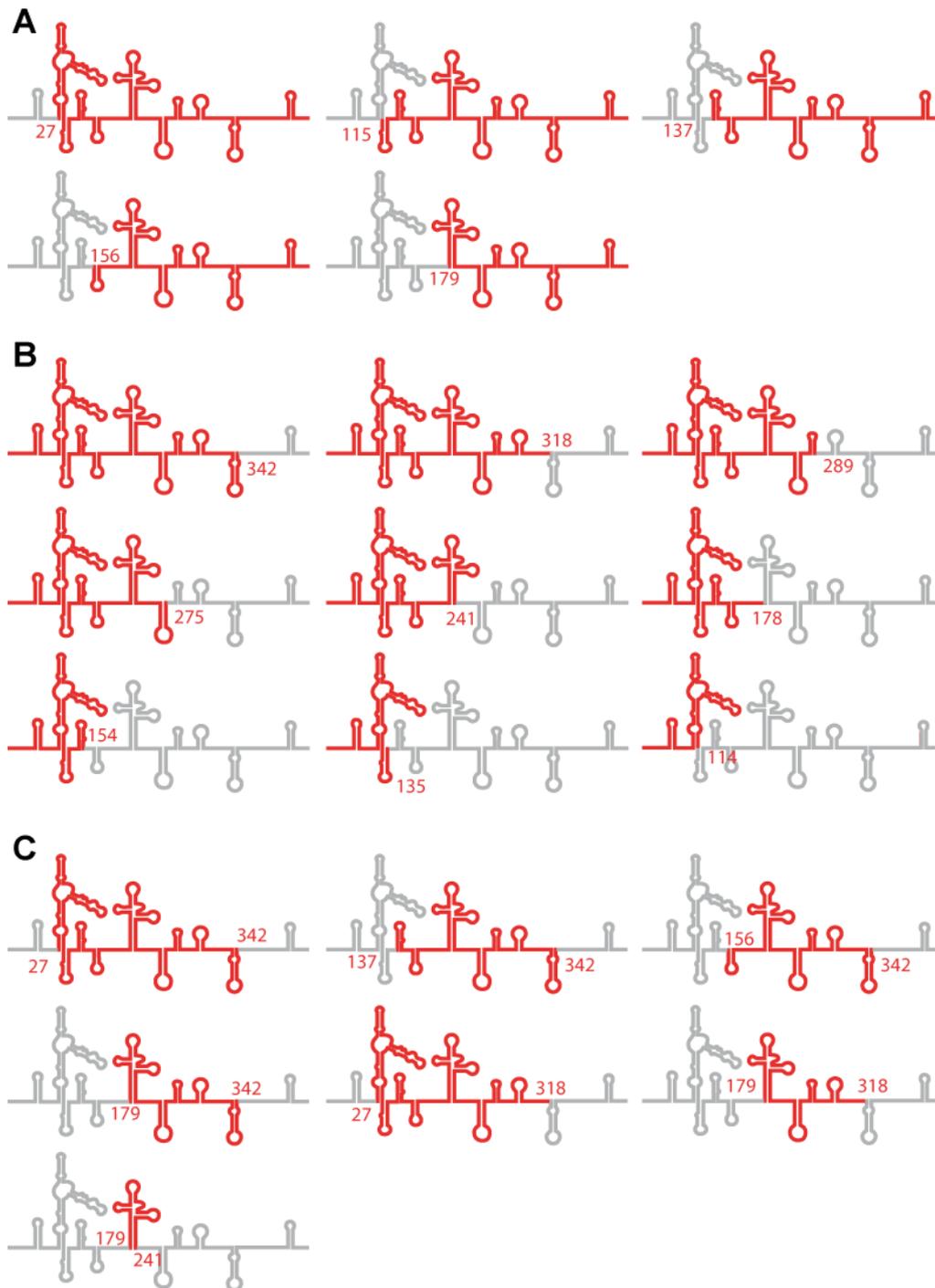


Figure 34. Fragments d'ARNm H4 sélectionnés pour réaliser des essais de cristallisation. Au total, 3 catégories de constructions ont été sélectionnées : les fragments tronqués en 5' (A), ou en 3' (B), ou en 5' et 3' (C). Les zones éliminées ont été représentées en gris.

5.2.2 Étude du repliement des ARN tronqués

Le repliement des ARN étant dépendant de leurs séquences, il convenait d'étudier le repliement des ARNm H4 tronqués qui pouvait différer de l'ARNm complet. Idéalement, les conditions de repliement auraient dues être déterminées pour chaque construction. Dans un

premier temps, seule la construction 1-135 a été étudiée. Cette construction a été choisie comme modèle car elle correspond au domaine minimum de fixation des particules ribosomiques et au site optimal de liaison de la coiffe. À nos yeux, cette construction présentait un fort intérêt et devait être cristallisée en priorité. De plus, sa petite taille était un atout supplémentaire. Pour déterminer ses conditions optimales de repliement, les protocoles développés précédemment ont été utilisés (**Figures 5, 9, 13 et 14**). Finalement, le protocole consistant à chauffer brièvement l'ARN dans l'eau à 95°C (1 min) puis à le refroidir brutalement en plongeant le tube dans la glace (5 min) a été retenu car il permettait d'obtenir très rapidement et très simplement une seule forme d'ARN. Cette forme présentait une migration rapide sur gel natif par rapport aux autres formes obtenues qui pouvaient correspondre à des dimères ou des repliements alternatifs (données non présentées).

Le même protocole a ensuite été appliqué à tous les fragments tronqués sélectionnés qui ont été transcrits puis purifiés. Les résultats ont été analysés sur gel ABA natif. En parallèle, les protocoles utilisés pour obtenir la forme H ou B de l'ARNm H4 complet ont été appliqués à tous les fragments sélectionnés afin d'observer leur capacité à se replier dans l'une ou l'autre des conformations. Les gels ABA natifs de la **figure 35** montrent que le protocole retenu pour replier la construction 1-135 est également efficace pour replier les constructions 115-376, 137-376, 156-37, 179-376, 1-241, 1-178, 1-154, 1-114, 137-342, 156-342, 179-342, 27-318, 179-318 et 179-241. Au total, sur les 21 fragments construits et purifiés, 18 ont été repliés correctement en une forme unique et ont été retenus pour réaliser des essais de cristallisation. La **figure 35** montre également que les séquences localisées entre les positions 27 et 115 sont impliquées dans la formation de la forme H qui n'est plus obtenue dès la délétion des 115 premiers résidus (délétant 115-376). Ce résultat est confirmé par les doubles délétants (137-342 et autres dérivés). Du côté 3', on note l'apparition de bandes diffuses lors des délétions ce qui suggère une instabilité de la conformation des ARN lors de la migration sur gel ABA natif. Cette instabilité nous a contraint à écarter un certain nombre de fragments tels les ARNm 1-318, 1-289 et 1-275. Des délétions plus importantes de l'extrémité 3' ont à nouveau conduit à des formes migrant de façon plus focalisée mais ne présentant plus la ségrégation en formes H et B dans les protocoles de repliement correspondants. Ces formes ont été testées en cristallisation.

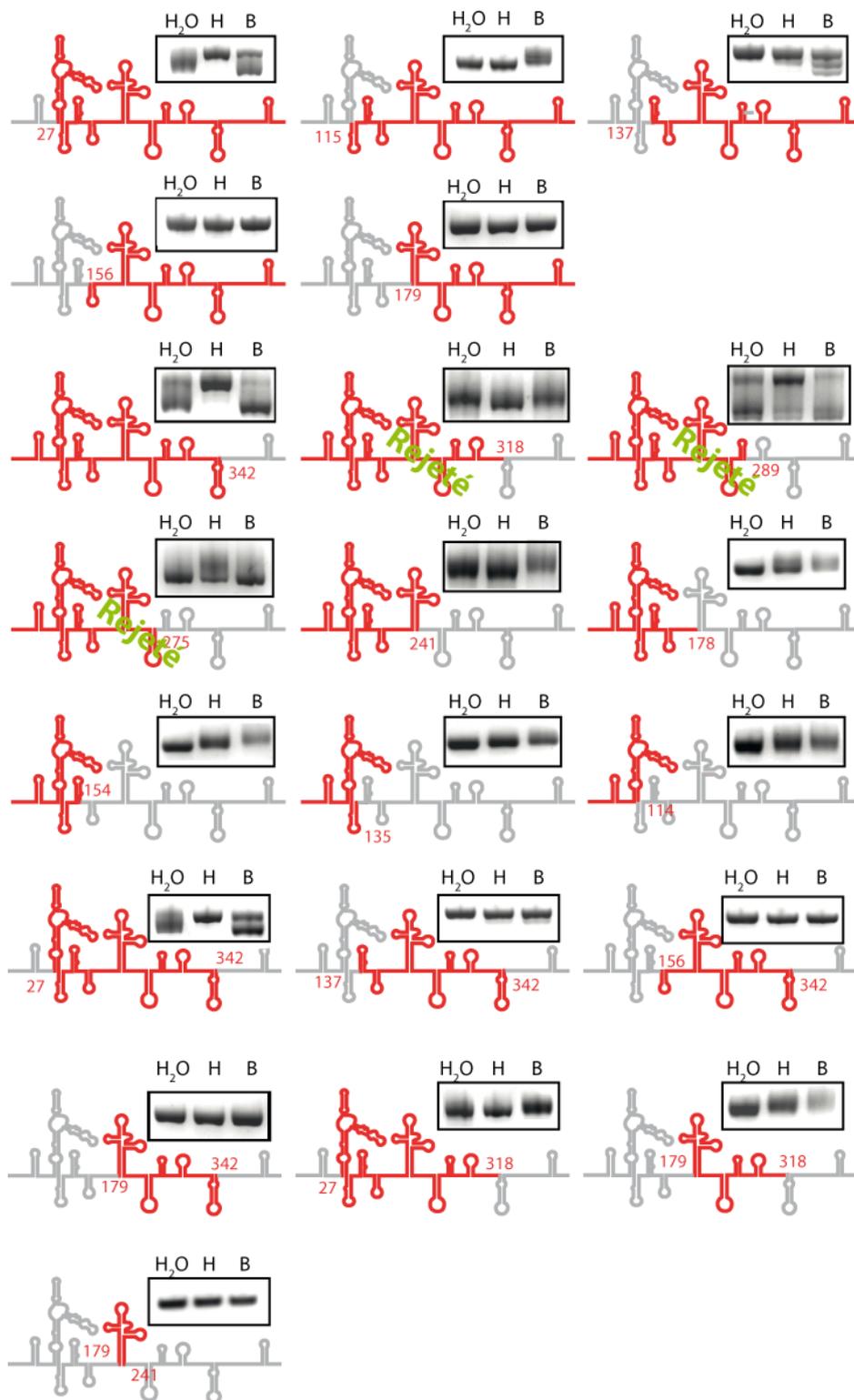


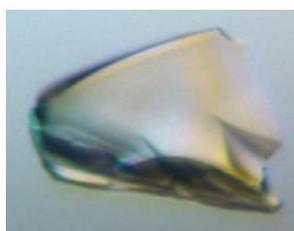
Figure 35. Analyse sur gel ABA natif des résultats des protocoles de repliement des fragments d'ARNm tronqués. Trois protocoles ont été utilisés : les pistes H ou B correspondent aux protocoles classiques exposés dans la figure 17 ; les pistes H₂O correspondent à une dénaturation de l'ARN dans l'eau à 95°C (1 min) puis refroidissement brutal dans la glace (5 min).

5.2.3 Inventaire et résultats des essais de cristallisation réalisés

Pour chaque construction sélectionnée, la concentration en ARN optimale pour réaliser des essais de cristallisation a été déterminée de la même manière que pour l'ARNm H4 complet. Pour les mêmes raisons évoquées précédemment, la construction 1-135 a fait l'objet d'une attention toute particulière et d'une étude plus approfondie (**Tableau 6**). Cependant, aucun cristal d'ARN n'a été obtenu à partir de cette construction, seuls quelques cristaux non reproductibles et ne diffractant par les rayons X ont été obtenus (**Figure 36**).

Construction	Matrice utilisée	Température (°C)	Concentration en ARN (mg/mL)	Nombre de précipités (%)
1-135	AmSO ₄	17	1,5	99
1-135	AmSO ₄	24	1,5	99
1-135	σ RNA	17	1,5	40
1-135	σ RNA	24	1,5	19
1-135	σ RNA	20	10	20
1-135	The Classics	17	1,5	55
1-135	The Classics	24	1,5	40
1-135	The PEGs	17	1,5	46
1-135	The PEGs	24	1,5	14
1-135	Index	17	1,5	78
1-135	Index	24	1,5	26
1-135	Nucleix	20	10	66
1-135 5' m7G	σ RNA	20	10	31
1-135 5' m7G	σ RNA	24	1,5	20
1-135 5' m7G	Nucleix	20	10	65
1-135 5' m7G	Nucleix	24	1,5	58
1-135 5' m7G	The Classics	24	1,5	48
1-135 5' m7G	Index	24	1,5	63
1-135 5' m7G	Nucleix 10% pH 7,5 + PEG 3350	20	1,5	20
1-135 5' m7G	Nucleix 10% pH 7,5 + PEG 2000	20	1,5	22
1-135 5' m7G	Nucleix 10% pH 7,5 + PEG 1500	20	1,5	25
1-135 5' m7G	Nucleix 10% pH 7,5 + MPD	20	1,5	30

Tableau 6. Résumé des essais de cristallisation réalisés à partir de matrices commerciales sur la construction 1-135. La matrice « Additif » a été utilisée suite à une dilution d'un facteur 10 dans un milieu contenant soit 22% (v/v) de PEG 3350/2000, ou 27% (v/v) de PEG 1500 ou bien 40% de MPD (v/v) (concentration finale). Le pH de la matrice « Additif » a été ajusté à l'aide de 100 mM Tris-HCl (pH 7,5). Pour chaque matrice utilisée, le pourcentage de conditions ayant conduit à une précipitation de la construction a été indiqué.



100 mM Tris-HCl pH 7,5
500 mM AmSO₄, 24°C



100 mM Tris-HCl pH 7,5
750 mM AmSO₄, 17°C

Figure 36. Exemples de cristaux obtenus lors des essais de cristallisation sur la construction 1-135.

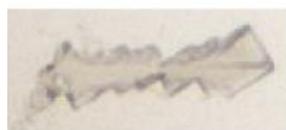
Construction	Matrice utilisée	Température (°C)	Concentration en ARN (mg/mL)	Nombre de précipités (%)
1-114	The PEGs	20	3	55
1-114	The MPD	20	3	60
1-114	Nucleix	20	3	70
1-342	Nucleix	20	1,5	40
1-342	Nucleix	20	5	94
1-342	The PEGs	20	1,5	48
1-342	The PEGs	37	1,5	35
1-342	The MPD	20	1,5	78
1-342	The MPD	37	1,5	65
1-318	Nucleix	20	1,5	66
1-318	Nucleix	20	5	83
1-318	The PEGs	20	1,5	41
1-318	The PEGs	37	1,5	35
1-318	The MPD	20	1,5	80
1-318	The MPD	37	1,5	59
27-376	Nucleix	20	1,5	69
27-376	Nucleix	20	5	82
27-376	The PEGs	20	1,5	42
27-376	The PEGs	37	1,5	30
27-376	The MPD	20	1,5	70
27-376	The MPD	37	1,5	52
27-342	Nucleix	20	1,5	55
27-342	Nucleix	20	5	81
27-342	The PEGs	20	1,5	42
27-342	The PEGs	37	1,5	36
27-342	The MPD	20	1,5	69
27-342	The MPD	37	1,5	45
27-318	Nucleix	20	1,5	65
27-318	Nucleix	20	5	94
27-318	The PEGs	20	1,5	56
27-318	The PEGs	37	1,5	42
27-318	The MPD	20	1,5	72
27-318	The MPD	37	1,5	49
179-376	Nucleix	20	3	70
179-376	Nucleix	20	2	69
179-376	Nucleix	20	1	61
179-376	The PEGs	20	1	44
179-376	The PEGs	37	1	48
179-376	The MPD	20	1	70
179-376	The MPD	37	1	50
179-376	The Classics	20	2	54
179-376	Index	20	2	62
179-342	Nucleix	20	3	74
179-342	Nucleix	20	1	60
179-342	The PEGs	20	1	44
179-342	The PEGs	37	1	36
179-342	The MPD	20	1	56
179-342	The MPD	37	1	44
179-318	Nucleix	20	3	76
179-318	Nucleix	20	1	56
179-318	The PEGs	20	1	54
179-318	The PEGs	37	1	44
179-318	The MPD	20	1	70
179-318	The MPD	37	1	56
179-241	The Classics	20	2	47
179-241	Index	20	2	51
179-241	Nucleix	20	2	62
1-289	Nucleix	20	1,5	80
1-289	The MPD	20	1,5	53
1-289	Index	20	1,5	65
1-289	σ RNA	20	1,5	32

Tableau 7. Résumé des essais de cristallisation réalisés à partir de matrices commerciales sur les différentes constructions. Pour chaque matrice utilisée, le pourcentage de conditions ayant conduit à une précipitation de la construction a été indiqué.

En cumulant tous les essais de cristallisations réalisés sur l'ensemble des fragments délétés, plus de 11 000 gouttes ont été réalisées (**Tableaux 7 et 8**), mais aucun cristal d'ARN n'a été obtenu, seuls des cristaux de sels ont été retrouvés dans les gouttes (**Figure 37**).

Construction	Matrice utilisée	Température (°C)	Concentration en ARN (mg/mL)	Nombre de précipités (%)
137-376	The Classics	20	2	55
137-376	Index	20	2	69
137-376	Nucleix	20	2	77
156-376	The Classics	20	2	27
156-376	Index	20	2	33
156-376	Nucleix	20	2	41
156-342	The Classics	20	2	44
156-342	Index	20	2	48
156-342	Nucleix	20	2	70
137-342	The Classics	20	2	39
137-342	Index	20	2	48
137-342	Nucleix	20	2	56
1-178	Index	20	1	34
1-178	Index	20	2	45
1-178	Nucleix	20	1	56
1-178	Nucleix	20	2	65
1-178	The MPD	20	1	45
1-178	σ RNA	20	1	30
1-178	The Classics	20	2	57
1-241	Index	20	1	54
1-241	Index	20	2	67
1-241	Nucleix	20	1	69
1-241	Nucleix	20	2	80
1-241	The MPD	20	1	58
1-241	σ RNA	20	1	43
1-241	The Classics	20	2	55
1-154	Index	20	1	24
1-154	Index	20	2	35
1-154	Nucleix	20	1	60
1-154	Nucleix	20	2	77
1-154	The MPD	20	1	28
1-154	σ RNA	20	1	33

Tableau 8. Résumé des essais de cristallisation réalisés à partir de matrices commerciales sur les constructions 137-376, 156-376, 137-342, 1-178, 1-241 et 1-154. Pour chaque matrice utilisée, le pourcentage de conditions ayant conduit à une précipitation de la construction a été indiqué.



179-376
200 mM K_2HPO_4
40% MPD, 20°C



27-342
100 mM Bicine pH 9
65% MPD, 20°C

Figure 37. Exemples de cristaux de sels obtenus lors des essais de cristallisation sur les constructions 179-376 et 27-342.

5.3 Conclusion

Au total, plus de 15 000 essais de cristallisation ont été réalisés sur l'ARNm H4 ou sur des fragments d'ARNm H4, mais malgré tous ces tests aucun cristal d'ARN n'a été obtenu après 1, 2, 4 ou 8 semaines d'incubation. Face à ces résultats, nous avons envisagé l'hypothèse que les fragments d'ARN pouvaient changer de conformation ou se dégrader dans les gouttes. Des tests de stabilité ont été réalisés sur certains de ces ARN afin d'observer leur comportement au cours du temps. Les résultats obtenus sur gel natif et par équilibre de sédimentation (données non présentées) ont montré que les constructions 1-178, 1-154, 1-135 et 1-114 sont stables dans l'eau, mais forment des dimères dès qu'elles entrent en contact avec des cations. Les constructions 115-376, 137-376, 156-376 et 179-376 sont stables durant plusieurs jours à 20°C, que se soit dans l'eau ou en présence de cations.

Cependant, pour contrer la capacité des molécules d'ARN à se repousser entre elles du fait de leurs charges nettes négatives, l'ajout de cations est nécessaire lors de la réalisation des expériences de cristallisation, sans quoi ces molécules seront incapables de s'empiler les unes sur les autres et donc de former des cristaux. Ce phénomène de répulsion est également préjudiciable aux expériences de SAXS et d'ultracentrifugation analytique car il complique fortement l'interprétation des données. Dans le contexte des constructions 1-178, 1-154, 1-135 et 1-114, il faut considérer que leur tendance à la dimérisation voir à l'agrégation en présence de cations a pu être un frein considérable pour l'obtention de cristaux.

Par ailleurs, d'autres expériences de sondage en solution sur des constructions tronquées à partir des extrémités 5' et 3' ont montré que leur profil de coupure semblait différer de celui de l'ARNm entier. Ceci suggérait que ces molécules ne se repliaient pas selon la prédiction initiale déduite du modèle de structure secondaire présenté dans la **figure 32**. Une des causes de ce phénomène pourrait être que le modèle présente des erreurs. Ceci nous a conduit à réexaminer le modèle de structure secondaire dans le travail présenté dans le chapitre suivant.

Chapitre 6 : Réexamen du modèle de structure secondaire de l'ARNm H4

Dans le cadre de contrôles visant à vérifier le bon repliement des fragments tronqués de l'ARNm H4, des expériences de sondage en solution ont été réalisées sur un certain nombre de mutants. Parmi les constructions tronquées, les fragments 137-376, 156-376 et 179-376 ont présenté des profils de coupures spontanées (par sondage « In Line ») inattendus compte tenu du modèle présenté dans la **figure 32**. Entre autre, une très forte coupure de l'ARNm se produisait aux environs de la position 240, suite à la délétion des 137 premiers nts (résultats non présentés). Dans le modèle de la **figure 32**, cette séquence devait se trouver dans l'hélice portant les tiges boucles VII, VIII et IX, c'est-à-dire à un endroit du modèle où la flexibilité n'était pas attendue. Ces données, ainsi que d'autres observations lors des expériences de repliement des formes tronquées, suggéraient que le modèle de structure secondaire présenté dans la **figure 32** n'était pas totalement correct et qu'il méritait d'être revu et ré analysé.

6.1 Ré analyse des modèles de structure secondaire

6.1.1 Région 1-135 présentant la jonction à trois hélices

Les profils de coupure aux RNases S1 et V1 présentés dans les **figures 27** et **28** montrent que globalement les formes H et B sont repliées de la même façon malgré les différences d'intensité observées à certains endroits. Ces différences de coupure reflètent certainement une différence d'accessibilité de l'ARNm aux RNases, mais ne permettent pas de conclure à un repliement complètement différent. Les différences de clivage, plus marquées en présence d'ions Pb^{2+} , pourraient s'expliquer par la petite taille de la sonde, capable de révéler des changements même subtils du repliement. L'analyse plus poussée des réactivités montre que la tige boucle III de la forme haute de la **figure 32** est bien plus réactive aux coupures S1, suggérant une déstabilisation de cette hélice déjà peu stable du fait de la présence de deux bulles internes. En parallèle, on peut noter que la boucle apicale de la tige boucle IV subit des variations de réactivité aux RNases V1, T1 et S1 qui indiquent éventuellement une réponse au changement de réactivité de la tige boucle III. Cette observation nous a conduit à noter la

présence d'une complémentarité possible entre les nt 94-97 et 125-128. Cette interaction potentielle pourrait ouvrir la tige de la région III et entraîner les réactivités à la nucléase S1 spécifiques de la forme haute et expliquer les réactivités double brin à la RNase V1 dans la boucle apicale IV. La forme haute pourrait donc présenter une interaction tertiaire de 4 paires de bases entre les nt 94-97 et 125-128 (**Figure 38**). Absente dans la forme basse, elle se traduirait par une meilleure accessibilité de la boucle IV aux sondes simple brin et un renforcement de la stabilité du bras III (moins de réactivités à la RNase S1). Logiquement, cette interaction tertiaire devrait se traduire par un rapprochement de la tige boucle IV du bras III et nous avons observé une diminution de l'accessibilité aux ions Pb^{2+} de la séquence 100-114 qui pourrait être masquée par la tige boucle IV.

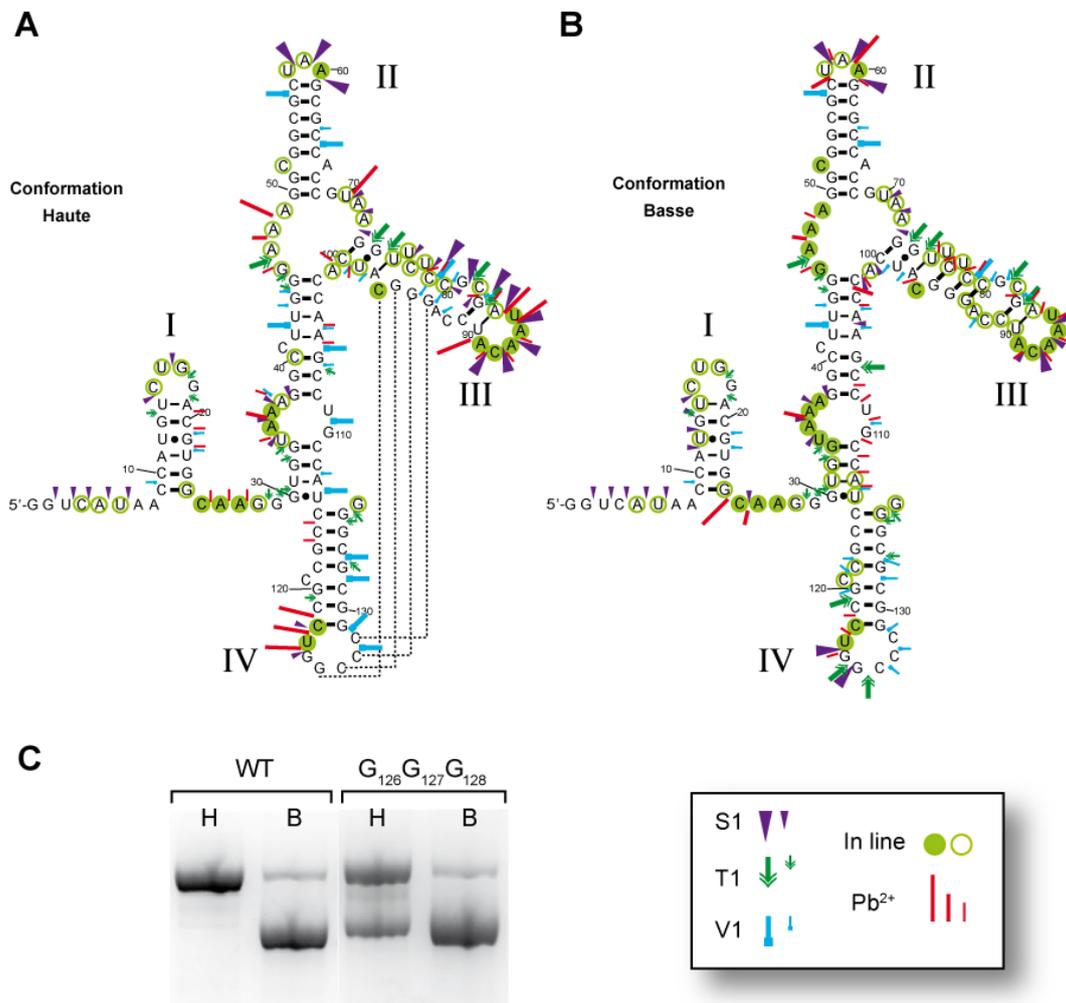


Figure 38. Modèle de structure secondaire de la partie 1 à 136 de l'ARNm H4 replié dans les formes haute et basse (**A**, **B**). Les réactivités aux RNases et ions Pb^{2+} ont été reportées sur le modèle. Les pointillés représentent l'interaction tertiaire potentielle. (**C**) Analyse sur gel ABA natif de l'effet du triple mutant sur le repliement de l'ARNm H4. Les protocoles de repliement permettant d'obtenir les conformères H ou B (pistes H et B correspondantes) ont été appliqués sur le triple mutant $G_{126}G_{127}G_{128}$ et sur l'ARNm H4 WT.

Pour appuyer cette hypothèse, j'ai étudié le comportement du triple mutant $C_{126}C_{127}C_{128} \rightarrow G_{126}G_{127}G_{128}$, qui avait été préalablement construit au laboratoire, dans les protocoles de repliement conduisant habituellement aux formes H et B. Le gel ABA natif de la figure 37 montre que le triple mutant $G_{126}G_{127}G_{128}$ n'est plus capable de se replier à 100% dans la forme haute mais dans un mélange composé à 50% de forme haute et 50% de forme basse. Ainsi, ce triplet pourrait participer à l'établissement d'une interaction tertiaire participant au repliement dans la conformation haute. Cette interaction n'est pas exclusive puisque le triple mutant est encore capable de se replier à plus de 50% dans la forme haute, mais elle contribue clairement au repliement dans la conformation haute.

6.1.2 Région 136-188

Les tiges boucles V et VI de la région 140-190 présentées dans la **figure 26** ne permettaient pas d'expliquer toutes les réactivités observées. Le modèle présentait notamment des réactivités simple brin dans la tige du domaine V alors que le domaine VI semblait correct. De plus, la particularité de cette région est qu'absolument aucune différence de réactivité n'est observée entre les deux conformations. Ceci indique que cette région est commune aux deux conformations et qu'elle n'intervient pas dans le repliement spécifique en forme H ou B. Pour corriger les réactivités incohérentes nous avons revu le modèle et nous

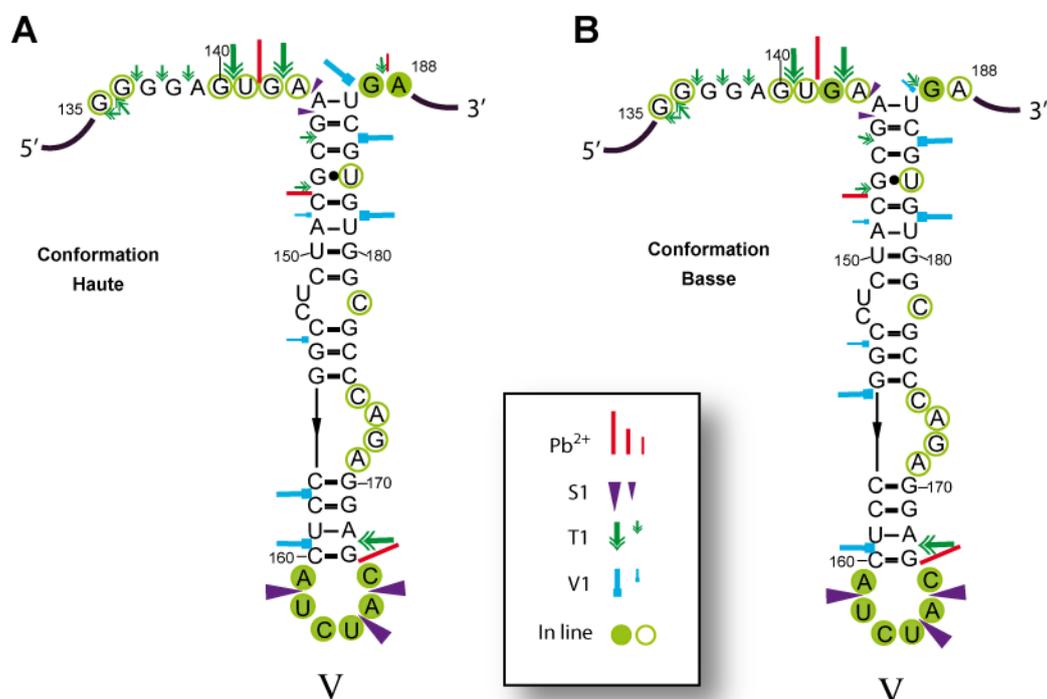


Figure 39. Modèle de structure secondaire de la région 135-188 de la forme haute (A) et basse (B). Les réactivités ont été reportées sur les modèles.

proposons de remplacer les deux tiges boucles originelles par une longue tige boucle dont la partie apicale est formée de l'ancienne tige boucle VI (**Figure 39**). Par rapport au modèle précédent, la nouvelle tige boucle V est plus allongée et une partie des réactivités simple brin se trouve désormais localisée dans la séquence de 8 nucléotides qui relie ce domaine au domaine précédent. Dans cette séquence de 8 nucléotides on trouve des coupures simple brin par la RNase T1 ainsi que des coupures in line plus marquées dans la forme haute.

6.1.3 Région 189-249 présentant l'élément 4E-SE

Cette région porteuse de l'élément 4E-SE fixant le facteur d'initiation eIF4E n'a pas été modifiée fondamentalement (**Figure 40**). Une petite tige boucle additionnelle pourrait prendre place au niveau des nucléotides 238 et 246, comme le suggère des réactivités au CMCT (Jaeger, 2005). Dans cette région, les différences de réactivité sont très importantes entre la forme haute et basse. Malheureusement, la localisation de ces signaux au centre de la molécule c'est-à-dire dans la partie supérieure des gels de polyacrylamide de séparation a rendu leur détermination plus délicate et quantitativement moins précise. Il conviendrait de sonder à nouveau les conformères par la méthode indirecte qui utilise l'extension d'amorce par la reverse transcriptase pour révéler les signaux de coupure ou de modification dans cette région. Nous avons été cependant frappés par la disparition des coupures S1 dans la boucle VIII et l'apparition de fortes coupures V1 dans le bras correspondant. Nous avons recherché des complémentarités possibles dans la séquence H4 et trouvé une interaction potentielle avec la séquence 28-34 située à la base de la tige boucle II (**Figure 38**). Notons que cette séquence correspond au site de « toe print » du ribosome et qu'elle présente aussi de forts arrêts de la reverse transcriptase lorsqu'un ARN inhibiteur antisens est apparié à l'élément 4E-SE (Jaeger, 2005). Ces forts arrêts pourraient être dus à l'interaction avec la boucle VIII décrite ci-dessus, interaction qui pourrait être renforcée par l'ARN antisens. Pour tester cette hypothèse, nous avons analysé le repliement d'un mutant préalablement construit au laboratoire : le triple mutant C₂₈C₂₉C₃₀ qui modifie 3 des nucléotides impliqués dans l'interaction potentielle. Nous avons ainsi constaté qu'avec le triple mutant, il n'était plus possible d'obtenir la forme basse de l'ARNm alors que la forme haute était obtenue normalement dans le protocole correspondant (**Figure 38**). Ce résultat montre que le triple mutant induit un effet qui perturbe le repliement dans la forme basse. La perte de l'interaction tertiaire avec la boucle VIII est une hypothèse qu'il faudrait confirmer par la construction de nouveaux mutants dans cette région de l'ARNm. L'existence de cette interaction tertiaire permettrait de rapprocher l'élément 4E-SE (et donc eIF4E fixé sur l'élément) de l'extrémité 5' de l'ARNm où se fixe le ribosome. Ce

rapprochement d'eIF4E permettrait de cibler l'activité ARN hélicase des facteurs qui lui sont associés (eIF4A dans le contexte d'eIF4F) sur les structures de l'ARNm afin de libérer le site de dépôt du ribosome.

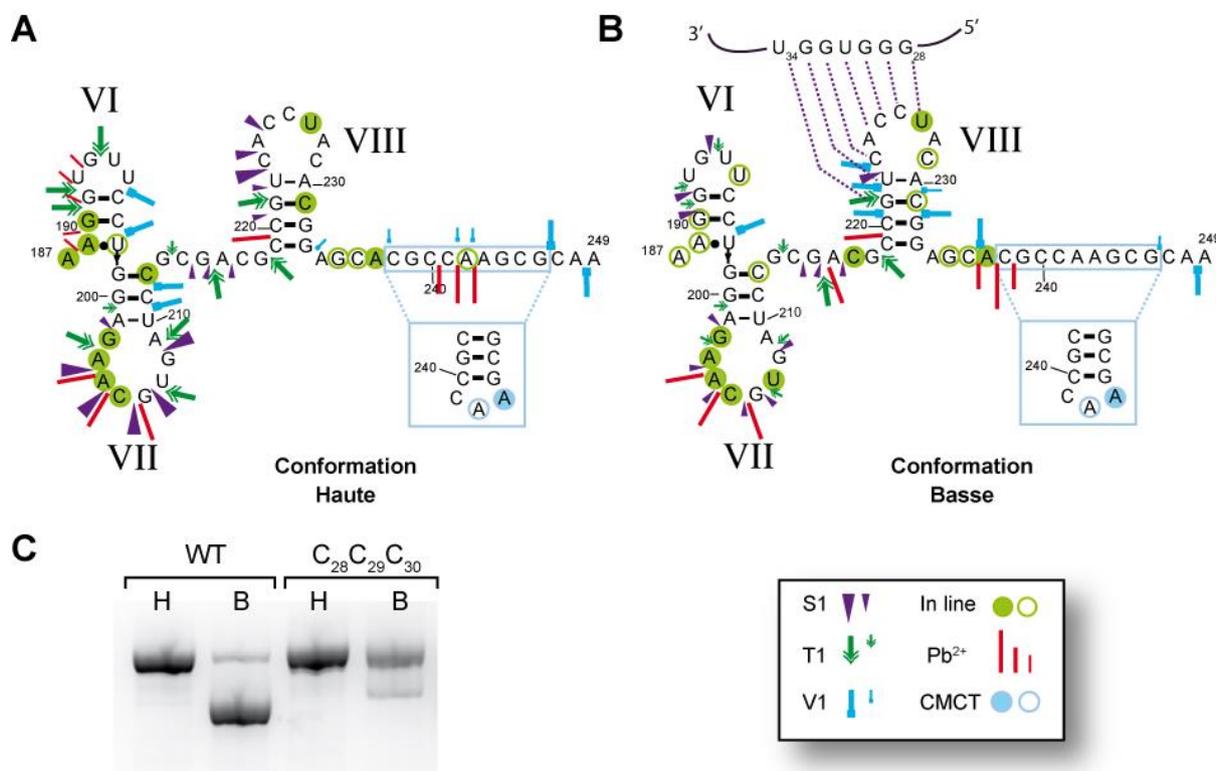


Figure 40. Modèle de structure secondaire de la région 187-249 de la forme haute (**A**) et basse (**B**). Les réactivités ont été reportées sur les modèles. Les pointillés représentent l'interaction tertiaire potentielle avec la séquence 28-34 de l'ARNm H4. (**C**) Analyse sur gel ABA natif de l'effet du triple mutant sur le repliement de l'ARNm H4. Les protocoles de repliement permettant d'obtenir les conformères H ou B (pistes H et B correspondantes) ont été appliqués sur le triple mutant $C_{28}C_{29}C_{30}$ et sur l'ARNm H4 WT.

6.1.4 Région 250-340

La région 250-340 a également été redessinée afin de mieux expliquer les différences de réactivité observées entre les 2 formes de l'ARNm H4. La tige boucle XI du modèle présenté dans la **figure 32** a été conservée car elle semble être extrêmement stable. En effet, la région qui constitue cette tige boucle n'est pas digérée par la RNase T1 en condition dénaturante (4,4 M urée à 50°C). Les précédentes tiges boucles X et XII ont été dépliées afin d'allonger la tige boucle XI de la **figure 32**, l'ensemble formant une longue tige boucle déjà proposée comme repliement alternatif (**Figure 26**, (Jaeger, 2005)). Cette structure inclut également les 4 premiers nucléotides de l'ancienne tige boucle XIII qui formaient le début d'une hélice et une bulle. Le

reste de l'ancienne tige boucle XIII a été maintenu en l'état. Le nouveau repliement de la région 250-340 et la nouvelle numérotation des tiges boucles sont présentés dans la **figure 41**.

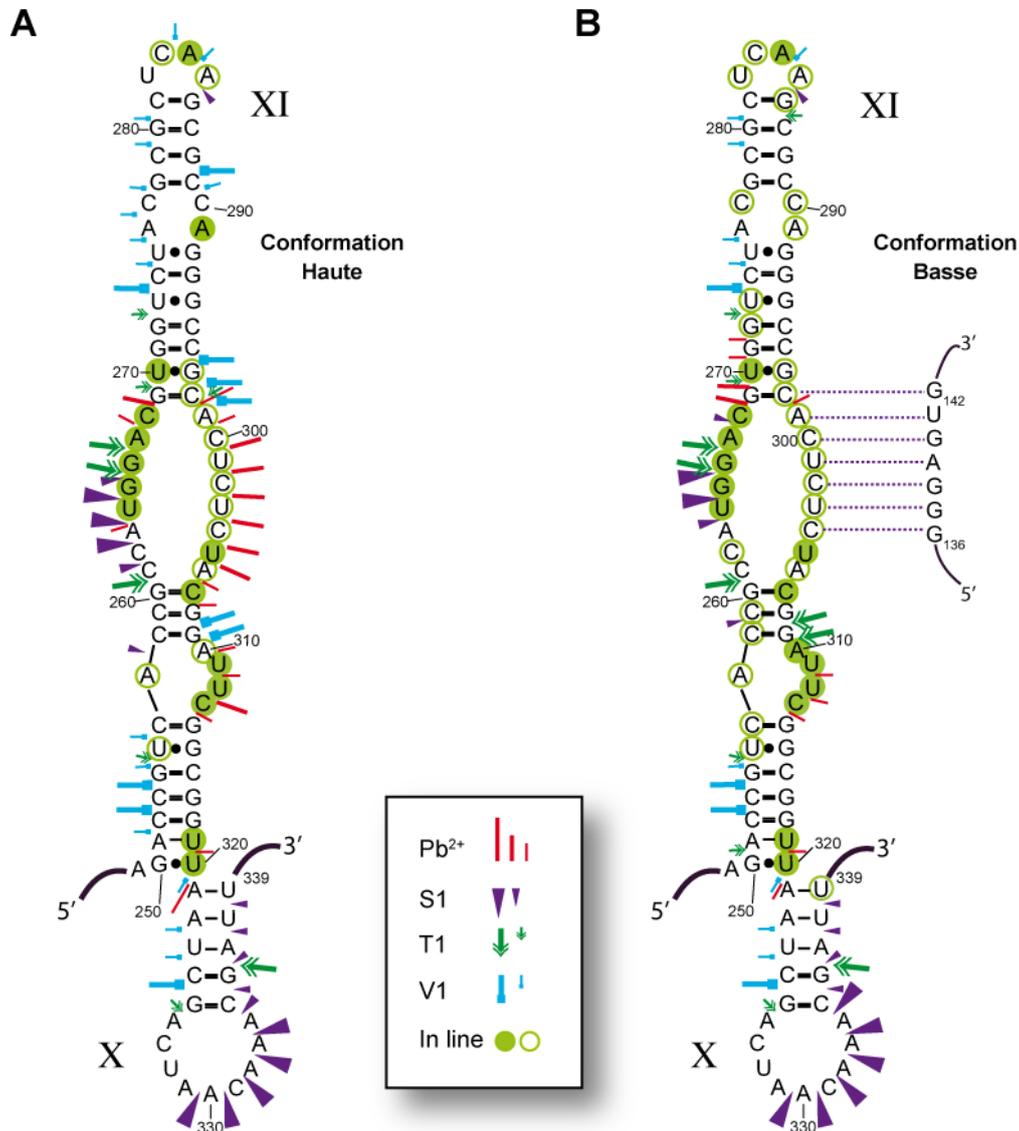


Figure 41. Modèle de structure secondaire de la région 249-339 de la forme haute (**A**) et basse (**B**). Les réactivités ont été reportées sur les modèles.

Il apparaît clairement que la tige boucle XI adopte des conformations ou établit des interactions différentes dans les formes H et B. De fortes réactivités à la RNase V1 sont observées dans les hélices de la forme haute ainsi que des fortes coupures aux ions Pb²⁺ dans la grande bulle interne aux environs de la position 300. L'absence de ces coupures au plomb dans la forme basse suggère fortement que cette région est appariée à d'autres nucléotides de l'ARNm. Cette interaction pourrait remodeler l'ensemble de la tige boucle XI et entraîner les différences de réactivité à la RNase V1 observés.

En examinant la séquence de l'ARNm H4, il apparaît qu'une interaction continue pourrait s'établir entre les nucléotides 298-304 et la région simple brin 136-142 qui sépare les tiges boucles IV et V. Ce pseudo nœud pourrait rapprocher ces deux parties de l'ARNm et expliquer les différences de réactivité observées. La forme basse qui est dépourvue de coupures aux ions Pb^{2+} à cet endroit présenterait le pseudo nœud ce qui pourrait expliquer la migration plus rapide sur gel (la structure étant plus compacte). Il aurait été intéressant de construire des mutants de ces deux séquences et d'observer leur comportement dans les protocoles de repliement. Nos résultats actuels suggèrent que ces mutants devraient perdre leur capacité à se replier en forme basse.

6.2 Modèle actuel de structure secondaire de l'ARNm H4 complet

Les différents modèles partiels ont finalement été assemblés en un modèle complet présenté dans la **figure 42**. Celui-ci diffère significativement des modèles précédents (**Figures 26 et 32**). Pour la première fois, ce modèle intègre des interactions tertiaires spécifiques aux formes H et B qui prennent en compte certaines différences de réactivité observées lors du sondage en solution ainsi que les résultats de repliement de l'ARNm. La conséquence principale de ce modèle est qu'il remet en cause le choix des fragments tronqués de l'ARNm H4. Le nouveau modèle suggère que certains fragments ne se terminent pas sur des hélices double brin comme initialement supposé et attendu, ce qui a pu être un facteur défavorable lors des expériences de cristallisation. Il conviendrait donc de renouveler les expériences de cristallisation en choisissant d'autres fragments d'ARN en accord avec la nouvelle structure secondaire.

Le présent modèle pourrait être affiné en poursuivant les expériences de sondage en solution en utilisant des sondes supplémentaires (DMS, CMCT, etc...) et en couplant la méthode de détection directe ici utilisée à une méthode de détection indirecte (par extension d'amorces par l'AMV reverse transcriptase) qui permet d'obtenir des données plus précises au centre de la molécule. Le modèle pourrait être validé en sondant des fragments tronqués représentatifs des domaines autonomes. L'existence des interactions tertiaires pourrait être vérifiée par mutagenèse dirigée suivie d'expériences de sondage en solution et d'études de repliement des ARNm mutés sur gel natif. Enfin, en l'absence de résultats cristallographiques, des investigations de type SAXS, déjà initiées durant ma thèse, pourraient être effectuées afin d'orienter la construction d'un modèle en trois dimensions.

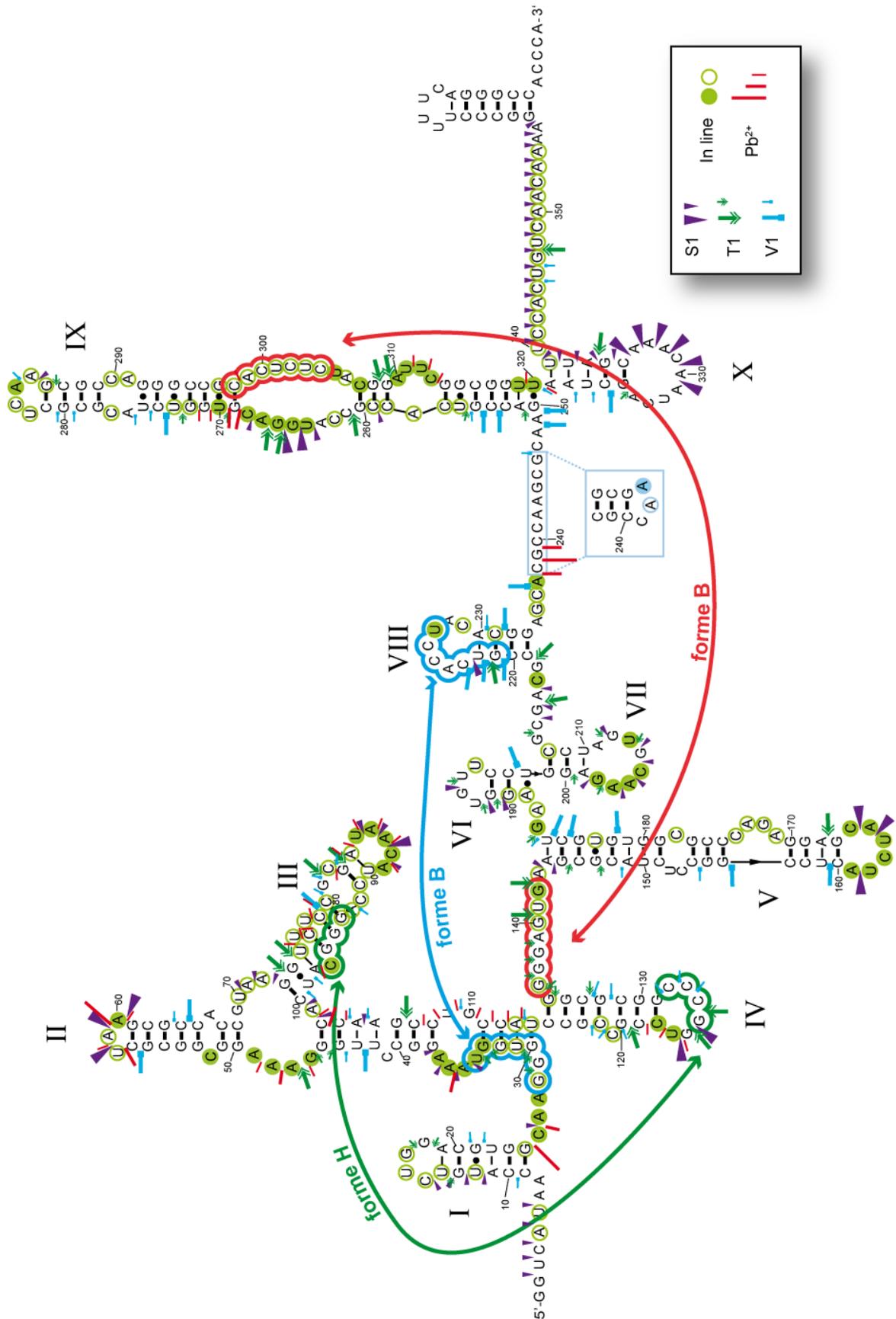


Figure 42. Modèle de structure secondaire de l'ARNm H4. Les réactivités de la forme basse ont été reportées sur le modèle. Les interactions présumées existant dans la forme haute (H) ou basse (B) sont indiquées.

Conclusion générale et perspectives

Les précédentes études portant sur la traduction de l'ARNm de l'histone H4 avaient montré la présence de structures dans l'ARNm participant au processus d'initiation en recrutant le facteur d'initiation eIF4E puis les ribosomes (Jaeger, 2005; Martin et al., 2011). En utilisant diverses méthodes d'investigation, il avait été montré que l'ARNm se repliait dans plusieurs conformations qui entraînaient un chevauchement de signaux contradictoires dans certaines régions de l'ARNm dont un premier modèle de repliement secondaire (2D) avait pu être proposé.

Après avoir amélioré les conditions de transcription de l'ARNm, la première tâche de mon travail de thèse a consisté à comprendre et à contrôler le repliement de l'ARNm afin d'obtenir une ou plusieurs formes repliées de manière homogène. Des dizaines de protocoles de repliement ont été testés, en jouant sur les paramètres de température, temps, ions divalents et force ionique. La recherche des conditions optimales de renaturation de l'ARNm H4 a conduit à la mise au point de deux protocoles originaux qui permettent d'isoler deux des trois conformations majeures observées : la conformation haute (H) et la conformation basse (B). Cette étude a permis de caractériser les paramètres ayant un impact sur le repliement de l'ARNm H4 et potentiellement sur d'autres ARNm. Au total, cinq paramètres se sont révélés importants : la concentration en ARN, la concentration en cations, la température d'ajout des cations, le type de cation utilisé et la cinétique de retour à 0°C. Jusqu'à présent, il n'avait pas été décrit que les trois premiers paramètres pouvaient avoir un impact sur le repliement des ARN. Il s'agit donc là d'une contribution originale qui pourra avoir un impact sur d'autres études structurales portant sur les ARN.

Une caractérisation fonctionnelle et structurale de ces deux formes de l'ARNm a ensuite été réalisée. Dans un premier temps, nous avons étudié la stabilité et la structure des deux formes d'ARNm en réalisant des expériences de diffusion dynamique de la lumière (DLS) de diffusion des rayons X aux petits angles (SAXS) et d'équilibre de sédimentation. Les tests de stabilité en fonction du temps et de la température ont permis de démontrer que les conformations H et B sont stables durant plusieurs jours à 20°C et pendant au moins 1 h à 4°C et 37°C. Les études structurales préliminaires réalisées par DLS et SAXS ont permis de définir le rayon hydrodynamique, le rayon de giration, la forme grossière de chaque conformation et de montrer que bien que la conformation H migre plus lentement que la conformation B sur gel ABA natif, elle semble être plus compacte que la conformation B. Puis les études par équilibre de sédimentation ont confirmé ces premiers résultats, en montrant que les deux conformations

H et B sont bien des monomères et que les différences de migration sur gel résultent bien de différences de repliement de leur structure et non de la formation de dimères.

Nous avons ensuite réalisé 4 tests destinés à observer d'éventuelles différences de fonction des deux repliements de l'ARNm H4. Dans un premiers temps, le test de fixation d'eIF4E et les expériences de pontage aux UV réalisées entre la coiffe et le reste de l'ARNm H4 n'ont pas révélé de différence majeure entre les deux formes d'ARNm H4. Puis les tests de traduction réalisés *in vitro* sur ces deux formes dans les extraits de germes de blé et de réticulocytes de lapin ont montré que ces deux formes sont traduites de manière équivalente ce qui suggérerait qu'elles étant reconnues de manière équivalente par la machinerie de traduction cellulaire. Étonnament, les expériences de « toe-print » ont montré que les ribosomes s'assemblent plus efficacement sur la forme H de l'ARNm. Il est donc probable qu'après l'étape d'assemblage des ribosomes, la structure de l'ARNm soit totalement dépliée par le passage des ribosomes en cours de traduction. Les ribosomes pourraient alors effacer les différences structurales existant entre les conformères d'ARN ce qui expliquerait qu'après plusieurs cycles de traduction, la même synthèse globale d'histone H4 est observée. Une autre possibilité serait que l'une ou les deux conformations aient un rôle biologique au cours de l'étape d'export dans le cytoplasme ou lors de l'étape de « processing ». De même, ces deux conformations peuvent refléter un mécanisme de régulation de la traduction de type « riboswitch » utilisé par la cellule pour s'assurer que la synthèse d'histone soit stœchiométrique. Il est également possible que ces deux conformations codent pour des produits différents ; en effet l'ARNm H4 est également utilisé par la cellule pour synthétiser dans certaines circonstances un petit peptide, l'OGP, dont la concentration peut être très élevée dans la moelle osseuse (« Osteogenic Growth Peptide ») (Bab et al., 1992; Bab et al., 1999; Smith et al., 2005).

Des expériences de sondage *in vitro* ont ensuite été réalisées à l'aide de sondes enzymatiques et chimiques. Les données obtenues indiquent que les deux formes d'ARNm présentent une majorité de signaux de coupure et de réactivité identique mais aussi des différences notables dans certaines régions. Malgré l'utilisation de 5 sondes différentes, il n'a pas été possible de déduire deux modèles de structure secondaire distincts et exclusifs. Il semble donc que les deux formes d'ARNm ne diffèrent que par des interactions à distance entre des régions éloignées de l'ARNm. Dans l'état actuel de nos investigations, plusieurs interactions à distance de type pseudo-nœud ont été prédites. Les données de sondage ont également permis de déterminer les zones flexibles sur ces deux conformations. Afin de progresser davantage dans la résolution des structures, il faudrait poursuivre les expériences de sondage en utilisant des sondes supplémentaires (DMS, CMCT, etc...) et en couplant la méthode de détection

directe à une méthode de détection indirecte (par extension d'amorces par la reverse transcriptase) afin d'obtenir des données plus précises au niveau du centre de la molécule. La connaissance exacte de la structure secondaire de chaque conformation permettra à terme de délimiter plus précisément les domaines fonctionnels et de synthétiser des fragments d'ARNm mieux structurés, moins flexibles et plus aptes à cristalliser. Évidemment, il serait intéressant de sonder la structure de l'ARNm H4 *in vivo* afin de détecter la présence de l'une ou l'autre de ces conformations et de confirmer leur existence dans la cellule.

En parallèle, des essais de cristallisation ont été réalisés sur ces formes de l'ARNm, mais malgré des tests de cristallisation nombreux, aucun cristal d'ARN n'a été obtenu. Nous avons donc synthétisé et purifié des fragments d'ARNm tronqués en 5' ou 3' ou des deux extrémités, afin d'isoler des domaines structurellement indépendants et homogènes d'un point de vue du repliement. Au total, 21 délétants ont été synthétisés et purifiés. Nous avons réussi à obtenir une population structurellement homogène en condition native pour 18 de ces 21 délétants mais malgré des tests de cristallisation nombreux, aucun cristal d'ARN n'a été obtenu. Pour poursuivre cette partie du projet, il conviendra de poursuivre les essais de cristallisation en utilisant de nouveaux fragments d'ARNm tronqués et utiliser l'approche SAXS pour générer des modèles de structures tridimensionnelles des fragments et de l'ARNm entier.

ANNEXE : Matériel et méthodes

A. Méthodes couramment utilisées lors de la manipulation des acides nucléiques

a. Solutions couramment utilisées

Tampon TAE :	40 mM Tris-Base 5 mM acétate de sodium 1 mM EDTA acide acétique qsp pH 7,9
Tampon TBE :	89 mM Tris-Base 89 mM acide borique 2,5 mM EDTA
Tampon EMSA :	50 mM glycine 50 mM Tris-Base
Tampon TE :	10 mM Tris-HCl pH 8 1 mM EDTA
TGS :	25 mM Tris-Base 192 mM glycine SDS 0,1% (p/v)
Formamide Dye :	formamide 95% (v/v) bleu de bromophénol 0,025% (p/v) xylène cyanol 0,025% (p/v)
Bleu de toluidine :	éthanol 40% (v/v) acide acétique 1% (v/v) bleu de toluidine 1 g/l

b. Extraction des protéines au phénol

L'extraction phénolique est une technique d'extraction liquide-liquide utilisée en biochimie pour extraire les protéines des solutions d'acides nucléiques (ARN ou ADN) (Kirby, 1956). A un volume de phénol (saturé en tampon TE) est ajouté un volume équivalent de la

solution aqueuse à extraire. Le mélange est ensuite agité vigoureusement à l'aide d'un vortex pendant une dizaine de secondes avant d'être centrifugé à température ambiante à $16\ 000 \times g$ pendant 2 min. Les protéines dénaturées forment alors un précipité blanc à l'interface des phases phénol et tampon TE, tandis que les acides nucléiques restent en solution dans la phase aqueuse supérieure. Après prélèvement de cette phase, une seconde extraction similaire, mais utilisant un mélange de phénol/chloroforme (v/v) saturé par du tampon TE est réalisée. Après centrifugation, la phase aqueuse est soumise à une précipitation à l'éthanol.

c. Précipitation à l'éthanol

Cette étape permet de concentrer les acides nucléiques et d'éliminer les substances solubles dans l'éthanol comme le phénol. Une solution d'acide nucléique est précipitée par addition de NaCl (0,25 M final) et de 3 volumes d'éthanol pur. Le tout est ensuite incubé à -20°C pendant une nuit. Une fois précipités, les acides nucléiques sont récupérés par centrifugation à $16\ 000 \times g$ pendant 10 min à 4°C . Le surnageant est aspiré, puis le culot est lavé par un volume équivalent d'éthanol 80% (v/v). Après une nouvelle centrifugation à $16\ 000 \times g$ pendant 5 min, le culot est séché sous vide pendant 10 min puis dissout dans de l' H_2O milliQ.

d. Détermination de la concentration de l'ARN en solution

La concentration de l'ARN en solution est déterminée par spectrophotométrie en mesurant l'absorbance à 260 nm de la solution, à l'aide d'un Nanodrop 1000 ou d'un Uvikon 943 selon la précision requise.

B. Synthèse et purification des ARN étudiés

a. Synthèse par amplification PCR d'une matrice à transcrire

Le gène de l'histone H4 et ses délétants ont été amplifiés par PCR afin d'obtenir une quantité suffisante de matrice ADN pour synthétiser les ARN correspondants par transcription *in vitro*, et, pour ajouter le promoteur de l'ARN polymérase du phage T7 en amont de ces matrices ADN. La PCR (Polymerase Chain Reaction) permet d'amplifier *in vitro* l'ADN à partir d'une matrice ADN, d'amorces oligonucléotidiques et d'une ADN polymérase thermostable par une succession de cycles de dénaturation, hybridation et élongation conduisant à une amplification exponentielle de l'ADN double brin (Mullis et al., 1986).

Toutes les réactions de PCR ont été réalisées à l'aide d'un thermocycleur « Veriti » (Applied Biosystems) à partir d'un plasmide pUC19 (Yanisch-Perron et al., 1985) contenant le gène H4-12 murin. Le milieu réactionnel d'amplification contenait : 0,2 ng/μL pUC19-H4, 0,3 mM MgCl₂, DMSO 3% (v/v), 5 mM TAPS-HCl pH 9,3, 10 mM KCl, 0,2 mM β mercaptoéthanol, 0,25 mM de chaque dNTP, 2 mM d'amorces (Sigma-Aldrich) et 20 U/mL de Phusion High Fidelity DNA polymerase (Finnzymes).

Le milieu PCR est soumis à 35 cycles d'amplification commençant par une étape de dénaturation de 15 sec à 95°C, suivie d'une étape d'hybridation de 15 sec à 55°C, puis d'une étape d'élongation de 1 min 30 à 72°C. En fin d'amplification, un gel de contrôle est réalisé. Des aliquotes de 5 μL sont alors prélevées, additionnées d'un volume de tampon de charge (bleu de bromophénol 0,7% (p/v), sarkosyl 0,1% (p/v), 25 mM EDTA, glycérol 20% (v/v)) et déposées sur gel d'agarose 1% (p/v). L'électrophorèse est réalisée à 50 V en présence de tampon TAE. La révélation est réalisée par coloration au bromure d'éthidium.

En fin de réaction de PCR, les milieux d'amplification sont soumis aux extractions au phénol avant d'être précipités à l'éthanol. Après centrifugation, les culots d'ADN sont lavés, séchés et dissous dans l'H₂O milliQ.

Amorces utilisées :

Les PCR ont été réalisées en utilisant une amorce sens contenant en amont un promoteur T7 de classe III (ATATTAATACGACTCACTATAG) (Coleman et al., 2004).

Au cours du processus de transcription *in vitro*, la T7 ARN polymérase incorpore de manière aléatoire 1 à 3 nucléotides sur l'extrémité 3' du transcrit, générant ainsi des extensions N+1, N+2 et/ou N+3 (Milligan et al., 1987). Afin d'obtenir des transcrits avec une extrémité 3' franche, nous avons synthétisé par PCR des matrices se terminant (du côté 3' du gène) par des nucléotides porteurs de groupements méthyles sur les derniers riboses (Kao et al., 1999), dont la particularité est d'entraîner un arrêt net de la T7 ARN polymérase sur le dernier nucléotide. Pour cela, nous avons commandé des amorces contenant au niveau des 2 premiers nucléotides en 5' un groupement méthyle sur 2'OH du ribose. Ces amorces ont été incorporées par PCR dans les matrices qui ont ensuite servi de substrat pour la T7 ARN polymérase.

T7-H4 : 5'-ATATTAATACGACTCACTATAGGTCATAACCATGTCTGGACG
 T7-H4 - 27 : 5'-ATATTAATACGACTCACTATAGGGTGGTAAAGGCCTTGG
 T7-H4 -115 : 5'-ATATTAATACGACTCACTATACCGCCGCTGGCCCCGGCGCGG
 T7-H4 -137 : 5'-ATATTAATACGACTCACTATAGGAGTGAAGCGCATCTCCGG
 T7-H4 -156 : 5'-ATATTAATACGACTCACTATAGCCTCATCTACGAGGAGACCCG

T7-H4 -179 : 5'-ATATTAATACGACTCACTATAGGTGTGCTGAAGGTGTTTCCTG
Rev-H4 : 5'-(2'O-méthyl rU)(2'O-méthyl rG)GGTGGCCCTGAAAAGGGC
Rev-H4 (1-375) : 5'-(2'O-méthyl rG)(2'O-méthyl rG)GTGGCCCTGAAAAGGGC
Rev-H4 (1-342) : 5'-GGAAATCGTTTGTAGTCGATTAAC
Rev-H4 (1-318) : 5'-CCGCCGAATCCGTAGAGAG
Rev-H4 (1-289) : 5'-GCGCTTGAGCGCGTAGACCAC
Rev-H4 (1-275) : 5'-AGACCACGTCCATGGCGGTG
Rev-H4 (1-241) : 5'-GGCGTGCTCCGTGTAGGTGAC
Rev-H4 (1-178) : 5'-GCGGGTCTCCTCGTAGATGAG
Rev-H4 (1-154) : 5'-GGAGATGCGCTTCACTCCCC
Rev-H4 (1-135) : 5'-CCGCGCCGGGCCAGGCGGCGG
Rev-H4 (1-134) : 5'-CGCGCCGGGCCAGGCGGCGG
Rev-H4 (1-133) : 5'-GCGCCGGGCCAGGCGGCGGATG
Rev-H4 (1-132) : 5'-CGCCGGGCCAGGCGGCGGATG
Rev-H4 (1-131) : 5'-GCAGGCTTGGTGATGCCCTG
Rev-H4 (1-130) : 5'-CCGGGCCAGGCGGCGGATGGC
Rev-H4 (1-111) : 5'-GCCGGGCCAGGCGGCGGATG

b. Transcription *in vitro* par la T7 ARN polymérase

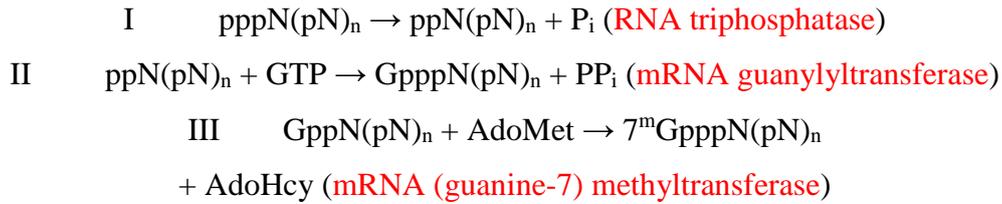
Pour synthétiser les transcrits correspondant aux ARNm étudiés, nous avons utilisé la T7 ARN polymérase. Toutes les réactions de transcription ont été réalisées dans des volumes de 500 µL du milieu suivant : 40 mM Tris-HCl pH 8,1, 100 µg/mL matrice ADN, 22 mM MgCl₂, 1 mM spermidine, 5 mM DTT, 0,01% Triton X-100 (v/v), 5 mM de chaque rNTP, 0,5 mg/mL T7 ARN polymérase, 0,4 U/µL RNasin (Promega) (Milligan et al., 1987). Après ajout de la T7 ARN polymérase, le milieu réactionnel est incubé à 37°C pendant 1 h. Lors de la réaction de transcription, des molécules de pyrophosphate sont générées. L'accumulation de ces molécules augmente le taux de transcription abortive et inhibe l'incorporation des rNTPs en formant un précipité blanc de pyrophosphate de magnésium (Kern and Davis, 1997). Pour libérer les ions Mg²⁺ indispensables à l'activité de l'ARN polymérase, la pyrophosphatase inorganique (Roche) est ajoutée (0,01 mg/mL final) après 1 h de transcription. Celle-ci hydrolyse le PPi (P₂O₇⁴⁻) en orthophosphate (2HPO₄²⁻) conduisant à la clarification du milieu de transcription. Après une nouvelle incubation de 30 min, la DNase I RNase free (Biolabs) est ajoutée (20 U/mL final) afin de digérer la matrice ADN et d'achever définitivement la réaction de transcription. La digestion se déroule pendant 1 h à 37°C puis la réaction est arrêtée par des extractions au phénol puis l'ARN est précipité. Le lendemain, les culots d'ARN sont lavés, séchés et repris dans de l'eau ou du Formamide Dye en cas de purification sur gel.

c. Purification des produits de transcription sur gel de polyacrylamide dénaturant

Afin de séparer, en fonction de leur taille et de leur charge nette, les différentes molécules présentes dans le milieu de transcription, les ARN ont été purifiés par électrophorèse sur gel de polyacrylamide 4% (p/v), urée 8 M, tampon TBE, de dimensions 33×44×0,2 cm. Deux poches de 13 cm de long permettent de déposer jusqu'à 16 mg d'ARN par gel, soit l'équivalent de 4 mL de transcription. Pour la séparation, les culots d'ARN sont préalablement dissous dans du Formamide Dye puis dénaturés par chauffage à 95°C durant 5 min. L'électrophorèse est réalisée à 60 W, en présence de tampon TBE, jusqu'à ce que le bleu de bromophénol arrive au bas du gel. Cette migration permet la séparation de l'ARN entier, des transcrits d'ARN abortifs et des nucléotides libres. Le gel est ensuite maintenu au dessus d'une plaque de silice afin de visualiser les ARN sous lumière UV (visualisation par ombrage). Puis, la bande correspondant à l'ARN d'intérêt est découpée et éluée par électrophorèse dans des cuves d'électroélution « Biotrap » (Schleicher & Schuell). Ces cuves sont munies de 2 membranes imperméables entre lesquelles est placée une membrane perméable aux ARN, formant ainsi 2 compartiments de capacités différentes dont la fonction est de séparer l'ARN des fragments de polyacrylamide. Le compartiment situé du côté de l'anode, d'une capacité de 11 mL, sert à placer la bande contenant l'ARN d'intérêt. Tandis que le compartiment situé du côté de la cathode, d'une capacité de 800 µL sert à récupérer l'ARN élué lors du processus d'électroélution. L'électroélution est réalisée à 120 V, en présence de tampon TBE, pendant 1 h. Les 800 µL d'ARN sont alors récupérés, puis remplacés par un volume équivalent et l'élution est répétée une deuxième fois dans les mêmes conditions. Les 2 fractions de 800 µL sont ensuite précipitées et le lendemain, les culots d'ARN sont lavés, séchés et repris dans l'H₂O milliQ.

d. Coiffage des transcrits purifiés

Les ARN étudiés ont été synthétisés par transcription *in vitro*, ils sont donc non coiffés et possèdent une extrémité 5' triphosphate. La coiffe étant importante dans le processus d'initiation de l'ARNm de l'histone H4, celui-ci devait être coiffé avant utilisation dans les tests fonctionnels. La coiffe d'un ARNm est constituée d'une 7-méthylguanosine reliée à l'extrémité 5' de l'ARNm par une liaison 5'-5' triphosphate (mG[5']ppp[5']NpN...) (Muthukrishnan et al., 1975). Cette structure est formée par une série de 3 réactions enzymatiques :



La réaction de coiffage est réalisée après purification de l'ARNm de l'histone H4 à l'aide d'un kit (ScriptCap™ m7G Capping System, Cambio) utilisant la VCE (Vaccinia Virus Capping Enzyme). Cette enzyme est un hétérodimère dont la particularité est de pouvoir catalyser successivement les 3 réactions nécessaires à la formation de la coiffe (Ensinger et al., 1975). Toutes les réactions de coiffage ont été réalisées à partir d'une solution d'ARN à la concentration finale de 0,5 µg/µL. L'ARN est préalablement replié par chauffage à 65°C durant 10 min, puis refroidissement à 0°C durant 5 min. Puis la solution d'ARN est diluée dans le milieu réactionnel final suivant : 50 mM Tris-HCl pH 8, 6 mM KCl, 1,5 mM MgCl₂, 1 mM GTP, 0,1 mM SAM (*S*-adénosyl-méthionine), 1 U/µL ScriptGuard RNase Inhibitor, 0,4 U/µL VCE. Après ajout de la VCE, le milieu réactionnel est incubé à 37°C pendant 1 h. La réaction est arrêtée par des extractions au phénol, puis l'ARN est précipité. Le lendemain, les culots d'ARN sont lavés, séchés et repris dans un tampon 10 mM Tris pH 7,5, 150 mM NaCl.

e. Ultrafiltration des produits de coiffage

L'ultrafiltration est une technique de séparation des molécules ou des ions contenus dans un liquide utilisant des membranes aux propriétés de filtration bien définies. Cette technique est utilisée après coiffage afin de réduire significativement la concentration des molécules et ions issus de cette réaction, en particulier le GTP et les ions Mg²⁺. Les culots d'ARN sont repris dans 200 µL de tampon de filtration (10 mM Tris pH 7,5, 150 mM NaCl) Puis, l'ARN est replié par traitement thermique (2 min à 95°C et 5 min à 0°C) puis dilué 150 fois dans le tampon de filtration avant d'être concentré par centrifugation jusqu'à un volume final de 200 µL à l'aide d'un Amicon® Ultra-15 (30K) (Merck millipore). La centrifugation est réalisée à 2 500 × g à 20°C. Ensuite, l'ARN est précipité, lavé, séché et repris dans de l'H₂O milliQ.

f. Elimination des agrégats par ultracentrifugation

Afin d'éliminer les agrégats d'ARN présents en solution, une étape d'ultracentrifugation est réalisée dans des microtubes de 1,5 mL à l'aide d'une micro-ultracentrifugeuse Sorvall Discovery M150 SE et dans un rotor S45A (Thermo Scientific). Les échantillons d'ARN sont centrifugés à 100 000 × g pendant 40 min à 20°C, puis les surnageants sont récupérés avec précaution.

C. Marquages au [³²P] des ARN étudiés et des amorces utilisées pour réaliser des transcriptions inverses

a. Marquage d'ARN au [³²P] en 5'

Le marquage en 5' de l'ARNm est réalisé en incorporant une guanine radioactive ([α -³²P]GTP, 3000 Ci/mmol, Amersham) à l'aide d'un kit de coiffage (ScriptCap™ m7G Capping System, Cambio) utilisant la VCE (Vaccinia Virus Capping Enzyme). La réaction est réalisée selon les mêmes modalités décrites plus haut mais en remplaçant le GTP non radioactif par du [α -³²P]GTP (0,25 μ Ci/ μ L final). La réaction est arrêtée par des extractions au phénol puis l'ARN est précipité. Le lendemain, les culots d'ARN obtenus après centrifugation sont lavés, séchés et repris dans du Formamide Dye.

Tous les transcrits marqués radioactivement sont ensuite purifiés par électrophorèse sur gel de polyacrylamide 4% (p/v), tampon TBE, 8 M urée. L'électrophorèse est réalisée à 60 W, en présence de tampon TBE, jusqu'à ce que le bleu de bromophénol soit à 10 cm du bas du gel. L'ARN marqué est ensuite localisé à l'aide d'un film autoradiographique (Fujifilm Life Sciences) développé dans une machine Kodak (M35-M X-Omat Processor). L'ARN marqué est ensuite découpé et élué par diffusion passive dans le tampon 10 mM Tris-HCl pH 7,5, 300 mM NaCl, 0,5 mM EDTA. L'élution est réalisée à 4°C pendant 14 h. L'ARN est ensuite précipité, lavé, séché et repris dans de l'H₂O milliQ. La quantité de radioactivité éluée est mesurée à l'aide d'un compteur à scintillation LS 6000 (Beckman).

b. Marquage d'ARN au [³²P] en 3'

Le marquage en 3' de l'ARNm est réalisé en incorporant une cytidine-3',5'-bis-phosphate radioactive ([5'-³²P]pCp, 3000 Ci/mmol, Amersham) à l'aide de la T4 RNA ligase (Biolabs). La fonction de la T4 RNA ligase est de catalyser la formation d'une liaison covalente entre l'extrémité 5' phosphate d'un ARN donneur et l'extrémité 3' hydroxyle d'un ARN accepteur (Kozak and Shatkin, 1978a). La T4 RNA ligase ne peut donc pas incorporer plus d'une molécule de [5'-³²P]pCp par molécule d'ARN.

Tous les marquages en 3' ont été réalisés à partir de 100 picomoles d'ARN incubées pendant 14 h, à 4°C, dans le milieu suivant : 50 mM Tris-HCl pH 7,8, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, DMSO 10% (v/v), 1,3 μ Ci/ μ L [5'-³²P]pCp, 0,65 U/ μ L T4 RNA ligase (Biolabs). La réaction est arrêtée par addition d'un volume de Formamide Dye. Le mélange est ensuite purifié sur gel de polyacrylamide 4% (p/v), tampon TBE, 8M urée selon les modalités décrites précédemment.

c. Marquage d'amorce au [³²P] en 5'

Le marquage en 5' des amorces est réalisé en incorporant une adénine radioactive ([γ -³²P]ATP, 3000 Ci/mmol, Amersham) à l'aide de la T4 polynucléotide kinase (T4 PNK) (Thermo Scientific). L'une des fonctions de la T4 PNK est de catalyser le transfert du phosphate γ de l'ATP sur une l'extrémité 5'-OH d'un ARN ou d'un ADN (Richardson, 1965).

Tous les marquages d'amorce ont été réalisés dans le milieu suivant : 1 μ M d'amorces (Sigma-Aldrich), 50 mM Tris-HCl pH7,6, 10 mM MgCl₂, 5 mM DTT, 0,2 mM spermidine, 1 μ Ci/ μ L [γ -³²P]ATP, 0,2 U/ μ L T4 PNK. Après ajout de T4 PNK, le milieu réactionnel est incubé à 37°C pendant 30 min. La réaction est ensuite arrêtée par chauffage à 65°C durant 10 min. Le mélange est ensuite purifié sur gel de polyacrylamide 4% (p/v), tampon TBE, 8M urée selon les modalités décrites précédemment.

D. Etude fonctionnelle de l'ARNm H4

a. Expériences de retard sur gel en présence d'eIF4E

Des tests de retard sur gel ont été réalisés sur différents conformères de l'ARNm de l'histone H4 (ARNm H4 ci-après) afin de détecter l'interaction avec eIF4E. Ce test consiste à comparer sur gel en conditions natives, le profil de migration de l'ARNm H4 isolé ou en complexe avec eIF4E. L'interaction ARN-protéine se traduit par un retard de la migration du complexe par rapport à l'ARN isolé.

Les complexes ARN-protéine sont formés en présence de 10 000 cpm d'ARNm H4 dans le milieu suivant : 10 mM Tris-HCl pH 7,5, 50 mM KCl, 1 mM DTT, glycérol 10% (v/v), 1 μ M eIF4E. Après l'ajout d'eIF4E, le mélange est incubé à 4°C pendant 20 min. Le mélange est ensuite fractionné sur gel de polyacrylamide natif 4% (p/v), tampon EMSA. L'électrophorèse est réalisée à 1,5 W, à 4°C, pendant 5 h en présence de tampon EMSA. Le gel est ensuite démoulé, transféré sur une feuille de papier Whatman 3MM, séché sous vide à 80°C, puis révélé par autoradiographie à l'aide d'une Imaging Plate (Fujifilm Life Sciences) et d'un lecteur Bioimager FLA5100 (Fujifilm Life Sciences).

b. Tests de traduction *in vitro*

Des tests de traduction *in vitro* ont été réalisés sur les différents conformères de l'ARNm H4, dans des lysats de réticulocytes de lapin et des extraits de germe de blé. L'efficacité de

traduction est mesurée en quantifiant la [³⁵S]-méthionine (40 µCi.ml⁻¹, 1000 Ci.mmol⁻¹, GE Healthcare) incorporée dans l'histone H4 synthétisée.

i. Traduction dans des lysats de réticulocytes de lapin

Nous avons utilisé des lysats de réticulocytes de lapin (RRL) traités à la nucléase micrococcalle (Promega). Les réticulocytes sont des globules rouges immatures, des cellules anucléées dépourvues de péroxysomes, qui possèdent un système de traduction efficace, fonctionnel et relativement proche du système utilisé par les cellules nucléées.

Chaque réaction de traduction a été réalisée à l'aide de 10 µL de RRL incubés dans 20 µL d'un milieu de composition finale 100 mM KCl, 0,5 mM MgCl₂, 75 µM en acides aminés (sauf méthionine), 0,125 µCi/µL [³⁵S]méthionine, 1 U/µL RNasin (Promega), 0,25 µM ARNm H4. Après ajout de l'ARNm H4, le milieu réactionnel est incubé à 30°C pendant 1 h. La réaction est arrêtée par addition d'un demi-volume de bleu dissociant : 40 mM Tris-HCl pH 8,8, bleu de bromophénol 0,05% (p/v), dodécylsulfate de sodium (SDS) 4% (p/v), β-mercaptoéthanol 4% (v/v), glycérol 30% (v/v).

Les produits de traduction sont ensuite séparés par électrophorèse sur un gel de polyacrylamide dénaturant subdivisé en une partie dite de concentration (polyacrylamide 4.5% (p/v), 125 mM Tris-HCl pH 6,8, SDS 0,1%(v/v)) et une partie dite de séparation (polyacrylamide 15% (p/v), 375 mM Tris-HCl pH 8,8, SDS 0,1%(v/v)). L'électrophorèse est réalisée à 200 V, pendant 35 min, en présence de tampon TGS. Puis le gel est démoulé et coloré pendant 40 min par une solution de bleu de Coomassie (éthanol 50% (v/v), acide acétique 10% (v/v), bleu de Coomassie 0,1% (p/v)). Le gel est ensuite décoloré par une solution d'acide acétique 7% (v/v), éthanol 5% (v/v). Afin que le gel ne craque pas lors du séchage, le gel est trempé durant 3 min dans une solution d'éthanol 40% (v/v), acide acétique 7,5% (v/v), glycérol 10% (v/v). Il est ensuite transféré sur une feuille de papier Whatman 3MM, séché sous vide à 65°C, puis révélé par autoradiographie à l'aide d'une Imaging Plate (Fujifilm Life Sciences) et d'un lecteur Bioimager FLA5100 (Fujifilm Life Sciences).

ii. Traduction dans des extraits de germe de blé

Les extraits de germe de blé (WGE) sont moins performants en terme d'efficacité de traduction que les RRL, mais ils présentent d'autres avantages (Van Herwynen and Beckler, 1995). En effet, ces extraits contiennent très peu d'ARN endogènes et ne nécessitent pas de traitement par une nucléase calcium dépendante comme la nucléase micrococcalle. Il en résulte

que les extraits sont davantage coiffe-dépendants et qu'ils reproduisent plus fidèlement les conditions de traduction cellulaire.

Chaque réaction de traduction *in vitro* contient dans un volume final de 20 μL : 10 μL de WGE, 100 mM acétate de K, 75 μM des 19 acides aminés (sans méthionine), 0,125 $\mu\text{Ci}/\mu\text{L}$ [^{35}S]méthionine, 1 U/ μL RNasin (Promega), 0,25 μM ARNm H4. Le milieu réactionnel est incubé à 25°C pendant 1 h. La réaction est arrêtée par addition d'un demi-volume de bleu dissociant. Les produits de traduction sont ensuite déposés sur gel de polyacrylamide et analysés comme décrit ci-dessus.

E. Etude structurale préliminaire de l'ARNm H4

a. Diffusion dynamique de la lumière

Dans le but de caractériser les différents conformères de l'ARNm-H4, des expériences de diffusion dynamique de la lumière (DLS) ont été réalisées afin de déterminer le rayon hydrodynamique de chaque conformation.

La DLS (Dynamic light scattering) est une technique d'analyse spectroscopique non destructive permettant de déterminer le coefficient de diffusion de translation D_t d'une particule en solution. Dans notre cas, ce coefficient D_t indique la facilité de déplacement d'une macromolécule au sein d'une solution aqueuse par le phénomène de diffusion. Pour les échantillons contenant une seule population de particules, le rayon hydrodynamique moyen de la particule (R_h) et la polydispersité de la population peuvent être calculés à partir de D_t . En effet, pour une solution de viscosité donnée et à température constante, le coefficient de diffusion D_t est inversement proportionnel à la taille de la particule, suivant l'équation de Stokes-Einstein :

$$D_t = \frac{K_b T}{6\pi\eta R_h}$$

ou K_b représente la constante de Boltzmann, T la température en degré kelvin, η la viscosité du solvant et R_h le rayon hydrodynamique de la particule. La polydispersité de la population est quant à elle donnée par l'écart type sur le coefficient de diffusion. Le principe de la DLS repose sur l'étude au cours du temps, des fluctuations d'intensité de lumière diffusées par des particules en suspension. Ces variations d'intensité lumineuse sont provoquées par le mouvement brownien des particules. Une des particularités de ce mouvement Brownien est que les petites particules bougent plus vite que les grosses. Ainsi, les changements d'intensités lumineuses les plus lents proviennent des particules les plus grandes et inversement. Lors d'une expérience de

DLS, un faisceau laser de couleur rouge irradie la solution étudiée et les variations au cours du temps de l'intensité de la lumière diffusée sont enregistrées. Dans le cas d'une solution de particules monodisperse, la fonction d'autocorrélation $G(\tau)$ de l'intensité de la lumière diffusée peut s'écrire sous la forme d'une exponentielle :

$$G(\tau) = 1 + b \exp(-2D_t q^2 \tau)$$

où b représente une constante dépendant de l'optique de l'instrument, D_t le coefficient de diffusion des particules et τ le temps de décroissance caractéristique. La norme du vecteur de diffusion q est obtenue à l'aide de l'équation suivante :

$$\|q\| = \frac{4\pi n_0}{\lambda_0} \sin\left(\frac{\theta}{2}\right)$$

où n_0 représente l'indice de réfraction du solvant, λ_0 la longueur d'onde du laser dans le vide et θ l'angle de diffusion de la lumière. La mesure expérimentale de $G(\tau)$ permet alors détermination du coefficient de diffusion D et donc de calculer le R_h de la particule suivant l'équation de Stokes-Einstein précédemment décrite.

Toutes les expériences de DLS ont été réalisées à 37°C, à partir d'une solution d'ARNm H4 concentrée à 1,5 mg/mL et à l'aide d'un Zetasizer Nano S (Malvern). Le faisceau incident est produit par un laser He-Ne (4 Milliwatts) fonctionnant à une longueur d'onde de 633 nm. La lumière diffusée est observée à 173° du faisceau incident. L'analyse des fonctions d'autocorrélation a été réalisée selon la méthode des cumulants à l'aide du logiciel DTS (Malvern Software).

b. Ultracentrifugation analytique

Afin de caractériser l'état d'oligomérisation des différents conformères de l'ARNm-H4, des expériences d'ultracentrifugation analytique ont été réalisées sur chaque conformation.

L'ultracentrifugation analytique est une technique d'analyse des macromolécules en solution permettant de déterminer le coefficient de sédimentation (s) d'une macromolécule. Le principe de cette technique repose sur la capacité des macromolécules à sédimenter lorsqu'elles sont centrifugées avec une grande vitesse angulaire. Au cours de ce processus de sédimentation, un front de macromolécules se forme et se déplace vers le fond de la cellule. La position de ce front en fonction du temps est déterminée par mesure de l'absorbance ou de l'indice de réfraction et permet de calculer le coefficient de sédimentation (s) de la macromolécule étudiée par la relation :

$$S = M \frac{(1 - \bar{v}\rho)}{N_A f}$$

où v représente le volume spécifique de la macromolécule, ρ la densité du solvant, N_A le nombre d'Avogadro, f le coefficient de friction et M la masse moléculaire.

Afin d'obtenir une mesure directe de la masse indépendante de l'estimation du rayon hydrodynamique de l'ARNm H4, des expériences d'équilibres de sédimentation ont été réalisées. Cette méthode consiste à centrifuger un échantillon avec une vitesse suffisamment élevée pour que la macromolécule étudiée forme un gradient de concentration, mais insuffisante pour que la macromolécule forme un culot. Après un certain temps, la vitesse de migration de la macromolécule est exactement compensée par sa diffusion, le gradient de concentration atteint alors un équilibre et la distribution de la macromolécule au sein de la cellule est analysée afin de déterminer sa masse. La masse est calculée à l'aide de l'équation suivante :

$$M = \frac{kT}{\omega^2 r c (1 - \bar{v}\rho)} \left(\frac{\partial c}{\partial r} \right)$$

où k représente la constante de Boltzman, T la température, r la distance à l'axe de rotation, c la concentration de la macromolécule et ω la vitesse angulaire. La masse obtenue à partir de cette technique possède une marge d'erreur généralement comprise entre 1 et 2%.

Toutes les expériences d'ultracentrifugation ont été réalisées à 4°C, à l'aide d'une ultracentrifugeuse Proteomelab XL-I (Beckman Coulter) et dans un rotor An-50 Ti (Beckman Coulter). Le déplacement des macromolécules était analysé par mesure de l'absorbance à 260 nm grâce à un système optique intégré.

c. Expériences de pontage aux UV

i. Préparation de l'ARNm H4 modifié

Afin de déterminer le rôle de la coiffe dans le repliement de l'ARNm H4 nous avons réalisé des expériences de pontage (cross-link) « zero length » à l'aide d'un groupement photoactivable (thio) placé en position 6 de la guanosine méthylée de la coiffe. Ce groupement est incorporé lors de l'étape de coiffage réalisée à l'aide d'un kit (ScriptCap™ m7G Capping System, Cambio) utilisant la VCE (Vaccinia Virus Capping Enzyme). L'incorporation est réalisée en remplaçant le GTP du kit par du 6-thio-GTP (Jena Bioscience). La réaction est arrêtée par des extractions au phénol puis l'ARN est précipité. Le lendemain, les culots d'ARN sont lavés, séchés et repris dans un tampon 10 mM Tris pH 7,5, 150 mM NaCl. L'ARNm H4 m⁷s⁶G coiffé est ensuite ultrafiltré selon les modalités décrites plus-haut.

ii. Pontage aux UV et détection des modifications par transcription inverse

Les expositions des ARN aux UV ont été réalisées dans des volumes de 4 µL, à 4°C et à l'aide d'un Stratalinker (Fischer Scientific) aux longueurs d'onde de 254 nm ou 365 nm et durant des temps variables en fonction des échantillons.

Après irradiation, les liaisons covalentes formées entre la coiffe et les nucléotides interagissant avec elle sont détectées par élongation d'amorce par la transcriptase inverse du virus de la myéloblastose aviaire (AMV-RT) (Life Science). En effet, les pontages covalents vont bloquer la progression de la transcriptase inverse lors du processus d'élongation. Cette élongation est réalisée à partir d'une amorce marquée au [³²P] sur son extrémité 5'.

Les réactions de transcription inverse ont été réalisées à partir d'une solution d'ARN (0,5 µM final) hybridée à 100 000 cpm d'amorce par un protocole de dénaturation/renaturation de 2 min à 95°C puis 5 min à 0°C. Le produit d'hybridation est ensuite ajouté au milieu suivant : 25 mM Tris-HCl pH 8,3, 50 mM KCl, 2 mM DTT, 5 mM MgCl₂, 2 mM dNTP, 2 U/µL RNasin (Promega), 0,15 U/µL AMV-RT (Life Science). Le milieu réactionnel est incubé à 37°C pendant 30 min. La réaction est arrêtée par précipitation à l'éthanol. Le lendemain, les culots d'ARN sont lavés, séchés et repris dans la solution de Formamide Dye.

iii. Séquençage de Sanger par transcription inverse

Afin de pouvoir identifier les nucléotides formant une liaison covalente avec la coiffe, un séquençage de l'ARN est réalisé selon la technique de Sanger (Sanger et al., 1977). Cette technique consiste à réaliser en parallèle 4 réactions d'élongation contenant chacune un didésoxynucléotide différent dont la concentration est ajustée de manière à arrêter l'élongation de l'amorce à chacune des bases correspondantes.

Toutes les réactions de séquençage par transcription inverse ont été réalisées à partir d'une solution d'ARN (0,13 µM finale) hybridée à 100 000 cpm d'amorce par chauffage à 95°C durant 2 min puis refroidissement à 0°C durant 5 min (dans 20 µL final). Le produit d'hybridation est ensuite fractionné en 4 et ajouté à 10 µL des milieux suivants :

-25 mM Tris-HCl pH 8,3, 50 mM KCl, 2 mM DTT, 5 mM MgCl₂, 0,03 U/µL AMV-RT (Life Science), 0,33 µM ddATP, 2,5 µM dATP, 10 µM dTTP/dCTT/dGTT.

- 25 mM Tris-HCl pH 8,3, 50 mM KCl, 2 mM DTT, 5 mM MgCl₂, 0,03 U/µL AMV-RT (Life Science), 0,33 µM ddTTP, 2,5 µM dTTP, 10 µM dATP/dCTT/dGTT.

- 25 mM Tris-HCl pH 8,3, 50 mM KCl, 2 mM DTT, 5 mM MgCl₂, 0,03 U/μL AMV-RT (Life Science), 0,33 μM ddGTP, 2,5 μM dGTP, 10 μM dTTP/dCTT/dATT.

- 25 mM Tris-HCl pH 8,3, 50 mM KCl, 2 mM DTT, 5 mM MgCl₂, 0,03 U/μL AMV-RT (Life Science), 0,33 μM ddCTP, 2,5 μM dCTP, 10 μM dTTP/dATT/dGTT.

Après ajout de l'AMV-RT, le milieu réactionnel est incubé à 37°C pendant 1 h. La réaction est arrêtée par précipitation à l'éthanol. Le lendemain, les culots d'ARN sont lavés, séchés et repris dans la solution de Formamide Dye.

iv. Analyse des résultats sur gel de polyacrylamide dénaturant

Les produits de transcription inverse sont ensuite séparés par électrophorèse sur gel de polyacrylamide 8% (p/v), 8 M urée, tampon TBE, de dimensions 33x44x0,03 cm. L'électrophorèse est réalisée à 60 W, en présence de tampon TBE. Après migration, le gel est trempé pendant 10 min dans une solution d'acide acétique 10% (v/v), éthanol 10% (v/v) puis transféré sur une feuille de papier Whatman 3MM, séché sous vide à 80°C, et révélé par autoradiographie à l'aide d'une Imaging Plate (Fujifilm Life Sciences) et d'un lecteur Bioimager FLA5100 (Fujifilm Life Sciences).

F. Etude du repliement de l'ARN en solution

La structure en solution des différents conformères de l'ARNm H4 a été étudiée à l'aide de différentes sondes enzymatiques et chimiques. Les sondes choisies permettent de différencier les zones simple brin de l'ARN et les régions repliées en double brin. De plus, l'analyse peut apporter des informations sur la dynamique des nucléotides et orienter le choix des fragments d'ARN à étudier par cristallographie.

Pour être interprétables, les coupures de l'ARN introduites par les sondes doivent être uniques et statistiques. La recherche et l'établissement des conditions de coupure est donc l'un des points clé de cette approche et nécessite une mise au point soignée

a. Sondes utilisées et conditions d'utilisation

i. Sondes enzymatiques

Les sondes enzymatiques ont une taille supérieure à 100 Å. Elles sont donc sensibles à l'encombrement stérique de l'ARN étudié et aux changements de conformation de l'ARN, mais elles ne permettent pas de réaliser une analyse structurale fine.

- RNase T1

La RNase T1 est une endoribonucléase extraite d'*Aspergillus orizae*, dont la fonction est de couper spécifiquement la liaison phosphodiester entre le 3'-P de la guanine et le 5'-OH du nucléotide adjacent, lorsque la guanine est non appariée (Takahashi et al., 1970). Cette enzyme est active dans un intervalle de pH compris entre 4,5 et 9. Elle peut être également utilisée en présence d'urée 7 M (Warrington, 1974).

Toutes les réactions de digestion avec la RNase T1 ont été réalisées dans le milieu suivant : 40 mM acétate de Na pH 4,5, 300 mM NaCl, 2 mM ZnSO₄, 10 000 cpm/μL et 0,5 μg/μL ARNm H4, 0,03 U/μL RNase T1 (Fermentas). Après ajout de la RNase, le milieu réactionnel est incubé à 20°C pendant 10 min. La réaction est ensuite arrêtée par addition d'un demi-volume de Formamide Dye et d'un volume d'une solution « Stop Mix » : 600 mM acétate de Na pH 6, 3 mM EDTA, 0,1 μg/μL ARNt (extrait total de levure).

- RNase V1

La RNase V1 est une endoribonucléase extraite du venin de cobra *Naja naja oxiana*, dont la particularité est de couper sans spécificité les bases des régions empilées ou doubles brins de tailles supérieures à 4 nucléotides (Lowman and Draper, 1986). Le produit de dégradation obtenu est un oligonucléotide avec une extrémité 5'-P. Cette enzyme est active en présence d'ions Mg²⁺ et dans un intervalle de pH compris entre 4 et 9.

Toutes les réactions de digestion avec la RNase V1 ont été réalisées dans le milieu suivant : 40 mM acétate de Na pH 4,5, 300 mM NaCl, 2 mM ZnSO₄, 10 000 cpm/μL et 0,5 μg/μL ARNm H4, 0,017 U/μL RNase V1 (Ambion). Après ajout de la RNase, le milieu réactionnel est incubé à 20°C pendant 20 min. La réaction est ensuite arrêtée par addition d'un demi-volume de Formamide Dye et d'un volume d'une solution « Stop Mix ».

- RNase S1

La nucléase S1 est une endonucléase extraite d'*Aspergillus orizae*, dont la fonction, à basse concentration, est de couper sans spécificité de base les régions ARN simple brin (Ando, 1966). Le produit de dégradation obtenu est un oligonucléotide avec une extrémité 5'-P. Cette enzyme est active en présence d'ions Zn²⁺ dans un intervalle de pH compris entre 4 et 5 (Vogt, 1973) ou en présence d'ions Mg²⁺ à pH 7.5 (Esteban et al., 1992).

Toutes les réactions de digestion avec la Nucléase S1 ont été réalisées dans le milieu suivant : 40 mM acétate de Na pH 4,5, 300 mM NaCl, 2 mM ZnSO₄, 10 000 cpm/μL et 0,5 μg/μL ARNm H4, 0,17 U/μL Nucléase S1 (Fermentas). Après ajout de la nucléase, le milieu

réactionnel est incubé à 20°C pendant 10 min. La réaction est ensuite arrêtée par addition d'un demi volume de Formamide Dye et d'un volume d'une solution « Stop Mix ».

ii. Sondage aux ions Pb²⁺

L'ion Pb²⁺ est une sonde de petite taille, inférieure à 10 Å, utile pour détecter les remaniements structuraux entre différents mutants ou entre différents conformères issus de la même molécule ARN. Au niveau de régions spécifiques, l'ion Pb²⁺ a la capacité de catalyser une réaction de transestérification générée par une attaque nucléophile du 2'-OH sur le groupement phosphate du nucléotide adjacent, créant une extrémité 2'3' phosphate cyclique et une extrémité 5' OH (Pan, 2001). Deux types de coupure ont été décrites dans la littérature : les coupures fortes au niveau de motifs structuraux stabilisés par un ion divalent (Brown et al., 1985; Rubin and Sundaralingam, 1983) et les coupures faibles au niveau de régions flexibles telles que les boucles (Ciesiolka et al., 1998; Gornicki et al., 1989).

Les réactions de coupure avec l'ion Pb²⁺ ont été réalisées à partir d'une solution d'ARN (6 000 cpm/μL et 0,3 μg/μL) incubée en présence de 0,5 mM d'acétate de plomb pendant 5 min et à 20°C (Pan, 2001). La réaction est arrêtée par addition d'un demi-volume d'EDTA (33 mM final) suivi d'une précipitation à l'éthanol. Le lendemain, les culots d'ARN sont lavés, séchés et repris dans du Formamide Dye.

iii. « In-line probing »

Cette technique exploite les propriétés auto-catalytiques de l'ARN, plus précisément sa capacité à s'autodégrader au niveau des régions peu structurées (Soukup and Breaker, 1999). Cet auto clivage est déclenché lorsque le groupement 2'-OH, le phosphate et l'oxygène en 5' partant adoptent une géométrie linéaire dite « In-Line ». Le 2'-OH va alors réaliser une attaque nucléophile sur le phosphate du nucléotide adjacent, libérant ainsi une extrémité 5'-OH et une extrémité 3' se terminant par un phosphate 2'-3' cyclique. Les nucléotides non appariés ayant potentiellement peu ou pas de contrainte géométrique, ceux-ci seront éventuellement libre d'adopter un large éventail de géométries incluant la géométrie « In-Line » (Regulski and Breaker, 2008). Cette méthode permet donc d'identifier les zones dynamiques à l'intérieur d'un ARN.

Toutes les réactions de « In-Line probing » ont été réalisées à partir d'une solution d'ARN (6 000 cpm/μL et 0,3 μg/μL) incubée pendant 40 h à 24°C, dans le tampon suivant : 50 mM Tris-HCl pH 8,3, 20 mM MgCl₂, 100 mM KCl (Regulski and Breaker, 2008). Un demi-volume de Formamide Dye est ensuite rajouté.

b. Réalisation d'échelles de coupure

- Réalisation d'une échelle d'hydrolyse alcaline

Pour identifier les fragments d'ARN clivés par les différentes sondes, des échelles de coupure de l'ARN sont effectuées. L'hydrolyse alcaline partielle de l'ARN permet d'obtenir une « échelle » complète de coupure permettant de visualiser tous les nucléotides de l'ARNm H4. Associée à l'hydrolyse dénaturante par la nucléase T1 (voir paragraphe suivant) qui permet de détecter les résidus G, il est généralement possible d'identifier toutes les coupures induites par les sondes.

Les réactions d'hydrolyse alcaline de l'ARNm H4 sont réalisées dans les conditions suivantes : 10 000 cpm/ μ L ARNm H4, 0,1 μ g/ μ L ARNt total, 40 mM NaCO₃/Na₂CO₃ pH 9.2. Le milieu réactionnel est incubé pendant 12 min à 95°C, puis directement placé sur la glace et dilué dans 1 demi-volume de Formamide Dye.

- Réalisation d'une échelle de coupure par la RNase T1

Cette échelle de coupure permet de détecter la plupart des résidus G contenus dans un ARN. Elle se déroule en conditions dénaturantes (4,4 M urée) à haute température. Malgré ces conditions, certains résidus G demeurent résistants à la coupure et n'apparaissent pas dans l'échelle.

« L'échelle T1 » est réalisée en présence de 12,5 mM citrate de sodium pH 4,5, 6000 cpm/ μ L ARNm H4, 0,25 μ g/ μ L ARNt (extrait total de levure), 4,4 M urée, 0,6 mM EDTA, bleu de bromophénol 0,06% (p/v), xylène cyanol 0,06% (p/v). Le milieu réactionnel est incubé à 50°C pendant 5 min, puis de la RNase T1 est ajoutée (0,125 U/ μ L) (Fermentas) et l'incubation est poursuivie durant 10 min à la même température. La réaction est stoppée en plaçant le tube sur la glace et en ajoutant un demi volume de Formamide Dye.

Les produits de digestion sont ensuite séparés par électrophorèse sur gel de polyacrylamide dénaturant avec des réticulations variant de 4 à 10% et des temps de migration variables afin de séparer les différentes tailles de fragments.

G. Essais de cristallisation

Afin de déterminer la structure de l'ARN H4 complet ou fragmentée, par la méthode de diffraction des rayons X sur des mono-cristaux, des essais de cristallisation ont été réalisés pour obtenir des cristaux d'ARN H4.

a. Principe

La cristallisation est un processus de transition de phase au cours duquel un composé passe d'un état en solution à un état ordonné solide, sous la forme d'un cristal. Un cristal est formé à partir d'une maille cristalline contenant un motif atomique ou moléculaire qui se répète par translations dans l'espace tridimensionnel de manière périodique ordonnée selon une loi mathématique décrite par un réseau ponctuel. La symétrie macroscopique des cristaux est décrite par 7 systèmes cristallins et 32 groupes ponctuels de symétrie. La symétrie microscopique est décrite par 230 groupes d'espace correspondant aux 14 réseaux de Bravais engendrés par les 7 systèmes cristallins.

La cristallisation peut-être schématiquement décomposée en 4 étapes: la formation d'un état de sursaturation, la nucléation, la croissance cristalline et l'arrêt de la croissance (McPherson, 2004). L'état de sursaturation, pré-requis pour l'obtention d'un cristal, est un état métastable où la concentration de la macromolécule est supérieure à sa solubilité. En théorie, cet état peut être atteint en faisant varier de nombreux paramètres tels que le pH, le tampon, la température, la concentration en ions et l'agent précipitant, etc... Bien que plusieurs milliers de macromolécules biologiques aient déjà été cristallisées, la recherche des conditions favorables est largement empirique et variable d'une molécule à l'autre, car il n'existe toujours pas de règle directement applicable à une nouvelle macromolécule. Un autre paramètre à prendre compte est la pureté de la macromolécule à cristalliser. En effet, les macromolécules biologiques possèdent divers sources potentielles d'hétérogénéité qui peuvent nuire fortement à leur capacité à cristalliser. Les sources d'hétérogénéité les plus courantes sont la présence de molécules contaminantes, de produits de dégradation, de différentes conformations et de domaines flexibles. La cristallisation est donc, un processus multiparamétrique complexe et un facteur limitant pour réaliser une étude structurale par diffraction des rayons X sur des monocristaux.

b. Techniques utilisées

Les tests de cristallisation ont été réalisés en utilisant la technique par diffusion de vapeur. Cette technique consiste à utiliser 2 solutions ayant des compositions et un volume différents :

- une solution de faible volume (une goutte de 100 nL à 10 µL), contenant la macromolécule, un tampon et un ou plusieurs agents dits « précipitants ». Le rôle d'un agent « précipitant » est de diminuer la solubilité de la macromolécule dans la goutte.
- une solution de grand volume par rapport au volume de la goutte (typiquement des

volumes de l'ordre de 35 μL à 1 mL), qui se comporte comme un réservoir au sens thermodynamique du terme. Ce « réservoir » contient un agent déshydratant dont la fonction est de diminuer la pression de vapeur d'eau au dessus du réservoir. Dans la pratique, on utilise le plus souvent la même espèce pour assurer le rôle d'agent « précipitant » et le rôle d'agent déshydratant.

Ces 2 solutions sont placées dans une enceinte fermée et étanche, sans les mettre en contact. L'objectif de cette technique est de déshydrater lentement la goutte, afin d'augmenter la concentration de l'agent précipitant et/ou de la macromolécule pour que celle-ci atteigne éventuellement un état de sursaturation nécessaire à la nucléation et puis à la formation de l'état cristallin. Dans notre cas, les gouttes ont été créées par ajout d'un volume du réservoir à un volume de la solution de macromolécule, afin que la pression de vapeur d'eau au-dessus de la goutte devienne supérieure à celle au-dessus du réservoir. Cette différence de tension de vapeurs, va induire un transfert d'eau de la goutte vers le réservoir jusqu'à ce que les 2 pressions de vapeur s'équilibrent. Le dispositif expérimental utilisé variait selon l'état d'avancement des essais. Lors des étapes de criblage, un montage expérimental dit de « goutte assise » a été utilisé, tandis que lors des étapes d'optimisation, un montage expérimental dit de « goutte suspendue » a été utilisé.

c. Matrices et approches utilisées

Afin de déterminer les conditions générales pour lesquelles la macromolécule étudiée atteint un état de sursaturation, une première étape dite de criblage est réalisée. Durant cette étape, la macromolécule est exposée à différentes solutions, issues de matrices commerciales, dont les compositions ont été choisies de façon à couvrir une large gamme de tampons, de pH, de sels et d'agent précipitant. La composition de ces solutions peut-être également issue d'études statistiques basées sur les conditions qui ont permis de générer des cristaux de macromolécules. Les matrices contenant peu de conditions avec un pH basique et peu de conditions avec une forte concentration en ion divalent (supérieur ou égale à 50 mM) ont été favorisées. Concrètement, lors des essais de cristallisation, les matrices suivantes ont été utilisées : The MPD (Qiagen), The Ammonium Sulfates (Qiagen), The Classics (Qiagen), The PEGs (Qiagen), The, Nucleix (Qiagen) et Index (Hampton Research).

Tous les criblages ont été réalisés sur des microplaques CrysAlQuick 96 (Greiner). La station de pipetage Miniprep 75 (Tecan) a été utilisée pour remplir de manière automatisée les 96 réservoirs d'une microplaque avec les différentes solutions de cristallisation, tandis que les robots de cristallisation HoneyBee 8+1 (Cartesian) et Mosquito (ttrplabtech) ont été utilisés pour

remplir de manière automatisée les 96 puits d'une microplaque avec la macromolécule et les différentes solutions de cristallisation. Les microplaques ont ensuite été fermées hermétiquement par un film adhésif transparent VIEWseal™(Greiner). Une partie des gouttes a ensuite été automatiquement et régulièrement photographiée à l'aide d'un robot Rock Imager (Formulatrix). Dès que des conditions prometteuses étaient identifiées, l'optimisation de ces conditions était réalisée.

Pour optimiser des conditions à priori prometteuses, 2 méthodologies ont été utilisées :

- la première consistait à modifier un par un chaque paramètre de la condition initiale. Cette étape était réalisée sur des plaques VDX 24 puits (Hampton Research) en réalisant un double gradient de sels et d'agent précipitant.

- la deuxième méthode consistait à réaliser un criblage sur la condition initiale à l'aide d'une matrice commerciale utilisée comme une solution stock (10X). Ces essais de cristallisation étaient réalisés selon les mêmes modalités qu'une étape de criblage initiale.

Bibliographie

- Acker, M.G., and Lorsch, J.R. (2008). Mechanism of ribosomal subunit joining during eukaryotic translation initiation. *Biochem Soc Trans* 36, 653-657.
- Algire, M.A., Maag, D., and Lorsch, J.R. (2005). Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol Cell* 20, 251-262.
- Allfrey, V.G., and Mirsky, A.E. (1964). Structural Modifications of Histones and their Possible Role in the Regulation of RNA Synthesis. *Science* 144, 559.
- Ando, T. (1966). A nuclease specific for heat-denatured DNA in isolated from a product of *Aspergillus oryzae*. *Biochim Biophys Acta* 114, 158-168.
- Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E., and Moudrianakis, E.N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A* 88, 10148-10152.
- Asano, K., Shalev, A., Phan, L., Nielsen, K., Clayton, J., Valasek, L., Donahue, T.F., and Hinnebusch, A.G. (2001). Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. *EMBO J* 20, 2326-2337.
- Bab, I., Gazit, D., Chorev, M., Muhlrud, A., Shteyer, A., Greenberg, Z., Namdar, M., and Kahn, A. (1992). Histone H4-related osteogenic growth peptide (OGP): a novel circulating stimulator of osteoblastic activity. *EMBO J* 11, 1867-1873.
- Bab, I., Smith, E., Gavish, H., Attar-Namdar, M., Chorev, M., Chen, Y.C., Muhlrud, A., Birnbaum, M.J., Stein, G., and Frenkel, B. (1999). Biosynthesis of osteogenic growth peptide via alternative translational initiation at AUG85 of histone H4 mRNA. *J Biol Chem* 274, 14474-14481.
- Banerjee, A.K. (1980). 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol Rev* 44, 175-205.
- Bedford, M.T., and Richard, S. (2005). Arginine methylation an emerging regulator of protein function. *Mol Cell* 18, 263-272.
- Belsham, G.J. (1992). Dual initiation sites of protein synthesis on foot-and-mouth disease virus RNA are selected following internal entry and scanning of ribosomes in vivo. *EMBO J* 11, 1105-1110.
- Bostrom, K., Wettsten, M., Boren, J., Bondjers, G., Wiklund, O., and Olofsson, S.O. (1986). Pulse-chase studies of the synthesis and intracellular transport of apolipoprotein B-100 in Hep G2 cells. *J Biol Chem* 261, 13800-13806.
- Brown, R.S., Dewan, J.C., and Klug, A. (1985). Crystallographic and biochemical investigation of the lead(II)-catalyzed hydrolysis of yeast phenylalanine tRNA. *Biochemistry* 24, 4785-4801.

- Ciesiolka, J., Michalowski, D., Wrzesinski, J., Krajewski, J., and Krzyzosiak, W.J. (1998). Patterns of cleavages induced by lead ions in defined RNA secondary structure motifs. *J Mol Biol* 275, 211-220.
- Coleman, T.M., Wang, G., and Huang, F. (2004). Superior 5' homogeneity of RNA from ATP-initiated transcription under the T7 phi 2.5 promoter. *Nucleic Acids Res* 32, e14.
- Culjkovic, B., Topisirovic, I., and Borden, K.L. (2007). Controlling gene expression through RNA regulons: the role of the eukaryotic translation initiation factor eIF4E. *Cell Cycle* 6, 65-69.
- Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M., and Borden, K.L. (2006). eIF4E is a central node of an RNA regulon that governs cellular proliferation. *J Cell Biol* 175, 415-426.
- Davies, M.V., Pelletier, J., Meerovitch, K., Sonenberg, N., and Kaufman, R.J. (1991). The effect of poliovirus proteinase 2Apro expression on cellular metabolism. Inhibition of DNA replication, RNA polymerase II transcription, and translation. *J Biol Chem* 266, 14714-14720.
- Dever, T.E. (2002). Gene-specific regulation by general translation factors. *Cell* 108, 545-556.
- Dominski, Z., and Marzluff, W.F. (2007). Formation of the 3' end of histone mRNA: getting closer to the end. *Gene* 396, 373-390.
- Doudna, J.A., Grosshans, C., Gooding, A., and Kundrot, C.E. (1993). Crystallization of ribozymes and small RNA motifs by a sparse matrix approach. *Proc Natl Acad Sci U S A* 90, 7829-7833.
- Ederly, I., and Sonenberg, N. (1985). Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc Natl Acad Sci U S A* 82, 7590-7594.
- Ensinger, M.J., Martin, S.A., Paoletti, E., and Moss, B. (1975). Modification of the 5'-terminus of mRNA by soluble guanylyl and methyl transferases from vaccinia virus. *Proc Natl Acad Sci U S A* 72, 2525-2529.
- Esteban, J.A., Salas, M., and Blanco, L. (1992). Activation of S1 nuclease at neutral pH. *Nucleic Acids Res* 20, 4932.
- Etchison, D., Milburn, S.C., Ederly, I., Sonenberg, N., and Hershey, J.W. (1982). Inhibition of HeLa cell protein synthesis following poliovirus infection correlates with the proteolysis of a 220,000-dalton polypeptide associated with eucaryotic initiation factor 3 and a cap binding protein complex. *J Biol Chem* 257, 14806-14810.
- Gamarnik, A.V., and Andino, R. (1996). Replication of poliovirus in *Xenopus* oocytes requires two human factors. *EMBO J* 15, 5988-5998.
- 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.

- Gingras, A.C., Svitkin, Y., Belsham, G.J., Pause, A., and Sonenberg, N. (1996). Activation of the translational suppressor 4E-BP1 following infection with encephalomyocarditis virus and poliovirus. *Proc Natl Acad Sci U S A* *93*, 5578-5583.
- Gornicki, P., Baudin, F., Romby, P., Wiewiorowski, M., Kryzosiak, W., Ebel, J.P., Ehresmann, C., and Ehresmann, B. (1989). Use of lead(II) to probe the structure of large RNA's. Conformation of the 3' terminal domain of E. coli 16S rRNA and its involvement in building the tRNA binding sites. *J Biomol Struct Dyn* *6*, 971-984.
- 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.
- Gray, N.K., and Hentze, M.W. (1994). Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs. *EMBO J* *13*, 3882-3891.
- Gunjan, A., Paik, J., and Verreault, A. (2005). Regulation of histone synthesis and nucleosome assembly. *Biochimie* *87*, 625-635.
- He, Y., Yan, W., Coito, C., Li, Y., Gale, M., Jr., and Katze, M.G. (2003). The regulation of hepatitis C virus (HCV) internal ribosome-entry site-mediated translation by HCV replicons and nonstructural proteins. *J Gen Virol* *84*, 535-543.
- Hellen, C.U., Witherell, G.W., Schmid, M., Shin, S.H., Pestova, T.V., Gil, A., and Wimmer, E. (1993). A cytoplasmic 57-kDa protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. *Proc Natl Acad Sci U S A* *90*, 7642-7646.
- Henriet, S., Sinck, L., Bec, G., Gorelick, R.J., Marquet, R., and Paillart, J.C. (2007). Vif is a RNA chaperone that could temporally regulate RNA dimerization and the early steps of HIV-1 reverse transcription. *Nucleic Acids Res* *35*, 5141-5153.
- Hentze, M.W. (1997). eIF4G: a multipurpose ribosome adapter? *Science* *275*, 500-501.
- Huang, Y., and Carmichael, G.G. (1997). The mouse histone H2a gene contains a small element that facilitates cytoplasmic accumulation of intronless gene transcripts and of unspliced HIV-1-related mRNAs. *Proc Natl Acad Sci U S A* *94*, 10104-10109.
- Hunt, S.L., and Jackson, R.J. (1999). Polypyrimidine-tract binding protein (PTB) is necessary, but not sufficient, for efficient internal initiation of translation of human rhinovirus-2 RNA. *RNA* *5*, 344-359.
- Huynen, M., Konings, D., 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.
- Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* *147*, 789-802.

- Jackson, R.J., Hellen, C.U., 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., Howell, M.T., and Kaminski, A. (1990). The novel mechanism of initiation of picornavirus RNA translation. *Trends Biochem Sci* *15*, 477-483.
- Jackson, R.J., and Kaminski, A. (1995). Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. *RNA* *1*, 985-1000.
- Jaeger, S. (2005). Etude de la maturation de l'extrémité 3' non traduite et de la traduction de l'ARN messager codant pour l'histone H4. (Strasbourg, Université de Strasbourg).
- Jaeger, S., Barends, S., Giege, R., Eriani, G., and Martin, F. (2005). Expression of metazoan replication-dependent histone genes. *Biochimie* *87*, 827-834.
- Jang, S.K., Krausslich, 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.
- 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 U S A* *101*, 16990-16995.
- Johannes, G., Carter, M.S., Eisen, M.B., Brown, P.O., and Sarnow, P. (1999). Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. *Proc Natl Acad Sci U S A* *96*, 13118-13123.
- Kaminski, A., Howell, M.T., and Jackson, R.J. (1990). Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning mechanism. *EMBO J* *9*, 3753-3759.
- Kao, C., Zheng, M., and Rudisser, S. (1999). A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA* *5*, 1268-1272.
- Kern, J.A., and Davis, R.H. (1997). Application of solution equilibrium analysis to in vitro RNA transcription. *Biotechnol Prog* *13*, 747-756.
- Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. *Cell* *116*, 259-272.
- Kieft, J.S., Grech, A., Adams, P., and Doudna, J.A. (2001a). Mechanisms of internal ribosome entry in translation initiation. *Cold Spring Harb Symp Quant Biol* *66*, 277-283.
- Kieft, J.S., Zhou, K., Jubin, R., and Doudna, J.A. (2001b). Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* *7*, 194-206.
- Kirby, K.S. (1956). A new method for the isolation of ribonucleic acids from mammalian tissues. *Biochem J* *64*, 405-408.

- Kodama, Y., Rothman, J.H., Sugimoto, A., and Yamamoto, M. (2002). The stem-loop binding protein CDL-1 is required for chromosome condensation, progression of cell death and morphogenesis in *Caenorhabditis elegans*. *Development* *129*, 187-196.
- Kolupaeva, V.G., Pestova, T.V., and Hellen, C.U. (2000). Ribosomal binding to the internal ribosomal entry site of classical swine fever virus. *RNA* *6*, 1791-1807.
- Kolupaeva, V.G., Pestova, T.V., Hellen, C.U., and Shatsky, I.N. (1998). Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. *J Biol Chem* *273*, 18599-18604.
- Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* *184*, 868-871.
- Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. *Science* *184*, 865-868.
- Kozak, M. (1989). The scanning model for translation: an update. *J Cell Biol* *108*, 229-241.
- Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem* *266*, 19867-19870.
- Kozak, M., and Shatkin, A.J. (1978a). Identification of features in 5' terminal fragments from reovirus mRNA which are important for ribosome binding. *Cell* *13*, 201-212.
- Kozak, M., and Shatkin, A.J. (1978b). Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. *J Biol Chem* *253*, 6568-6577.
- 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. Implications for cap-dependent and cap-independent translational initiation. *J Biol Chem* *270*, 21975-21983.
- Lee, J.H., Pestova, T.V., Shin, B.S., Cao, C., Choi, S.K., and Dever, T.E. (2002). Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. *Proc Natl Acad Sci U S A* *99*, 16689-16694.
- 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.
- Lorsch, J.R., and Herschlag, D. (1999). Kinetic dissection of fundamental processes of eukaryotic translation initiation in vitro. *EMBO J* *18*, 6705-6717.
- Lowman, H.B., and Draper, D.E. (1986). On the recognition of helical RNA by cobra venom V1 nuclease. *J Biol Chem* *261*, 5396-5403.
- Luger, K., Mader, 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.
- Macejak, D.G., and Sarnow, P. (1991). Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature* *353*, 90-94.

- Maiti, T., Das, S., and Maitra, U. (2000). Isolation and functional characterization of a temperature-sensitive mutant of the yeast *Saccharomyces cerevisiae* in translation initiation factor eIF5: an eIF5-dependent cell-free translation system. *Gene* 244, 109-118.
- Majumdar, R., Bandyopadhyay, A., and Maitra, U. (2003). Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. *J Biol Chem* 278, 6580-6587.
- Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1997). Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* 89, 951-961.
- Martin, F. (2012). Exploring the molecular mechanisms required for S-phase expression of metazoan replication-dependent histone genes (Strasbourg, Université de Strasbourg).
- 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., 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.
- 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.
- McPherson, A. (2004). Introduction to protein crystallization. *Methods* 34, 254-265.
- Meeks-Wagner, D., and Hartwell, L. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. *Cell* 44, 43-52.
- Meerovitch, K., Pelletier, J., and Sonenberg, N. (1989). A cellular protein that binds to the 5'-noncoding region of poliovirus RNA: implications for internal translation initiation. *Genes Dev* 3, 1026-1034.
- Meerovitch, K., Svitkin, Y.V., Lee, H.S., Lejbkowitz, F., Kenan, D.J., Chan, E.K., Agol, V.I., Keene, J.D., and Sonenberg, N. (1993). La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *J Virol* 67, 3798-3807.
- Merrick, W.C. (1979). Evidence that a single GTP is used in the formation of 80 S initiation complexes. *J Biol Chem* 254, 3708-3711.
- Merrick, W.C., and Hershey, J.W.B. (1996). The pathway and mechanism of eukaryotic protein synthesis. In *In Translational Control*, J.W.B. Hershey, M.B. Mathews, and N. Sonenberg, eds. (Cold Spring Harbor Laboratory Press).
- Milligan, J.F., Groebe, D.R., Witherell, G.W., and Uhlenbeck, O.C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res* 15, 8783-8798.

- Mitchell, S.F., and Lorsch, J.R. (2008). Should I stay or should I go? Eukaryotic translation initiation factors 1 and 1A control start codon recognition. *J Biol Chem* 283, 27345-27349.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* 51 Pt 1, 263-273.
- Muthukrishnan, S., Both, G.W., Furuichi, Y., and Shatkin, A.J. (1975). 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature* 255, 33-37.
- Niepmann, M. (2009). Internal translation initiation of picornaviruses and hepatitis C virus. *Biochim Biophys Acta* 1789, 529-541.
- Oh, S.K., Scott, M.P., and Sarnow, P. (1992). Homeotic gene Antennapedia mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding. *Genes Dev* 6, 1643-1653.
- 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.
- Pahlich, S., Zakaryan, R.P., and Gehring, H. (2006). Protein arginine methylation: Cellular functions and methods of analysis. *Biochim Biophys Acta* 1764, 1890-1903.
- Pan, T. (2001). Probing RNA structure by lead cleavage. *Curr Protoc Nucleic Acid Chem Chapter 6*, Unit 6.3.
- Passmore, L.A., Schmeing, T.M., Maag, D., Applefield, D.J., Acker, M.G., Algire, M.A., Lorsch, J.R., and Ramakrishnan, V. (2007). The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell* 26, 41-50.
- Pause, A., Belsham, G.J., Gingras, A.C., Donze, O., Lin, T.A., Lawrence, J.C., Jr., 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.
- Pestova, T.V., Borukhov, S.I., and Hellen, C.U. (1998a). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* 394, 854-859.
- Pestova, T.V., and Hellen, C.U. (2000). The structure and function of initiation factors in eukaryotic protein synthesis. *Cell Mol Life Sci* 57, 651-674.
- Pestova, T.V., Hellen, C.U., and Shatsky, I.N. (1996a). Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol Cell Biol* 16, 6859-6869.

- Pestova, T.V., Hellen, C.U., and Wimmer, E. (1991). Translation of poliovirus RNA: role of an essential cis-acting oligopyrimidine element within the 5' nontranslated region and involvement of a cellular 57-kilodalton protein. *J Virol* *65*, 6194-6204.
- 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. (1998b). A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev* *12*, 67-83.
- Pestova, T.V., Shatsky, I.N., and Hellen, C.U. (1996b). Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol Cell Biol* *16*, 6870-6878.
- Pettitt, J., Crombie, C., Schümperli, D., and Müller, B. (2002). The *Caenorhabditis elegans* histone hairpin-binding protein is required for core histone gene expression and is essential for embryonic and postembryonic cell division. *J Cell Sci*, 857-866.
- Piontkivska, H., Rooney, A., and Nei, M. (2002). Purifying selection and birth-and-death evolution in the histone H4 gene family. *Mol Biol Evol* *19*, 689-697.
- Prongidi-Fix, L., Schaeffer, L., Simonetti, A., Barends, S., Menetret, J.F., Klaholz, B.P., Eriani, G., and Martin, F. (2013). Rapid purification of ribosomal particles assembled on histone H4 mRNA: a new method based on mRNA-DNA chimaeras. *Biochem J* *449*, 719-728.
- Pyronnet, S., and Sonenberg, N. (2001). Cell-cycle-dependent translational control. *Curr Opin Genet Dev* *11*, 13-18.
- Regulski, E.E., and Breaker, R.R. (2008). In-line probing analysis of riboswitches. *Methods Mol Biol* *419*, 53-67.
- 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.
- Richardson, C.C. (1965). Phosphorylation of nucleic acid by an enzyme from T4 bacteriophage-infected *Escherichia coli*. *Proc Natl Acad Sci U S A* *54*, 158-165.
- Rozen, F., Edery, I., Meerovitch, K., Dever, T.E., Merrick, W.C., and Sonenberg, N. (1990). Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Mol Cell Biol* *10*, 1134-1144.
- Rubin, J.R., and Sundaralingam, M. (1983). Lead ion binding and RNA chain hydrolysis in phenylalanine tRNA. *J Biomol Struct Dyn* *1*, 639-646.
- Salzler, H.R., Davidson, J.M., Montgomery, N.D., and Duronio, R.J. (2009). Loss of the histone pre-mRNA processing factor stem-loop binding protein in *Drosophila* causes genomic instability and impaired cellular proliferation. *PLoS One* *4*, e8168.

- Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74, 5463-5467.
- Smith, E., Meyerrose, T.E., Kohler, T., Namdar-Attar, M., Bab, N., Lahat, O., Noh, T., Li, J., Karaman, M.W., Hacia, J.G., *et al.* (2005). Leaky ribosomal scanning in mammalian genomes: significance of histone H4 alternative translation in vivo. *Nucleic Acids Res* 33, 1298-1308.
- 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., Morgan, M.A., Merrick, W.C., and Shatkin, A.J. (1978). A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proc Natl Acad Sci U S A* 75, 4843-4847.
- Soukup, G.A., and Breaker, R.R. (1999). Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* 5, 1308-1325.
- Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41-45.
- Sullivan, E., Santiago, C., Parker, E.D., Dominski, Z., Yang, X., Lanzotti, D.J., Ingledue, T.C., Marzluff, W.F., and Duronio, R.J. (2001). Drosophila stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. *Genes Dev* 15, 173-187.
- Svitkin, Y.V., Gradi, A., Imataka, H., Morino, S., and Sonenberg, N. (1999). Eukaryotic initiation factor 4GII (eIF4GII), but not eIF4GI, cleavage correlates with inhibition of host cell protein synthesis after human rhinovirus infection. *J Virol* 73, 3467-3472.
- Svitkin, Y.V., Meerovitch, K., Lee, H.S., Dholakia, J.N., Kenan, D.J., Agol, V.I., and Sonenberg, N. (1994a). Internal translation initiation on poliovirus RNA: further characterization of La function in poliovirus translation in vitro. *J Virol* 68, 1544-1550.
- Svitkin, Y.V., Pause, A., and Sonenberg, N. (1994b). La autoantigen alleviates translational repression by the 5' leader sequence of the human immunodeficiency virus type 1 mRNA. *J Virol* 68, 7001-7007.
- Takahashi, K., Uchida, T., and Egami, F. (1970). Ribonuclease T1, Structure and function. *Adv Biophys* 1, 53-98.
- Tarun, S.Z., Jr., and Sachs, A.B. (1996). Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J* 15, 7168-7177.
- Trachsel, H., Erni, B., Schreier, M.H., and Staehelin, T. (1977). Initiation of mammalian protein synthesis. II. The assembly of the initiation complex with purified initiation factors. *J Mol Biol* 116, 755-767.

- Trachsel, H., and Staehelin, T. (1978). Binding and release of eukaryotic initiation factor eIF-2 and GTP during protein synthesis initiation. *Proc Natl Acad Sci U S A* 75, 204-208.
- Trachsel, H., and Staehelin, T. (1979). Initiation of mammalian protein synthesis. The multiple functions of the initiation factor eIF-3. *Biochim Biophys Acta* 565, 305-314.
- Tsukiyama-Kohara, K., Iizuka, N., Kohara, M., and Nomoto, A. (1992). Internal ribosome entry site within hepatitis C virus RNA. *J Virol* 66, 1476-1483.
- Vagner, S., Gensac, M.C., Maret, A., Bayard, F., Amalric, F., Prats, H., and Prats, A.C. (1995). Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes. *Mol Cell Biol* 15, 35-44.
- van der Meijden, C.M., 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.
- Van Herwynen, J.F., and Beckler, G.S. (1995). Translation using a wheat-germ extract. *Methods Mol Biol* 37, 245-251.
- Vogt, V.M. (1973). Purification and further properties of single-strand-specific nuclease from *Aspergillus oryzae*. *Eur J Biochem* 33, 192-200.
- Wang, C., Sarnow, P., and Siddiqui, A. (1993). Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *J Virol* 67, 3338-3344.
- Warrington, R.C. (1974). Ribonuclease T1 catalyzed degradation of transfer RNA: an unusual alteration induced by urea. *Biochim Biophys Acta* 353, 63-68.
- Wells, S.E., Hillner, P.E., Vale, R.D., and Sachs, A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* 2, 135-140.
- Wilson, J.E., Pestova, T.V., Hellen, C.U., and Sarnow, P. (2000a). Initiation of protein synthesis from the A site of the ribosome. *Cell* 102, 511-520.
- Wilson, J.E., Powell, M.J., Hoover, S.E., and Sarnow, P. (2000b). Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol Cell Biol* 20, 4990-4999.
- Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103-119.

Arnaud D'ORCHYMONT

Etudes structurales de l'ARN messager de l'histone H4

Chez les Eucaryotes, l'étape d'initiation est de loin la plus complexe et la plus lente du processus de traduction. Elle nécessite l'intervention de 12 facteurs protéiques, d'une coiffe $m^7G_{ppp}N$ située à l'extrémité 5' des ARNm et d'une queue poly(A) en 3'. Les ARNm des histones « réplication-dépendantes » sont particuliers car dépourvus d'extrémité 3' polyadénylée et dotés d'une extrémité 5' non traduite extrêmement courte, de 9 nt seulement chez l'ARNm H4 de la souris. Pour traduire ces ARNm, un processus d'initiation non conventionnel a été décrit au laboratoire.

L'objectif de ma thèse a été d'établir les bases structurales de ce mécanisme en combinant différentes approches expérimentales. Deux protocoles originaux de repliement ont été mis au point afin d'isoler l'ARNm H4 dans deux conformations distinctes et stables. La stabilité et la structure de ces deux formes ont été étudiées par DLS, par SAXS et par équilibre de sédimentation. Puis, nous avons étudié la capacité de ces deux formes d'ARNm H4 à fixer le facteur d'initiation eIF4E et les ribosomes assemblés sur le codon d'initiation ainsi que leur aptitude à être traduits *in vitro*. Un modèle de repliement de la structure secondaire de l'ARNm H4 a été construit après sondage enzymatique et chimique des deux formes de l'ARNm. Ce modèle a servi de base pour le travail d'ingénierie de l'ARNm H4 qui a conduit à son découpage en sous-domaines. Des essais de cristallisation ont porté sur 18 de ces fragments ainsi que sur les deux formes de l'ARNm H4 complet.

Mots clés : ARNm, Repliement, Structure, Traduction, Histone

In eukaryotes, the initiation step is far more complex and the slowest inside the translation process. It requires the intervention of 12 protein factors, an $m^7G_{ppp}N$ cap located to the 5' end of the mRNA and a poly (A) tail at the 3'. Histone mRNAs "replication-dependent" are specific because they lack of polyadenylated tail at the 3' end and have a 5' end untranslated extremely short (9 nt only in the mouse). To translate these mRNAs, a process of unconventional initiation has been described in the laboratory.

The aim of my thesis was to establish the structural basis of the mechanism by combining different experimental approaches. Two original refolding protocols have been developed to isolate mRNA H4 in two distinct and stable conformations. The stability and the structure of these two shapes have been studied by DLS, SAXS and sedimentation equilibrium. Then, we studied the ability of these two conformations of H4 mRNA to bind the eIF4E initiation factor and ribosomes assembled on the start codon as well as their ability to be translated *in vitro*. Models of the secondary structure has been constructed after enzymatic and chemical probing of the two shapes of the mRNA. This model was the basis for the engineering of the H4 mRNA that led to its division into sub-domains. Crystallization trials focused on 18 of these fragments as well as on both H4 mRNA shapes.

Keywords : mRNA, Refolding, Structure, Translation, Histone