



Etude du mécanisme d'hyperméthylation de la coiffe des ARNm de sélénoprotéines et impact sur leur traduction

Anne-Sophie Gribling-Burrer

► To cite this version:

Anne-Sophie Gribling-Burrer. Etude du mécanisme d'hyperméthylation de la coiffe des ARNm de sélénoprotéines et impact sur leur traduction. Biochimie, Biologie Moléculaire. Université de Strasbourg, 2015. Français. NNT: 2015STRAJ051 . tel-01599035

HAL Id: tel-01599035

<https://theses.hal.science/tel-01599035>

Submitted on 1 Oct 2017

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 SANTÉ
UPR9002 DU CNRS

THÈSE présentée par :

Anne-Sophie GRIBLING-BURRER

soutenue le : 25 Septembre 2015

pour obtenir le grade de : **Docteur de l'université de Strasbourg**
Science du Vivant / Aspects Moléculaires et Cellulaires de la Biologie

**Etude du mécanisme d'hyperméthylation
de la coiffe des ARNm de sélénoprotéines
et impact sur leur traduction**

THÈSE dirigée par :

Dr ALLMANG-CURA Christine

Chargée de Recherche du CNRS, Strasbourg

RAPPORTEURS :

Dr NAMY Olivier

Chargé de Recherche du CNRS, Paris

Dr BORDONNE Rémy

Directeur de Recherche du CNRS, Montpellier

AUTRES MEMBRES DU JURY :

Dr ROMBY Pascale

Directeur de Recherche du CNRS, Strasbourg

Dr MASSENET Séverine

Chargée de Recherche du CNRS, Nancy

Remerciements

Je souhaite tout d'abord remercier les Drs. Olivier Namy, Severine Massenet et Rémy Bordonné d'avoir accepté d'évaluer mon travail de thèse. Un merci particulier à Dr. Pascale Romby qui, en plus d'avoir bien voulu examiner mon travail, a toujours été d'un grand soutien tant scientifique qu'humain.

Je remercie le Pr. Eric Westhof de m'avoir accueilli au sein de son unité de recherche.

J'exprime ma gratitude à Gilbert Eriani et Alain Krol, de m'avoir accepté au sein de leur équipe. Travailler dans ces deux équipes a été très enrichissant et m'a permis d'échanger avec des personnes exceptionnelles et d'apprendre beaucoup de méthodes et de techniques différentes. Ils m'ont également permis de communiquer mes résultats à travers plusieurs congrès internationaux et m'ont apporté de nombreux conseils scientifiques.

Mon plus grand merci va à Christine Allmang. Elle m'a soutenu durant toutes ces années et a fait de moi la scientifique que je suis aujourd'hui. J'ai réalisé mon premier stage volontaire avec Christine il y a sept ans, elle a toujours été disponible et son appui m'a permis de réaliser mon master avec succès, d'obtenir une bourse de thèse de l'école doctorale et même une bourse de 4^{ème} année de thèse auprès de la Ligue contre le Cancer. Sa confiance m'a permis de devenir indépendante et de m'épanouir scientifiquement. Je pense qu'à nous deux, nous formions une bonne équipe ! C'était chouette de travailler avec toi pendant ces quatre années.

Je remercie toute l'équipe Eriani pour la bonne ambiance de travail durant ces deux dernières années. Merci à Franck pour les nombreux conseils scientifiques qu'il m'a apporté. Je remercie particulièrement Laure et Angelita, mes deux folles préférées, qui m'ont donné le sourire chaque jour. Je m'excuse d'avoir été un peu trop souvent tête en l'air et, d'avance, pour les anniversaires que je vais oublier de souhaiter ! Merci à Quentin pour ses conseils sur le post-doc et la graisse de canard dans les gâteaux. Enfin, merci à toute l'équipe pour les labmeeting, scientifiquement très enrichissants et les pots du vendredi, autrement instructifs !

J'aimerais également remercier l'ancienne équipe Krol, Anne et Stéphanie pour leur assistance technique et leur soutien de tous les jours, Philippe et Evelyne pour leurs conseils scientifiques. Un grand merci à Patryk qui a été mon mentor et mon doctorant modèle. Merci

à ma petite Olga avec qui j'ai passé de chouettes moments à savourer une petite gourmandise, cherchée à la boulangerie du coin.

Je remercie également nos collaborateurs de l'équipe de Jean Cavarelli (IGBMC, Strasbourg), de l'équipe de Séverine Massenet (IMoPa, Ingénierie Moléculaire et Physiopathologie Articulaire, Nancy) et Yaser Hashem (IBMC, Strasbourg).

Je remercie la Ligue contre le Cancer et ses donateurs de m'avoir accordé leur confiance et de m'avoir permis de continuer mes travaux de thèse en finançant cette quatrième année.

J'exprime ma gratitude à Joern Putz et Daniela Lener qui m'ont permis de faire mes premiers pas dans l'enseignement en tant que monitrice à l'université de Strasbourg. Cette expérience exceptionnelle m'a permis de gagner de l'assurance et de développer des compétences pédagogiques.

Un grand merci à tous les amis merveilleux de l'IBMC. Mélo, Hagen, Anja, Sarah, Paola, Angy, Stephano, Jana, Franzi, Patryk, Erika, Lee, JC, Arnaud, Olga, Micka, Sémi, Megha, Pierre, Tina, Franck, Santi, Chouchou (Alexis), Ludo, Mathieu... avec qui les fêtes et les sorties du labo sont inoubliables. Une petite pensée particulière pour Vijay, qui a apporté beaucoup de soleil au 3^{ème} étage et que je n'oublierai pas.

J'aimerais particulièrement remercier Mélodie et Morgane. Nous nous sommes soutenues pendant les cinq années de la fac et les quatre années de thèse. Votre écoute et vos conseils ont été essentiels pour moi pendant toutes ces années. Trois copines, trois thèses !

Je remercie mes amis d'enfance Nico, Simon, Sam, Charles, Ben et Anaïs d'avoir toujours été là pour moi.

Ma plus grande pensée va à ma famille. Merci à mes parents, mes petites sœurs Margot et Léa, mon frère Jérôme, ma cousine Fanny, mon filleul Julien, mes grands parents, ma marraine et mon parrain, qui ont toujours cru en moi et m'ont donné les moyens et la force de réussir ma thèse.

Je remercie de tout mon cœur Redmond. Je te suis infiniment reconnaissante pour ton soutien de tous les jours, ta compréhension et ta patience. Merci pour tous les moments de bonheur qu'on a passé et qu'on passera encore ensemble.

Sommaire

SOMMAIRE.....	1
ABREVIATIONS.....	5
INTRODUCTION	9
I LES SELENOPROTEINES	9
I.1 SELENIUM ET SELENOPROTEINES	9
I.1.1 <i>Le sélénium</i>	9
I.1.2 <i>La sélénocystéine</i>	11
I.1.3 <i>Les sélénoprotéines</i>	12
I.1.4 <i>Sélénoprotéines, stress oxydant et cancer</i>	18
I.2 LA SYNTHÈSE DES SELENOPROTEINES	21
I.2.1 <i>La biosynthèse de la sélénocystéine</i>	21
I.2.1.1 L'ARN de transfert sélénocystéine (ARNt ^{[Ser]Sec})	22
I.2.1.2 Le séryl-ARNt ^{[Ser]Sec} et la transformation de la sérine en phosphosérine	24
I.2.1.3 La transformation du O-phosphoséryl-ARNt ^{[Ser]Sec} en sélénocystéinyl-ARNt ^{[Ser]Sec}	25
I.2.1.3.a La Sélénocystéine synthase	25
I.2.1.3.b La sélénophosphate synthétase et le donneur de sélénium	26
I.2.2 <i>Incorporation de sélénocystéine dans les sélénoprotéines</i>	27
I.2.2.1 Les ARNm de sélénoprotéines	28
I.2.2.1.a L'élément SECIS (SElenoCysteine Insertion Sequence)	29
I.2.2.1.b L'élément SRE (Selenocysteine Redefinition Element)	32
I.2.2.2 Les Protéines impliquées dans le mécanisme de synthèse des sélénoprotéines.....	34
I.2.2.2.a Le facteur d'élongation spécifique eEFSec.....	34
I.2.2.2.b SBP2 (SECIS binding protein 2)	35
I.2.2.2.b.1 SBP2 est une protéine de la famille L7Ae	37
I.2.2.2.b.2 Interaction SBP2-ARN SECIS	40
I.2.2.2.b.3 Interaction SBP2-eEFSec.....	40
I.2.2.2.b.4 Interaction SBP2-ribosome.....	41
I.2.2.2.c La protéine ribosomique L30	42
I.2.3 <i>Modèles du mécanisme de synthèse des sélénoprotéines</i>	43
I.2.4 <i>Régulation de l'expression des sélénoprotéines</i>	45
I.2.4.1 Efficacité du mécanisme de recodage	45
I.2.4.1.a Relation entre terminaison et recodage	46
I.2.4.1.b Relation entre initiation et recodage	47
I.2.4.2 Expression hiérarchique des sélénoprotéines	48
I.2.4.2.a Les facteurs de régulation de l'expression des sélénoprotéines	49
I.2.4.2.a.1 Le sélénium et l'ARNt ^{[Ser]Sec}	49
I.2.4.2.a.2 eIF4A3	51
I.2.4.2.a.3 Nucléoline	52
I.2.4.2.a.4 SECp43	52
I.2.4.2.a.5 NSEP1	53
I.2.4.2.a.6 SBP2L.....	53
I.2.4.2.a.7 GAPsec.....	54
I.2.4.2.b Régulation de la dégradation des ARNm de sélénoprotéines par le NMD.....	54
I.2.4.3 L'assemblage des mRNP de sélénoprotéines.....	56
I.2.4.3.a Un assemblage nucléaire des mRNP de sélénoprotéines.....	57
I.2.4.3.b Assemblage des ARNm de sélénoprotéines : similitudes avec l'assemblage des sn/snoRNP.....	60
II MECANISME DE BIOGENESE DES SN ET SNORNP A BOITE C/D	61

II.1	BIOGENESE DES snRNP	61
II.1.1	<i>Transcription et export dans le cytoplasme.....</i>	63
II.1.2	<i>Maturation et assemblage des snRNP dans le cytoplasme.....</i>	65
II.1.3	<i>Réimport, assemblage et modification des snRNP dans le noyau</i>	65
II.1.4	<i>Le complexe SMN et le méthylosome fonctionnent de manière coordonnée dans l'assemblage des snRNP</i>	66
II.1.4.1	Le méthylosome.....	67
II.1.4.2	Le complexe SMN	67
II.1.4.3	Rôle du méthylosome et du complexe SMN dans l'assemblage des protéines Sm	70
II.1.4.4	SMN et la pathologie SMA.....	73
II.2	BIOGENESE DES snORNP A BOITE C/D.....	73
II.3	L'HYPERMETHYLATION DE LA COIFFE DES snARN ET DES snoARN.....	75
III	MECANISMES DE MATURATION ET D'INITIATION DE LA TRADUCTION DES ARNm CHEZ LES EUKARYOTES.....	77
III.1	MATURATION DES ARNm EUKARYOTES.....	77
III.1.1	<i>Les ARNm portent une coiffe monométhylée m⁷G.....</i>	<i>78</i>
III.1.1.1	Biogenèse de la coiffe m ⁷ G.....	78
III.1.1.2	Reconnaissance de la coiffe m ⁷ G par le facteur CBC.....	79
III.1.2	<i>Polyadénylation de l'extrémité 3' des ARNm.....</i>	<i>81</i>
III.2	LA TRADUCTION DES ARNm.....	82
III.2.1	<i>Mécanisme d'initiation de la traduction canonique, coiffe dépendant.....</i>	<i>83</i>
III.2.1.1	Recrutement de eIF4F au niveau de la coiffe m ⁷ G.....	83
III.2.1.2	Le complexe 43S	85
III.2.1.3	Formation du complexe 48S sur l'ARNm	86
III.2.1.4	Assemblage du complexe 80S	86
III.2.2	<i>Mécanismes d'initiation coiffe-indépendant induits par les IRES</i>	<i>87</i>
III.2.2.1	IRES de type I	88
III.2.2.2	IRES de type II.....	89
III.2.2.3	IRES de type III	89
III.2.2.4	IRES de type IV.....	89
III.2.3	<i>Mécanisme d'initiation de la traduction coiffe-dépendant par "tethering".....</i>	<i>90</i>
III.2.4	<i>Mécanismes d'initiation coiffe-indépendant chez les virus de plantes.....</i>	<i>92</i>
IV	OBJECTIFS ET REALISATIONS DU TRAVAIL DE THESE	95
	RESULTATS.....	96
I	LA COIFFE DES ARNm DE SELENOPROTEINES EST HYPERMETHYLEE CHEZ LES MAMMIFERES	96
I.1	RESUME DE LA PUBLICATION	96
I.2	PUBLICATION	101
I.3	RESULTATS SUPPLEMENTAIRES	122
I.3.1	<i>Les ARNm de sélénoprotéines n'interagissent pas avec la protéine snurportine-1.....</i>	<i>122</i>
I.3.2	<i>Les protéines Sm ne sont pas assemblées sur les ARNm de sélénoprotéines..</i>	<i>123</i>
I.3.3	<i>Impact des facteurs de liaison à la coiffe sur la synthèse de la sélénoprotéine HA-GPx1</i>	<i>124</i>
I.3.3.1	Le complexe de liaison à la coiffe CBC	124
I.3.3.2	L'hélicase DDX3.....	127
I.4	CONCLUSION	128

II LE CHAPERON D'ASSEMBLAGE SMN ET LE METHYLOSOME SONT IMPLIQUES DANS LA SYNTHÈSE DES SELENOPROTEINES	133
II.1 OBJECTIF	133
II.2 RESULTATS	133
II.2.1 <i>Le complexe SMN et le méthylosome interagissent avec la protéine SBP2.....</i>	<i>133</i>
II.2.2 <i>L'inhibition des complexes SMN et du méthylosome a un impact fonctionnel sur la synthèse de la sélénoprotéine HA-GPx1 in vivo.....</i>	<i>139</i>
II.2.3 <i>Rôle du complexe SMN dans la synthèse des sélénoprotéines - origine du stress oxydant observé dans la pathologie SMA.....</i>	<i>143</i>
II.2.4 <i>A la recherche du rôle du Méthylosome.....</i>	<i>146</i>
III ETUDE DU MECANISME DE TRADUCTION DES ARNM DE SELENOPROTEINES	149
III.1 OBJECTIFS.....	149
III.2 RESULTATS.....	150
III.2.1 <i>L'initiation de la traduction des sélénoprotéines est atypique.....</i>	<i>150</i>
III.2.1.1 Analyse du rôle de la coiffe m ³ G sur la traduction des ARNm de sélénoprotéines in vitro.....	150
III.2.1.1.a La coiffe m ³ G inhibe le recrutement des ribosomes sur l'ARNm de SelM <i>in vitro</i>	150
III.2.1.1.b La coiffe m ⁷ G des ARNm en cours de traduction est accessible et son hyperméthylation n'induit pas la dissociation des ribosomes	156
III.2.1.2 Analyse du mécanisme d'initiation et de recodage de la traduction des ARNm de sélénoprotéines in vitro.....	158
III.2.1.2.a La présence du codon UGA ^{Sec} induit l'accumulation de ribosomes 80S sur l'ARNm de SelM.....	159
III.2.1.2.b Un mécanisme atypique d'initiation de la traduction pour l'ARNm de SelM	162
III.2.1.2.c Localisation du ribosome sur l'ARNm	163
III.2.1.2.d La région 3'UTR de l'ARNm de SelM est impliquée dans le recrutement du ribosome.....	166
III.2.1.2.e La protéine SBP2 se lie en amont du codon UGA ^{Sec} de l'ARNm de SelM.....	169
III.2.2 <i>Vers l'identification de nouveaux facteurs des complexes de traduction des ARNm de sélénoprotéines.....</i>	<i>175</i>
III.2.3 <i>Vers une analyse structurale des complexes de traduction des ARNm de sélénoprotéines</i>	<i>182</i>
DISCUSSION GENERALE ET PERSPECTIVES	187
CONCLUSION	219
MATERIEL ET METHODES	221
I MATERIEL.....	221
I.1 LIGNEES CELLULAIRES EUCARYOTES.....	221
I.2 PLASMIDES	221
I.3 ANTICORPS	223
I.4 OLIGONUCLEOTIDES	224
II METHODES	229
II.1 SURPRODUCTION ET PURIFICATION DE PROTEINES (HSBP2-HIS ET DERIVES) DANS E. COLI.....	229
II.2 CULTURE CELLULAIRE : LIGNEES HEK 293FT ET FLP-IN™ T-REX™ 293.....	230
II.3 TRANSFECTION DES LIGNEES CELLULAIRES PAR LA LIBRAIRIE DE SIARN ET ANALYSE DE L'EXPRESSION DE HA-GPX1 PAR WESTERN BLOT.....	232

II.4	IMMUNOPRECIPITATION DE PARTICULES RIBONUCLEOPROTEIQUES (RNP).....	233
II.5	DETECTION DES ARN PAR RT-PCR ET RT-PCR QUANTITATIVE	234
II.6	EXPERIENCES D'IMMUNOPRECIPITATION D'ARN A L'AIDE DES ANTICORPS ANTI-M ₃ G.....	235
II.7	EXPERIENCES DE CO-IMMUNOPRECIPITATION.....	236
II.8	ANALYSE DE PROTEINES SUR GEL DE POLYACRYLAMIDE SDS-PAGE	236
II.9	IMMUNO-DETECTION DE PROTEINES PAR WESTERN BLOT	237
II.10	TEST D'INTERACTION SBP2-METHYLOSOME A PARTIR DU SYSTEME D'EXPRESSION BACULOVIRUS	237
II.11	TEST DE METHYLATION <i>IN VITRO</i> DE SBP2 PAR LES PROTEINES PRMT5/Mep50 DU METHYLOSOME.....	238
II.12	EXPÉRIENCES DE DOUBLE HYBRIDE	239
II.13	TRANSCRIPTION <i>IN VITRO</i> DES ARN SELM ET GPx1 PAR LA T7 ARN POLYMERASE.....	239
II.14	ADDITION DE LA COIFFE DES ARN	240
II.15	TRADUCTION <i>IN VITRO</i> DES ARNm.....	241
II.16	CINETIQUES DE TRADUCTION	241
II.17	ASSEMBLAGE DE COMPLEXES TRADUCTIONNELS SUR UN ARNm A PARTIR D'EXTRAITS RRL.....	242
II.18	ASSEMBLAGE DE COMPLEXES TRADUCTIONNELS SUR UN ARNm A PARTIR D'EXTRAITS DE CELLULES HeLa	242
II.19	PURIFICATION DE COMPLEXES TRADUCTIONNELS EN VUE D'ANALYSE PAR CRYO- MICROSCOPIE ELECTRONIQUE	243
II.20	EXPERIENCES DE TOEPRINT.....	245
II.21	EXPERIENCES DE RETARD SUR GEL.....	246
II.22	PURIFICATION DES COMPLEXES ASSOCIES A L'ARNm DE HA-GPx1 <i>IN VIVO</i>	246
BIBLIOGRAPHIE.....		251

Abréviations

Å	Angström
A, C, G, T, U, N	adénine, cytosine, guanine, thymine, uracile, N : l'un des 4 nucléotides
ADN	acide désoxyribonucléique
ADNc	ADN complémentaire
ARE	AU-Rich Element
ARN	acide ribonucléique
ARNm	ARN messenger
ARNr	ARN ribosomique
ARNt	ARN de transfert
ARNt ^{[Ser]Sec}	ARN de transfert sélénocystéine
ATP	adénosine 5'-triphosphate
CBC	cap binding complex
CBP	Cap binding pocket
<i>C.elegans</i>	<i>Caenorhabditis elegans</i>
3'CITE	3' cap-independent translation enhancer
cpm	coups par minute
CRD	cystein rich domain
CRM1	chromosome region maintenance-1
C-terminal	carboxy-terminal
DIO	iodothyronine désiodase
DTT	dithiothreitol
<i>E.coli</i>	<i>Escherichia coli</i>
eEF	facteur d'élongation eucaryotique
eEFSec	facteur d'élongation spécifique de la sélénocystéine
4E-SE	eIF4E Sensitive Element
eIF	facteur d'initiation eucaryotique
EJC	exon junction complex
eRF	facteur de terminaison eucaryotique
g	gramme
GAP	GTPase-activating protein
GFP	green fluorescent protein
GPx	glutathion peroxydase
GSH	glutathion
GST	glutathion-S-transférase
GTP	guanosine 5'-triphosphate
HA	hémagglutinine
IDP	intrinsically disordered protein
IP	immunoprécipitation
IRES	internal ribosomal entry site
kDa	kilo Dalton
μ	micro
M	molaire

m ⁷ G	coiffe monométhylguanosine
m ₃ G	coiffe triméthylguanosine (m ₃ ^{2,2,7} G)
MEP50	methylosome protein 50
Mg	magnésium
min	minute
mRNP	RNP d'ARN messagers
NES	nuclear export signal
NLS	nuclear localisation signal
NMD	nonsense mediated decay
NOP58	Nucleolar protein 58
NSEP1	nuclease sensitive element binding protein 1
Nt	nucléotide
N-terminal	amino-terminal
PAGE	polyacrylamide gel electrophoresis
Pb	paire de bases
PCR	polymerase chain reaction
PHAX	phosphorylated adaptor for RNA export
PH-GPx	phospholipide-hydroperoxyde glutathion peroxydase
pICln	chloride conductance regulatory protein
pol II	ARN polymérase II
pré-ARNm	précurseur des ARN messagers
PRMT5	Protein arginine N-methyl transferase 5
PSTK	O-phosphoséryl-ARNtSec kinase
PTC	premature translation termination codon
R	purine
RE	réticulum endoplasmique
RIP	immunopurification de complexes ribonucléoprotéiques
RNP	particule ribonucléoprotéique
RRL	rabbit reticulocyte lysate
RT	transcriptase inverse
<i>S.cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SAM	S-adenosyl-methionine
SBP2	SECIS binding protein 2
scaRNP	RNP des corps de Cajal
sDMA	diméthylarginines symétriques
SDS	sodium dodécyl sulfate
Se	sélénium
Sec	sélénocystéine
Sec-ARNt ^{[Ser]Sec}	sélénocystéinyl-ARNtSec
SECIS	selenocysteine insertion sequence
SelD	sélénophosphate synthétase bactérienne
SelN	sélénoprotéine N
Sep-ARNt ^{[Ser]Sec}	phosphoséryl-ARNtSec
SepSecS	sélénocystéine synthase humaine
Ser	sérine
SerRS	séryl-ARNt-synthétase

siARN	petit ARN interférent
SID	sec incorporation domain
SMA	atrophie musculaire spinale
SMN	survival of motor neurons
snARN	ARN nucléaire
snoARN	ARN nucléolaire
snoRNP	RNP nucléolaire
snRNP	RNP nucléaire
SPS	sélénophosphate synthétase
SRE	selenocysteine redefinition element
SRP	signal recognition particle
TAP-tag	tandem affinity purification-tag
TBE	Tris borate EDTA
TCA	acide trichloroacétique
Tgs1	trimethylguanosine synthase 1
TMG	trimethylated guanosine (coiffe)
TrxR	thiorédoxine réductase
U	unité
UTR	région non-traduite
VRC	vanadyl ribonucleotide complex
YFP	yellow fluorescent protein
°C	degré Celsius
Y2H	test de double hybride chez la levure

Introduction

I Les sélénoprotéines

I.1 Sélénium et sélénoprotéines

I.1.1 Le sélénium

Découvert en 1817 par Jöns Berzelius, le sélénium doit son nom à la déesse grecque de la lune, Sêlênê. Le sélénium est un élément chimique, appartenant à la famille des chalcogènes dans la classification périodique des éléments de Mendeleïv. Longtemps considéré comme toxique, il se révèle également être un oligoélément essentiel (pour revue, Lee and Jeong, 2012). Le sélénium, présent dans le sol, est assimilé par les micro-organismes et les végétaux, qui sont à l'origine de son incorporation dans la chaîne alimentaire. La teneur des sols en sélénium varie beaucoup selon les zones géographiques, ce qui se traduit par une inégalité dans l'apport de sélénium par l'alimentation. Ainsi, dans certaines régions d'Asie, on observe, chez l'homme, des carences en sélénium pouvant conduire à de graves problèmes cardiaques (maladie de Keshan), osseux (maladie de Kashin-Beck), ou neuro-musculaires (pour revue, voir Bellinger et al., 2009; Roman et al., 2014). Des fortes carences en sélénium sont également associées à diverses pathologies telles que des retards mentaux, des déficiences immunitaires, des myopathies et des cancers (Chen et al., 2013b; Lescure et al., 2008; Rayman and Rayman, 2002). Le sélénium est essentiel pour la santé humaine et l'apport journalier conseillé est de 100 à 120 μg (pour revues, Hatfield et al., 2009; Rederstorff et al., 2006; Roman et al., 2014; Stoffaneller and Morse, 2015). Le métabolisme du sélénium chez l'homme est représenté sur la Figure 1. Le sélénium provient de la nourriture, principalement d'origine végétale, mais également des viandes rouges et des poissons, et est assimilé sous différentes formes (sélénate, selenométhionine, sélénocystéines, γ -glutamyl-CH₃-sélénocystéine...). Lorsque le sélénium est en excès, il peut réagir avec l'oxygène, formant des radicaux oxygénés toxiques. Des voies d'excrétions sont alors mises en place par les cellules pour procéder à l'élimination de composés sélénés.

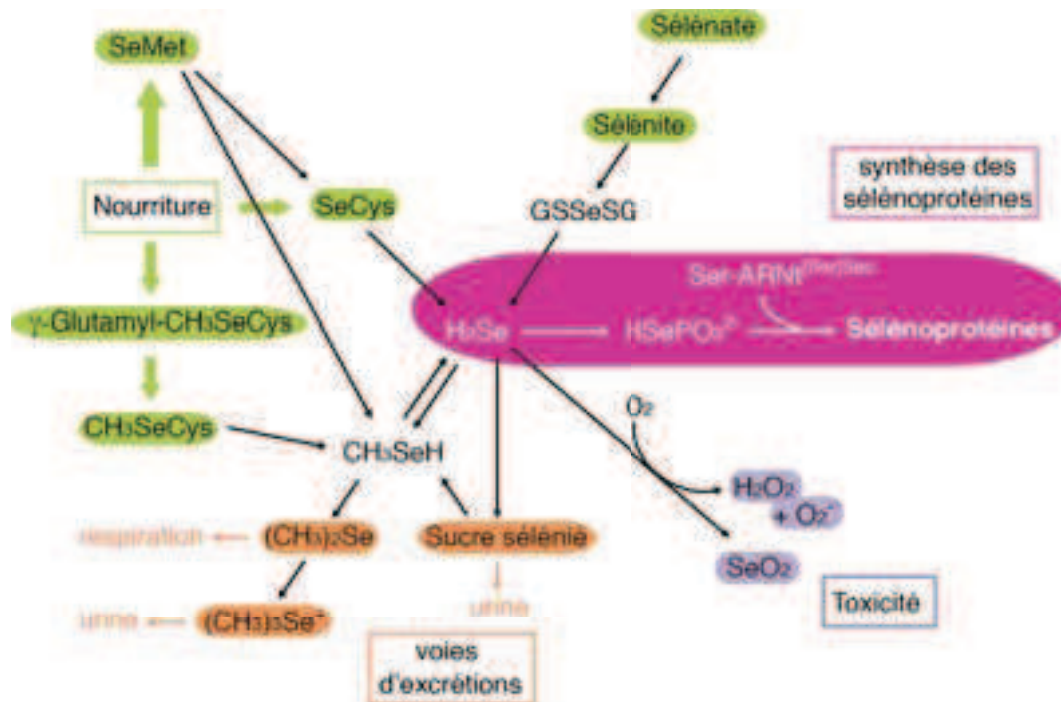


Figure 1. Métabolisme du sélénium provenant de l'alimentation chez l'homme (d'après Chen et al., 2013)

Les éléments sélénés provenant de l'alimentation sont représentés en vert. La voie de synthèse des sélénoprotéines est indiquée en rose et les voies d'excrétions des composés sélénés sont en orange. L'oxydation du sélénure d'hydrogène est toxique pour la cellule (violet). SeMet, sélénométhionine ; SeCys, sélénocystéine ; GSSeSG, séléno-diglutathion ; γ -glutamyl-CH₃SeCys, γ -glutamyl-méthyl-sélénocystéine ; H₂Se, sélénure d'hydrogène ; HSePO₃²⁻, sélénophosphate ; CH₃SeCys, méthylsélénocystéine ; CH₃SeH, méthylsélénol ; (CH₃)₂Se, diméthyl-sélénite ; SeO₂, dioxyde de sélénium ; (CH₃)₃Se⁺, ion triméthylsélénium.

La forme biologique majeure du sélénium est l'acide aminé sélénocystéine (Figure 1 et 2) qui est incorporé par un mécanisme spécifique dans les sélénoprotéines (Flohe et al., 2000), dont les fonctions sont essentielles pour les cellules (Korotkov et al., 2002a).

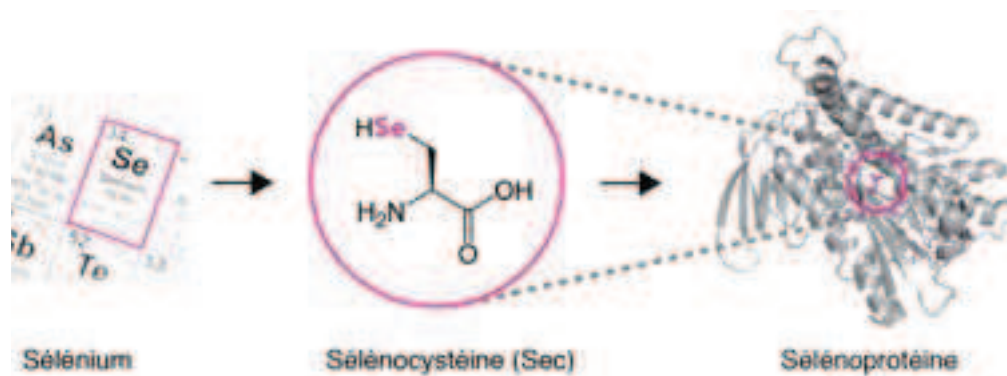


Figure 2. Le sélénium est incorporé dans les sélénoprotéines sous forme de sélénocystéine.

Le sélénium (Se) est un élément chimique de la famille des chalcogènes, de numéro atomique 34 et de masse atomique 78,96. L'acide aminé sélénocystéine est incorporé spécifiquement dans le site actif des sélénoprotéines.

1.1.2 La sélénocystéine

La sélénocystéine, considérée comme le 21^{ème} acide aminé, est la forme biologique majeure du sélénium. Il s'agit d'un analogue de la cystéine dont le groupement thiol est remplacé par un groupement sélénol (Figure 2 et 3). La sélénocystéine est présente dans le site catalytique des sélénoprotéines et y est incorporée de manière co-translationnelle. A pH physiologique, le groupement sélénol de la sélénocystéine isolée ($pK_a=5,2$) est ionisé, contrairement au thiol de la cystéine ($pK_a=8,5$). En raison de sa configuration électronique, le sélénolate est un meilleur nucléophile que le thiol, ce qui lui confère un potentiel réducteur beaucoup plus important. Ceci se traduit par une activité catalytique accrue des sélénoprotéines en comparaison de celle de leurs homologues à cystéine (pour revue, Muttenthaler and Alewood, 2008), et le remplacement de la sélénocystéine par la cystéine conduit à une diminution drastique de l'activité enzymatique des sélénoprotéines (Gasdaska et al., 1999; Kim et al., 2006, 2015; Lee et al., 2000). La puissante activité réductrice que confère la sélénocystéine semble justifier l'origine des sélénoprotéines dont la synthèse est très complexe et coûteuse en énergie pour la cellule. En effet, l'acide aminé sélénocystéine est incorporé dans les sélénoprotéines en réponse à un codon UGA, reconnu habituellement par la machinerie traductionnelle comme l'un des trois codons stop (Chambers et al., 1986). La synthèse des sélénoprotéines requiert un mécanisme de recodage traductionnel qui implique le recrutement de plusieurs facteurs spécifiques vers les ARNm de sélénoprotéines (Allmang et al., 2009).

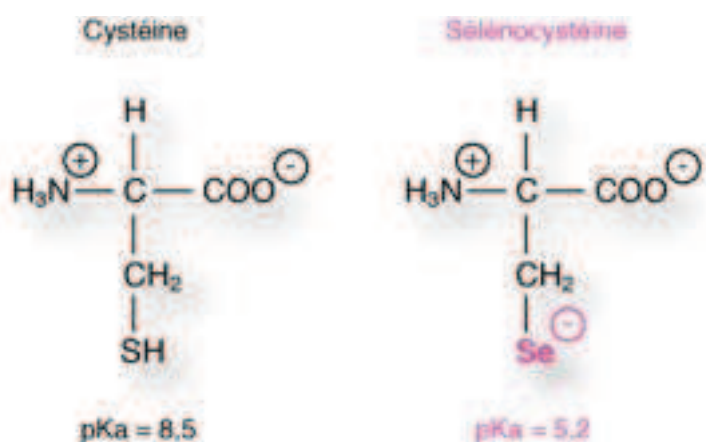


Figure 3. L'acide aminé sélénocystéine est un analogue de la cystéine.

Le groupement thiol de la cystéine est remplacé par un groupement sélénol au niveau de l'analogue sélénocystéine. Ces deux acides aminés ont un pK_a différent.

1.1.3 Les sélénoprotéines

Les sélénoprotéines sont retrouvées dans les trois domaines de la vie, les bactéries, les archées et les eucaryotes, elles sont également présentes chez certains virus (Mix et al., 2007; Shisler et al., 1998). Cependant, toutes les espèces n'ont pas de sélénoprotéines. Ainsi, chez les levures, les plantes supérieures et certains insectes, les facteurs nécessaires à la synthèse des sélénoprotéines sont absents et ces espèces expriment des orthologues remplaçant l'acide aminé sélénocystéine par la cystéine (Johansson et al., 2005; Lobanov et al., 2008). Le nombre de sélénoprotéines codées par les génomes des différents organismes (ou sélénoprotéomes) est très variable. Le plus grand sélénoprotéome est celui de l'algue pélagophyte *Aureococcus anophagefferens* qui comprend 59 sélénoprotéines (Gobler et al., 2013) alors que le nématode *C. elegans* ne possède qu'une sélénoprotéine unique (Taskov et al., 2005). L'analyse des sélénoprotéomes a fourni des informations quant à l'évolution de l'utilisation de sélénocystéine. De manière générale, les plus grands sélénoprotéomes sont retrouvés chez les organismes aquatiques (poissons et algues), tandis que les organismes terrestres ont réduit le leur en remplaçant la sélénocystéine par son homologue cystéine. Ainsi, l'évolution des sélénoprotéines semble dépendre de la disponibilité du sélénium dans l'environnement dans lequel les organismes se développent (Lobanov et al., 2009). A ce jour, 25 gènes de sélénoprotéines ont été identifiés dans le génome humain (pour revue, Bellinger et al., 2009; Hatfield et al., 2014; Labunskyy et al., 2014; Lobanov et al., 2009) (Figure 4). Peu d'homologies de séquences sont retrouvées entre les différentes sélénoprotéines et, pour la majorité, leur fonction exacte n'est pas connue. La sélénocystéine est soit localisée dans la région N-terminale de la protéine, où elle fait partie du site actif, soit dans la partie C-terminale où elle est impliquée dans des transferts d'électrons (Hatfield and Gladyshev, 2002; Kryukov et al., 2003a). La sélénocystéine de nombreuses sélénoprotéines est incluse dans un motif CXXU (C=Cys, U=Sec et X=un acide aminé quelconque), similaire au motif CXXC retrouvé dans le site actif de la thioredoxine (Dikiy et al., 2007; Ferguson et al., 2006; Ladenstein et al., 1979; Lu and Holmgren, 2009; Novoselov et al., 2007). Ce motif est impliqué dans l'activité d'oxydo-réduction des sélénoprotéines (Hatfield and Gladyshev, 2002; Kryukov et al., 2003a). Les structures de plusieurs sélénoprotéines ont été déterminées par cristallographie ou résonance magnétique nucléaire (RMN), elles sont représentées

dans la Figure 5. Plusieurs sélénoprotéines sont caractérisées par un repliement de type thioredoxine (Figure 5B et C), tandis que d'autres présentent des repliements différents, similaires aux glutathion réductases par exemple (Figure 5A).

seleno-protein	Function	Schematic representation
DIO1	thyroid hormone maturation (Conversion of T4 to T3 and T4 to rT3)	
DIO2	thyroid hormone maturation (Conversion of T4 to T3)	
DIO3	thyroid hormone catabolism (Conversion of T4 and T3 to rT3)	
GPx1	Antioxidant protection, cytosolic	
GPx2	Antioxidant protection, gastrointestinal	
GPx3	Maintenance of cellular redox status, plasma	
GPx4	Detoxification of lipid hydroperoxides, male fertility	
GPx6	Antioxidant protection, olfactory	
TrxR1	Part of the thioredoxin system, cytosolic Antioxidant defense, redox regulation, cell signaling	
TrxR2	Part of the thioredoxin system, testis specific Antioxidant defense, redox regulation, cell signaling	
TrxR3	Part of the thioredoxin system, mitochondrial Antioxidant defense, redox regulation, cell signaling	
SPS2	Selenocysteine synthesis	
15kDa	Role in cell apoptosis and quality control of protein folding in the ER	
SelN	Muscle development. Associated with muscular diseases	
SelW	Antioxidant protection Associated with cardiac calcification	
SelP	Plasma selenium transport Antioxidant defense	
SelH	Unknown	
SelI	Unknown	
SelJ	Present in fish and sea urchin	
SelK	Unknown, ER	
SelM	Unknown	
SelO	Unknown	
SelS	Unknown, involved in cellular redox balance, ER	
SelT	Role in regulation of Ca2+ homeostasis and neuroendocrine secretion	
SelU	Unknown Present in fish, birds, sea urchin, green algae and diatoms only	
SelV	Unknown, testis	
MsrA	Reduction of oxidized methionine residues Selenocysteine in C. reinhardtii only	
MsrB	Reduction of oxidized methionine residues	

Figure 4. Sélénoprotéome eucaryote.

La séquence en acide aminé des différentes sélénoprotéines est représentée par un rectangle gris et la position de l'acide aminé sélénocystéine par un trait noir. DIO: iodothyronine désiodase, GPx: glutathion peroxydase, TrxR: thioredoxine réductase, SPS: sélénophosphate synthétase, Sel: sélénoprotéine, Msr: méthionine sulfoxyde réductase. (pour revues, Hatfield et al., 2014; Labunsky et al., 2014; Roman et al., 2014).

Les sélénoprotéines, de part leur importante activité redox, sont impliquées dans des mécanismes variés tels que la défense cellulaire contre les stress oxydants

(Steinbrenner and Sies, 2009), la fertilité masculine (Guerriero et al., 2014; Imai et al., 2009), le métabolisme thyroïdien (Beckett and Arthur, 2005; Köhrle, 2013), la réponse immunitaire (Steinbrenner et al., 2015), le développement des muscles (Castets et al., 2012; Rederstorff et al., 2006) et la protection des télomères (Squires et al., 2009). La fonction des 25 sélénoprotéines du sélénoprotéome eucaryote, lorsqu'elle est connue, est indiquée dans la Figure 4. Parmi les sélénoprotéines les mieux caractérisées, on retrouve 3 familles d'enzymes: les glutathion peroxydases (GPx), les thioredoxine réductases (TrxR) et les iodothyronine désiodases (DIO).

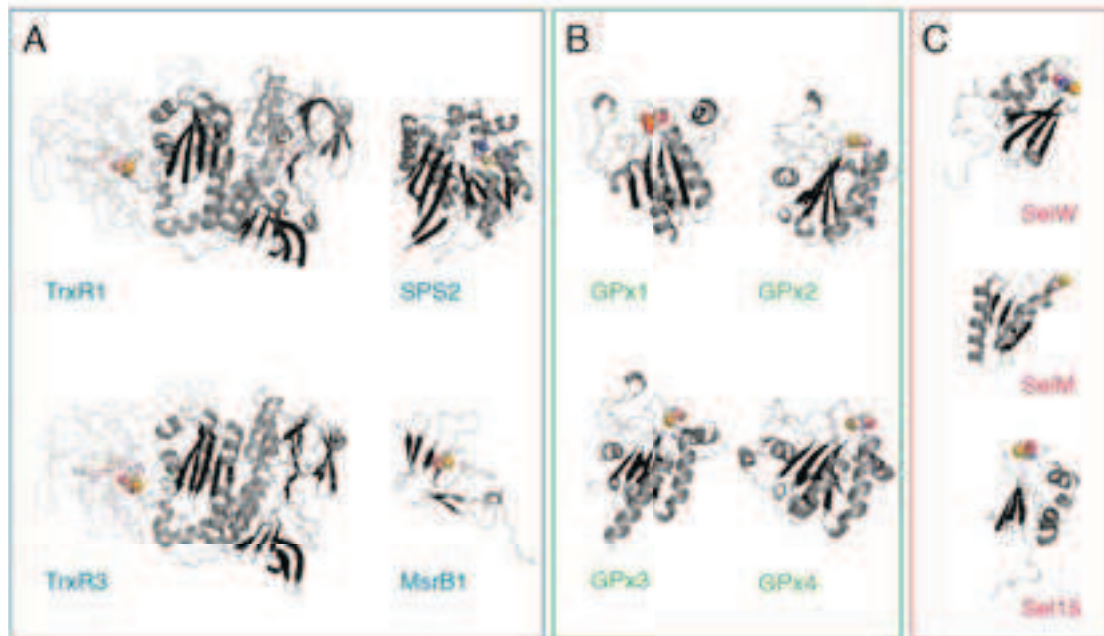


Figure 5. Structures de sélénoprotéines de mammifères. (d'après Labunsky et al., 2014)

Les structures cristallographiques de GPx1 de boeuf (Epp et al., 1983), de GPx4 de souris (Savaskan et al., 2007), de TrxR1 de rat (cytosolique) (Cheng et al., 2009; Sandalova et al., 2001), de TrxR3 de souris (mitochondriale) (Biterova et al., 2005) ainsi que les structures RMN de MsrB1 (Aachmann et al., 2010), de SelW de souris (Aachmann et al., 2007), et des domaines thioredoxine de Sel15 de drosophile et de SelM de souris (Ferguson et al., 2006) sont représentées. TrxR1 et TrxR3 sont des homodimères (pour l'un des deux monomères, seul le squelette de la chaîne carbonnée est représenté). GPx1, GPx2 et GPx3 forment des homotétramères dans les cellules. Les hélices α sont représentées en gris foncé, les brins β en noir et les boucles en gris clair. Pour chaque protéine, l'acide aminé Sec/Cys catalytique est représenté en modèle atomique. (A) Sélénoprotéines sans repliement de type thioredoxine. (B) et (C) Sélénoprotéines présentant un repliement de type thioredoxine. (B) Famille des glutathion peroxydase (GPx). (C) Autres sélénoprotéines.

Chez les mammifères, on retrouve 8 paralogues de glutathion peroxydases (GPx) parmi lesquelles 5 (GPx1, GPx2, GPx3, GPx4 et GPx6) sont des sélénoprotéines et contiennent un résidu sélénocystéine dans leur site actif. Les GPx ont un large éventail de fonctions physiologiques et sont impliquées dans la signalisation cellulaire liée au peroxyde d'hydrogène (H_2O_2), les mécanismes de détoxification des

hydroperoxides, et le maintien de l'homéostasie redox dans les cellules (pour revue, Brigelius-Flohé and Maiorino, 2013). Après réduction des hydroperoxides, les enzymes GPx oxydées sont réduites par deux molécules de glutathion (GSH) (Flohe et al., 1973; Kraus et al., 1980). GPx1 est la sélénoprotéine la plus abondante chez les mammifères et la première à avoir été identifiée (Flohe et al., 1973; Rotruck et al., 1973 et pour revue, Flohé, 2009). La réaction catalysée par GPx1 est représentée Figure 6A, et sa structure, caractérisée par un repliement de type Thioredoxin (Figure 5B). GPx1 est une enzyme ubiquitaire du cytosol, présente dans tous les types cellulaires, avec une plus forte expression dans le foie et les reins où elle participe à la détoxification du H_2O_2 . Des polymorphismes dans les gènes de GPx1 sont associés à des risques de cancer (pour revue, Chen et al., 2013; Zhuo and Diamond, 2009), cet aspect sera développé dans le paragraphe I.1.4. Des souris invalidées pour le gène de GPx1 sont bien plus sensibles au stress oxydant, lequel entraîne un taux de mortalité plus important chez ces souris (Fu et al., 1999).

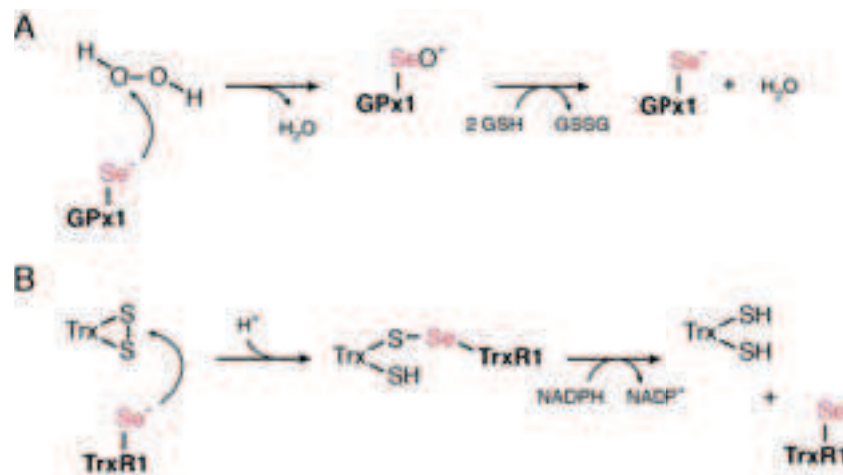


Figure 6. Réactions d'oxydo-réduction catalysées par les sélénoprotéines. (d'après Labunskyy et al., 2014; Lubos et al., 2011; Mustacich and Powis, 2000)

Cycles catalytiques de la glutathion peroxydase 1 (GPx1) (A) et de la thiorédoxine réductase 1 (TrxR1) (B). GSH, glutathion réduit ; GSSG, glutathion oxydé ; Trx, thiorédoxine ; NADPH/NADP⁺, Nicotinamide adénine dinucléotide phosphate.

Toutes les GPx tétramériques (GPx1, GPx2, GPx3 et GPx6) ont une spécificité de substrat pour le H_2O_2 ou des hydroperoxydes organiques de faible poids moléculaire (hydroperoxydes d'acides gras par exemple). GPx4 (ou PHGPx, Phospholipide Hydroperoxyde GPx) quant à elle, est une enzyme monomérique impliquée dans la réduction des phospholipides hydroperoxydés complexes tels que ceux dérivés de la phosphatidylcholine ou du cholestérol, associés aux membranes. GPx4 est un gène

essentiel requis à un stade précoce du développement embryonnaire, indiquant que sa fonction antioxydante est fondamentale (Conrad, 2009; Yant et al., 2003). La protéine GPx4 est également impliquée dans le mécanisme de maturation des spermatozoïdes et la fertilité mâle. Ainsi, GPx4 joue un rôle structural dans la formation de la pièce intermédiaire des spermatozoïdes, ce qui ne nécessite pas la fonction enzymatique de la protéine (Imai et al., 2009; Ursini et al., 1999).

Les thiorédoxine réductases (TrxR) sont des enzymes d'oxydo-réduction qui, avec la thiorédoxine, constituent le système de réduction des ponts disulfures majeur de la cellule (Arnér and Holmgren, 2000). Ces enzymes utilisent le NADPH pour réduire la thiorédoxine qui est impliquée dans un grand nombre de processus cellulaires indispensables tels que la prolifération cellulaire, la défense antioxydante et les voies de signalisation (Kondo et al.; Mukherjee and Martin, 2008). Chez les mammifères, trois isoenzymes TrxR (TrxR1, TrxR2, TrxR3) sont retrouvées et contiennent toutes une sélénocystéine dans leur domaine C-terminal, en avant-dernière position (pour revue, Arnér, 2009; Labunskyy et al., 2014; Lu and Holmgren, 2009). De nombreux variants de ces protéines sont générés par épissage alternatif permettant leur adressage vers différents organites (Rundlöf and Arnér, 2004; Sun et al., 2001a). TrxR1 est principalement localisée dans le cytoplasme et le noyau. La thiorédoxine est le substrat majeur de TrxR1 et la réaction catalytique de l'enzyme est représentée Figure 6B. L'invalidation des gènes de TrxR1 et TrxR2 est létale au stade embryonnaire, en accord avec l'implication de ces protéines dans des processus cellulaires indispensables.

La famille des iodothyronine désiodases (DIO) contient trois paralogues chez les mammifères (DIO1, DIO2, DIO3), qui sont impliqués dans la régulation de l'activité des hormones thyroïdiennes. L'activation de ces hormones est catalysée par DIO1 et DIO2 et leur inactivation par DIO3, par clivage de liaisons carbone-iode (désiodation réductive). Les hormones thyroïdiennes régulent différents processus physiologiques tels que la croissance, l'audition et influencent l'humeur et les mécanismes cognitifs (Bellinger et al., 2009; Papp et al., 2007; Ritchie and Yeap, 2015). Elles sont indispensables au développement du cerveau à l'état fœtal (Rovet, 2014). Ces protéines ont une localisation subcellulaire et une expression tissulaire distincte. Les

DIO sont des sélénoprotéines homodimériques intégralement membranaires, caractérisées par un repliement de type thiorédoxine (voir Figure 5B et C).

La méthionine-R-sulfoxide réductase 1 (MsrB1, SelR ou SelX) est une sélénoprotéine contenant un atome de zinc (Kryukov et al., 1999; Lescure et al., 1999), dont la fonction est bien caractérisée. Elle catalyse la réparation de l'énantiomère R oxydé de la méthionine dans les protéines (Kryukov et al., 2002). Localisée dans le cytoplasme et dans le noyau, son activité est plus importante dans le foie et le rein (Kim et al., 2006). La méthionine est l'un des deux acides aminés contenant un atome de soufre, et est ainsi sensible à l'oxydation. L'oxydation des méthionines dans les protéines peut conduire à des altérations significatives de leur structure perturbant la fonction des protéines. MsrB1 est donc impliquée dans les mécanismes de protection des protéines contre les dommages oxydants (Kim et al., 2006). La baisse du niveau de MsrB1, associée à des déficiences en sélénium, se traduit par une réduction de la réponse immunitaire (Lee et al., 2013; Rayman, 2012), contribuant ainsi au déclin du système immunitaire lié au vieillissement.

La sélénoprotéine P est exprimée de manière abondante et sécrétée dans le plasma sanguin où elle y représente 50% du sélénium total (Burk and Hill, 2005). Cette sélénoprotéine est particulière puisqu'elle contient 10 acides aminés sélénocystéine. SelP joue un rôle important dans le transport du sélénium vers les tissus périphériques à travers le plasma sanguin, particulièrement vers le cerveau et les testicules, préservant ainsi les fonctions de ces organes en conditions limitantes en sélénium (Burk and Hill, 2009, 2015; Hill et al., 2003; Saito and Takahashi, 2002).

Les sélénoprotéines M (SelM) et 15KDa (Sel15) sont des protéines homologues qui partagent 31% d'identité. Elles sont toutes deux présentes dans le réticulum endoplasmique (RE) et présentent un repliement de type thiorédoxine (Figure 5C) mais sont exprimées dans des tissus différents. Ces deux sélénoprotéines seraient impliquées dans la réduction ou le réarrangement des ponts dissulfures dans les protéines localisées dans ou sécrétées par le RE (Labunskyy et al., 2007). Sel15 est impliquée dans la régulation de l'homéostasie redox du RE (Labunskyy et al., 2009). SelM est hautement exprimée dans le cerveau et plusieurs études sont focalisées sur

son rôle potentiel de neuro-protecteur. La surexpression de SelM dans les neurones prévient les dommages oxydants induits par l' H_2O_2 , tandis que son inhibition par shARN induit une baisse dramatique de la viabilité des cellules (Reeves et al., 2010). De plus, sa surexpression inhibe l'agrégation de peptides β -amyloïdes lorsque ces peptides sont co-transfectés avec SelM dans les cellules HEK293T, suggérant un rôle possible de SelM dans la prévention de la maladie d'Alzheimer (Chen et al., 2013a). Enfin, SelM semble également impliquée dans la régulation du Ca^{2+} en réponse au H_2O_2 dans les neurones (Reeves et al., 2010).

La sélénoprotéine N (SelN) a été identifiée dans notre laboratoire et constitue l'une des premières sélénoprotéines identifiées par un crible bioinformatique, basé sur la recherche d'un motif structural d'ARN conservé qui caractérise l'ensemble des ARNm de sélénoprotéines (Lescure et al., 1999). Il s'agit, à ce jour, de la seule sélénoprotéine connue pour être impliquée dans une maladie génétique (Moghadaszadeh et al., 2001; Rederstorff et al., 2006). Différentes mutations dans le gène de SelN sont associées à des types de dystrophies musculaires distincts (Lescure et al., 2008), indiquant que SelN est impliquée dans le développement musculaire. Cependant, la fonction biochimique précise de SelN reste à être déterminée. SelN est glycosylée, localisée dans la membrane du RE présente un repliement de type thiorédoxine et possède un motif « EF-hand » (motif de liaison au calcium). Des avancées récentes proposent que SelN participe à l'homéostasie oxydative et calcique avec un rôle potentiel dans la régulation de l'activité du récepteur ryanodine du RE (Castets et al., 2012). Ainsi, SelN serait impliquée dans le transport de calcium cellulaire, requis pour le développement des muscles (Lescure et al., 2008).

1.1.4 Sélénoprotéines, stress oxydant et cancer

Les stress oxydants jouent un rôle majeur dans la carcinogenèse en induisant des mutations dans l'ADN ainsi qu'en affectant les voies de transduction intracellulaires (Frederiks et al., 2010) (Figure 7). Les espèces réactives oxygénées (ROS) sont des messagers secondaires fondamentaux qui sont impliqués dans des cascades de signalisation et régulent l'état redox des cellules. En excès, ils sont à l'origine des stress oxydants et peuvent induire et maintenir le phénotype oncogène des cellules cancéreuses (Sengottuvelan et al., 2009) ou activer les mécanismes de sénescence et

d'apoptose. De plus, les ROS peuvent induire des dommages à l'ADN qui ont été rapportés à des dysfonctions de gènes impliqués dans la genèse et le développement de cancers. Ainsi, les cellules ont mis en place des mécanismes de régulation ; des enzymes aux fonctions antioxydantes qui les protègent notamment des stress oxydants induits par les ROS et leur permettent de rétablir ou maintenir l'homéostasie redox (Zhuo and Diamond, 2009b) (Figure 7).

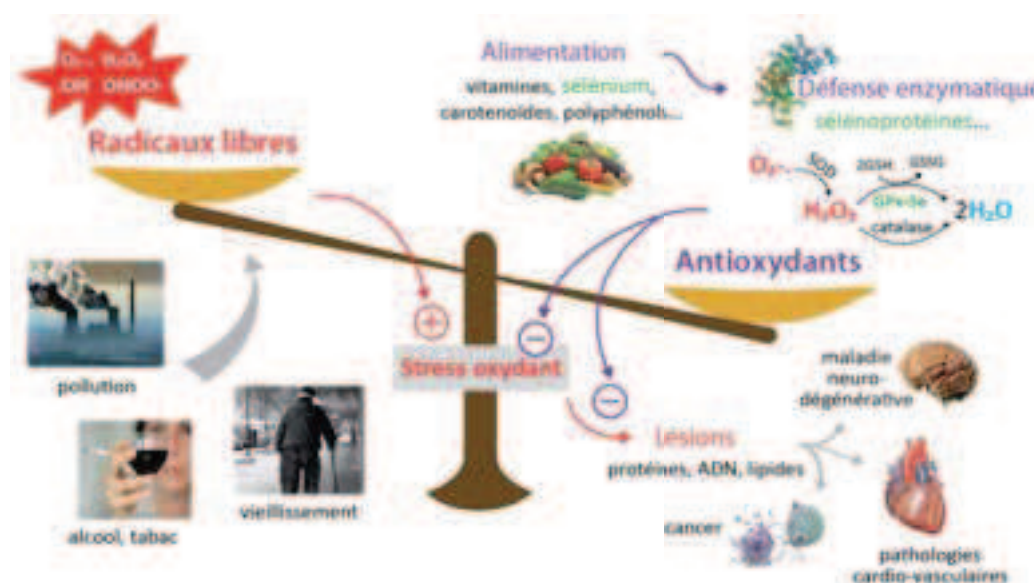


Figure 7. Stress oxydant et sélénoprotéines

Les facteurs environnementaux tels que la pollution, l'alcool et le tabac induisent l'accumulation de radicaux libres oxygénés dans les cellules menant au stress oxydant. Les stress oxydants peuvent provoquer des mutations dans l'ADN ou induire des lésions dans les protéines et les lipides, à l'origine de diverses pathologies neuro-dégénératives, cardio-vasculaires ou de cancers par exemple. Une alimentation équilibrée permet de fournir les nutriments nécessaires à la réponse antioxydante des cellules. Les sélénoprotéines, de par leur puissante activité redox permettent de contrecarrer les stress oxydants et de maintenir le potentiel redox des cellules.

Les sélénoprotéines, de par leur important pouvoir réducteur, constituent l'une des classes d'enzymes qui jouent un rôle essentiel dans la protection antioxydante des cellules (Allmang et al., 2009; Rayman, 2000; Zhuo and Diamond, 2009b). Ces dernières années, de nombreuses méta-analyses faisant le lien entre le sélénium dans l'alimentation et les risques de cancer ont été réalisées. D'après certaines de ces analyses, le sélénium, présenterait des propriétés préventives contre plusieurs types de cancers (Roman et al., 2014) tels que le cancer de la prostate (Brinkman et al., 2006; Hurst et al., 2012), le cancer colorectal (Ou et al., 2012), le cancer du poumon (Zhuo et al., 2004), le cancer de la vessie (Amaral et al., 2010) et les cancers de manière générale (Bardia et al., 2008; Lee et al., 2011; Persson-Moschos et al., 2000). L'action

antioxydante du sélénium est liée à son incorporation dans les selenoprotéines, et de nombreuses études génétiques et biochimiques analysant l'expression de ces enzymes dans les tissus cancéreux et les lignées cellulaires ont été réalisées (Hatfield et al., 2014; Zhuo and Diamond, 2009b). Des changements de niveau d'expression des selenoprotéines dans les tumeurs en comparaison aux tissus sains ont fréquemment été observés (Gonzalez-Moreno et al., 2011; Yoo et al., 2006). Cependant, il est difficile de faire la distinction entre les différences d'expression qui sont réellement à l'origine du développement de cancers (facteurs de causalité), de celles qui sont uniquement une conséquence de l'état transformé des cellules, typiquement associé à la dérégulation d'un large nombre de gènes. En revanche, des données génétiques impliquent certaines selenoprotéines dans des risques de cancer. Le tableau 1, adapté de Zhuo and Diamond (2009b), résume ces études.

Sélénoprotéine	Position (chromosome)	Variation allélique	Localisation	Cancer associé	Références bibliographiques
SePP	5q31	G/A (rs3877899) G/A (rs7579) C/G A/G (rs12055266) A/G (rs3797310) C/T (rs2972994)	CS 3'NC 5'NC 3'NC 3'NC 3'NC	Prostate Prostate Colorectal Colorectal Colorectal Colorectal	(Cooper et al., 2008) (Méplan et al., 2007) (Peters et al., 2008) (Peters et al., 2008) (Peters et al., 2008) (Peters et al., 2008)
GPx1	3p21.3	C/T (rs1050450) Répétition de GCG	CS CS	Poumon Sein Vessie Foie Lymphome Prostate Sein Prostate Peau	(Lee et al., 2006; Raaschou-Nielsen et al., 2007; Ratnasinghe et al., 2000; Yang et al., 2004) (Cox et al., 2006; Ravn-Haren et al., 2006) (Ichimura et al., 2004; Zhao et al., 2005a) (Sutton et al., 2006) (Lightfoot et al., 2006) (Arsova-Sarafinovska et al., 2009) (Knight et al., 2004) (Kote-Jarai et al., 2002) (Jefferies et al., 2005)
GPx4	19p13.3	C/T (rs713041)	3'NC	Colorectal	(Bermano et al., 2007; Méplan et al., 2008)
Sel15	1p31	C/T G/A	3'NC 3'NC	Poumon Poumon	(Jablonska et al., 2008) (Jablonska et al., 2008)
SelS		rs34713741	5'NC (prom)	Colorectal	(Méplan et al., 2010)
SelS		G/A (rs28665122)	5'NC (prom)	Estomac	(Shibata et al., 2009)

Tableau 1. Etudes révélant une association entre des polymorphismes dans les gènes de selenoprotéines et le risque de cancer. CS : séquence codante, NC : séquence non-codante

D'autres études lient sélénoprotéines et cancer. En effet, des modifications épigénétiques au niveau du promoteur du gène de GPx3, conduisant à l'inhibition de son expression, sont fréquemment retrouvées dans le cancer de la prostate (Jones and Baylin, 2002; Lodygin et al., 2005). La sélénoprotéine GPx2 est surexprimée dans certains types de cancer ; particulièrement les cancers colorectaux (Emmink et al., 2014). Enfin, TrxR1 est probablement la sélénoprotéine la plus étudiée pour son lien avec le cancer : elle régule l'état redox de la cellule et active le suppresseur de tumeur p53 (Cassidy et al., 2006). Elle est surexprimée dans plusieurs types cellulaires cancéreux et lorsque son expression est affectée, elle altère la morphologie des cellules (Yoo et al., 2006). Par ailleurs, il a été proposé qu'en plus de son rôle préventif contre le cancer, le sélénium pourrait être un élément anti-métastatique (Chen et al., 2013b). Des composants sélénés ont spécialement été conçus pour cibler des cellules tumorales (Misra et al., 2015) et le sélénium pourrait potentiellement être utilisé dans certains traitements oncologiques.

I.2 La synthèse des sélénoprotéines

Parce que le codon sélénocystéine est incorporé en réponse à un codon UGA, habituellement reconnu comme un codon stop, la traduction des sélénoprotéines fait appel à un mécanisme complexe de recodage. Ce mécanisme requiert la présence d'éléments d'ARN agissant en *cis*, et une machinerie spécialisée agissant en *trans* (Allmang et al., 2009; Papp et al., 2007; Squires and Berry, 2008). La biogenèse des sélénoprotéines comporte deux étapes : la synthèse de la sélénocystéine et son incorporation dans les sélénoprotéines.

I.2.1 La biosynthèse de la sélénocystéine

La voie de biosynthèse de la sélénocystéine est un processus original puisqu'il s'agit de l'unique acide aminé directement synthétisé sur son ARNt, appelé ARNt^{[Ser]Sec} (Lee et al., 1989). Ce processus est réalisé en deux étapes : l'ARNt^{[Ser]Sec} est d'abord amino-acylé par une sérine lors d'une réaction catalysée par la séryl-ARNt synthétase (SerRS) pour former un Ser-ARNt^{[Ser]Sec}, puis l'acide aminé sérine est transformé en sélénocystéine (Sheppard et al., 2008).

1.2.1.1 L'ARN de transfert sélénocystéine (ARNt^{[Ser]Sec})

L'ARNt spécifique de la sélénocystéine, l'ARNt^{[Ser]Sec}, reconnaît le codon UGA^{Sec} par l'intermédiaire de son anticodon complémentaire UCA. Des modèles de structures secondaires et tertiaires des ARNt^{[Ser]Sec} bactériens, d'archées et d'eucaryotes ont été proposés, basés sur des données expérimentales (Baron and Böck, 1991; Hubert et al., 1998; Sturchler et al., 1993). Ces modèles révèlent des caractéristiques structurales communes à l'ensemble des l'ARNt^{[Ser]Sec} qui les distinguent des autres ARNt (pour revue, Allmang and Krol, 2006). Chez les eucaryotes, l'ARNt^{[Ser]Sec} est le plus long des ARNt avec 90 nucléotides, comparé aux 76 nucléotides des ARNt canoniques. Il présente un taux de bases modifiées relativement faible et une boucle variable exceptionnellement longue (Diamond et al., 1981, 1993). Les ARNt^{[Ser]Sec} sont caractérisés par une tige du bras D de 6 paires de bases (pb), au lieu de 3-4 pb dans les ARNt canoniques ; cette différence est requise pour la phosphorylation de la sérine, un intermédiaire lors de la synthèse de la sélénocystéine chez les eucaryotes (Carlson et al., 2004; Itoh et al., 2009a; Kaiser et al., 2005; Wu and Gross, 1994). Ces deux paires de bases additionnelles remplacent deux interactions tertiaires (8-14 et 15-48) universellement présentes dans tous les autres ARNt cytosoliques. Elles apportent de la flexibilité et de la stabilité à l'ARNt^{[Ser]Sec}, mais ne seraient pas essentielles à sa fonction (Ishii et al., 2013). Le bras accepteur de l'acide aminé (A-T) de l'ARNt^{[Ser]Sec} se distingue également par une taille plus importante, 13 pb (9+4) contre 12 pb (7+5) dans les autres ARNt (Hubert et al., 1998; Sturchler et al., 1993) (Figure 8), ce qui représente un déterminant pour la reconnaissance de l'ARNt^{[Ser]Sec} par le facteur d'élongation spécifique de la sélénocystéine chez les bactéries (Baron & Bock, 1991). Les structures cristallographiques de l'ARNt^{[Ser]Sec} humain (Itoh et al., 2009a; Palioura et al., 2009) et bactérien (Itoh et al., 2013) ont été résolues et révèlent des structures secondaires inhabituelles et des interactions tertiaires uniques. Ces données structurales démontrent que le bras D inhabituellement long n'interagit pas avec la boucle variable, ce qui résulte en une cavité ouverte au lieu d'un cœur hydrophobe (Itoh et al., 2009a). Le bras variable possède une orientation particulière ce qui induit une conformation unique de l'ARNt^{[Ser]Sec}. Ces caractéristiques structurales permettent à l'ARNt^{[Ser]Sec} d'être reconnu par les facteurs de biogenèse de la sélénocystéine et les facteurs d'incorporation de la sélénocystéine dans les sélénoprotéines (Itoh et al., 2009a).

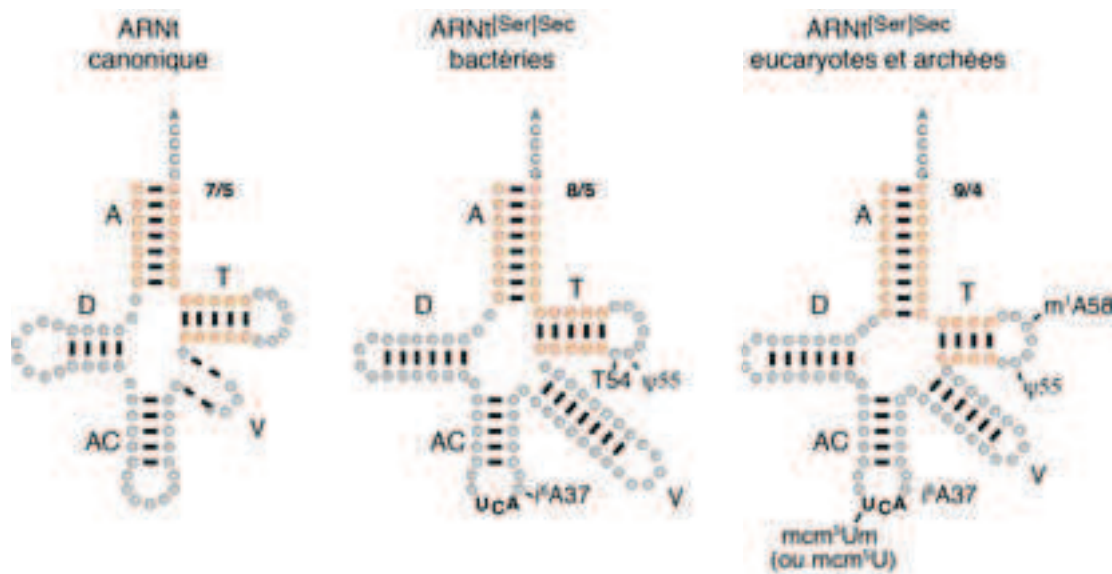


Figure 8. Structure secondaire d'un ARNt canonique, d'un ARNt^{[Ser]Sec} de bactérie et d'un ARNt^{[Ser]Sec} d'eucaryote ou d'archée. (d'après Allmang and Krol, 2006)

A: bras accepteur ; D: bras dihydrouracile ; AC: bras de l'anti-codon ; T: bras de la thymine. 7/5, 8/5 et 9/4: nombre de paires de bases qui forment les bras A/T (en orange). Les pointillés dans le bras additionnel des ARNt canoniques reflètent la variabilité de sa longueur. Les bases modifiées sont indiquées : i6A37, T54, Ψ55, m1A58 et mcm5Um.

Chez les eucaryotes, des modifications post-transcriptionnelles n'ont été trouvées que sur 4 bases de l'ARNt^{[Ser]Sec}, tandis qu'une moyenne de 11 à 13 modifications par ARNt canonique est habituellement trouvée (Diamond et al., 1993; Sturchler et al., 1994). Une pseudouridine (pseudo-U55) et une 1-méthyl-adénosine (m¹A58) sont trouvées dans la boucle du bras T. Une isopentényl-adénosine (i⁶A37) et une 5-méthylcarboxyméthyluridine-2'-O-méthylribose (mcm⁵Um34) sont localisées en position 37 et 34 dans la boucle de l'anticodon. La modification mcm⁵Um34 est une caractéristique unique de l'ARNt^{[Ser]Sec}, puisqu'elle n'a été observée dans aucun autre ARNt (Diamond et al., 1993) (Figure 8). Deux isoformes majeurs de l'ARNt^{[Ser]Sec}, contenant chacun soit la modification mcm⁵U soit la modification mcm⁵Um en position 34 existent chez les mammifères, et leur distribution relative est influencée par le niveau de sélénium (Chittum et al., 1997; Diamond et al., 1993). La protéine SECp43, retrouvée associée à l'ARNt^{[Ser]Sec} semble être à l'origine de la modification en position 34 (Ding and Grabowski, 1999; Xu et al., 2005). La présence ou non de cette modification serait un moyen de régulation de l'expression des sélénoprotéines puisqu'il destinerait l'ARNt^{[Ser]Sec} à la traduction d'ARNm différents (Carlson et al., 2007; Howard et al., 2013). L'inactivation du gène de l'ARNt^{[Ser]Sec} dans les souris est

létale au stade embryonnaire, révélant l'importance de cet ARNt et le caractère essentiel des sélénoprotéines (Bosl et al., 1997; Carlson et al., 2005).

1.2.1.2 Le séryl-ARNt^{[Ser]Sec} et la transformation de la sérine en phosphosérine

La biosynthèse de la sélénocystéine à partir de la sérine se fait directement sur l'ARNt^{[Ser]Sec} et est un processus en 2 étapes. Lors de la première étape de la synthèse, la sérine-ARNt-synthétase conventionnelle (SerRS) charge l'acide aminé sérine sur l'ARNt^{[Ser]Sec}. Les boucles variables des ARNt^{Ser} et ARNt^{[Ser]Sec} présentent une orientation similaire, ce qui permet leur reconnaissance par la SerRS (Itoh et al., 2009a). Chez les bactéries, la sélénocystéine synthase (SelA) convertit ensuite directement le Ser-ARNt^{[Ser]Sec} en Sec-ARNt^{[Ser]Sec}. Chez les archées et les eucaryotes, une voie alternative en deux étapes est employée pour synthétiser la sélénocystéine (figure 9).

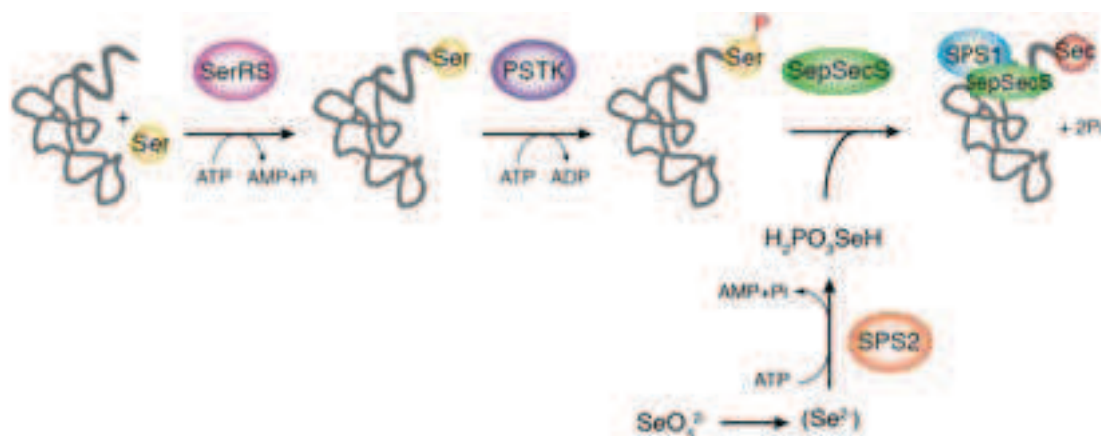


Figure 9. Voie de biosynthèse de la sélénocystéine chez les eucaryotes.

L'acide aminé sérine (Ser) est chargé sur l'ARNt^{[Ser]Sec} par la sérine-ARNt-synthétase (SerRS). La phosphoséryl-ARNt^{[Ser]Sec} kinase (PSTK) phosphoryle le résidu séryl, qui est ensuite transformé en sélénocystéine par la sélénocystéine synthase (SepSecS) aboutissant à la formation du Sec-ARNt^{[Ser]Sec}. Pour cela, la SepSecS utilise du monosélénophosphate ($\text{H}_2\text{PO}_3\text{SeH}$), un donneur de sélénium généré à partir de sélénure (Se^{2-}) par la sélénophosphate synthétase 2 (SPS2). L'ensemble de ces enzymes utilise de l'ATP lors des réactions catalytiques.

Dans un premier temps, le résidu séryl du Ser-ARNt^{[Ser]Sec} est phosphorylé en présence d'ATP et de Mg^{2+} par la O-phosphoséryl-ARNt^{[Ser]Sec} kinase (PSTK) qui reconnaît spécifiquement le Ser-ARNt^{[Ser]Sec} (Carlson et al., 2004; Kaiser et al., 2005; Sherrer et al., 2008). La PSTK reconnaît la conformation unique du bras D, qui, contrairement aux bras D des ARNt canoniques, est déplacé vers l'extérieur de l'axe formé par le

bras de la boucle de l'anticodon et le bras T (Itoh et al., 2009a). Le phosphoséryl-ARNt^{[Ser]Sec} (Sep-ARNt^{[Ser]Sec}) est ensuite converti en Sec-ARNt^{[Ser]Sec}.

I.2.1.3 La transformation du O-phosphoséryl-ARNt^{[Ser]Sec} en sélénocystéinyl-ARNt^{[Ser]Sec}

I.2.1.3.a La Sélénocystéine synthase

La sélénocystéine synthase (SepSecS) catalyse la dernière étape de formation de la sélénocystéine : la conversion du Sep-ARNt^{[Ser]Sec} en Sec-ARNt^{[Ser]Sec} (Araiso et al., 2008; Xu et al., 2007b; Yuan et al., 2006). Elle substitue le phosphate de la phosphosérine présent dans l'intermédiaire Sep-ARNt^{[Ser]Sec}, par un groupement sélénol, en utilisant le phosphate de pyridoxal comme cofacteur (Araiso et al., 2008; Ganichkin et al., 2008; Palioura et al., 2009). La structure cristallographique de la SepSecS humaine en complexe avec le Sep-ARNt^{[Ser]Sec} et du thiophosphate, un analogue du monosélénophosphate, a été résolue à 2.8 Å (Palioura et al., 2009) et révèle le mécanisme de reconnaissance de l'ARNt^{[Ser]Sec} par l'enzyme. La SepSecS forme un homotétramère liant 2 ARNt^{[Ser]Sec} au niveau des 13 pb formées par la co-axialité des hélices du bras accepteur et du bras T (structure 9/4 spécifique) (Figure 10). L'oligomérisation de la SepSecS sous forme d'homotétramère constitue le complexe physiologique et catalytiquement actif, et suggère une régulation allostérique qui pourrait jouer un rôle important dans la régulation de la synthèse de sélénocystéine (French et al., 2014). La SepSecS est un membre de la famille des enzymes à phosphate de pyridoxal qui adoptent un repliement de type 1 similaire à celui de la sélénocystéine synthase bactérienne (SelA) (Araiso et al., 2008; Ganichkin et al., 2008; Palioura et al., 2009). L'inactivation de la SepSecS chez les souris conduit à une létalité au stade embryonnaire (Wirth et al., 2010). Chez l'homme, deux mutations homozygotes dans le gène de SepSecS qui ont un impact sur le repliement et l'activité catalytique de la protéine ont été observés et sont associés à une atrophie cérébrale progressive (Agamy et al., 2010). De manière intéressante, il a récemment été montré que SPS2 et SepSecS étaient également impliquée dans la biosynthèse de la cystéine et sont capables de transformer le Sep-ARNt^{[Ser]Sec} en Cys-ARNt^{[Ser]Sec}, menant à une incorporation de cystéine à la place de la sélénocystéine dans la sélénoprotéine TrxR1 de souris, lorsque les conditions en sélénium sont limitantes (Xu et al., 2010b).

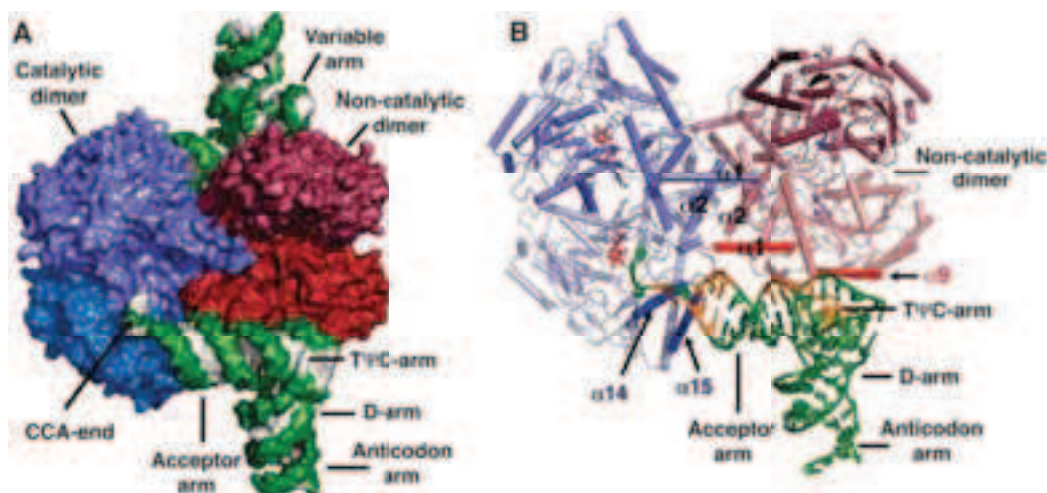


Figure 10. Structure cristallographique de la sélénocystéine synthase (SepSecS) en complexe avec l'ARNt^{[Ser]Sec} non aminoacylé. (Figure de Palioura et al., 2009)

(A) Représentation du complexe physiologique, sous forme d'homotétramère. Les sous-unités du dimère catalytique de SepSecS sont en bleu foncé et bleu clair, celles du dimère non-catalytique en rouge foncé et rouge clair. Le squelette sucre-phosphate et les bases de l'ARNt^{[Ser]Sec} sont en vert et gris respectivement. (B) Le dimère catalytique interagit avec le bras accepteur de l'ARNt^{[Ser]Sec} via les hélices α14 and α15 de la SepSecS (bleu). L'hélice α1 d'une SepSecS du dimère non-catalytique interagit avec le bras accepteur et le bras T. La région de l'ARNt^{[Ser]Sec} en interaction avec la SepSecS est indiquée en orange.

I.2.1.3.b La sélénophosphate synthétase et le donneur de sélénium

Chez les bactéries, la sélénocystéine synthase (SelA) utilise du monosélénophosphate comme source de sélénium (Böck, 2001) pour convertir le Ser-ARNt^{[Ser]Sec} en Sec-ARNt^{[Ser]Sec} (Ganichkin et al., 2008). Ce donneur de sélénium est produit par la sélénophosphate synthétase (SelD) à partir d'ATP et de sélénium réduit (Se²⁻) (Itoh et al., 2009b). Deux homologues de SelD ont été identifiés chez les eucaryotes : les sélénophosphate synthétase 1 et 2 (SPS1 et SPS2) (Guimaraes et al., 1996; Low et al., 1995). SPS2 catalyse la synthèse de sélénophosphate *in vitro* et, contrairement à SPS1, son inhibition par siARN dans des cellules NIH3T3 entrave la biosynthèse des sélénoprotéines (Xu et al., 2007a, 2007b). Ainsi, SPS2 est essentielle à la génération de donneur de sélénium pour la biosynthèse de la sélénocystéine, tandis que SPS1 pourrait avoir un rôle plus spécialisé et non-essentiel dans le métabolisme des sélénoprotéines (Figure 9). SPS2 est elle-même une sélénoprotéine et pourrait servir d'autorégulateur dans la synthèse des sélénoprotéines (Guimaraes et al., 1996).

1.2.2 Incorporation de sélénocystéine dans les sélénoprotéines

L'incorporation co-translationnelle de sélénocystéine dans la chaîne peptidique est dictée par la présence d'un codon UGA^{Sec} présent dans la séquence codante de l'ARNm de sélénoprotéine. Ce mécanisme complexe requiert des facteurs protéiques spécialisés et le Sec-ARNt^{[Ser]Sec} agissant en *trans*, ainsi que la présence d'une structure en tige-boucle, appelée SECIS (pour SElenoCysteine Insertion Sequence), dans l'ARN agissant en *cis*. Lorsque le ribosome rencontre le codon UGA^{Sec}, qui habituellement signale la terminaison de la traduction, la machinerie d'incorporation de la sélénocystéine interagit avec la machinerie traductionnelle canonique pour permettre le recodage du codon UGA^{Sec} et empêcher une terminaison prématurée. Ce mécanisme est différent chez les bactéries et les eucaryotes.

Chez les bactéries, l'élément SECIS est localisé en aval du codon UGA^{Sec} dans la région codante des ARNm de sélénoprotéines (Böck et al., 2006; Stock and Rother, 2009; Yoshizawa and Böck, 2009). Un facteur d'élongation de la traduction spécialisé, SelB, interagit spécifiquement avec le Sec-ARNt^{[Ser]Sec} et avec l'élément SECIS. De cette façon, le Sec-ARNt^{[Ser]Sec} est amené directement vers le codon UGA^{Sec} et sera placé dans le site A du ribosome au moment du recodage ce qui permettra l'incorporation de la sélénocystéine dans la chaîne peptidique (pour revue, Böck et al., 2006; Stock and Rother, 2009; Yoshizawa and Böck, 2009). Un modèle de ce mécanisme est représenté sur la Figure 11A.

Chez les eucaryotes, l'ensemble des facteurs nécessaires au recodage du codon UGA^{Sec} sont plus nombreux que chez les bactéries et ne sont pas encore tous identifiés (Allmang et al., 2009). Toutefois, il est déjà clairement établi que l'assemblage des mRNP de sélénoprotéines et la régulation de leur expression requiert un nombre plus élevé de facteurs qu'initialement anticipé (pour revue Allmang et al., 2009). Contrairement aux bactéries, l'élément SECIS est localisé dans la région 3'UTR (untranslated region) des ARN messagers chez les eucaryotes (Figure 11B). De plus, le mode de recrutement du facteur d'élongation eucaryote eEFSec est différent de celui de SelB bactérien, puisque eEFSec ne peut interagir directement avec l'élément SECIS (pour revue, Allmang and Krol, 2006; Allmang et al., 2009). En revanche, le SECIS est reconnu par la protéine SBP2 (SECIS Binding Protein 2) qui permet de

recruter eEFSec lorsque celui-ci est en complexe avec le Sec-ARNt^{[Ser]Sec}. La liaison de SBP2 au SECIS induirait des changements conformationnels au sein de eEFSec permettant une interaction fonctionnelle avec le ribosome, ce qui conduirait à l'accommodation du Sec-ARNt^{[Ser]Sec} dans le site A du ribosome (Gonzalez-Flores et al., 2012). La protéine ribosomique L30 semble également jouer un rôle dans ce mécanisme ; en compétition avec SBP2 pour la liaison à l'ARN SECIS, elle permettrait de rapprocher le SECIS du site A du ribosome durant le mécanisme de recodage (Chavatte et al., 2005). D'autres protéines seraient impliquées dans la régulation de ce mécanisme (voir paragraphe I.2.4.2.a).

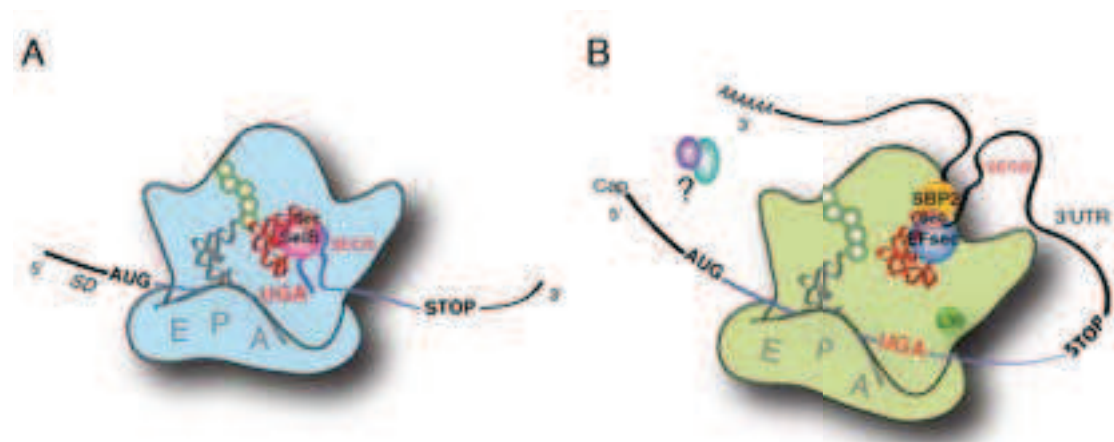


Figure 11. Facteurs *cis* et *trans* nécessaires à la synthèse des sélénoprotéines chez les bactéries et les eucaryotes.

Les ARNm de sélénoprotéines comportent un codon UGA^{Sec} et une structure secondaire en tige-boucle appelée SECIS (SEC Insertion Sequence). (A) Chez les bactéries, l'élément SECIS est localisé dans la séquence codante, juste en aval du codon UGA^{Sec}. L'ARN SECIS est spécifiquement reconnu par le facteur SelB qui interagit également avec le Sec-ARNt^{[Ser]Sec}. Ceci permet de positionner le Sec-ARNt^{[Ser]Sec} directement au niveau du codon UGA^{Sec} et dans le site A du ribosome. (B) Chez les eucaryotes, l'élément SECIS est situé dans la région 3'UTR des ARNm. La protéine SBP2 reconnaît l'élément SECIS et recrute le facteur d'élongation spécifique eEFSec associé au Sec-ARNt^{[Ser]Sec}. La protéine ribosomique L30 joue également un rôle dans le mécanisme et sans doute d'autres facteurs encore inconnus.

I.2.2.1 Les ARNm de sélénoprotéines

Les ARNm de sélénoprotéines se caractérisent par la présence du codon UGA^{Sec} et de l'élément SECIS indispensable au mécanisme de recodage (Berry et al., 1991). Cet élément diffère au niveau de la séquence et de la structure chez les bactéries, les archées et les eucaryotes. De plus, chez les bactéries, l'élément SECIS se trouve directement en aval du codon UGA^{Sec} (Böck et al., 2006), tandis qu'il est localisé dans la région 3' non-traduite (3'UTR) des ARN messagers des eucaryotes et des archées (Berry et al., 1991) (Figure 11). Chez ces organismes, la distance entre l'élément SECIS et le codon UGA^{Sec} peut varier de 300 à plusieurs milliers de nucléotides

(Berry et al., 1993). Une distance minimale de 51-111 nucléotides et cependant requise entre le codon UGA^{Sec} et l'élément SECIS pour que l'évènement de recodage se produise (Martin et al., 1996). La localisation du SECIS en 3'UTR, c'est à dire hors de la région codante, permet aux eucaryotes et archées de s'affranchir des contraintes de codage, ce qui présenterait un avantage évolutif (Berry et al., 1993). Pour compenser l'éloignement entre le SECIS et le codon UGA^{Sec} et apporter le Sec-ARNt^{[Ser]^{Sec}} au site A du ribosome ces organismes ont développé un mécanisme complexe. La majorité des ARNm de sélénoprotéines ne présente qu'un seul élément SECIS dans la région 3'UTR. Cependant, deux éléments SECIS situés en 3'UTR sont capables d'assurer le recodage de plus de 10 codons UGA^{Sec} localisés dans l'ARNm de la sélénoprotéine P (Hill et al., 1991, 1993; Shetty et al., 2014). Il a été montré que le premier codon UGA^{Sec} est décodé par le second élément SECIS et que le premier SECIS est requis pour le recodage des autres UGA^{Sec} en aval (Gupta and Copeland, 2007; Shetty et al., 2014; Stoytcheva et al., 2006a). De manière intéressante, un élément SECIS a été identifié dans la phase codante de l'ARN messager de la sélénoprotéine GPx4 du virus de la variole aviaire, et permet l'incorporation de sélénocystéine dans les cellules de mammifères (Mix et al., 2007).

I.2.2.1.a L'élément SECIS (SElenoCysteine Insertion Sequence)

Les éléments SECIS eucaryotes présentent des séquences très variables mais adoptent une structure secondaire en tige-boucle conservée. Les caractéristiques structurales spécifiques des éléments SECIS eucaryotes dérivées d'analyses conformationnelles en solution et de mutagenèse dirigée (Berry et al., 1993; Martin et al., 1998; Shen et al., 1995; Walczak et al., 1996, 1998) sont représentées Figure 12. Les ARN SECIS adoptent une structure en tige-boucle formée de deux hélices séparées d'une boucle interne. Le pied de l'hélice II est constitué de quatre paires de bases non Watson-Crick consécutives (le quartet de paires de bases) ; ce motif constitue le cœur du SECIS et est essentiel pour l'incorporation de sélénocystéine *in vivo* (Walczak et al., 1996, 1998) (Figure 12). Deux éléments SECIS de structures légèrement différentes sont retrouvés chez les eucaryotes : la forme 1 et la forme 2.

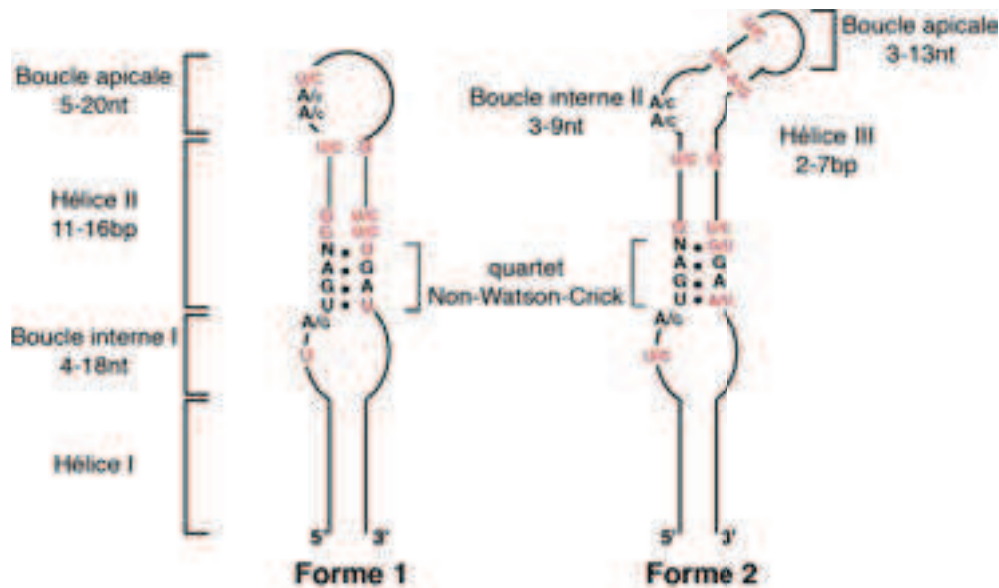


Figure 12. Modèles de structure secondaire de l'élément SECIS de forme 1 et de forme 2 chez les eucaryotes.

Les éléments de structure et de séquences conservés sont représentés. Les résidus conservés identifiés par le programme SECISaln sont indiqués en rouge (Chapple et al., 2009a). N: A, T, G ou C. A/G et A/C indiquent que A est la base majoritaire retrouvée à cette position. (Fagegaltier et al., 2000a; Grundner-Culemann et al., 1999; Walczak et al., 1996, 1998).

La forme 1 est composée de deux hélices (I et II) séparées par une boucle interne et surmontée d'une boucle apicale. La forme 2 se distingue de la forme 1 par la présence additionnelle d'une hélice (III) et d'une boucle interne (II) et par une boucle apicale plus courte (Fagegaltier et al., 2000a; Grundner-Culemann et al., 1999). 62 éléments SECIS de forme 1 contre 224 de forme 2 sont retrouvés dans le sélénoprotéome eucaryote par alignement de séquences (Chapple et al., 2009a). Parmi les SECIS de forme 1 on peut citer ceux des ARNm de GPx1 et DIO1 par exemple, les SECIS de SelM, TrxR2, SelT, SelO sont quant à eux de forme 2. Les deux éléments de l'ARN SECIS essentiels à l'incorporation de sélénocystéine dans les sélénoprotéines *in vivo* sont le quartet de paires de bases non Watson-Crick, et un motif conservé AAR dans la boucle apicale (Walczak et al., 1996, 1998). Les deux paires de bases G-A/A-G en tandem, au centre du quartet de paires de bases non Watson-Crick, sont parfaitement conservées, adoptent une conformation de type "sheared" (Fagegaltier et al., 2000a; Walczak et al., 1996, 1998) et représentent le site de liaison de la protéine SBP2, indispensable au mécanisme de recodage. Après sa découverte dans les éléments SECIS, l'existence des tandems G-A/A-G a été établie dans les ARN ribosomiques, les snARN et les snoARN. Ce motif adopte une structure tertiaire conservée appelée "kink-turn" ou "K-turn" (Klein et al., 2001). Le "K-turn" canonique est un motif

structural d'environ 15 nucléotides. Il se caractérise par la présence d'une hélice-boucle interne-hélice, un résidu protubérant et deux paires de bases non-Watson-Crick G-A/A-G qui adoptent une conformation de type "sheared" (Lescoute et al., 2005). Ces éléments induisent un coude important dans le squelette sucre-phosphate et une différence d'orientation de 120° entre les axes des hélices adjacentes (Vidovic et al., 2000). Les structures en K-turn sont capables de lier les protéines d'une famille très spécifique appelée L7Ae dont SBP2 fait partie et qui fait l'objet du paragraphe I.2.2.2.b.1. L'élément SECIS présente un repliement secondaire très similaire et nous avons proposé qu'il s'agit vraisemblablement d'un variant de K-turn (Allmang & Krol, 2006). Une mutation dans le quartet de paires de bases non Watson-Crick de l'élément SECIS de l'ARNm de la sélénoprotéine N, provoque chez l'homme une dystrophie musculaire congénitale (Allamand et al., 2006). Dans ce cas, le U de la première paire de base non-Watson-Crick est muté en C, ce qui inhibe l'interaction SECIS-SBP2, provoquant une diminution de la synthèse de la sélénoprotéine N. Jusqu'à présent, ceci constitue le seul exemple de mutation dans l'élément SECIS impliqué dans une maladie génétique humaine. Le deuxième élément de l'ARN SECIS essentiel au mécanisme de recodage du codon UGA^{Sec} est le motif conservé AAR, localisé dans la boucle apicale du SECIS de forme 1 et dans la boucle interne de la forme 2 (Figure 12). Ces deux A conservés sont accessibles au solvant ce qui laisse suggérer qu'ils interagissent avec un ligand qui reste à être identifié (Fagegaltier et al., 2000a). Des expériences de mutagenèse dirigée ont montré qu'ils constituent un élément important pour la synthèse des sélénoprotéines *in vivo* (Berry et al., 1993). De manière intéressante, certains ARNm de sélénoprotéines tels que SelM et SelO de mammifère, contiennent un élément SECIS avec des nucléotides CC à la place des AA dans le motif AAR (Korotkov et al., 2002a; Kryukov et al., 2003b). Ceci suggère que la boucle apicale ne serait pas impliquée dans la liaison à des protéines mais pourrait aider au recodage en assistant la liaison de eEFSec au ribosome ou l'accommodation du Sec-ARNt^{[Ser]Sec} par exemple (Labunskyy et al., 2014). L'ARN SECIS présente une série de déterminants qui permettent sa reconnaissance spécifique par la protéine SBP2 (Cléry et al., 2007) et influencent l'efficacité de recodage (Latreche et al., 2009), ils seront discutés plus en détail dans le paragraphe I.2.2.2.b.2. L'élément SECIS interagirait directement avec les protéines ribosomiques (Chavatte et al., 2005; Kossinova et al., 2013).

Le développement d'outils bioinformatiques a permis de grandes avancées dans l'identification des sélénoprotéines grâce aux données de conservation de séquences et de structure des éléments SECIS (Castellano et al., 2008; Chapple et al., 2009b; Kryukov et al., 1999, 2003a; Lescure et al., 1999, 2002a). Le logiciel SECISearch a permis d'identifier le sélénoprotéome complet des mammifères (Kryukov et al., 2003a) et de prédire la présence d'éléments SECIS dans une séquence (Mariotti et al., 2013). SelenoDB (<http://www.selenodb.org>) est une base de donnée très utile, qui fournit des annotations de gènes de sélénoprotéines et d'élément SECIS (Castellano et al., 2008; Romagné et al., 2014). Enfin, le logiciel SECISaln (Chapple et al., 2009a) permet de prédire, par des alignements de séquence basés sur les données structurales, la structure secondaire d'élément SECIS d'une séquence donnée. Ces outils ont permis de prédire les structures secondaires des SECIS de sélénoprotéines nouvellement identifiées et ont mis en évidence l'existence d'autres nucléotides conservés dans les éléments SECIS chez les eucaryotes (Figure 12).

1.2.2.1.b L'élément SRE (Selenocysteine Redefinition Element)

Un autre élément, appelé SRE (Selenocysteine Redefinition Element), initialement identifié dans l'ARNm de SelN, agirait en *cis* et serait impliqué dans l'incorporation de sélénocystéine dans certaines sélénoprotéines eucaryotes (Howard et al., 2005a, 2007a). Il s'agit d'une structure en tige-boucle de séquence conservée localisée dans la région codante, environ 6 nucléotides en aval du codon UGA^{Sec}, un peu à la manière des éléments SECIS bactériens. Cet élément est suffisant pour stimuler à lui-seul le recodage du codon UGA^{Sec} de SelN bien que cette activité soit doublée en présence de l'élément SECIS en 3'UTR, laissant suggérer que ces deux éléments contribuent ensemble au mécanisme de recodage. Des éléments SRE ont également été prédits dans d'autres ARNm de sélénoprotéines tels que SPS2, SelH, SelO et SelT, bien qu'aucune conservation de structure n'ait été observée (Howard et al., 2005a) (Figure 13).



Figure 13. Prédiction de structures de type SRE dans la région codante en aval du codon UGA^{Sec} de différents ARNm de sélénoprotéines eucaryotes. (d'après Howard et al., 2005)

Le codon UGA^{Sec} est encadré. Les structures 2D ont été prédites à l'aide du logiciel mfold v3.1, par Howard et al., 2005.

Les variations de structures des éléments SRE pourraient jouer un rôle dans la régulation de l'expression des sélénoprotéines en contrôlant l'efficacité de recodage du codon UGA^{Sec} (Howard et al., 2007b). Très récemment, une autre tige-boucle, appelée SL1, présente 9 nucléotides en aval du codon UGA^{Sec} de l'ARNm de SelS a été identifiée et présente une activité similaire à l'élément SRE de SelN. Dans cet ARNm, le codon UGA^{Sec} se trouve à l'extrémité 3' de la séquence codante, et l'élément SL1 est ainsi localisé dans la région 3'UTR (Bubenik et al., 2013).

De manière intéressante, 4 mutations dans l'élément SRE de l'ARNm de SelN seraient à l'origine de myopathies musculaires (Maiti et al., 2008). Ces mutations affaibliraient la structure secondaire de la tige-boucle SRE, réduisant l'efficacité de recodage du codon UGA^{Sec}. Ceci prouve l'importance de l'élément SRE dans la synthèse de SelN.

I.2.2.2 Les Protéines impliquées dans le mécanisme de synthèse des sélénoprotéines

I.2.2.2.a Le facteur d'élongation spécifique eEFSec

Le facteur d'élongation eucaryote spécifique eEFSec est responsable du recrutement de Sec-ARNt^{[Ser]^{Sec}} et permet l'incorporation de la sélénocystéine dans la chaîne peptidique naissante en réponse au codon UGA^{Sec} (Fagegaltier et al., 2000b; Tujebajeva et al., 2000) (Figure 14). Similaire au facteur d'élongation canonique eIF1A, qui est impliqué dans l'incorporation des 20 autres acides aminés, eEFSec a une activité GTPase. Cependant, contrairement à eIF1A, il a une forte spécificité pour Sec-ARNt^{[Ser]^{Sec}} mais ne lie ni le Sep-ARNt^{[Ser]^{Sec}}, ni les autres ARNt aminoacylés (Carlson et al., 2004). eEFSec présente une structure en « pseudo-calice » constituée de 4 domaines (Leibundgut et al., 2005). Les domaines I, II et III, qui forment la coupe du calice ont une forte similarité avec le facteur eEF1A. Le domaine IV est une extension C-terminale unique à eEFSec et absente dans eEF1A. Ce domaine a été proposé comme impliqué dans l'interaction avec SBP2 et le long bras variable de l'ARNt^{[Ser]^{Sec}} (Itoh et al., 2009a; Zavacki et al., 2003). Chez les bactéries, la fonction de eEFSec et de SBP2 est réalisée par un seul facteur protéique appelé SelB, qui est à la fois impliqué dans la liaison au SECIS, et dans la distribution du Sec-ARNt^{[Ser]^{Sec}} au ribosome (Böck et al., 2006). La présence de deux protéines différentes chez les eucaryotes s'explique par la localisation distante de l'élément SECIS (en 3'UTR) par rapport au codon UGA^{Sec}, alors que celui des bactéries est directement situé en aval du UGA^{Sec}. Pour éviter que le Sec-ARNt^{[Ser]^{Sec}} puisse agir comme un ARNt supprimeur de codons UGA stop, la structure de eEFSec (ou SelB bactérien) ne lui permet pas, dans les conditions habituelles, d'avoir accès au site A du ribosome en cours de traduction. Pour décoder le codon UGA^{Sec}, le facteur eEFSec subit des changements conformationnels une fois lié à l'élément SECIS via SBP2, qui stimulent son interaction fonctionnelle avec le ribosome (Gonzalez-Flores et al., 2012). En effet, il a été montré que SelB est capable de catalyser l'hydrolyse ribosome-dépendante du GTP uniquement en présence de l'élément SECIS (Hüttenhofer and Böck, 1998).

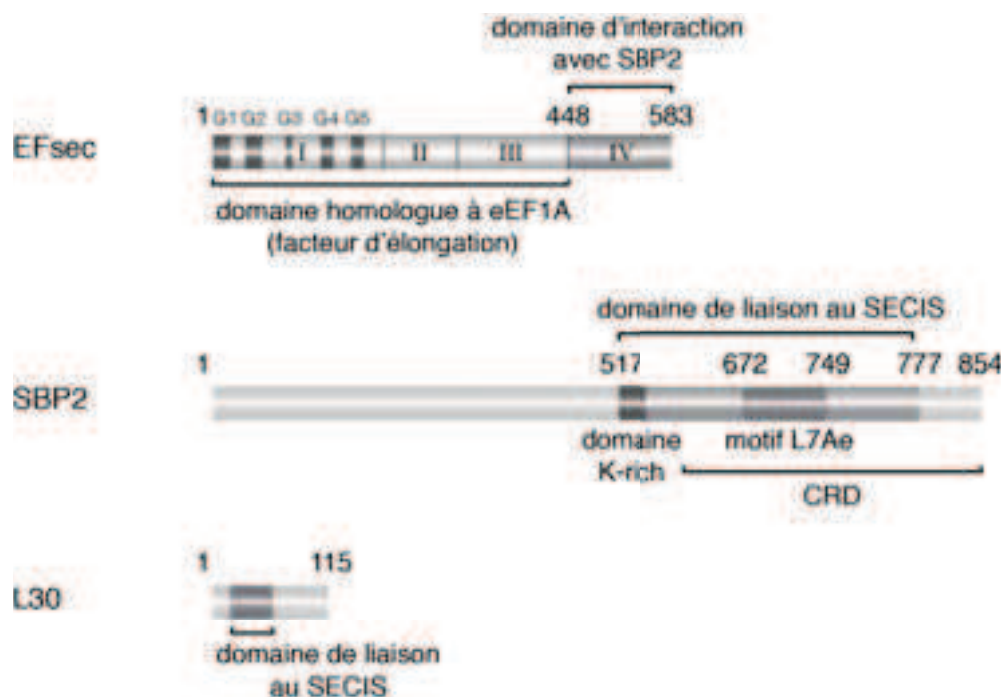


Figure 14. Représentation schématique des facteurs eEFSec, SBP2 et L30, impliqués dans la synthèse des sélénoprotéines.

Les domaines fonctionnels importants et leur position dans la séquence en acide aminés sont représentés. CRD (cysteine rich domain) : domaine riche en cystéine. G1-G5 : domaines de liaison au GTP (Allmang and Krol, 2006a; Bifano et al., 2013; Chavatte et al., 2005; Copeland et al., 2000; Itoh et al., 2009a; Lescure et al., 2002b; Takeuchi et al., 2009; Zavacki et al., 2003).

I.2.2.2.b SBP2 (SECIS binding protein 2)

La protéine SBP2 est au cœur du mécanisme de synthèse des sélénoprotéines en interagissant avec l'élément SECIS et avec les ribosomes durant le mécanisme de recodage (Caban et al., 2007; Copeland et al., 2000, 2001; Kinzy et al., 2005; Lescure et al., 2002a; Allmang 2002; Takeuchi 2009). Des mutations dans le gène de SBP2 humain conduisent à un défaut de synthèse de l'ensemble des sélénoprotéines et sont associées à des troubles multisystémiques tels que des défauts dans la maturation des spermatozoïdes, des dystrophies musculaires, une augmentation de la quantité de radicaux oxygénés intracellulaires se traduisant par une photosensibilité cutanée accrue aux rayons ultra-violets, des anomalies dans le métabolisme thyroïdien et du système immunitaire et un raccourcissement des télomères (Di Cosmo et al., 2009; Dumitrescu et al., 2005; Schoenmakers et al., 2010). Ceci met en évidence le rôle fondamental de la protéine SBP2 dans la synthèse des sélénoprotéines, elles-mêmes impliquées dans divers processus biologiques. La protéine SBP2 a été isolée et caractérisée fonctionnellement chez le rat, l'homme et la drosophile (Copeland et al.,

2000; Lescure et al., 2002b; Takeuchi et al., 2009). La protéine humaine fait 854 acides aminés et est représentée Figure 14. Le domaine N-terminal qui représente les deux tiers de la protéine (1-516), n'a pas de fonction attribuée et n'est pas nécessaire au mécanisme de recodage des sélénoprotéines dans les extraits de réticulocytes de lapin (Copeland et al., 2000). Le domaine C-terminal (517-854) comporte le domaine de liaison à l'ARN et au ribosome (Allmang et al., 2002a; Copeland et al., 2001; Lescure et al., 2002b). Le motif L7Ae présent dans ce domaine C-terminal est à la fois impliqué dans la liaison à l'ARN SECIS et au ribosome (Allmang et al. 2002; Clery et al. 2007; Caban et al., 2007) (Figure 14). SBP2 ne peut donc pas interagir de manière simultanée avec ces deux éléments. La liaison avec l'ARN SECIS requiert, en plus du motif L7Ae un domaine riche en lysines appelé « K-rich » ou SID (Sec Incorporation Domain) (Allmang et al., 2002a; Bubenik and Driscoll, 2007; Copeland et al., 2000; Donovan et al., 2008; Takeuchi et al., 2009). Plusieurs techniques biophysiques ont permis de révéler que 70% de la séquence en acides aminés de SBP2 présente un aspect désordonné. Seul le motif de liaison à l'ARN L7Ae apparaît être structuré. Ainsi, la protéine SBP2 ferait partie de la famille de protéines intrinsèquement déstructurées (Intrinsically Disordered Proteins, IDP) (Olieric et al., 2009). A cause de leur plasticité inhérente, les protéines déstructurées adoptent fréquemment des structures complètement différentes lorsqu'elles se lient à des partenaires différents et sont souvent impliquées dans des interactions multiples au cœur de complexes (Berlow et al., 2015; Uversky, 2015). Elles peuvent également être prises en charge par des chaperons qui peuvent les aider à adopter un repliement correct. Le repliement de SBP2 dépend du complexe de chaperons HSP90/R2TP (Boulon et al., 2008). Les IDP permettent d'augmenter le répertoire d'interactions disponibles dans la cellule, sans nécessité d'exprimer plusieurs protéines différentes. Cette caractéristique ajoute de la précision dans les mécanismes de régulation, puisque la disponibilité d'un domaine déstructuré donné dépend de la compétition entre plusieurs partenaires pouvant être impliqués dans des processus cellulaires et des voies de régulations variées. Des événements d'épissage alternatifs dans la région 5' du pré-ARNm de SBP2 sont à l'origine de 8 isoformes différents dont l'un contenant un signal d'adressage à la mitochondrie (Papp et al., 2008). Des séquences conservées impliquées dans la régulation de l'expression de SBP2 ont été mises en évidence dans la région 3'UTR de l'ARNm de SBP2 (Bubenik et al., 2009). Ces

séquences sont reconnues par les protéines CUG-BP1 et HuR, impliquées dans la dégradation et la régulation de la traduction d'ARNm cellulaires, suggérant que l'expression de SBP2 serait régulée post-transcriptionnellement (Bubenik et al., 2009).

I.2.2.2.b.1 SBP2 est une protéine de la famille L7Ae

SBP2 fait partie de la famille des protéines L7Ae. Cette famille comprend plusieurs protéines de fonctions variées qui partagent un motif conservé d'interaction avec l'ARN d'environ 50 acides aminés (motif L7Ae) (Figure 15A) (Allmang et al., 2002a; Copeland et al., 2000). Ces protéines se lient aux ARN au niveau de structures en "K-turn" (détaillé dans le paragraphe I.2.2.1.a). Ces protéines sont importantes pour la structure du ribosome, l'assemblage du spliceosome (snRNP) et des complexes guides de modifications site-spécifiques des ARN (snoRNP) (Huang and Lilley, 2013). La famille L7Ae inclut les protéines ribosomiques L30 et L7A qui génèrent des interactions tertiaires et stabilisent la structure du ribosome. Dans les snoRNP à boîte C/D, la liaison de la protéine 15.5kD (Snu13p chez la levure) au K-turn du snoARN est la première étape de l'assemblage de la particule (Kuhn et al., 2002; Watkins et al., 2002). Chez les mammifères, cette même protéine lie le snoRNP U3 (Marmier-Gourrier et al., 2003) ainsi que le K-turn du snARN U4 dans le complexe d'épissage (Schultz et al., 2006; Vidovic et al., 2000; Watkins et al., 2000). Les protéines NHP2 (Nhp2p chez la levure) et L7Ae sont des composants critiques des snoRNP à boîte H/ACA et de la télomérase (Antal et al., 2002; Kiss et al., 2010; Rozhdestvensky et al., 2003). Des protéines L7Ae font également partie de la RNase P d'archée (Cho et al., 2010), requise pour la maturation des ARNt. Des protéines bactériennes de la famille L7Ae telles que YbxF ont également été identifiées, leur fonction biologique n'est pas encore connue (pour revue, Huang and Lilley, 2013). La protéine SBP2 au cœur des mRNP de sélénoprotéines (Allmang et al., 2002a) joue un rôle essentiel dans la fonction et l'assemblage des mRNP de sélénoprotéines.

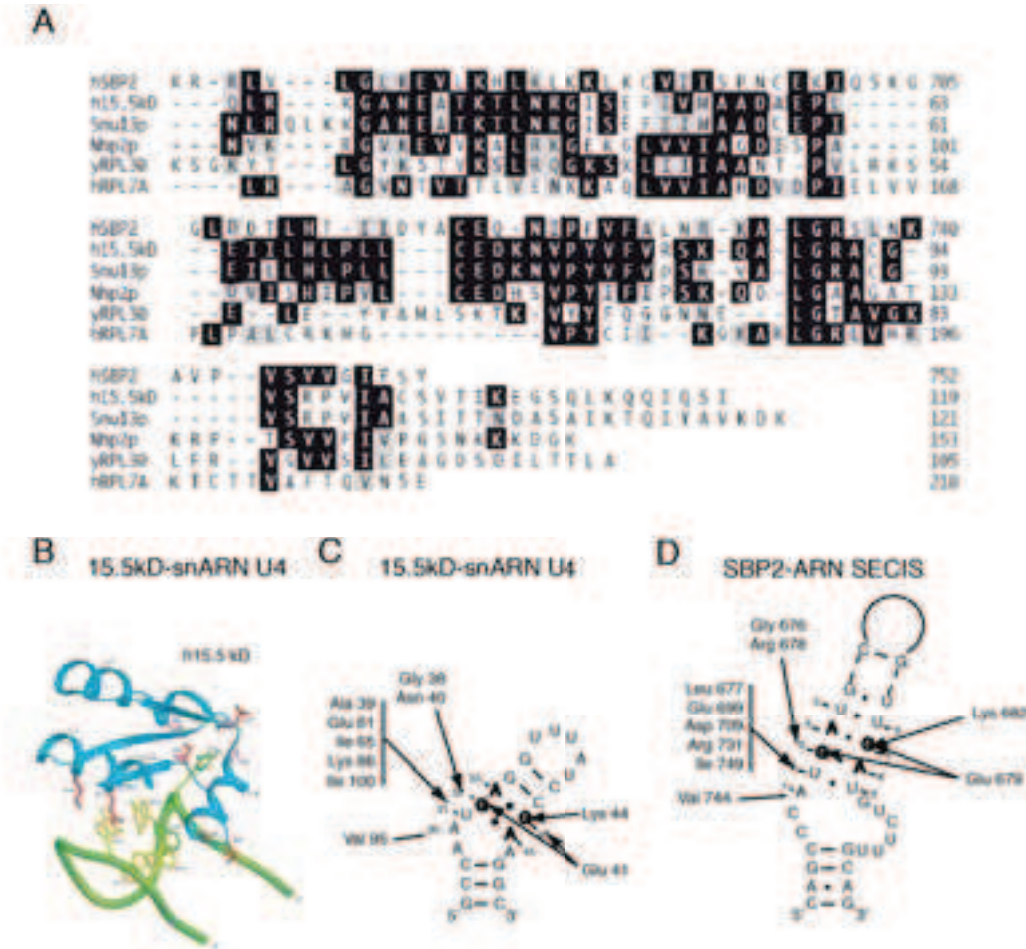


Figure 15. Modèle d'interaction entre SBP2 et l'élément SECIS. (d'après Allmang and Krol, 2006)
 (A) Alignement de séquences des domaines L7Ae des protéines de la famille L7Ae dont SBP2 (Allmang et al., 2002b). SBP2 présente le plus d'homologie avec la protéine 15.5kD. (B) Structure cristallographique du complexe de la protéine 15.5kD (membre de la famille L7Ae) avec sa cible en K-turn présente dans le snARN U4 humain (Vidovic et al., 2000). Représentation des acides aminés impliqués dans l'interaction entre 15.5kD et le snARN U4, à partir des données de la structure cristallographique. (C) Les acides aminés de SBP2 qui interagissent avec l'élément SECIS ont été déterminés expérimentalement d'après le modèle 15.5kD-snARN U4 (Allmang et al., 2002b). Les modes d'interaction des protéines SBP2 et 15.5kD avec leur ARN cible sont très similaires.

Les structures cristallographiques des complexes 15.5kD-snARN U4 humain (Figure 15B), de L30e-L30 pré-ARNm de *S. cerevisiae* et de L7Ae-snARN archée ont été résolues et révèlent une interface ARN-protéine commune (Chao and Williamson, 2004; Moore et al., 2004; Vidovic et al., 2000). Cette interface se caractérise d'une part, par des contacts spécifiques d'un nombre limité d'acides aminés avec les deux paires de bases non-Watson-Crick G-A du quartet de paires de bases et, d'autre part, par l'interaction du nucléotide en bulge U (snARN U4, ARNr, snARN) ou A (pré-ARNm de L30e) avec une poche d'acides aminés hydrophobes au sein de la protéine (Figure 15B). Le modèle de structure 2D du SECIS et la modélisation de sa structure

3D, en comparaison avec les autres K-turn indique que l'ARN SECIS serait un variant de K-turn, qui comporte une boucle interne plus grande (Allmang and Krol, 2006b; Cléry et al., 2007; Walczak et al., 1996) (Figure 16). La fixation des protéines L7Ae aux structures en K-turn induit un changement de structure de l'ARN cible favorable à l'interaction de protéines supplémentaires. De ce fait, les protéines L7Ae associées à l'ARN cible, créent une plateforme pour recruter d'autres protéines "cœur" lors de l'assemblage des RNP (Vidovic et al., 2000). SBP2 pourrait jouer un rôle similaire lors de l'assemblage des mRNP de sélénoprotéines en recrutant les facteurs de biogenèse, de traduction et de recodage au niveau de ces mRNPs. En plus de SBP2, la protéine L30 interagit également avec le K-turn de l'élément SECIS présent dans les ARNm de sélénoprotéines (Chavatte et al., 2005). Les protéines L7Ae présentent donc trois caractéristiques majeures : 1) elles sont présentes au sein de diverses RNP de fonctions essentielles à la cellule, 2) elles interagissent avec des ARN au niveau de structures en K-turn, créant une plateforme pour recruter d'autres protéines cœur des RNP, 3) un membre donné de la famille L7Ae peut faire parti de plusieurs RNP différentes.

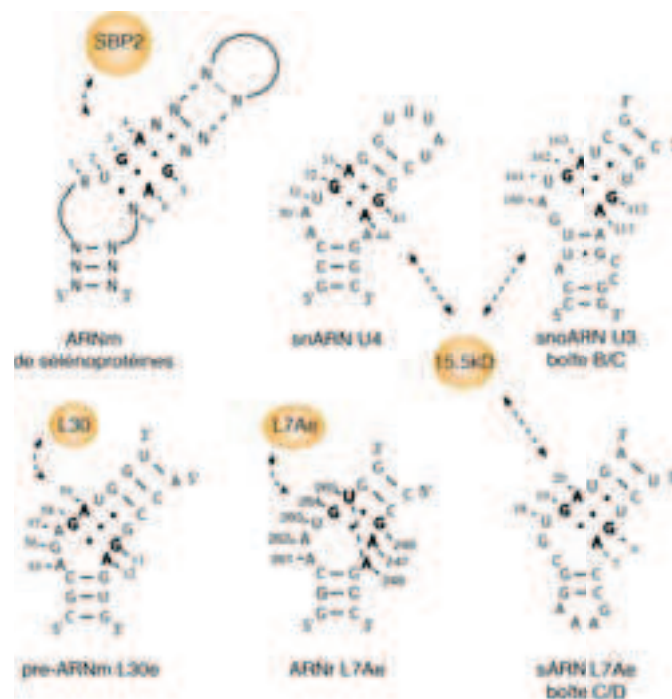


Figure 16. Structures secondaires d'ARN en K-turn dont l'ARN SECIS associés à des protéines L7Ae . Représentation des structures secondaires des K-turn du snARN U4, du pré-ARNm de L30e, de l'ARNr L7Ae et de la boîte C/D du sARN L7Ae obtenues à partir des données cristallographiques des complexes ARN-protéines correspondants (Chao and Williamson, 2004; Moore et al., 2004; Vidovic et al., 2000). Les structures secondaires de la boîte B/C du snoARN U3 et de l'élément SECIS ont été déterminées par des expériences de cartographie en solution (Fagegaltier et al., 2000a; Marmier-Gourrier et al., 2003; Walczak et al., 1996, 1998). Les appariements G-A/A-G de type "sheared" sont annotés en gras. Les protéines capables de lier ces ARN sont représentées.

I.2.2.2.b.2 Interaction SBP2-ARN SECIS

La protéine SBP2 présente le plus de similarité avec la protéine 15.5kD de la famille L7Ae. Les données de la structure cristallographique du complexe 15.5kD/snARN U4 (Figure 15B) (Schultz et al., 2006; Vidovic et al., 2000) ont permis de prédire les acides aminés du domaine L7Ae de SBP2 impliqués dans l'interaction avec le SECIS et de démontrer, par mutagenèse dirigée, que les acides aminés conservés entre les protéines 15.5kD et SBP2, sont essentiels à la liaison au SECIS (Allmang et al., 2002b). Le modèle d'interaction proposé (Figure 15C) montre qu'un nombre restreint d'acides aminés interagit avec les G des 2 paires de bases non-Watson-Crick et avec le nucléotide U en bulge (Figure 15C) (Allmang et al., 2002b).

Des expériences de SELEX (Systematic Evolution of Ligands by EXponential enrichment) et de mutagenèse dirigée ont montré que SBP2 est capable de lier un K-turn présentant un résidu U protubérant et qu'elle possède une plus grande affinité pour les boucles internes plus larges (Cléry et al., 2007). D'autres déterminants de spécificité ont également été détectés dans l'hélice II du SECIS et la taille des hélices I et II joue un rôle dans cette interaction (Figure 12). L'ensemble de ces éléments contribue à la spécificité de l'interaction SBP2-SECIS. De manière intéressante, un second domaine de SBP2, absent dans les autres protéines L7Ae, serait également impliqué dans la spécificité de reconnaissance du SECIS. Ce domaine de liaison à l'ARN a été identifié en N-terminal du motif L7Ae, il est appelé K-rich ou SID (Sec Incorporation Domain) (Figure 14) (Bubenik and Driscoll, 2007; Donovan et al., 2008; Takeuchi et al., 2009). De manière intéressante, SBP2 présente des affinités variables pour les éléments SECIS des différents ARNm des sélénoprotéines (Squires et al., 2007a; Takeuchi et al., 2009; Taylor et al., 2015).

I.2.2.2.b.3 Interaction SBP2-eEFSec

Le domaine IV de eEFSec, une extension C-terminale unique à eEFSec, est impliqué dans l'interaction avec SBP2 et le long bras variable de l'ARNt^{[Ser]Sec} (Gonzalez-Flores et al., 2012; Itoh et al., 2009a; Zavacki et al., 2003) (Figure 14). Le domaine C-terminal de SBP2 qui comprend le domaine de liaison au SECIS est impliqué dans l'interaction avec eEFSec (Donovan et al., 2008; Tujebajeva et al., 2000; Zavacki et al., 2003). La formation du complexe entre SBP2 et eEFSec est dépendante de la présence de l'élément SECIS mais pas du Sec-ARNt^{[Ser]Sec}, ni du GTP (Donovan et al.,

2008; Gonzalez-Flores et al., 2012). La présence de Sec-ARNt^{[Ser]^{Sec}} semble cependant stabiliser l'interaction SBP2-eEFSec (Itoh et al., 2009a) et la co-expression de SECp43 la stimule (Small-Howard et al., 2006).

I.2.2.2.b.4 Interaction SBP2-ribosome

Une autre fonction de SBP2 est de communiquer avec le ribosome durant le mécanisme d'incorporation de la sélénocystéine (Caban et al., 2007; Copeland et al., 2001). En effet, il a été montré que SBP2 co-sédimente avec les ribosomes dans des gradients de glycérol, et plus particulièrement avec l'ARNr 28S. Il a ainsi été proposé que SBP2 interagisse avec le ribosome via l'une des 7 structures en K-turn de l'ARNr 28S (Copeland and Driscoll, 2001; Kinzy et al., 2005). Par ailleurs, l'interaction SBP2-ribosome impliquerait le module L7Ae de SBP2 (Caban et al., 2007). La protéine SBP2 humaine se lierait à la sous-unité ribosomique 60S mais pas à la sous-unité 40S *in vitro* (Takeuchi et al., 2009). Les domaines d'interaction de SBP2 avec le ribosome et le SECIS se chevauchent et se trouvent dans le motif L7Ae (Caban et al., 2007). Un motif supplémentaire requis pour l'interaction avec le ribosome mais pas avec le SECIS a été identifié et est localisé entre les acides aminés 647 et 651. De plus, des expériences de mutagenèse dirigée de la protéine SBP2 ont révélé que le domaine K-rich de SBP2 serait d'importance primordiale pour l'interaction SBP2-ribosome (Takeuchi et al., 2009). Des résultats récents révèlent que le site de liaison de SBP2 sur le ribosome humain est principalement au niveau de l'ARNr 28S (Kossinova et al., 2014). Des expériences d'empreinte aux radicaux libres hydroxyle sur l'ARNr 28S révèlent que la liaison de SBP2 au ribosome (80S ou 60S) protège l'hélice ES7L-E localisée dans le segment d'extension 7 de l'ARNr 28S, ce qui a été confirmé par des expériences de pontage chimique (Kossinova et al., 2014) (Figure 17). De manière intéressante, la liaison de SBP2 au niveau de l'hélice ES7L-E induirait des changements conformationnels dans l'ARNr 28S, qui pourraient permettre l'incorporation de sélénocystéine en facilitant l'accommodation du Sec-ARNt^{[Ser]^{Sec}} dans le site A du ribosome ou de stimuler l'activité GTPase de eEFSec (Caban and Copeland, 2012; Kossinova et al., 2014). L'interaction SBP2-ribosome est inhibée en présence du SECIS, ce qui indique que SBP2 n'est pas capable d'interagir simultanément avec le ribosome et l'ARN SECIS. Cette compétition n'est

plus observée lorsque le SECIS est muté dans le quartet de paires de bases non-Watson-Crick (Donovan et al., 2008; Kinzy et al., 2005).

Chez les eucaryotes la machinerie d'incorporation de la sélénocystéine doit communiquer à la fois avec l'élément SECIS en 3'UTR et avec la machinerie traductionnelle, pour assurer le recodage. Ce rôle pourrait être exécuté par SBP2. Cependant, peu de choses sont connues quant à ce mécanisme moléculaire. Des modèles seront discutés dans le paragraphe I.2.3.

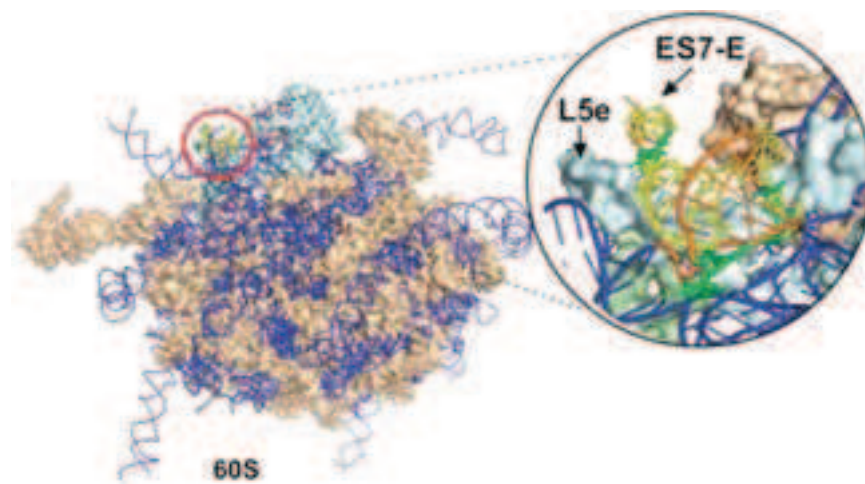


Figure 17. Localisation de SBP2 sur la structure de la sous-unité 60S du ribosome humain obtenu par cryo-microscopie électronique. (Figure de Kossinova et al., 2014)

Localisation du segment d'expansion ES7L-E de l'ARNr 28S sur la sous-unité 60S (Anger et al., 2013). La structure du site de liaison de SBP2 est présentée dans l'encart. Les protéines ribosomiques sont représentées en brun clair, les ARNr sous forme de rubans bleus. Dans l'encart, ES7L-E est indiqué en orange, les nucléotides protégés des clivages aux radicaux hydroxyles par SBP2 sont en jaune et les nucléotides pontés à SBP2 en vert.

I.2.2.2.c La protéine ribosomique L30

Il a été proposé que la protéine ribosomique L30, qui n'a pas d'homologue chez les bactéries, soit impliquée dans le mécanisme d'incorporation de la sélénocystéine chez les eucaryotes (Chavatte et al., 2005). L30 est capable d'entrer en compétition avec SBP2 pour la liaison à l'ARN SECIS *in vitro* (Chavatte et al., 2005). L30, comme SBP2, fait partie de la famille des protéines L7Ae qui partagent le même motif de liaison à l'ARN (Figure 14). L30 est une protéine ribosomique (Dabeva and Warner, 1987) localisée à l'interface entre les sous-unités 60S et 40S, d'après les structures cristallographique et de cryo-microscopie électronique récentes du ribosome 80S (Armache et al., 2010; Ben-Shem et al., 2010; Chandramouli et al., 2008; Halic et al., 2005; Jenner et al., 2012). Cette protéine est également retrouvée à l'état libre où elle

est capable de réguler sa propre synthèse en se liant à une structure en K-turn en 5' d'un site d'épissage de son pré-ARNm, inhibant ainsi l'épissage (Macías et al., 2008). Une étude récente a permis d'identifier les nucléotides et acides aminés impliqués dans l'interaction L30-SECIS et a établi que L30 et SBP2 lient des sites chevauchants main non-identiques sur le SECIS (Bifano et al., 2013). De plus, les résidus requis pour l'interaction de L30 avec le SECIS diffèrent de ceux impliqués dans la liaison de L30 à son pré-ARNm (Bifano et al., 2013). L30 stimule la synthèse des sélénoprotéines dans des cellules transfectées et est en compétition avec SBP2 pour l'élément SECIS. Ainsi, il a été proposé que L30 déplace de façon transitoire SBP2 et permette ainsi de rapprocher le SECIS du site A du ribosome durant le mécanisme de recodage (Chavatte et al., 2005).

Certaines protéines sont impliquées dans la régulation du mécanisme d'incorporation de la sélénocystéine, elles seront discutées dans le paragraphe I.2.4.2.a.

1.2.3 Modèles du mécanisme de synthèse des sélénoprotéines

Jusqu'à présent, aucun modèle unificateur n'a été proposé pour le processus moléculaire d'incorporation de la sélénocystéine. Trois modèles ont cependant été suggérés, basés sur les données expérimentales actuelles, mais qui ne répondent cependant pas à toutes les questions mécanistiques (Figure 18). Dans le premier modèle, SBP2 est associée au ribosome qui se déplace sur l'ARNm pendant la traduction des sélénoprotéines (Figure 18A). Le ribosome est ralenti au niveau du codon UGA^{Sec}, ce qui permet à l'élément SECIS d'interagir avec SBP2. La protéine SBP2 recrute alors le facteur d'élongation eEFSec associé au Sec-ARNt^{[Ser]Sec} et le dépose au niveau du site A du ribosome (Donovan et al., 2008; Kinzy et al., 2005). Ce modèle est fondé sur l'observation que SBP2 interagit de manière quantitative avec le ribosome mais ne peut se lier simultanément avec l'ARN SECIS. D'après le second modèle, la liaison SBP2-SECIS constitue le point de départ du mécanisme (Figure 18B). SBP2 recrute le complexe eEFSec-Sec-ARNt^{[Ser]Sec} au niveau de l'élément SECIS. Lorsque le ribosome arrive au niveau du codon UGA^{Sec}, la protéine ribosomique L30 déplace SBP2 du SECIS par compétition. Le complexe SBP2-eEFSec-Sec-ARNt^{[Ser]Sec} est libéré et le Sec-ARNt^{[Ser]Sec} est chargé dans le site ribosomique A et permet l'incorporation de la sélénocystéine dans la chaîne peptidique.

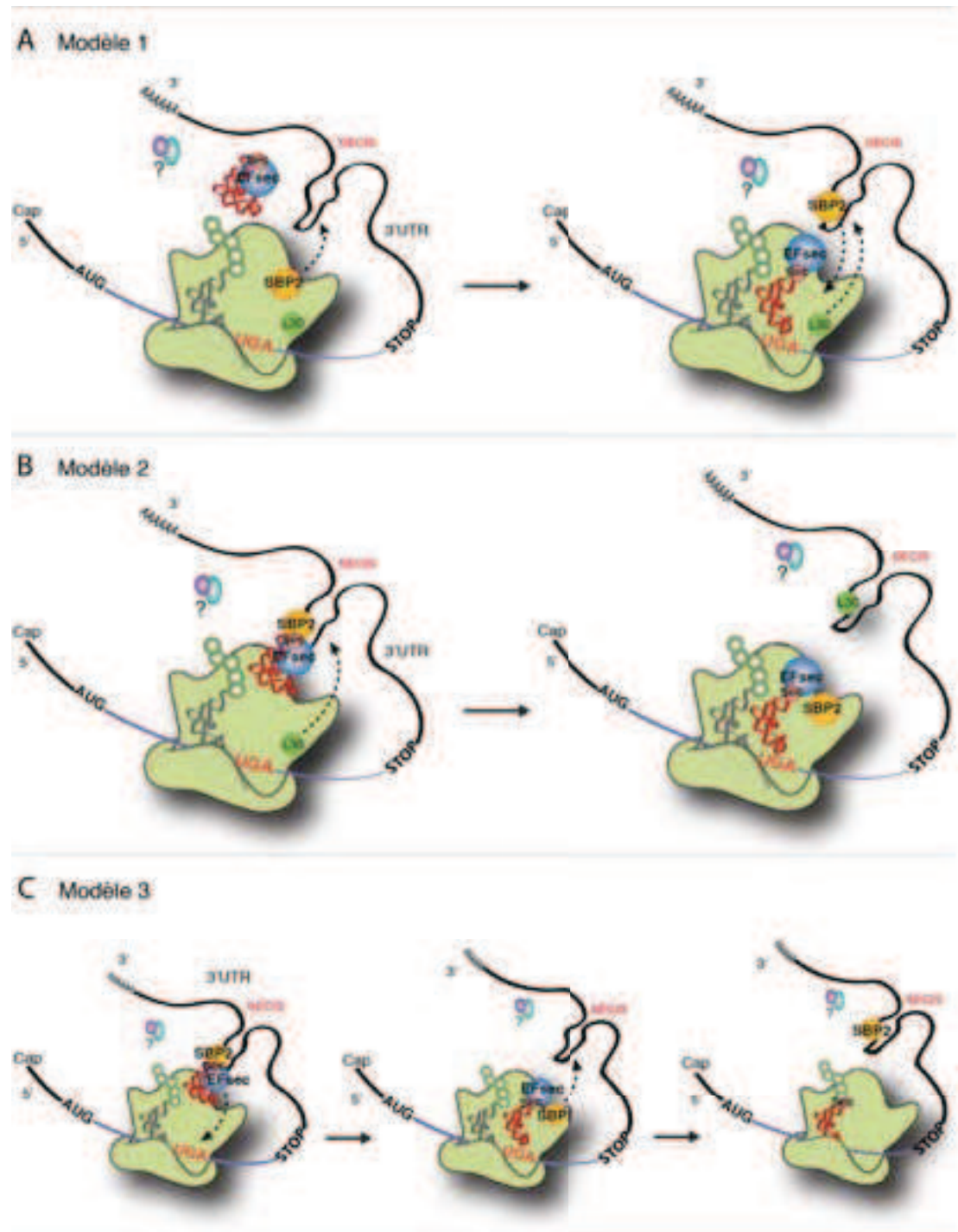


Figure 18. Modèles d'incorporation de la sélénocystéine dans les sélénoprotéines eucaryotes. (d'après Allmang and Krol, 2006)

Trois modèles du mécanisme ont été proposés chez les eucaryotes. (A) Modèle 1: SBP2 est liée au ribosome, puis interagit avec l'élément SECIS et le complexe eEFSec-Sec-ARNt^{[Ser]Sec}, ce qui permet de positionner l'ARNt dans le site A du ribosome. L30 déplace alors la protéine SBP2 liée à l'élément SECIS (Donovan et al., 2008; Kinzy et al., 2005). (B) Modèle 2: SBP2 est, dès le départ, liée à l'élément SECIS et y recrute le facteur eEFSec associé au Sec-ARNt^{[Ser]Sec}. L'échange entre la protéine SBP2 et la protéine ribosomique L30 permet d'apporter le Sec-ARNt^{[Ser]Sec} dans le site A du ribosome (Chavatte et al., 2005). (C) Modèle 3: Ce modèle ne fait pas intervenir la protéine L30. Comme pour le modèle 2, le complexe SBP2-eEFSec-Sec-ARNt^{[Ser]Sec} est, dès le départ, liée à l'élément SECIS. SBP2 resterait liée au ribosome après l'accommodation du Sec-ARNt^{[Ser]Sec} dans le site A du ribosome, se dissocierait durant l'étape de transpeptidation et se lierait à nouveau à l'élément SECIS, permettant une nouvelle étape de recodage.

Une étude récente a permis de proposer un troisième modèle (Kossinova et al., 2013) (Figure 18C). Il a été montré que SBP2 est associée à l'élément SECIS durant l'étape de formation du complexe d'initiation 48S. SBP2 reste liée au ribosome après l'accommodation du Sec-ARNt^{[Ser]^{Sec}} dans le site A du ribosome et se dissocie durant l'étape de transpeptidation pour se lier à nouveau à l'élément SECIS pour lequel elle possède une forte affinité. L'élément SECIS contacte la sous-unité ribosomique 60S après l'accommodation du Sec-ARNt^{[Ser]^{Sec}}, l'incorporation de sélénocystéine est donc gouvernée par des réarrangements conformationnels à la fois dans le SECIS et dans l'ARNr 28S (Kossinova et al., 2013). De nombreux aspects de ce mécanisme restent cependant à être élucidés. Le mode de communication entre le codon UGA^{Sec} et l'élément SECIS n'est pas bien compris. De même, on ignore ce qui empêche les facteurs de terminaison eRF1/3 de se lier au site A du ribosome au niveau du codon UGA^{Sec}. Enfin, il a été proposé que d'autres facteurs pourraient participer à ce mécanisme (Allmang et al., 2009; Copeland and Driscoll, 1999; Gupta et al., 2013; Hirosawa-Takamori et al., 2009). L'assemblage particulier des mRNP de sélénoprotéines joue un rôle important dans le mécanisme de recodage (Boulon et al., 2008; Wurth et al., 2014) et de nombreuses études pointent vers l'existence d'un mécanisme complexe de régulation de la traduction des ARNm de sélénoprotéines.

1.2.4 Régulation de l'expression des sélénoprotéines

1.2.4.1 Efficacité du mécanisme de recodage

Plusieurs études indiquent que le mécanisme d'incorporation de la sélénocystéine est inefficace. En effet, cela est le cas chez les bactéries où l'efficacité maximum d'incorporation de la sélénocystéine rapportée est de 7-10% (Suppmann et al., 1999). Chez les mammifères, plusieurs sélénoprotéines endogènes sont exprimées de manière abondantes dans les cellules, particulièrement PHGPx (GPx4) dans les testicules (Ursini et al., 1999), GPx1 dans le foie et les reins (Cheng et al., 1997), et SelP dans le foie et le plasma sanguin (Yang et al., 1987) ce qui suggère que l'efficacité pourrait être bien plus importante que celle trouvée chez les procaryotes. Pourtant, l'efficacité de recodage dans les extraits de traduction de réticulocytes de lapin est également de 10% (Mehta et al., 2004a). Un système de traduction *in vitro* dans lequel l'incorporation de sélénocystéine a été reconstitué à partir de composants purifiés a montré que les facteurs connus du mécanisme de recodage des

sélénoprotéines (eEFSec, Sec-ARNt^{[Ser]Sec} et SBP2) sont suffisants pour l'incorporation de sélénocystéine *in vitro* mais que l'efficacité de traduction se limite dans ce cas à 3% (Gupta et al., 2013). Plusieurs études ont tenté de déterminer l'efficacité du recodage dans les cellules transfectées et montré qu'elle était 6 à 10 fois inférieure à celle des réticulocytes, même dans des conditions où SBP2 n'est pas limitante (Berry et al., 1994; Low et al., 2000; Weiss Sachdev and Sunde, 2001; Mehta et al., 2004). L'existence de facteurs augmentant l'efficacité du mécanisme de recodage et d'incorporation de la sélénocystéine a été proposé (Mehta et al., 2004b). Enfin, l'efficacité d'incorporation de la sélénocystéine est augmentée lorsqu'un ARNm naturel est utilisé plutôt qu'un rapporteur luciférase, indiquant que l'ARNm pourrait contenir des éléments *cis* autres que le SECIS impliqués dans l'efficacité de recodage (Berry et al., 1994; Gupta and Copeland, 2007). Ces éléments pourraient se trouver à proximité du codon UGA^{Sec} ou au niveau de la région 3'UTR. La sélénoprotéine P contient 10 résidus sélénocystéine et est la sélénoprotéine la plus abondante du plasma sanguin avec une concentration de 30 µg/mL en condition de sélénium idéal chez le rat (Read et al., 1990). Sa synthèse et sa dégradation *in vivo* sont rapides et efficaces (Hill et al., 1991) indiquant que le mécanisme d'incorporation de la sélénocystéine est également efficace. Il a été montré que dans les lysats de réticulocytes de lapin, le recodage du premier UGA^{Sec} de SelP est peu efficace mais augmente d'environ 10 fois pour les codons UGA^{Sec} situés en aval (Fixsen and Howard, 2010). Il s'agit d'un mécanisme de recodage processif sélénium-dépendant (Shetty et al., 2014). La régulation de l'efficacité de recodage des sélénoprotéines pourrait être un déterminant essentiel de leur niveau et de leur mode d'expression tissu-spécifique, elle semble faire appel à des facteurs encore inconnus.

I.2.4.1.a Relation entre terminaison et recodage

Des mécanismes de compétition entre la terminaison de la traduction et le recodage ont été mis en évidence au niveau du codon UGA^{Sec} *in vitro* et *in vivo* (Berry et al., 1994; Gupta and Copeland, 2007; Low et al., 2000a; Weiss Sachdev and Sunde, 2001). De manière intéressante, une surexpression du facteur de terminaison de la traduction eRF1 diminue le recodage au niveau du codon UGA^{Sec} d'environ 50% dans les lysats de réticulocytes de lapin (Gupta and Copeland, 2007). Cependant, en présence de la protéine SBP2, eRF1 n'est plus capable d'induire une terminaison aussi efficace. Ceci indique que la machinerie d'incorporation de la sélénocystéine

associée à SBP2 pourrait être en compétition directe avec eRF1 pour l'accès au site A du ribosome. La question de la compétition entre les mécanismes d'incorporation de sélénocystéine et de terminaison a récemment été résolue chez les bactéries. En effet, il a été montré que le facteur de terminaison RF2 est, de manière inattendue, un faible compétiteur de l'incorporation de sélénocystéine en conditions optimales en sélénium (Kotini et al., 2015). Seul 40% des ribosomes actifs qui atteignent le codon UGA^{Sec} sont capables d'incorporer la sélénocystéine, même en absence de RF2. La machinerie de recodage recrutée au niveau du SECIS bloquerait l'accès aux facteurs de terminaison et favoriserait ainsi le recodage plutôt que la terminaison au niveau des codons UGA^{Sec}. Le rôle de RF2 se limiterait à terminer la traduction des ARNm qui ne sont pas engagés dans un processus de recodage.

I.2.4.1.b Relation entre initiation et recodage

Les études concernant l'incorporation de sélénocystéine se sont majoritairement focalisée sur la relation entre le recodage et la terminaison au niveau du codon UGA^{Sec} et le rôle de l'initiation de la traduction n'a été que faiblement analysé. L'extrémité 3' des ARNm communique avec la machinerie d'initiation de la traduction par le biais de la queue poly(A), ce qui peut suggérer que le complexe SBP2/ARN SECIS en 3'UTR pourrait participer au processus d'initiation (Copeland, 2003). Une étude récente a montré que la présence de l'IRES (Internal Ribosome Entry Site) du virus CrPV (CRicket Paralysis Virus) en 5' d'un ARNm rapporteur permet l'incorporation de sélénocystéine, mais diminue son efficacité de manière significative (Donovan and Copeland, 2010). L'initiation de la traduction dépendante de cette IRES s'affranchit de tous les facteurs d'initiation et même de l'ARNt initiateur, en recrutant de manière directe les sous-unités ribosomiques 60S et 40S (Pisarev et al., 2005). Ainsi, la fonction canonique des facteurs d'initiation de la traduction ne serait pas essentielle au mécanisme d'incorporation de la sélénocystéine. L'efficacité d'incorporation de la sélénocystéine est cependant fortement diminuée en présence de l'IRES du CrPV. De plus, le mécanisme de recodage semble indépendant de la présence d'une coiffe m⁷G en 5' ou de la queue poly(A) en 3' dans les lysats de réticulocytes de lapin (Donovan and Copeland, 2010).

Les données actuelles semblent indiquer que l'efficacité d'incorporation de la sélénocystéine est intrinsèquement inefficace (Mehta et al., 2004a; Suppmann et al.,

1999). L'inefficacité de ce mécanisme pourrait constituer un atout évolutif majeur pour les organismes car il permettrait la mise en place de processus de régulation pour un meilleur contrôle de l'expression tissu-spécifique des différentes sélénoprotéines et en réponse aux variations environnementales.

1.2.4.2 Expression hiérarchique des sélénoprotéines

De nombreuses études utilisant des modèles animaux ou cellulaires ont montré que l'expression des sélénoprotéines est régulée de manière différentielle en fonction de la disponibilité du sélénium (Hill et al., 1992; Lei et al., 1995; Schomburg and Schweizer, 2009; Sunde and Raines, 2011). Tandis que l'expression de certaines sélénoprotéines telles que GPx1, MsrB1, SelW, SelH (appelées sélénoprotéines sensibles au stress, ou sélénoprotéines non-essentiels) diminue fortement en condition limitantes de sélénium, l'expression d'autres protéines incluant GPx4, TrxR1 et TrxR3 (appelées sélénoprotéines essentielles, ou insensibles au stress) est moins affectée par les variations de sélénium (pour revue, Labunsky et al., 2014; Schomburg and Schweizer, 2009). Les stress oxydants ont également une influence directe sur la surexpression sélective de plusieurs sélénoprotéines impliquées dans la défense antioxydante telles que GPx1, GPx4, TrxR1, SelS, SelK et SPS2 (Touat-Hamici et al., 2014); cette réponse est encore plus efficace dans des conditions limitantes en sélénium.

Cette régulation "hiérarchique" de l'expression des sélénoprotéines provient soit du contrôle de l'abondance des ARNm, soit d'une régulation traductionnelle (Howard et al., 2013; Sunde and Raines, 2011). L'intervention de facteurs agissant en *cis* et en *trans* apparaît central au processus de régulation différentiel de l'incorporation de sélénocystéine (Figure 19). Cependant, les mécanismes moléculaires sous-jacents restent encore largement incompris et feraient parti d'un processus multifactoriel complexe permettant d'aboutir à une régulation fine et tissu-spécifique des différentes sélénoprotéines (Figure 19).

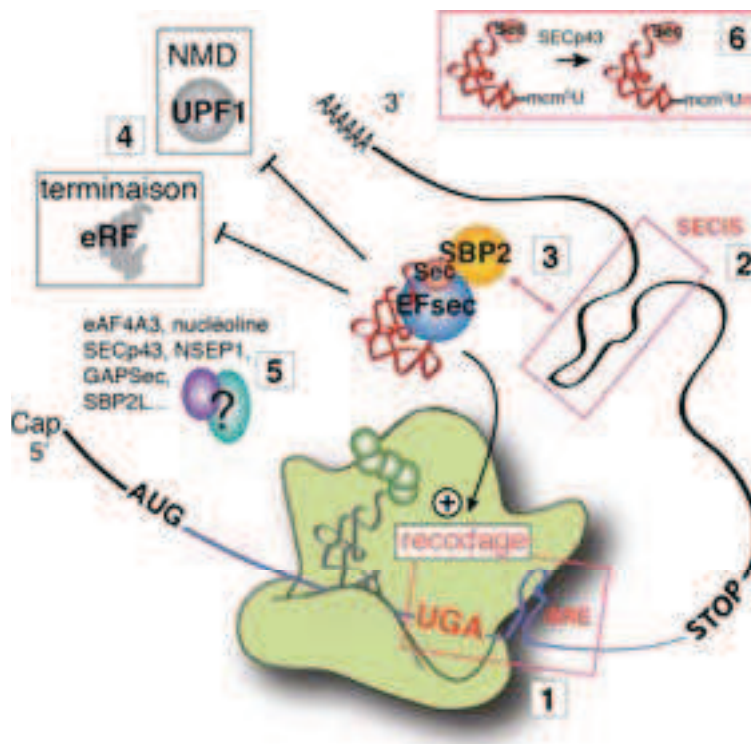


Figure 19. Facteurs impliqués dans la régulation de l'expression hiérarchique des sélénoprotéines. (d'après Schomburg and Schweizer, 2009)

(1) Le codon UGA^{Sec} peut être situé à proximité d'un élément SRE (Sec redefinition element) (Gupta and Copeland, 2007; Howard et al., 2007a). (2) L'élément SECIS dans la région 3'UTR de chaque ARNm de sélénoprotéine est unique (Krol, 2002; Latreche et al., 2009; Latreche et al., 2012; Low et al., 2000b). (3) L'affinité de SBP2 pour les différents éléments SECIS est variable (Bubenik et al., 2009; Squires et al., 2007b; Takeuchi et al., 2009). (4) SBP2 permet le recrutement des facteurs d'incorporation de la sélénocystéine, qui empêchent la terminaison de la traduction et la dégradation de l'ARNm par le mécanisme du NMD (Seyedali and Berry, 2014; Shetty and Copeland, 2015). (5) De nombreux facteurs agissant en *trans* semblent impliqués dans la synthèse des sélénoprotéines (Budiman et al., 2009a; Carlson et al., 2005; Donovan and Copeland, 2009; Hirosawa-Takamori et al., 2009; Howard et al., 2013; Miniard et al., 2010; Shen et al., 2006). (6) Deux isoformes de l'ARNt^{[Ser]Sec} sont impliqués dans l'incorporation de sélénocystéine. L'ARNt^{[Ser]Sec} mcm⁵U34 est impliqué dans le recodage des sélénoprotéines essentielles et l'isoforme mcm⁵U34m dans celui des sélénoprotéines sensibles au stress (Carlson et al., 2005; Howard et al., 2013).

I.2.4.2.a Les facteurs de régulation de l'expression des sélénoprotéines

I.2.4.2.a.1 Le sélénium et l'ARNt^{[Ser]Sec}

La variation d'expression des différentes sélénoprotéines en fonction de la disponibilité du sélénium, serait due à une régulation directe du mécanisme d'incorporation de sélénocystéine dans les différentes sélénoprotéines (Figure 19). Si l'expression des sélénoprotéines essentielles telles que GPx4 et Sel15 reste peu sensible aux variations du taux de sélénium, des conditions limitantes en sélénium réduisent de manière importante le nombre de ribosomes, détectés par ribosome

profiling, en aval du codon UGA^{Sec} sur les ARNm de sélénoprotéines non-essentiels telles que Gpx1, MsrB1, SelW, et SelH, indiquant une diminution de l'efficacité d'incorporation de la sélénocystéine (Howard et al., 2013). Cette régulation sélénium-dépendante serait en partie expliquée par le degré de méthylation de la position Um34 du Sec-ARNt^{[Ser]Sec} (Chittum et al., 1997; Diamond et al., 1993; Howard et al., 2013).

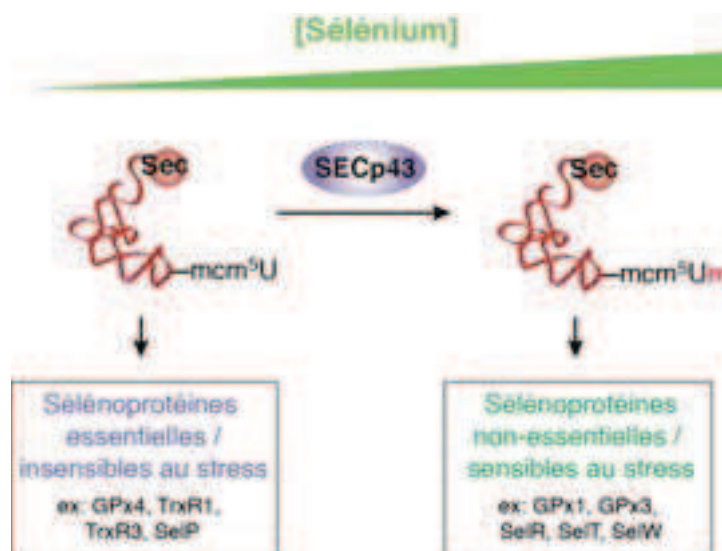


Figure 20. Rôle des deux isoformes du Sec-ARNt^{[Ser]Sec} dans la régulation de l'expression hiérarchique des sélénoprotéines.

Les deux isoformes du Sec-ARNt^{[Ser]Sec} interviennent soit dans la synthèse des sélénoprotéines essentielles (isoforme mcm⁵U34), soit dans celle des sélénoprotéines sensibles au stress (isoforme mcm⁵U34m) (Carlson et al., 2005; Howard et al., 2013). La biosynthèse de l'isoforme mcm⁵U34m dépend de la protéine SECp43 et est dépendante du sélénium (Ding and Grabowski, 1999; Small-Howard et al., 2006; Xu et al., 2005). La méthylation du Sec-ARNt^{[Ser]Sec} serait impliquée dans la régulation sélénium-dépendante de l'incorporation de sélénocystéine et l'expression hiérarchique des sélénoprotéines.

En effet, deux isoformes majeurs de l'ARNt^{[Ser]Sec}, contenant chacun soit la modification mcm⁵U soit la modification mcm⁵Um en position 34 existent chez les mammifères, et leur distribution relative est influencée par le niveau de sélénium (Chittum et al., 1997; Diamond et al., 1993). De manière intéressante, il a été montré que les sélénoprotéines essentielles telles que TrxR1 et TrxR3 sont synthétisées par l'isoforme mcm⁵U34, tandis que les sélénoprotéines sensibles au stress (GPx1, GPx3, MsrB1, SelT et SelW) sont synthétisées par l'isoforme mcm⁵Um34 (Carlson et al., 2005; Howard et al., 2013) (Figure 20). La méthylation mcm⁵Um en position 34 est assurée par la protéine SECp43 et le taux de méthylation est dépendant du niveau de sélénium (Ding and Grabowski, 1999; Xu et al., 2005). Cette régulation par SECp43

serait en partie responsable de l'incorporation différentielle de sélénocystéine et de l'expression hiérarchique des sélénoprotéines (Carlson et al., 2005; Howard et al., 2013). La régulation différentielle de l'expression des sélénoprotéines en réponse au stress oxydant se situerait également au niveau de l'incorporation de la sélénocystéine, plutôt qu'au niveau post-transcriptionnel (Touat-Hamici et al., 2014).

I.2.4.2.a.2 eIF4A3

Le facteur d'initiation de la traduction eIF4A3 a récemment été identifié comme étant un important régulateur de la synthèse des sélénoprotéines. Il contribuerait à la traduction préférentielle des sélénoprotéines essentielles en condition limitante de sélénium et à l'inhibition de la traduction des sélénoprotéines sensibles au stress (Budiman et al., 2009a). eIF4A3 est un membre de la famille des hélicases ATP dépendantes à motif "DEAD", majoritairement impliquées dans l'épissage des pré-ARNm dans le noyau (Li et al., 1999). Elle lie les ARNm épissés durant la formation du complexe EJC (Exon Junction Complex) et participe au mécanisme de dégradation des ARNm par le mécanisme du NMD (Chan et al., 2004; Gehring et al., 2005; Palacios et al., 2004). En plus de son activité canonique, eIF4A3 est impliquée dans la régulation négative de l'incorporation de sélénocystéine (Budiman et al., 2009a) (Figure 19). Cette protéine se lie préférentiellement aux éléments SECIS en 3'UTR des ARNm de sélénoprotéines sensibles au stress tels que GPx1 et MsrB1 mais a une faible affinité pour les SECIS des ARNm des sélénoprotéines essentielles (TrxR1 et GPx4). eIF4A3 se lie au niveau de la boucle interne et de la boucle apicale du SECIS et discrimine les différents éléments SECIS en fonction de la taille de la boucle apicale (Budiman et al., 2011). Ainsi, il reconnaît préférentiellement les éléments SECIS de type 1 dont la boucle apicale est large et non-structurée (GPx1 et MsrB) (Budiman et al., 2011). Puisque eIF4A3 et SBP2 ont des sites de liaison qui se chevauchent, ces deux protéines sont en compétition pour la liaison à l'ARN SECIS (Morley and Willett, 2009). Il a également été montré que le niveau de protéine eIF4A3 est augmenté de manière importante en condition limitante de sélénium (Budiman et al., 2009a). Ainsi, il a été proposé que eIF4A3 se lie de manière sélective à une sous-classe d'éléments SECIS appartenant aux ARNm de sélénoprotéines non-essentielles et empêche l'incorporation de sélénocystéine lorsque le sélénium est limitant en masquant le site de liaison de SBP2 (Figure 19). Il a été proposé que cela

puisse également faciliter la dégradation de l'ARNm par le mécanisme du NMD (Morley and Willett, 2009).

I.2.4.2.a.3 Nucléoline

La nucléoline est également une protéine de liaison au SECIS qui influence l'efficacité de recodage du codon UGA^{Sec} (Figure 19). Il s'agit de la phosphoprotéine la plus abondante du nucléole et qui est impliquée dans diverses fonctions dont la synthèse des ARNr et la biogenèse des ribosomes (pour revue, Tajrishi et al., 2011). Cette protéine joue également un rôle dans la régulation de la transcription et le remodelage de la chromatine et contient 4 motifs de reconnaissance de l'ARN. La nucléoline a été identifiée comme une protéine de liaison au SECIS (Wu et al., 2000), cependant, les données concernant sa spécificité d'interaction pour les éléments SECIS et son rôle dans la régulation de la synthèse des sélénoprotéines sont conflictuelles. Une première étude a montré que l'affinité de la nucléoline pour des éléments SECIS distincts variait peu et qu'elle ne contribuait donc probablement pas à l'expression hiérarchique des sélénoprotéines (Squires et al., 2007b). Une seconde étude a, quant à elle, montré que la nucléoline liait de manière sélective les SECIS des ARNm de sélénoprotéines essentielles et présentait une faible affinité pour ceux des sélénoprotéines sensibles au stress (Miniard et al., 2010). L'inactivation de l'expression de la nucléoline par siARN, a cependant démontré que la nucléoline agissait comme un régulateur positif de la traduction des ARNm codant pour les sélénoprotéines essentielles mais pas pour celles sensibles au stress. Puisque l'expression de la nucléoline n'est pas régulée par le sélénium, il a été proposé qu'elle régula la synthèse des sélénoprotéines en stimulant le recrutement de SBP2 vers certains ARNm (Miniard et al., 2010).

I.2.4.2.a.4 SECp43

La protéine SECp43 est impliquée dans la méthylation de l'hydroxyle en 2' du ribose de l'uridine 34 de l'ARNt^{[Ser]^{Sec}} (Ding and Grabowski, 1999; Xu et al., 2005). Cette modification a une influence sur l'expression différentielle des sélénoprotéines (Carlson et al., 2005; Howard et al., 2013) et l'invalidation du gène codant pour SECp43 inhibe la synthèse des sélénoprotéines (Ding and Grabowski, 1999; Small-Howard et al., 2006; Xu et al., 2005). Il a également été montré que SECp43 interagit

avec le complexe eEFSec-Sec-ARNt^{[Ser]^{Sec}} *in vitro* et que la co-expression de SECp43 promeut l'interaction eEFSec-SBP2 *in vivo*. De plus, SECp43 augmente le taux d'incorporation de sélénocystéine (voir paragraphe I.2.4.2.a.1) et le niveau d'ARNm de sélénoprotéines probablement en déjouant le mécanisme du NMD. Elle semble également jouer un rôle dans la localisation cellulaire des facteurs de recodages (voir paragraphe I.2.4.3.a) (Small-Howard et al., 2006). Ainsi, SECp43 pourrait jouer un rôle clé dans l'orchestration des interactions et la localisation d'autres facteurs impliqués dans la biosynthèse des sélénoprotéines (Figure 19 et 24).

I.2.4.2.a.5 NSEP1

La protéine NSEP1 (Nuclease Sensitive Element Binding Protein 1), connue également sous les noms dbpB (DNA-binding protein 1B) ou YB-1 (Y-box binding protein 1), interagit avec l'ADN et a un rôle de facteur de transcription. NSEP1 a été identifiée comme protéine de liaison à l'élément SECIS (Fagegaltier et al., 2000c; Shen et al., 1998). Cette protéine est associée *in vivo* aux ARNm de sélénoprotéines et son épuisement par siARN inhibe la synthèse d'un rapporteur luciférase SECIS-dépendant, dont l'expression dépend du recodage d'un codon UGA. Ces données laissent suggérer que NSEP1 joue un rôle lors du mécanisme de recodage et pourrait être impliquée dans le mécanisme de régulation de l'expression des sélénoprotéines (Shen et al., 2006).

I.2.4.2.a.6 SBP2L

SBP2L est un paralogue de la protéine SBP2 chez les vertébrés. Bien que capable de se lier à l'élément SECIS, SBP2L ne permet pas de stimuler l'incorporation de sélénocystéine. En effet, SBP2 et SBP2L comprennent toutes les deux le domaine de liaison à l'ARN L7Ae qui est impliqué dans une interaction avec l'élément SECIS. Cependant, le domaine K-rich de SBP2L, contrairement à celui de SBP2, n'est pas impliqué dans la liaison au SECIS (Donovan and Copeland, 2012), suggérant un mode de liaison différent. Chez les invertébrés deutérostomes, SBP2L est en revanche l'unique protéine de liaison au SECIS et permet le recodage du codon UGA^{Sec} et donc la synthèse des sélénoprotéines (Donovan and Copeland, 2009, 2012). SBP2L est en compétition avec SBP2 pour la liaison à l'ARN SECIS et les deux protéines sont associées aux ARNm de GPx1 et GPx4 *in vivo*, indiquant que SBP2L pourrait jouer

un rôle dans leur régulation post-transcriptionnelle. Cependant, la surexpression transitoire de SBP2L dans les cellules de mammifères n'a pas d'effet sur la synthèse des sélénoprotéines, ce qui n'a pas permis de démontrer clairement la participation de SBP2L dans l'incorporation de sélénocystéine chez les mammifères (Donovan and Copeland, 2012).

I.2.4.2.a.7 GAPsec

La recherche de nouveaux facteurs impliqués dans le recodage du codon UGA^{Sec} a permis l'identification d'une protéine impliquée dans l'activation des GTPases (GTPase-activating protein, GAP) appelée GAPsec et conservée à travers l'évolution (Hirosawa-Takamori et al., 2009). GAPsec interagit avec eEFSec, stimulerait son activité GTPase et serait impliquée dans le mécanisme de translecture du codon UGA^{Sec}. GAPsec ne semble pas avoir d'effet direct sur le mécanisme d'incorporation de sélénocystéine (Hirosawa-Takamori et al., 2009). Cependant, la surexpression de GAPsec chez la drosophile et la souris augmente l'activité de redéfinition SECIS-dépendante du codon UGA d'un rapporteur luciférase. GAPsec semble avoir une fonction plus générale et être impliquée dans la redéfinition de codons stop autres que les codons sélénocystéines.

L'ensemble de ces protéines participent à la régulation fine et contrôlée de l'expression des sélénoprotéines.

I.2.4.2.b Régulation de la dégradation des ARNm de sélénoprotéines par le NMD

La diminution de l'abondance des ARNm des sélénoprotéines sensibles au stress, en condition limitante en sélénium a été attribuée à leur élimination par le mécanisme de dégradation des ARNm non-sens ou NMD (Nonsense Mediated Decay) (pour revue, Shetty and Copeland, 2015), plutôt qu'à une régulation transcriptionnelle (Moriarty et al., 1998; Weiss and Sunde, 1998) (Figure 19). Chez les eucaryotes, le NMD est un mécanisme de contrôle-qualité qui assure l'élimination de transcrits aberrants contenant un codon stop prématuré (PTC), évitant ainsi la synthèse de protéines tronquées et inactives par la cellule. Les ARNm qui contiennent un PTC situé à une distance supérieure à 50 nucléotides en amont du dernier intron en 3' sont des cibles

potentielles du NMD (Chang et al., 2007; Conti and Izaurralde, 2005; Nagy and Maquat, 1998; Popp and Maquat, 2014). Lors de l'épissage des pré-ARNm, un complexe multiprotéique ou EJC (Exon Junction Complex) est déposé 25 nucléotides en amont de la jonction exon-exon de l'ARNm. Les complexes EJC sont déplacés par les ribosomes lors du premier cycle de traduction. En cas de codon stop prématuré, les complexes EJC restent associés à l'ARNm jusque dans le cytoplasme où ils sont impliqués dans le recrutement des facteurs de dégradation du NMD (Le Hir and Séraphin, 2008). Parce qu'ils contiennent un codon UGA^{Sec} dans leur région codante, 14 des 25 ARNm de sélénoprotéines humains sont des cibles potentielles du NMD (Squires et al., 2007a; Moriarty et al., 1998; Sun et al., 2001; Weiss and Sunde, 1998). Certains ARNm de sélénoprotéines sont plus sensibles au NMD que d'autres, et cette sensibilité est tissu spécifique (Sun et al., 2001b). Différents critères tels que la position relative du UGA^{Sec} par rapport à l'élément SECIS, le contexte local du codon UGA^{Sec} et l'affinité de SBP2 pour les différents SECIS, ont été proposés comme étant impliqué dans la sensibilité des ARNm de sélénoprotéines au mécanisme du NMD (Low et al., 2000a; Squires et al., 2007b; Wen et al., 1998). Des expériences d'immunoprécipitation du facteur UPF1 essentiel au NMD ont révélé que les ARNm de sélénoprotéines les plus sensibles aux variations de sélénium sont également ceux qui sont les plus associés à UPF1 dans des conditions limitantes en sélénium (Seyedali et al., 2014) (Figure 21). De même, il apparaît que la délétion de SMG1, une autre protéine impliquée dans le mécanisme du NMD, abolit la dégradation des transcrits liés au stress (Seyedali et al., 2014). Enfin, il a été montré que l'expression du facteur eIF4A3, composant cœur du complexe EJC, était régulée par le niveau de sélénium. En conditions limitantes en sélénium eIF4A3 est capable de se lier à l'élément SECIS de GPx1, mais pas de GPx4, et d'entrer en compétition avec SBP2 pour inhiber l'incorporation de sélénocystéine (Budiman et al., 2009b). Il a été proposé que cela faciliterait également la dégradation de l'ARNm de GPx1 par le mécanisme du NMD (Morley and Willett, 2009). Ces résultats démontrent le rôle crucial du NMD dans la régulation du niveau des ARNm lorsque le sélénium est limitant et sa participation à l'expression hiérarchique des sélénoprotéines (Figure 19 et 21).

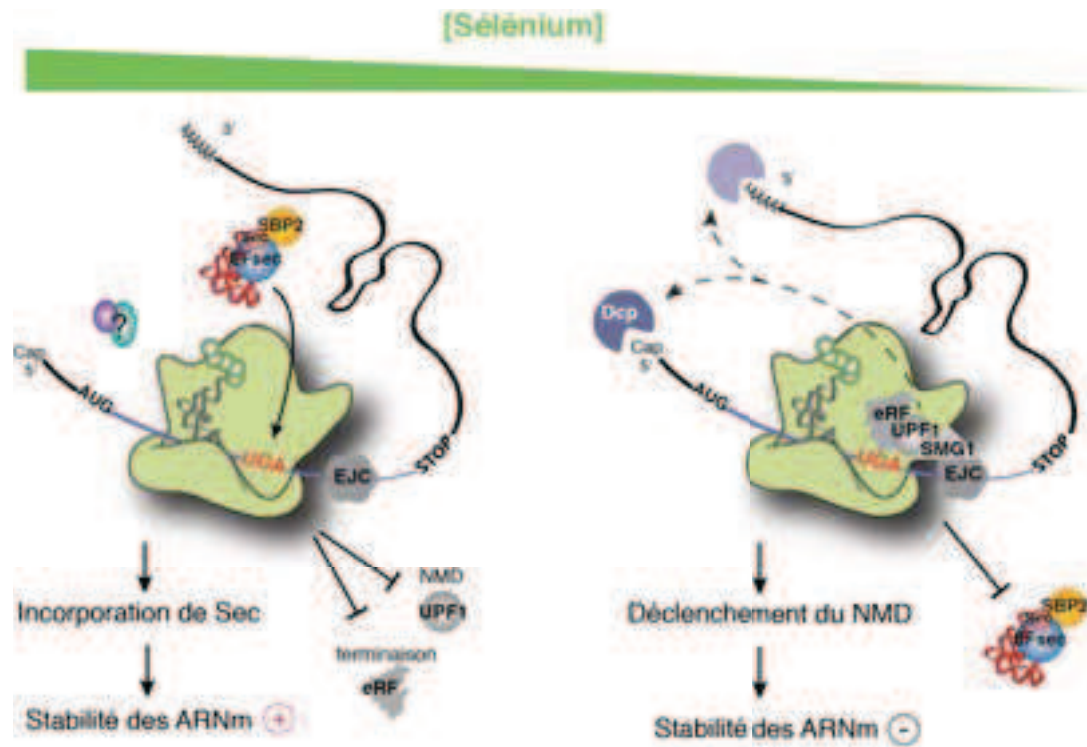


Figure 21. Le mécanisme du NMD est impliqué dans la régulation post-transcriptionnelle sélénium-dépendante des ARNm de sélénoprotéines.

Les ARNm de sélénoprotéines sensibles au niveau de sélénium sont la cible du mécanisme du NMD (Moriarty et al., 1998; Seyedali and Berry, 2014; Squires et al., 2007a; Sun et al., 2001b; Weiss and Sunde, 1998). En conditions de sélénium optimales le Sec-ARNt^{[Ser]^{Sec}} permet l'incorporation de sélénocystéine et bloque l'accès aux facteurs de terminaison empêchant ainsi le déclenchement du mécanisme du NMD. Lorsque le sélénium est limitant, le niveau du Sec-ARNt^{[Ser]^{Sec}} diminue ce qui augmente la possibilité d'accès des facteurs de terminaison au niveau du codon UGA^{Sec} et l'activation du mécanisme du NMD entraînant la dégradation des ARNm de sélénoprotéines. Cette dégradation se fait via le recrutement d'enzymes d'élimination de la coiffe (Dcp), de déadénylation, puis d'exonucléases (Popp and Maquat, 2013).

1.2.4.3 L'assemblage des mRNP de sélénoprotéines

L'assemblage des particules ribonucléoprotéiques (RNP) est une étape essentielle pour l'obtention de particules matures et fonctionnelles. Ce processus, souvent complexe, requiert l'intervention de nombreux facteurs d'assemblage et a parfois lieu dans un compartiment différent de celui où la RNP mature exerce sa fonction. Il en est ainsi pour les mRNP de sélénoprotéines, dont l'assemblage est un prérequis pour leur traduction.

I.2.4.3.a Un assemblage nucléaire des mRNP de sélénoprotéines

La protéine SBP2 est principalement associée aux ribosomes et aux ARNm dans le cytoplasme. Cependant, deux sites NES (Nuclear Export Signal) et un site NLS (Nuclear Localisation Signal) ont été identifiés dans sa séquence peptidique (Figure 22).

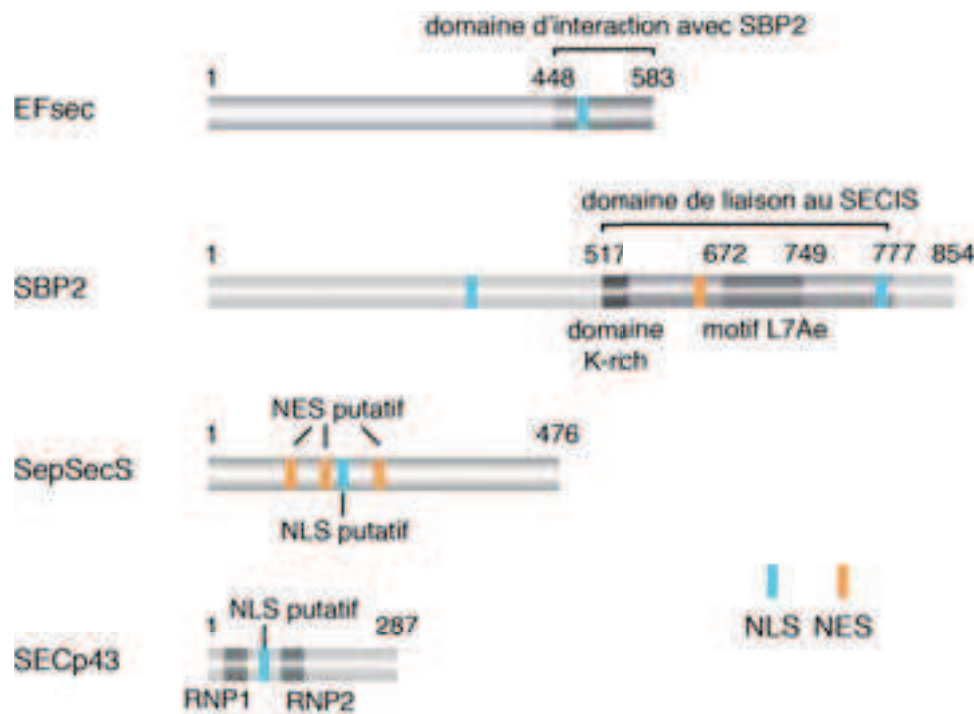


Figure 22. Représentation schématique de facteurs impliqués dans l'assemblage nucléaire des mRNP de sélénoprotéines.

Les signaux de localisation nucléaire (NLS) sont représentés en bleu et les signaux d'export nucléaire (NES) en orange. Certains domaines fonctionnels importants sont représentés. RNP: domaine d'interaction avec les particules ribonucléoprotéiques (RNP binding domain) qui constitue un motif d'interaction avec l'ARN (Papp et al., 2007; Small-Howard and Berry, 2005).

De manière intéressante, ces signaux de localisation cellulaire permettent à la protéine SBP2 de transiter entre le noyau et le cytoplasme (de Jesus et al., 2006; Papp et al., 2006) (Figure 23). Ce transport nucléo-cytoplasmique est dépendant du facteur d'export CRM1 (Papp et al., 2006) qui interagit directement avec le signal NES (Fornerod et al., 1997a; Stade et al., 1997). De façon intéressante, il a été montré que le stress oxydant a un effet sur la localisation subcellulaire de SBP2 (Papp et al., 2006; Touat-Hamici et al., 2014). Ainsi, lors d'un stress oxydant, les cystéines présentes dans les séquences NES et dans le domaine de liaison à l'ARN de SBP2 (Figure 22), forment des ponts disulfures (S-S et S-SG) (Figure 23). Ceci induirait un

changement conformationnel de SBP2, l'empêchant d'interagir avec CRM1 et entraînant son accumulation dans le noyau (Papp et al., 2006). Ces modifications peuvent être réduites *in vitro* par les systèmes thiorédoxine et glutathion (impliquant des sélénoprotéines), indiquant qu'elles pourraient être réversibles *in vivo*. Ces modifications seraient un moyen de régulation négative (Papp et al., 2006) ou différentielle (Touat-Hamici et al., 2014) de l'expression des sélénoprotéines.

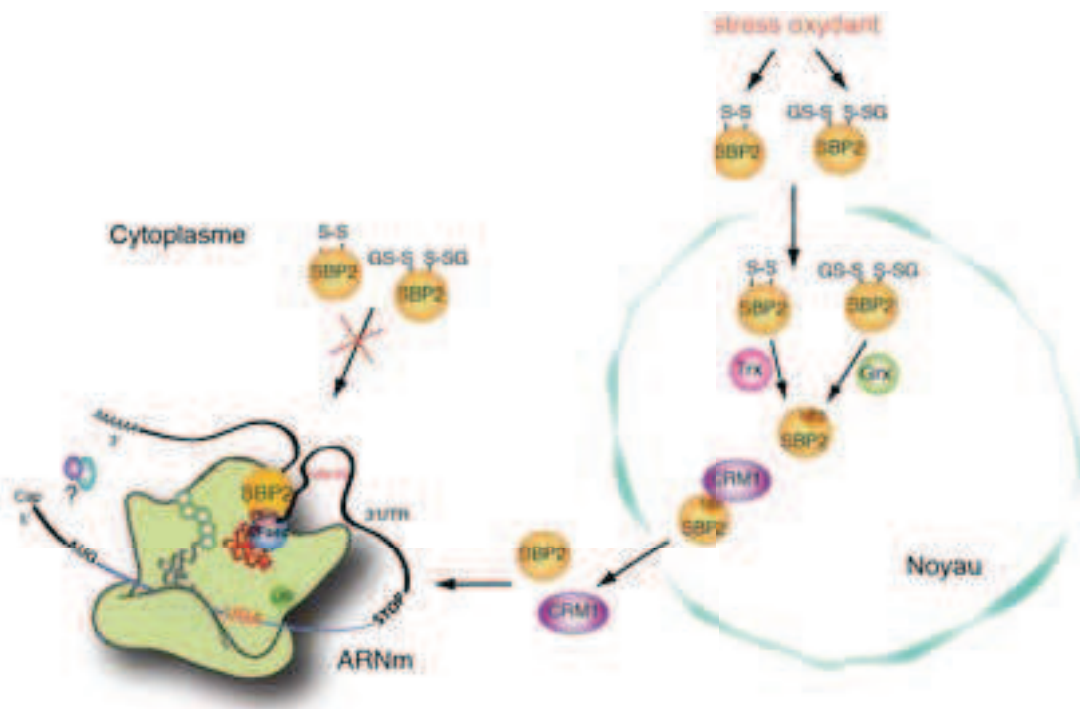


Figure 23. Régulation de la localisation sub-cellulaire de SBP2 en réponse à un stress oxydant. (d'après Allmang et al., 2009; Papp et al., 2006)

Des résidus cystéines de SBP2 sont oxydés lors d'un stress oxydant, entraînant la formation de ponts disulfures (S-S et S-SG). Ceci induit un changement de conformation de la protéine qui provoque une dissimulation des sites NES empêchant l'interaction de SBP2 avec le facteur d'export nucléaire CRM1. SBP2 s'accumule dans le noyau et la synthèse de certaines sélénoprotéines est inhibée. Cet effet est réversible puisque la thiorédoxine et la glutathion peuvent réduire les résidus Cys et SBP2 peut à nouveau être transportée dans le cytoplasme, où a lieu la synthèse des sélénoprotéines.

Le facteur d'élongation eEFSec comporte également des signaux NES et NLS (Figure 22) et sa localisation, comme pour SBP2, varie selon le type cellulaire (de Jesus et al., 2006). eEFSec est localisé majoritairement dans le cytoplasme mais la cotransfection de SBP2 stimule sa localisation dans le noyau (Figure 24). Enfin, la protéine SECp43, également localisée dans le noyau, interagit avec l'ARNt^{[Ser]Sec} et guiderait la localisation de SPS1 et de SepSecS dans le noyau (Small-Howard et al., 2006) (Figure 24). Puisque SECp43 stimule également l'interaction entre SBP2 et eEFSec, cette

protéine jouerait un rôle important dans le mécanisme de synthèse des sélénoprotéines en stimulant les interactions entre les facteurs impliqués dans l'incorporation de sélénocystéine et en redistribuant plusieurs sous-complexes vers le noyau.

La localisation nucléaire des facteurs impliqués dans la synthèse des sélénoprotéines suggère que l'assemblage des mRNP de sélénoprotéines puisse avoir lieu dans ce compartiment (Figure 24). Ceci permettrait aux ARNm de sélénoprotéines d'être associés aux facteurs de recodage avant d'être exportés dans le cytoplasme et pourrait ainsi constituer un moyen d'échapper au mécanisme du NMD (voir paragraphe I.2.4.2.b). Ainsi, l'assemblage des mRNP de sélénoprotéines dans le noyau pourrait constituer un moyen de régulation différentielle de l'expression des sélénoprotéines en protégeant les ARNm de sélénoprotéines contre le NMD. Les facteurs resteraient associés à l'ARN lors du transport nucléo-cytoplasmique et permettraient ainsi la synthèse des sélénoprotéines dans le cytoplasme.

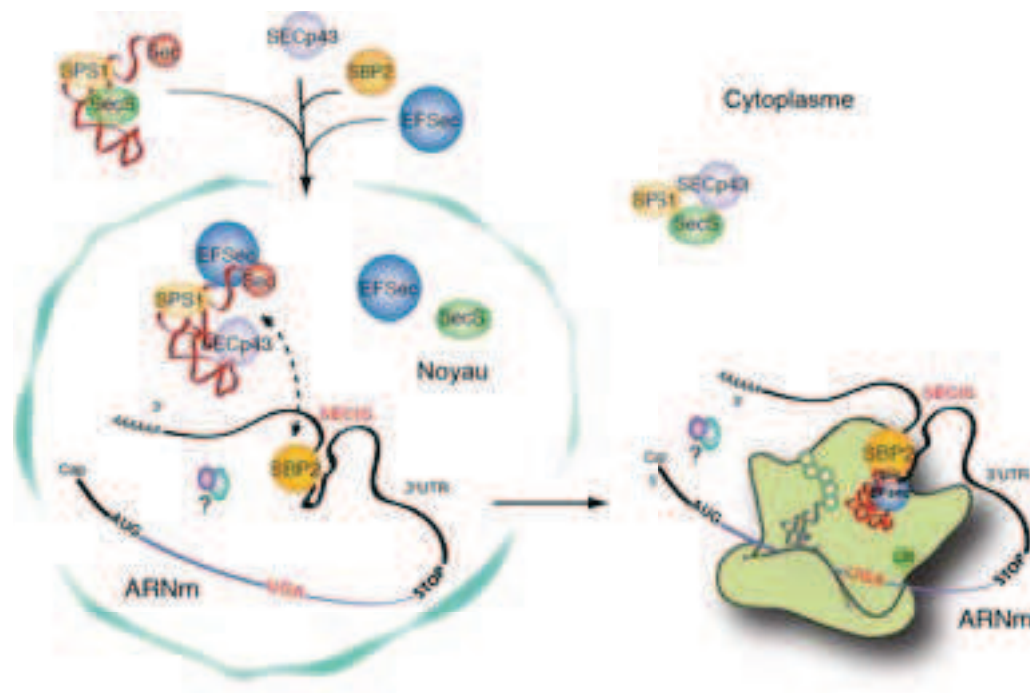


Figure 24. Modèle de l'assemblage nucléaire des mRNP de sélénoprotéines (d'après Allmang et al., 2009)

L'ensemble des facteurs nécessaires à l'incorporation de sélénocystéine sont importés dans le noyau. Secp43 stimule l'interaction entre SBP2 et eEFSec et les relocalise dans le noyau tout comme SPS1 et SepSecS associées au Sec-ARNt^{[Ser]Sec} (de Jesus et al., 2006; Small-Howard et al., 2006). Dans le noyau, SBP2 se fixe à l'élément SECIS et recrute les autres facteurs. Ce modèle propose que les mRNP de sélénoprotéines soient pré-assemblées dans le noyau, puis exportées dans le cytoplasme où ils sont traduits.

I.2.4.3.b Assemblage des ARNm de sélénoprotéines : similitudes avec l'assemblage des sn/snoRNP

Le domaine de liaison à l'ARN de SBP2 appartient à la famille L7Ae (voir paragraphe I.2.2.2.b.1). Les protéines de cette famille sont trouvées dans de nombreuses RNP essentielles (Watkins et al., 2000). Toutes les protéines L7Ae lient une structure d'ARN commune appelée K-turn (voir paragraphe I.2.2.1.a). L'interaction des protéines L7Ae avec leur ARN cible permet le recrutement de facteurs additionnels (protéines cœur, facteurs d'assemblage) au niveau de la RNP. L'assemblage des mRNP de sélénoprotéines obéit aux mêmes règles que celui des sn/snoRNP. En effet, notre laboratoire a montré que cet assemblage faisait appel à une machinerie moléculaire liée au chaperon protéique HSP90, conservée de la levure à l'homme et d'importance fondamentale car impliquée dans le contrôle de la prolifération cellulaire (Boulon et al., 2008). HSP90 intervient également dans la réponse aux chocs thermiques et au stress oxydant. Cette machinerie moléculaire est commune à l'assemblage de toutes les RNP de la famille L7Ae, incluant la snRNP U4, les snoRNP à boîte C/D et H/ACA, la télomérase et les mRNP de sélénoprotéines (Boulon et al., 2008; Zhao et al., 2008). Elle est composée d'un complexe de co-chaperons associé à HSP90 et d'une protéine adaptatrice Nufip (Rsa1 chez la levure). Le complexe de co-chaperons, appelé R2TP (Kakihara and Houry, 2012), est composé de deux ATPases AAA+ (Rvb1 et Rvb2) (Nano and Houry, 2013) et des protéines hSpagh (Tah1 chez la levure) et hPih1 (Gribun et al., 2008; Zhao et al., 2005b, 2008). Nous avons montré que HSP90 contrôlait le repliement correct de SBP2 et des protéines L7Ae des sn(o)RNP lors de la formation de complexes ARN/protéines. Nufip interagit avec SBP2 et toutes les protéines L7Ae, elle joue le rôle d'adaptateur en permettant le recrutement à la fois du complexe R2TP/HSP90 et des autres protéines cœur vers les RNP (Boulon et al., 2008). Une interaction directe a également été observée entre le co-chaperon Pih1 et SBP2 (Boulon et al., 2008) (Figure 25). HSP90 est impliquée dans des processus de cancérisation chez l'homme et contrôle le repliement des récepteurs nucléaires et des protéines kinases (Pearl & Prodromou, 2006). HSP90 semble donc également jouer un rôle dans la prolifération cellulaire par le biais de l'assemblage de complexes RNP dont les mRNP des sélénoprotéines.

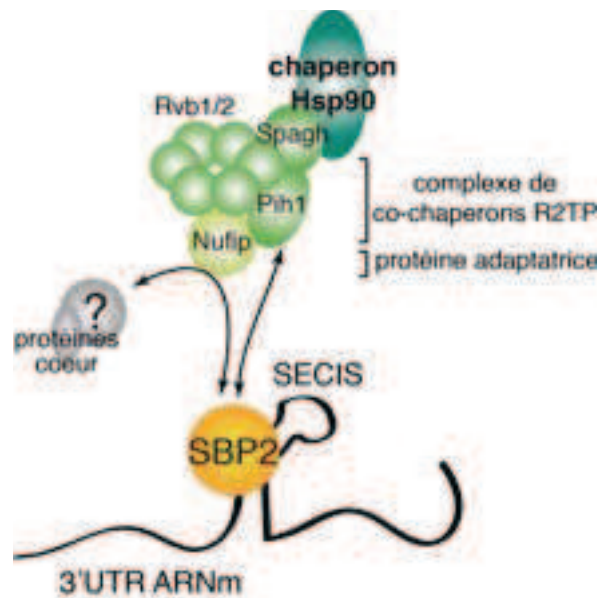


Figure 25. Représentation schématique du complexe d'assemblage impliqué dans la biogenèse des mRNP de sélénoprotéines. (d'après Boulon et al., 2008)

La stabilité et le repliement de la protéine SBP2 sont dépendants de son association avec le chaperon protéique Hsp90 et ses co-facteurs, le complexe R2TP. Celui-ci est composé des ATPases Rvb1/Rvb2, des protéines Spagh et Pih1. La protéine Nufip joue le rôle d'adaptateur. Elle interagit avec la protéine SBP2, mobilise cette machinerie d'assemblage au niveau des mRNP de sélénoprotéines et des RNP L7Ae de façon générale et recrute d'autres protéines cœur des RNP.

II Mécanisme de biogenèse des sn et snoRNP à boîte C/D

Des similitudes inattendues ont été révélées entre les mRNP de sélénoprotéines et les snRNP et les snoRNP à boîte C/D, du point de vue de la composition et de l'assemblage des particules (voir paragraphes I.2.2.2.b.1 et I.2.4.3.b). Au cours de ma thèse, nous avons mis en évidence d'autres points communs au niveau du mécanisme de biogenèse de ces RNP.

II.1 Biogenèse des snRNP

La biogenèse des snRNP est un processus complexe qui implique plusieurs étapes de maturation et d'assemblage et requiert plus de facteurs d'assemblages que de protéines présentes dans la particule mature fonctionnelle (Fischer et al., 2011; Neuenkirchen et al., 2008, 2015). Les snARN sont très abondants, non-polyadénylés et forment un groupe de petits ARN non-codants. A l'exception de U7, impliqué dans la maturation de l'extrémité des pré-ARNm d'histone, la majorité des snARN forment le complexe du spliceosome et sont responsables de l'épissage des pré-ARNm dans le

noyau (pour revue, Chen and Moore, 2014; Matera and Wang, 2014; Matera et al., 2007) (Figure 26). Il existe deux formes du spliceosome. Le spliceosome majeur est responsable de l'épissage de la majorité des introns (U2-dépendants) et est composé des snARN U1, U2, U4, U5 et U6. La forme mineure quant à elle est responsable de l'épissage des introns U12-dépendants ou ATAC, représentant moins de 0,5% de la totalité des introns, et est composée des snARN U11, U12, U4atac, U5 et U6atac (pour revue, Patel and Steitz, 2003; Turunen et al., 2013). Les snARN peuvent être divisés en deux classes, sur la base de caractéristiques de séquences et de cofacteurs protéiques communs. La classe Sm des snARN présente une coiffe triméthylée $m_3^{2,2,7}G$ (méthyl-2,2,7guanosine ou m_3G) en 5', une tige-boucle en 3' et un site de liaison riche en uridine (PuAU4-6GPu) pour 7 protéines Sm (B/B', D1, D2, D3, E, F et G) qui forment une structure en anneau heptamérique (Figure 27) (Branlant et al., 1982; Neuenkirchen et al., 2008, 2015). Cette classe comprend les snARN U1, U2, U4, U4atac U5, U7, U11 et U12. La classe des Lsm, représentée par les snARN U6 et U6atac, est caractérisée par une coiffe monométhylée m^7G et une tige-boucle en 3' et se terminant par un prolongement d'uridines qui forme le site de liaison d'un anneau heptamérique distinct de protéines Lsm (2-8) (He and Parker, 2000; Khusial et al., 2005).

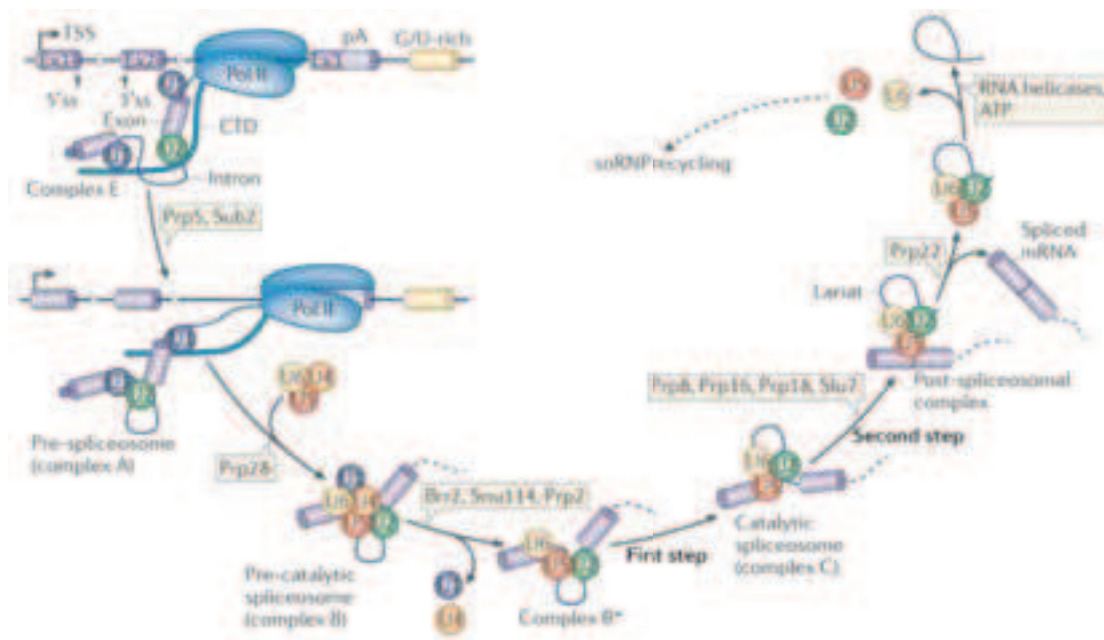


Figure 26. Assemblage du spliceosome majeur et épissage des pré-ARNm. (d'après Matera and Wang, 2014)

Les snRNP U1 et U2 s'assemblent sur le pré-ARNm de manière co-transcriptionnelle et reconnaissent les sites d'épissage 5' (5'ss) et 3' (3'ss). Ceci est modulé par le domaine CTD (C-terminal) de l'ARN polymérase II. Les snRNP U1 et U2 interagissent ensemble et forment le complexe A. Le complexe tri-snRNP U4-U6-U5 est alors recruté pour former le complexe B. Ce complexe va subir une série de

réarrangements pour former le complexe catalytiquement actif (complexe B*), ceci requiert de nombreuses hélicases et protéines de liaison à l'ARN et résulte en la libération des snRNP U4 et U1. Le complexe B* effectue alors la première étape catalytique de l'épissage, générant le complexe C qui contient l'exon 1 libre (Ex1) et l'intermédiaire intron-exon 2 en lasso. Le complexe C subit des réarrangements additionnels et réalise la seconde étape catalytique, résultant en un complexe post-spliceosome qui contient l'intron sous forme de lasso et les exons épissés. Finalement les snRNP U2, U5 et U6 sont libérés de la mRNP et recyclés. L'intron en forme de lasso est débranché et dégradé. Les noms des facteurs d'épissage sont indiqués dans les encarts.

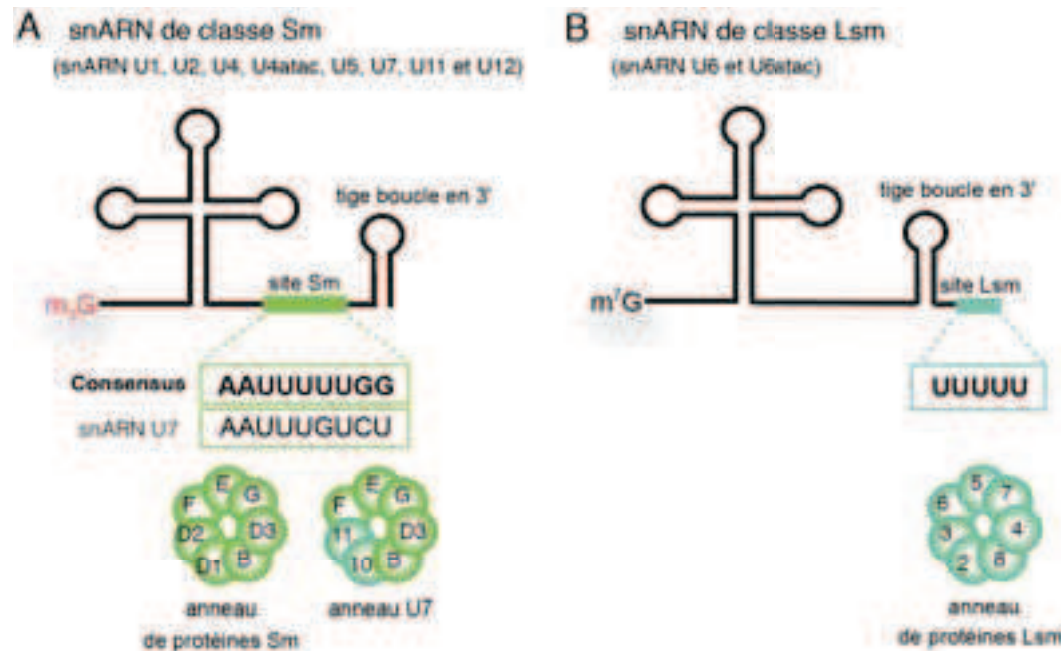


Figure 27. Caractéristiques des snARN de classes Sm et Lsm. (d'après Matera et al., 2007)

(A) Les snARN de la classe Sm sont transcrits par l'ARN polymérase II et contiennent un site de liaison aux protéines Sm. Durant leurs étapes de maturation dans le cytoplasme, ils acquièrent une coiffe m⁷G hyperméthylée ; l'anneau des 7 protéines Sm est assemblé au niveau du site Sm. L'extrémité 3' des ARN se termine par une structure en tige-boucle. Dans la snRNP U7, les deux protéines SmD1 et D2 se remplacent par Lsm10 et Lsm11. La séquence consensus du site Sm et de celle du snARN U7 sont indiquées. (B) Contrairement aux snARN de la classe Sm, les snARN Lsm sont transcrits par l'ARN polymérase III, contiennent un site de liaison aux protéines Lsm en aval d'une structure en tige-boucle en 3' et portent une coiffe monométhylée m⁷G.

II.1.1 Transcription et export dans le cytoplasme

Comme les ARNm, les précurseurs (pré-snARN) des snARN U1, U2, U4, U4atac, U5, U11 et U12 sont transcrits par l'ARN polymérase II et sont dotés d'une coiffe monométhylée (m⁷G), ajoutée de manière co-transcriptionnelle à leur extrémité 5' (Matera et al., 2007b; Westin et al., 1984). Ils ne sont cependant pas polyadénylés en 3' (Figure 28). Leur coiffe m⁷G est reconnue par le complexe CBC (Cap Binding Complex) (Izaurrealde et al., 1994). Les particules sont ensuite exportées dans le cytoplasme via le récepteur d'export CRM1 (chromosome region maintenance-1)

associé à RanGTP (Ras-related nuclear protein) et à l'aide d'un adaptateur spécifique, PHAX, (Askjaer et al., 1999; Fornerod et al., 1997b; Ohno et al., 2000). Les snARN U6 et U6atac quant à eux sont transcrits par l'ARN polymérase III, possèdent une coiffe atypique (γ -monométhyl-phosphate) et ne sont pas exportés dans le cytoplasme (Kunkel et al., 1986; Reddy et al., 1987; Singh and Reddy, 1989).

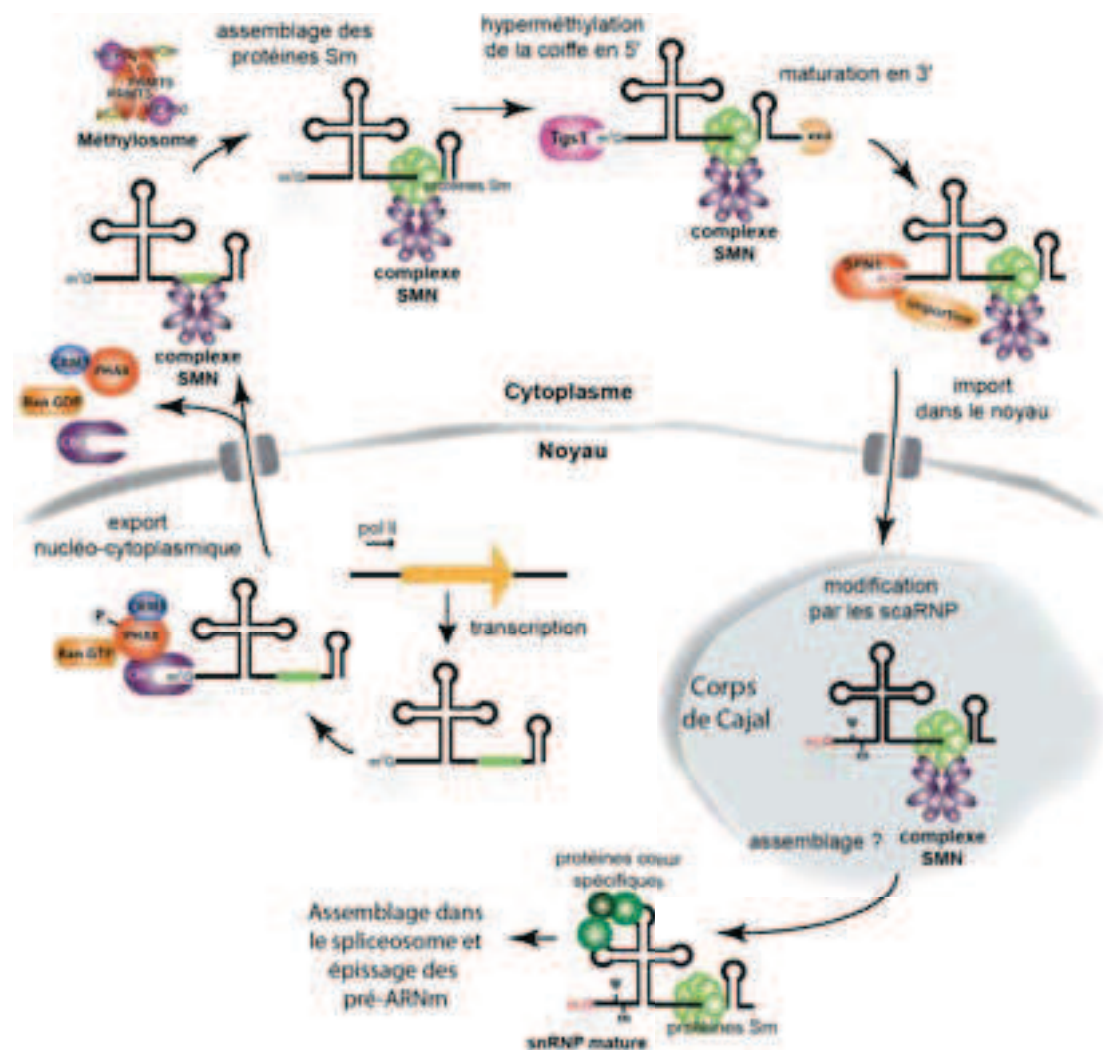


Figure 28. Biogenèse des snRNP de la classe Sm chez les vertébrés. (d'après Neuenkirchen et al., 2008; Patel et Bellini, 2008; Paushkin et al., 2002)

Dans le noyau, les pré-snARN sont transcrits par l'ARN polymérase II et portent une coiffe monométhylée m⁷G reconnue par le complexe CBC (Cap Binding Complex). Ils sont pris en charge par l'adaptateur PHAX, le facteur CRM1 et RanGTP afin d'être exportés vers le cytoplasme. Dans ce compartiment, le complexe SMN assisté du méthylosome assemble les protéines Sm sur les snARN. L'enzyme TGS1 est recrutée par le complexe SMN associé aux protéines Sm et hyperméthyle la coiffe des snARN en m_{3,2,7}G (ou m₃G). L'extrémité 3' des snARN est maturée par les exonucléases de l'exosome. Les snRNP sont réimportés dans le noyau par la snurportine-1 et l'importine β et s'accumulent dans les corps de Cajal où elles seront modifiées par les scaRNP (ψ: pseudouridylation, m: 2'-O-méthylation) et recruteront d'autres protéines cœur spécifiques. Les snRNP matures constituent le spliceosome.

II.1.2 Maturation et assemblage des snRNP dans le cytoplasme

Une fois dans le cytoplasme, les snRNP subissent des étapes de maturation, de modification et d'assemblage. Les 7 protéines Sm, préalablement méthylées par le méthylosome, sont transférées sur les snARN dans une conformation en anneau par le complexe SMN (Survival of Motor Neuron) (Kambach et al., 1999; Neuenkirchen et al., 2015) (Figure 28 et paragraphe II.1.4). La coiffe monométhylée m⁷G des snARN est alors hyperméthylée en coiffe m₃G par la triméthylguanosine synthase 1 (Tgs1) (Hausmann and Shuman, 2005; Lemm et al., 2006; Mouaikel et al., 2002; Plessel et al., 1994) (Figure 28 et paragraphe II.3). Enfin, l'extrémité 3' des snARN est maturée par l'exosome, un complexe d'exonucléases 3'-5' (Allmang et al., 1999; Dahlberg et al., 1990; van Hoof et al., 2000). L'ensemble de ces événements permet le réimport des snRNP dans le noyau (Huang and Pederson, 1999; Rollenhagen and Panté, 2006).

II.1.3 Réimport, assemblage et modification des snRNP dans le noyau

Le réimport des snRNP dans le noyau nécessite un signal de localisation nucléaire (NLS) bipartite (Fischer and Lührmann, 1990; Fischer et al., 1993). En effet, la coiffe m₃G constitue la première partie du NLS et est spécifiquement reconnue par la protéine snurportine-1 (SPN1) (Huber et al., 1998; Strasser et al., 2004) (Figure 28). Le second site de fixation serait constitué du complexe SMN associé aux protéines Sm, et serait reconnu par l'importine β (Meister et al., 2002; Narayanan et al., 2002; Palacios et al., 1997). Lorsque les deux facteurs snurportine-1 et importine β sont associés aux signaux NLS respectifs, les particules sont alors importées dans le noyau. Les snRNP, associés au complexe SMN s'accumulent au niveau d'organelles appelées corps de Cajal (Cajal, 1903) qui constituent des sites d'assemblage de diverses RNP (Lamond and Sleeman, 2003; Morris, 2008; Sleeman and Lamond, 1999). Les snARN y sont alors pseudouridylés et 2'-O-méthylés par des petits ARN appelés scaARN (small cajal RNA) (Figure 28) (Darzacq et al., 2002; Kiss et al., 2004). Des protéines cœur spécifiques des snRNP sont également assemblées sur les snARN dans ce compartiment (Patel and Bellini, 2008; Schaffert et al., 2004; Staněk et al., 2003). Les snRNP matures et fonctionnelles sont stockées dans le noyau, dans des structures dynamiques appelées "speckles" (Spector, 1993; Spector and Lamond, 2011) et dans lesquelles l'épissage des pré-ARNm aurait lieu.

II.1.4 Le complexe SMN et le méthylosome fonctionnent de manière coordonnée dans l'assemblage des snRNP

L'auto-assemblage des 7 protéines Sm sur les snARN est possible *in vitro* (Hermann et al., 1995; Raker et al., 1996), mais dans les cellules, ce processus requiert la participation de deux complexes protéiques en *trans* : le méthylosome et le complexe SMN (Meister et al., 2002; Paushkin et al., 2002) (Figure 29).

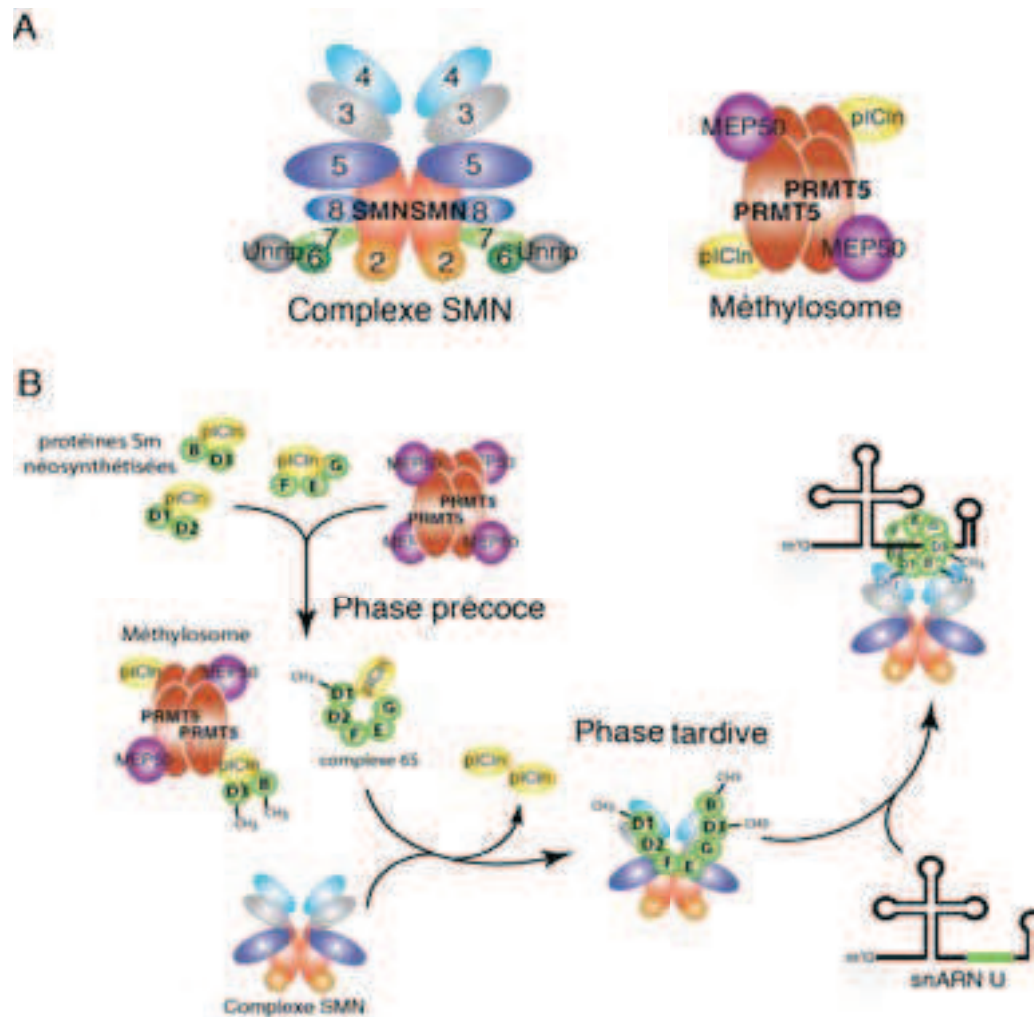


Figure 29. Modèle du mécanisme d'assemblage cytoplasmique des protéines Sm sur les snARN (d'après Chari et al., 2008; Grimm et al., 2013; Li et al., 2014; Neuenkirchen et al., 2008; Stopa et al., 2015; Yong et al., 2004)

(A) Représentation du complexe SMN et du méthylosome. Le complexe SMN est composé des protéines SMN, des Gemin 2 à 8 et de la protéine Unrip. Le méthylosome comprend la méthyltransférase PRMT5, son cofacteur MEP50 et le chaperon pICln. Le complexe SMN et le méthylosome sont représentés de manière non-stoechiométrique (Pellizzoni, 2007; Stopa et al., 2015). (B) Le complexe SMN et le méthylosome fonctionnent de manière coordonnée pour assembler les protéines Sm sur les snARN. Les protéines Sm néosynthétisées sont prises en charge par pICln qui forme des complexes intermédiaires avec les protéines Sm D1/D2, E/F/G ainsi qu'avec Sm B et D3. PRMT5 catalyse ensuite la formation de diméthylarginines symétriques (sDMA) en C-terminal des protéines Sm B, D1 et D3. Les protéines Sm modifiées sont ensuite transférées par pICln sous forme de complexe 6S (pICln/D1/D2/F/E/G) et pICln/B/D3 vers le complexe SMN qui les assemble sur les snARN. Le complexe SMN reste associé au snARN pour guider d'autres événements de maturation.

II.1.4.1 Le méthylosome

Le méthylosome est un complexe protéique composé de l'enzyme PRMT5 (Protein arginine N-methyl transferase 5), de MEP50 (methylosome protein 50) et de pICln (chloride conductance regulatory protein) (Friesen et al., 2001; Meister et al., 2001a; Pu et al., 1999) (Figure 29A). Le méthylosome est impliqué dans la méthylation des protéines Sm au niveau de sites RG qui est prérequis pour leur assemblage sur les snARN (Thandapani et al., 2013). PRMT5 est une protéine arginine méthyltransférase de type II qui catalyse la formation de ω -N^G-monométhylarginine (MMA) et de ω -N^G, N^G-diméthylarginine (sDMA) symétrique. La protéine MEP50 se lie directement à PRMT5 et augmente de manière significative son activité de méthyltransférase en ayant un effet allostérique positif sur la liaison de PRMT5 à ses autres cofacteurs ou à son substrat, la S-adénosyl-méthionine (SAM) (Antonysamy et al., 2012; Ho et al., 2013). Le cœur du méthylosome est organisé sous forme de structure hétéro-octamérique composée de 4 protéines PRMT5 centrales et de 4 protéines MEP50 périphériques (Stopa et al., 2015). La protéine pICln fonctionne comme un adaptateur et un chaperon qui va sélectionner et apporter les protéines cibles vers PRMT5/MEP50 et permettre ainsi leur méthylation (Barbarossa et al., 2014; Pesiridis et al., 2009). En plus de leur rôle dans l'assemblage des snRNP (voir ci-dessous), PRMT5 et MEP50 peuvent être associés à d'autres co-facteurs et sont responsables de la méthylation de diverses protéines au niveau de sites RG (Branscombe et al., 2001; Thandapani et al., 2013). Ils sont ainsi impliqués dans une variété de processus cellulaires (Karkhanis et al., 2011) incluant la biogénèse des ribosomes (Ren et al., 2010), l'assemblage de l'appareil de Golgi (Zhou et al., 2010) ou encore la différenciation cellulaire (Ancelin et al., 2006; Dacwag et al., 2007, 2009). PRMT5 joue un rôle important dans la méthylation des histones H2AR3, H4R3, H3R2, et H3R8 *in vivo*, ce qui implique l'enzyme dans des événements de régulation de la transcription (Ancelin et al., 2006; Pal et al., 2004; Seth-Vollenweider et al., 2014; Xu et al., 2010a).

II.1.4.2 Le complexe SMN

La protéine SMN est une protéine ubiquitaire conservée à travers l'évolution, qui est localisée à la fois dans le cytoplasme et le noyau au niveau des corps de Cajal (Liu and Dreyfuss, 1996 ; pour revue, Li et al., 2014a). SMN s'associe avec 8 autres

protéines (Gemin2-8 et Unrip) pour former un complexe macromoléculaire : le complexe SMN (Figure 29A) (Otter et al., 2007; Pellizzoni, 2007). Une des caractéristiques du complexe SMN est sa capacité d'oligomérisation conduisant à la formation de particules dont la taille varie de 20S à 80S (Battle et al., 2007; Carissimi et al., 2006; Martin et al., 2012; Pellizzoni et al., 1999). La protéine SMN contient un domaine conservé à travers l'évolution impliqué dans les interactions protéines-protéines, incluant un domaine Tudor qui lie des arginines diméthylées présentes dans de nombreux partenaires de liaison à SMN (Tripsianes et al., 2011). Bien qu'essentiels à sa fonction, ni la stoechiométrie exacte du complexe SMN, ni son mécanisme d'assemblage ne sont élucidés (Borg et al., 2015; Gupta et al., 2015; Zhang et al., 2011). Le rôle le mieux compris et le plus étudié du complexe SMN est celui de son implication dans la biogenèse des snRNP (Meister et al., 2001b; Pellizzoni, 2002) (Figures 28 et 29B). Il est indispensable à la formation des spliceosomes majeur et mineur (Boulisfane et al., 2011) et donc à l'expression des gènes contenant des introns chez les eucaryotes (voir paragraphe ci-dessous). La fonction des protéines Gemin 5, Gemin2 et Gemin3 du complexe SMN, dans l'assemblage des protéines Sm sur les snARN commence à être comprise. Gemin5 semble être le facteur responsable de la reconnaissance des snARN dans le cytoplasme (Battle et al., 2006; Yong et al., 2010) qu'il lie via son domaine "WD-repeat" localisé en N-terminal (Lau et al., 2009). Les snARN sont pris en charge par Gemin5 qui les délivre au complexe SMN, probablement à proximité de Gemin2, pour l'assemblage des protéines cœurs Sm (Yong et al., 2010). La majorité des protéines Sm sont reconnues par Gemin2 qui stabilise le pentamère de protéines Sm du complexe 6S sous forme de croissant (Figure 29B). La partie N-terminale de Gemin2 lie la poche de liaison aux snARN du pentamère de protéines Sm, pour empêcher leur interaction avec des ARN aspécifiques (Grimm et al., 2013; Zhang et al., 2011). L'activité chaperonne du complexe SMN ainsi que l'hydrolyse de l'ATP durant la réaction d'assemblage est probablement exécutée par la protéine Gemin3, une hélicase à motif "DEAD" (Campbell, 2000; Charroux et al., 1999; Yan et al., 2003).

Des travaux récents montrent que le complexe SMN joue également un rôle dans la biogenèse de nombreuses RNP autres que les snRNP (Figure 30). Ce complexe est

impliqué dans l'assemblage des anneaux hybrides Sm/Lsm de la snRNP U7 essentielle à la maturation de l'extrémité 3' des ARNm d'histones (Pillai, 2003; Tisdale et al., 2013a); de l'anneau Lsm (2-8) des snARN U6 et U6atac, constituants du spliceosome ; du complexe cytoplasmique Lsm (1-7) qui fonctionne dans la régulation des ARNm via son association avec l'extrémité 3' de transcrits cibles et le recrutement de facteurs d'élimination de la coiffe pour leur dégradation 5'-3' (Brahms et al., 2001a; Burghes and Beattie, 2009a; Friesen, 2000; Tharun, 2008) (Figure 30).

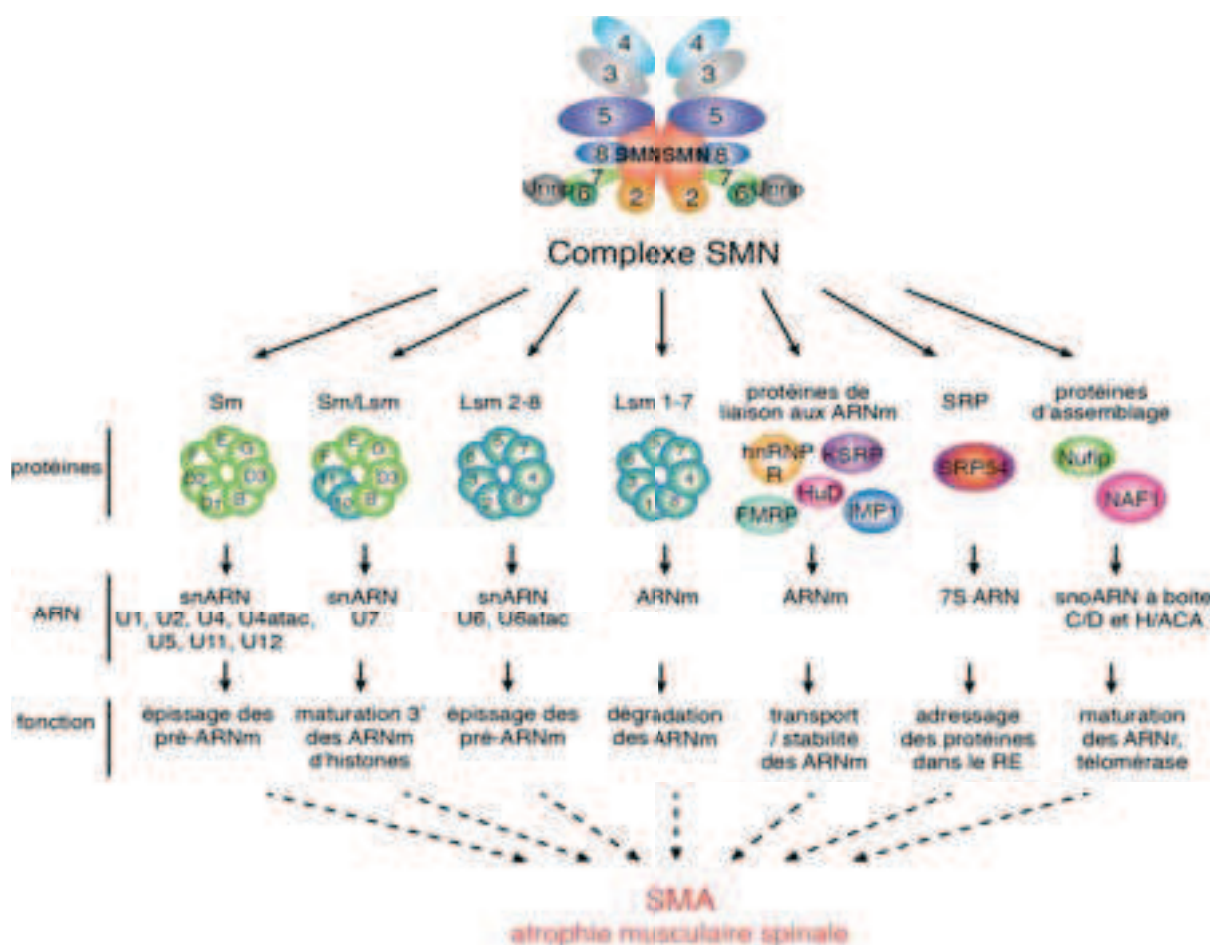


Figure 30. Les voies d'assemblage de RNP dépendantes du complexe SMN. (d'après Li et al., 2014) Les composants protéiques assemblés par le complexe SMN sur les différentes RNP cibles sont représentés au-dessus des ARN auxquels ils sont associés (Akten et al., 2011; Boulisfane et al., 2011; Brahms et al., 2001b; Fallini et al., 2011, 2012a, 2014; Friesen and Dreyfuss, 2000; Hubers et al., 2011; Piazzon et al., 2008, 2013; Pillai et al., 2003; Tadesse et al., 2008; Tisdale et al., 2013b et Thèse de M. Dodré, équipe de S. Massenet, données non-publiées). SRP, Signal Recognition Particle ; RE, réticulum endoplasmique.

Le complexe SMN est également requis pour l'assemblage de la protéine SRP54 sur l'ARN 7S de la particule SRP (signal recognition particle), une RNP qui adresse les protéines en cours de traduction dans le réticulum endoplasmique (RE) (Piazzon et

al., 2013). L'assemblage des snoRNP à boîte C/D et H/ACA, incluant la télomérase, ferait également intervenir le complexe SMN (données non-publiées, Thèse de M. Dodré, S. Massenet, C. Allmang). Le complexe SMN interagit avec de nombreuses autres protéines de liaison à l'ARN (Fallini et al., 2012b) dont des protéines de liaison aux ARNm, qui seraient impliquées dans de multiples aspects de la régulation post-transcriptionnelle des gènes, incluant le transport et la stabilité des ARNm, ainsi que la traduction localisée dans les neurones (Akten et al., 2011; Fallini et al., 2011, 2014; Hubers et al., 2011; Tadesse et al., 2008). SMN jouerait un rôle dans la stabilité des ARN tels que ceux contenant des motifs de liaison ARE (AU-Rich Element) en modulant leur association avec les RBP (Figure 30). Cependant les fonctions moléculaires précises de SMN dans la biogenèse des mRNP est encore peu connue.

II.1.4.3 Rôle du méthylosome et du complexe SMN dans l'assemblage des protéines Sm

La voie d'assemblage des snRNP peut être divisée en deux étapes (Figure 29B). Dans une phase précoce, les protéines Sm sont prises en charge par le méthylosome. Celui-ci est composé de l'enzyme PRMT5, de MEP50 et de pICln (Friesen et al., 2001; Meister et al., 2001a ; Meister et al., 2001b; Pu et al., 1999 ; Barbarossa et al., 2014; Pesiridis et al., 2009). Le chaperon pICln interagit avec les protéines Sm nouvellement synthétisées et les délivre au cœur du complexe du méthylosome composé des protéines PRMT5 et MEP50. L'enzyme PRMT5 catalyse alors la dimétylation symétrique des résidus arginine (sDMA) dans les protéines Sm B-B', D1 et D3 (Brahms et al., 2001a; Friesen et al., 2001) (Figure 29B). La méthylation des protéines Sm est requise pour la biogenèse des snRNP (Gonsalvez et al., 2007) et augmente leur affinité pour le complexe SMN par lequel elles seront prises en charge dans un second temps (Brahms et al., 2001a; Friesen et al., 2001; Matera and Wang, 2014; Meister et al., 2001b). Deux sous-complexes impliquant des protéines du méthylosome sont formés séparément. Un anneau hexamérique intermédiaire, composé de pICln, et des protéines Sm D1, D2, F, E et G, appelé complexe 6S, est formé et pré-organise les protéines Sm dans les positions spatiales qu'elles adopteront dans le complexe final (Chari et al., 2008; Grimm et al., 2013; Neuenkirchen et al., 2008) (Figure 29B). Un second complexe comprend pICln et les protéines Sm D3 et B (Chari et al., 2008; Neuenkirchen et al., 2008, 2015). La phase tardive du

mécanisme d'assemblage cytoplasmique des snRNP est assurée par le complexe SMN qui catalyse l'association des deux sous-complexes de protéines Sm sur leur sites de liaison spécifiques riches en uridines sur au niveau des snARN (Chari et al., 2008; Grimm et al., 2013; Li et al., 2014b; Zhang et al., 2011) (Figure 29B).

Des structures cristallographiques et de microscopie électronique du complexe 6S associé à deux protéines du complexe SMN, les protéines Gemin2 et SMN ont récemment été résolues et permettent une meilleure compréhension de ce mécanisme (Grimm et al., 2013) (Figure 31). En effet, la structure du complexe 6S, montre que le repliement de pICln mime celui d'une protéine Sm, ce qui permet une organisation topologique du pentamère Sm en anneau fermé incluant pICln (Grimm et al., 2013).

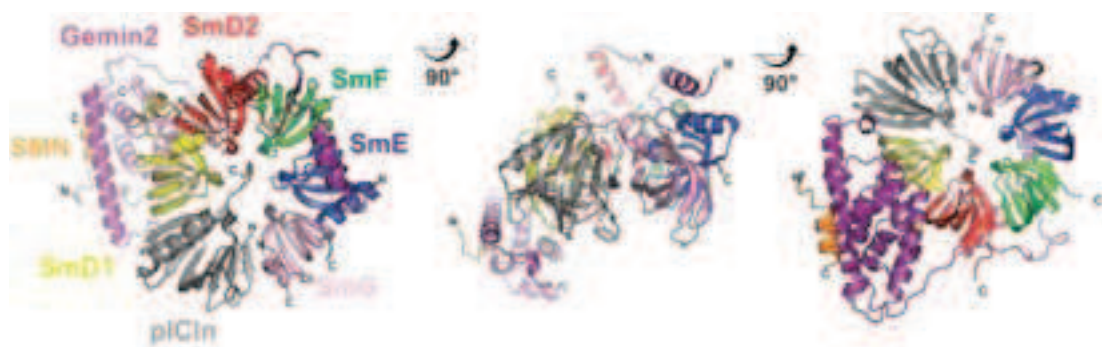


Figure 31. Structure cristallographique du complexe 6S (pICln/Sm D1/D2/F/E/G) associé à SMN et Gemin2. (d'après Grimm et al., 2013).

Structure globale du complexe 8S (complexe 6S associé à SMN et Gemin2) (PDB ID 4F77). Trois vues différentes de l'anneau 6S sont représentées. Les extrémités N et C terminales des protéines sont indiquées. Une couleur différente est attribuée à chaque sous-unité. La protéine SMN est une protéine tronquée SMN(1-120).

L'association du complexe 6S aux protéines du complexe SMN permettrait l'élimination de pICln et le transfert des protéines Sm sur les snARN. Par ailleurs, une autre étude structurale propose un modèle de liaison de la protéine SMN entière, de Gemin2 et de l'anneau Sm heptamérique permettant de comprendre la prise en charge des protéines Sm par le complexe SMN et les changements structuraux impliqués dans leur assemblage sur les snARN (Seng et al., 2015) (Figure 32). Contrairement aux protéines SMN et Gemin2 qui sont au cœur de l'interaction et de l'assemblage des protéines Sm (Fischer et al., 1997; Grimm et al., 2013; Kroiss et al., 2008; Meister et al., 2001b; Zhang et al., 2011), la fonction des autres sous-unités (Gemin3-8 et Unrip) du complexe SMN est moins claire (Chari and Fischer, 2010). Cependant, il a été montré que l'inactivation par siARN des différentes protéines de ce complexe perturbe l'assemblage des snRNP dans les cellules en culture (Feng et al., 2005;

Grimmler et al., 2005; Shpargel and Matera, 2005). Récemment, la première reconstitution complète de la machinerie d'assemblage des snRNP a été réalisée *in vitro* à partir de protéines recombinantes (Neuenkirchen et al., 2015). Ces travaux ont montré que le complexe SMN assemblait les protéines Sm de façon ATP-indépendante sur les snARN et corrigeait les défauts d'assemblages. Ni la protéine Gemin 3, qui comprend un domaine putatif d'hélicase ATP-dépendante, ni la protéine de liaison aux snARN Gemin5, ne sont requis pour la réaction d'assemblage *in vitro*. Le complexe SMN semblerait donc stimuler des changements conformationnels et des mécanismes d'assemblages spontanés. Le méthylosome et le complexe SMN forment une unité fonctionnelle coordonnée qui chaperone, régule et surveille l'assemblage des protéines Sm sur les snARN.

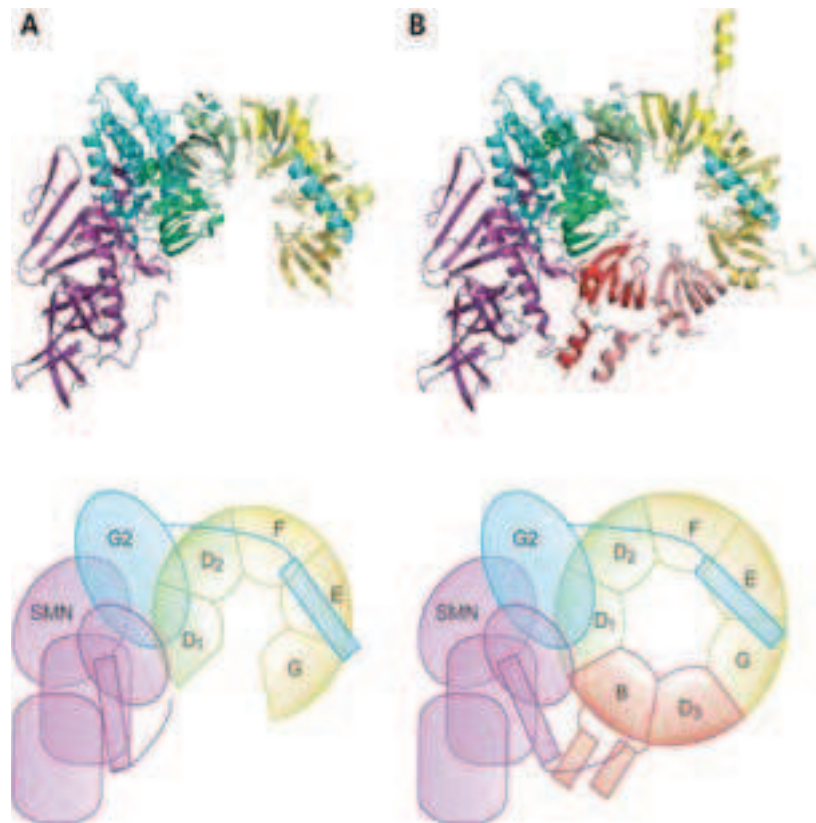


Figure 32. Modèle de structure tridimensionnelle de l'anneau Sm en complexe avec les protéines SMN et Gemin2. (d'après Seng et al., 2015)

Modèles réalisés à partir des données cristallographiques des différentes protéines et des sous-complexes (Grimm et al., 2013; Seng et al., 2015). (A) Modèle de SMN (violet), Gemin2 (G2) (cyan), le dimère Sm D1-D2 (vert) et le trimère Sm F-E-G (jaune). Le trimère Sm F-E-G est maintenu en place par le domaine C-terminal de Gemin2 constitué d'une hélice α et d'une région flexible. (B) Modèle de l'anneau heptamérique Sm fonctionnel associé à SMN et Gemin2. La liaison du dimère Sm B-D3 (rouge) induit un mouvement du trimère Sm F-E-G et de la région hélice α /flexible de Gemin2. Il est proposé que cette liaison fasse également intervenir le domaine C-terminal de la protéine SMN.

II.1.4.4 SMN et la pathologie SMA

L'atrophie musculaire spinale ou SMA (Spinal Muscular Atrophy) est une maladie génétique autosomique récessive, constituant une des causes de mortalité infantile les plus élevées. La prévalence de cette maladie est estimée à une naissance sur 6000 environ (Pearn, 1978). Cette pathologie est caractérisée par une dégénérescence des motoneurones de type α de la moelle épinière, menant progressivement à l'atrophie des muscles des bras, des cuisses et du tronc (Jablonka et al., 2000; Li et al., 2014b; Nicole et al., 2002). Des mutations dans le gène de *SMN*, menant à des défauts d'expression de la protéine SMN, sont à l'origine de la pathologie SMA (Lefebvre et al., 1995). Le génome humain contient deux gènes codants pour la protéine SMN (*SMN1* et *SMN2*). Ces deux gènes sont transcrits en ARN pré-messagers composés de 8 exons séparés par 7 introns. La conversion d'une cytosine (C) en thymine (T) en position 6 de l'exon 7 du gène *SMN2* modifie cependant l'épissage alternatif du pré-ARNm, entraînant une forte exclusion de l'exon 7 dans les ARNm produits par le gène *SMN2* (Lefebvre et al., 1995; Wirth et al., 2006). Les transcrits dépourvus d'exon 7 produisent une protéine SMN tronquée, instable *in vivo* et rapidement dégradée (Burnett et al., 2009; Vitte et al., 2007). Chez 95% des patients atteints de la pathologie SMA, le gène *SMN1* est absent, tronqué ou présente des mutations faux-sens ou non-sens empêchant l'expression d'une protéine SMN fonctionnelle (Lefebvre et al., 1995). Le gène *SMN2* est cependant présent et la gravité de la maladie est dépendante du nombre de copies du gène *SMN2* (Campbell et al., 1997; Feldkötter et al., 2002; Mailman et al., 2002). Pour l'ensemble des organismes modèles de la pathologie SMA (*S. pombe*, *C. elegans*, *D. melanogaster*, *D. rerio*, *S. scrofa domesticus* et *M. musculus*), l'absence totale de protéine SMN induit une létalité confirmant son rôle indispensable dans la survie cellulaire (Beebe et al., 2012; Borg et al., 2015; Burghes and Beattie, 2009b; Campion et al., 2010; Monani et al., 2000).

II.2 Biogenèse des snoRNP à boîte C/D

Les snoARN (small nucleolar ARN) constituent le groupe le plus abondant d'ARN non-codants (Lui and Lowe, 2013). Ils sont présents dans le nucléole où ils sont majoritairement impliqués dans la maturation des ARN ribosomiques (Boisvert et al.,

2007; Watkins and Bohnsack, 2012). La classe des snoRNP à boîte C/D catalyse la 2'O-méthylation des riboses des ARN ribosomiques et d'autres ARN cellulaires (Decatur and Fournier, 2003; Kiss-László et al., 1996; Watkins and Bohnsack, 2012). Les snoARN à boîte C/D sont caractérisés par deux motifs C et D conservés essentiels à leur biogénèse et leur fonction (Figure 33). La boîte C est localisée à l'extrémité 5' du snoARN mature et la boîte D est située près de l'extrémité 3'. D'autres séquences conservées, appelées boîtes C' et D' sont souvent présentes dans la région centrale de ces snoARN (Figure 33).

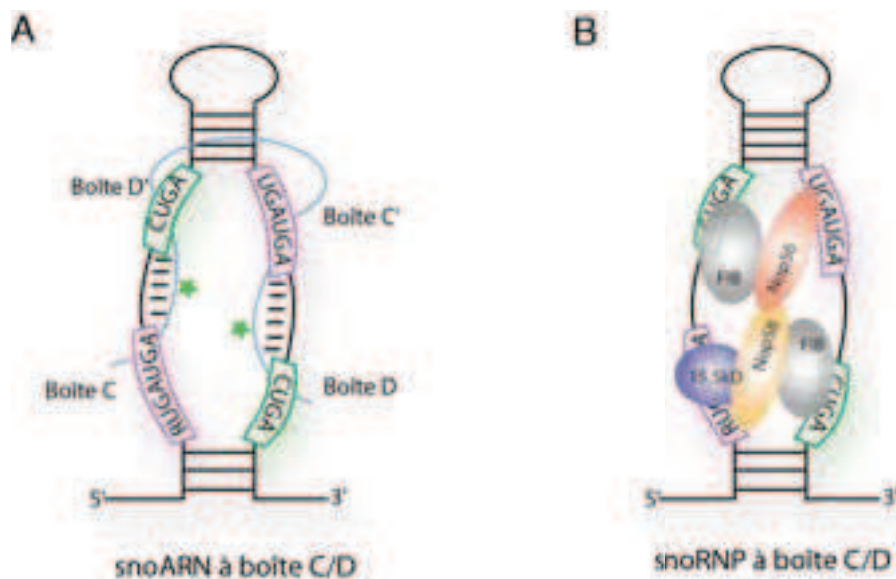


Figure 33. Représentation schématique des snoRNP à boîte C/D. (d'après Reichow et al., 2007)
 (A) Représentation d'un snoARN à boîte C/D (noir) en interaction avec son ARN cible (bleu). Les séquences consensus des boîtes C, D, C' et D' sont indiquées. N= A, U, C ou G. La position du ribose 2'O-méthylé sur l'ARN cible est symbolisée par une étoile. (B) Représentation d'un snoARN à boîte C/D associé de ces protéines cœurs. La protéine 15.5kD contact la boîte C, Nop56 et Nop58 les boîtes C et C' respectivement, et deux fibrillarines interagissent avec les boîtes D et D'.

Les snoARN à boîte C/D adoptent une structure en K-turn (Kiss-László et al., 1996, 1998; Matera et al., 2007a). Les snoRNP à boîte C/D matures comportent 4 protéines cœurs, la fibrillarine, (Tollervey et al., 1991, 1993), Nop56, Nop58 (Gautier et al., 1997; Lafontaine and Tollervey, 1999, 2000) et 15.5kD, une protéine de la famille L7Ae (Watkins et al., 2000) (voir paragraphe I.2.2.2.b.1). La fibrillarine est l'enzyme fonctionnelle des snoRNP qui catalyse la 2'O-méthylation des riboses des ARNr (Kiss-László et al., 1996; Tollervey et al., 1993). Les autres protéines cœurs sont indispensable à cette activité enzymatique et sont impliqués dans la stabilité et la

localisation nucléolaire des snoRNP à boîte C/D (Watkins et al., 2000). Contrairement aux snRNP, la biogenèse des snoRNP à boîte C/D est entièrement nucléaire (Pradet-Balade et al., 2011). Les snoARN sont généralement codés dans des séquences introniques et leur synthèse est initiée par l'épissage des pré-ARNm qui les contiennent (pour revue, Hesselberth, 2013; Kiss, 2006). Ils sont alors soumis à des mécanismes de maturation par des endonucléases et exonucléases (Allmang et al., 1999; van Hoof et al., 2000; Kufel et al., 2003). Certains snoARN dont les snoARN U3, U8 et U14 sont transcrits par l'ARN polymérase II à partir de leur propre promoteur et portent ainsi une coiffe m⁷G (Rodor et al., 2010). Durant la biogenèse de ces snoRNP dans le noyau, leur coiffe m⁷G est hyperméthylée par l'enzyme Tgs1 en coiffe m₃G (Girard et al., 2008; Mouaikel et al., 2002)(voir paragraphe suivant). La boîte C/D est essentielle à l'assemblage des protéines cœurs et le recrutement des facteurs de biogenèse des snoRNP C/D. L'ensemble de ces étapes de maturation et d'assemblage a lieu dans les corps de Cajal. Les snoRNP matures sont ensuite pris en charge par le transporteur CRM1 pour être localisés dans le nucléole, où ils exercent leur fonction dans la biogenèse des ARN ribosomiques (Kiss et al., 2006; Pradet-Balade et al., 2011; Samarsky et al., 1998; Verheggen and Bertrand, 2012).

II.3 L'hyperméthylation de la coiffe des snARN et des snoARN

Chez les vertébrés, après transcription par l'ARN polymérase II, les snARN U1, U2, U4, U4atac, U5, U11 et U12 portent une coiffe m⁷G monométhylée et sont exportés dans le cytoplasme où ils seront maturés et assemblés (Cougot et al., 2004; Hernandez, 2001). L'hyperméthylation de la coiffe en 5' des snRNP est une des étapes essentielles de la maturation des snRNP dans le cytoplasme. En effet, la coiffe m₃G des snARN joue un rôle important dans leur stabilité, leur assemblage (Reddy et al., 1992; Speckmann et al., 2000) et est requise pour l'import des particules dans le noyau, où ils forment le spliceosome (Figure 28) (Huber et al., 1998; Narayanan et al., 2002; Strasser et al., 2004). Tgs1 (triméthyl guanosine synthase 1) est l'enzyme responsable de l'hyperméthylation de la coiffe des sn et des snoARN (Mouaikel et al., 2002). Cette enzyme, conservée de la levure à l'homme, utilise pour substrat la S-adénosylméthionine (SAM) et catalyse l'ajout de deux groupements méthyl sur les NH₂ exocycliques de la guanosine déjà méthylée en position N7 (m⁷G) (Figure 34) (Lemm et al., 2006; Monecke et al., 2009; Mouaikel et al., 2003a).

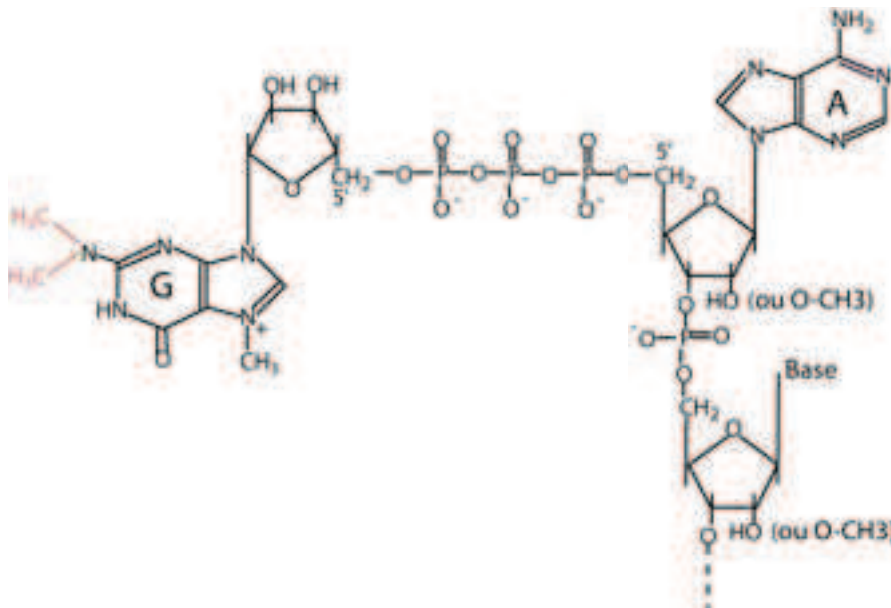


Figure 34. Structure de la coiffe triméthylée m₃G en 5' des snARN.

Une guanosine méthylée en position N7 est reliée par une liaison 5'-5' phosphodiester triphosphate au premier nucléotide du snARN, donnant une coiffe monométhylée m⁷G. Deux groupements méthyl (en rouge) sont alors ajoutés en position N2 de la guanosine par l'enzyme TGS1. G : guanosine, A : adénosine

L'assemblage correct des protéines Sm sur les snARN et la présence du groupement monométhyl en N7 de leur coiffe sont nécessaires à la triméthylation de la coiffe en 5' par Tgs1 (Massenet et al., 2002; Mattaj, 1986; Plessel et al., 1994; Raker et al., 1996). C'est le complexe SMN, associé aux protéines Sm qui est responsable du recrutement de Tgs1 vers les snARN (Mouaikel et al., 2003b). Deux isoformes de Tgs1 sont retrouvés chez l'homme (Girard et al., 2008). La forme longue, Tgs1 LF, est majoritairement cytoplasmique et est capable d'interagir avec SmB suggérant son rôle dans la triméthylation de la coiffe des snRNP (Girard et al., 2008; Mouaikel et al., 2002). Tgs1 LF est également présente dans le noyau, où, associée à CRM1, elle permettrait le transport des snoRNP vers le nucléole (Girard et al., 2008; Pradet-Balade et al., 2011; Verheggen et al., 2002). La forme courte de Tgs1 (Tgs1 SF), obtenue par clivage protéolytique de Tgs1 LF, ne contient pas la région N-terminale mais conserve son activité de méthyltransférase. Elle est localisée dans le noyau, où elle se concentre dans les corps de Cajal, interagit avec les protéines NOP56 et NOP58 et serait responsable de la triméthylation de la coiffe des snoRNP à boîtes C/D et H/ACA (Girard et al., 2008; Mouaikel et al., 2002). Une étude récente propose que l'enzyme Tgs1 soit fonctionnelle uniquement sous forme d'oligomère (Boon et

al., 2015). Après hyperméthylation de leur coiffe, les snARN sont pris en charge par la snurportine-1 et l'importine β afin d'être réimportés dans le noyau (Narayanan et al., 2002; Strasser et al., 2004) (Figure 28). Les snoARN à boîte C/D, une fois leur coiffe modifiée, sont transportés dans le nucléole via CRM1 (Girard et al., 2008; Pradet-Balade et al., 2011; Verheggen et al., 2002). Il a été proposé que la coiffe m₃G puisse moduler des interactions ARN-protéines, par analogie à la méthylation dans le cas des protéines (Mouaikel et al., 2002).

III Mécanismes de maturation et d'initiation de la traduction des ARNm chez les eucaryotes

De nombreuses études ont contribué à l'élucidation du mécanisme d'incorporation de la sélénocystéine dans les sélénoprotéines et à la compréhension du rôle de facteurs co-traductionnels ou d'assemblage spécifiques. Le déficit majeur est de comprendre comment ceux-ci s'inscrivent dans le contexte général de la traduction. En effet, bien que la traduction des ARNm de sélénoprotéines utilise la même machinerie traductionnelle que les ARNm non-sélénoprotéines, certaines différences mécanistiques peuvent être relevées. Durant ma thèse, nous avons montré que les ARNm de sélénoprotéines possédaient une coiffe atypique (voir chapitre I des résultats). Ceci suggère que la traduction de ces ARNm puisse faire intervenir un mécanisme d'initiation non-canonique. Dans ce chapitre nous présenterons les mécanismes de maturation et d'initiation de la traduction chez les eucaryotes ainsi que les différents mécanismes non-conventionnels répertoriés à ce jour.

III.1 Maturation des ARNm eucaryotes

Les ARNm synthétisés par l'ARN polymérase II subissent de multiples étapes de maturation avant d'être exportés dans le cytoplasme pour y être traduits. Les nouveaux transcrits sont modifiés à leur extrémité 5' par addition d'une coiffe m⁷G, à leur extrémité 3' par polyadénylation, et de manière interne par élimination des séquences introniques et parfois par un processus d'édition. Ces étapes sont décrites en détail dans diverses revues (Braun and Young, 2014; Haimovich et al., 2013; Katahira, 2015; Pérez-Ortín et al., 2013).

III.1.1 Les ARNm portent une coiffe monométhylée m⁷G

III.1.1.1 Biogenèse de la coiffe m⁷G

La coiffe m⁷G en 5' des ARNm (m⁷GpppN, où m⁷G est la guanosine méthylée en position N7 et N le premier nucléotide de l'ARNm, Banerjee, 1980) joue un rôle critique dans le cycle de vie des ARNm eucaryotes et est nécessaire à l'expression des gènes et à la viabilité des cellules de la levure à l'homme (pour revue, Cowling, 2010; Topisirovic et al., 2011). La coiffe est requise pour la traduction optimale de la majorité des ARNm cellulaires et est également caractéristique des ARNm de la plupart des virus et parasites eucaryotes (Shuman, 2002). De plus, la coiffe stabilise les ARNm en les protégeant contre les exonucléases et promeut la transcription, la polyadénylation, l'épissage et l'export nucléaire des ARNm (Gu and Lima, 2005; Lewis and Izaurralde, 1997). La coiffe est ajoutée de manière co-transcriptionnelle à l'extrémité 5' du pré-ARNm naissant lorsque celui-ci atteint une longueur de 22-25 nucléotides et émerge du canal de sortie de l'ARN polymérase II (Pol II) (Shatkin, 1976; Shatkin and Manley, 2000) (Figure 35). C'est la phosphorylation des sérines du domaine C-terminal (ou CTD) de la grande sous-unité de l'ARN polII, par le facteur de transcription TFIIF, qui permet de recruter les enzymes de biogenèse de la coiffe (Ho et al., 1998; McCracken et al., 1997). La guanosine méthylée en position N7 est liée par un pont 5'-5' phosphodiester triphosphate au premier nucléotide en 5' du transcrit naissant (Figure 35). L'addition de la coiffe est réalisée en 3 étapes enzymatiques. Durant la première étape, le groupement γ -phosphate du premier nucléotide en 5' de l'ARNm est éliminé par une ARN-triphosphatase (RT). Cette réaction est suivie du transfert d'une guanine monophosphate (GMP) à l'extrémité 5' diphosphate du transcrit par une guanylyltransférase (GT). Ceci résulte en la formation d'une coiffe guanosine (GpppN) qui est alors méthylée par une ARN-(guanine N7)-méthyltransférase (RNMT) pour produire la coiffe 7-méthylguanosine (m⁷GpppN) (Figure 35). Ces trois activités catalytiques, collectivement dénommées "enzymes d'addition de la coiffe" sont codées par des gènes séparés chez la levure, tandis que chez les vertébrés, les deux premières étapes sont catalysées par une même enzyme dénommée RNGTT (ARN guanylyltransferase et 5' phosphatase) constituée de deux domaines fonctionnels : le domaine N-terminal doté d'une activité ARN-triphosphatase et le domaine C-terminal qui correspond à la guanylyltransférase (Pillutla et al., 1998; Yamada-Okabe et al., 1998; Yue et al., 1997). La coiffe peut

subir des modifications supplémentaires par des enzymes ARN 2'-O-méthyltransférases spécifiques dans le noyau et le cytoplasme qui ajoutent un groupement méthyl sur les positions 2'hydroxyles des riboses du premier et du second nucléotide de l'ARNm, aboutissant aux structures m⁷GpppNm et m⁷GpppNmNm (Cowling, 2010). A peine synthétisée, la coiffe est directement reconnue par des protéines de liaison à la coiffe qui définissent sa localisation et régulent sa fonction dans la cellule. La coiffe et ses facteurs de liaison sont un moyen de contrôler l'expression des gènes en régulant entre-autre l'initiation de la traduction des ARNm (Cowling, 2010; Topisirovic et al., 2011).

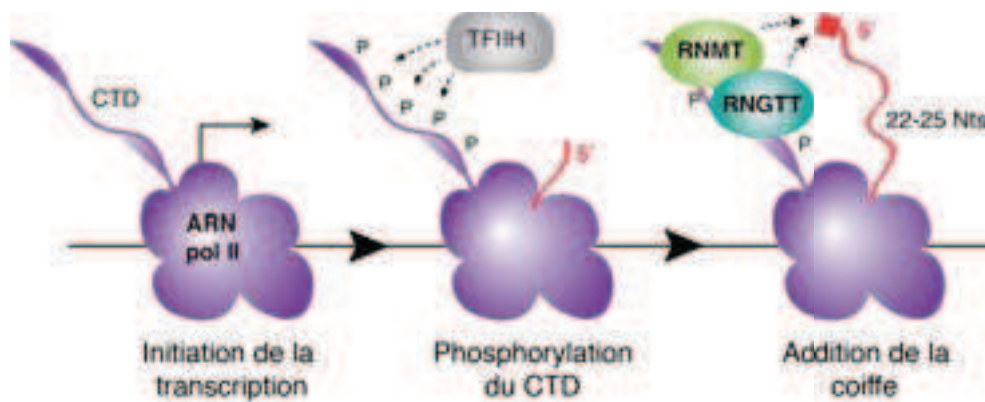


Figure 35. Mécanisme d'addition co-transcriptionnelle de la coiffe des ARNm. (d'après Cowling, 2010)

L'ARN polymérase II est recrutée au niveau du promoteur ; son domaine C-terminal (CTD) est hypophosphorylé. Après initiation de la transcription, le facteur TFIIF phosphoryle des sérines du CTD, ce qui permet de recruter les enzymes de biogenèse de la coiffe RGT (RNA guanylyltransferase and 5' phosphatase) et RMT (RNA guanine-7 methyltransferase). Lorsque le transcrit naissant atteint une longueur de 22 à 25 nucléotides, son extrémité 5' se situe à proximité des deux enzymes qui catalysent la formation de la coiffe 7-méthylguanosine.

III.1.1.2 Reconnaissance de la coiffe m⁷G par le facteur CBC

Le complexe CBC (Cap Binding Complex) est le premier complexe recruté sur la coiffe m⁷G lorsque celle-ci est ajoutée de manière post-transcriptionnelle en 5' des ARN dans le noyau. CBC est un complexe multifonctionnel, essentiel pour l'expression des gènes et impliqué dans l'assemblage, la maturation, l'export nucléaire et la traduction des ARN (Gonatopoulos-Pournatzis and Cowling, 2014; Izaurralde et al., 1994; Worch et al., 2005) (Figure 36). CBC est composé des protéines CBP20 et CBP80 qui lient la coiffe m⁷G de manière synergique. Le site de liaison de la coiffe réside dans la sous-unité CBP20 (Mazza et al., 2001), cependant, l'interaction avec CBP80 est requise pour induire un changement conformationnel

permettant la liaison de haute affinité du complexe à la coiffe m^7G . Dans les cellules de mammifère, l'inhibition de CBP80 par shARN résulte en une dérégulation de l'expression d'environ 400 gènes et une réduction significative du taux de prolifération des cellules (Narita et al., 2007). CBP20 possède également un domaine de reconnaissance de l'ARN (Pabis et al., 2010) qui pourrait augmenter l'affinité pour certaines séquences d'ARN ou motifs et qui expliquerait les effets gènes-spécifiques de CBC (Gonatopoulos-Pournatzis and Cowling, 2014).



Figure 36. Le complexe de liaison à la coiffe CBC est multifonctionnel (d'après Gonatopoulos-Pournatzis and Cowling, 2014).

Le complexe CBC est composé des sous-unités CBP20 et CBP80 et se lie à la coiffe m^7G des ARNm directement après leur transcription dans le noyau. CBC interagit avec de nombreux facteurs, modulant ainsi le métabolisme des ARN et le processus de traduction dans les cellules. Les processus dans lesquels CBC intervient sont indiqués dans les encarts et le nom des facteurs intervenant dans ces mécanismes sont représentés. NMD, Nonsense Mediated Decay ; SMD, Staufen Mediated Decay.

CBC permet l'export des ARNm dans le cytoplasme et assure le premier cycle de traduction de l'ARNm (connu sous le nom de « pioneer round of translation ») (Figure 36). Il a été proposé qu'une des fonctions de ce premier cycle de traduction CBC-dépendant est de contrôler la qualité de l'expression des gènes, en permettant l'activation du mécanisme du NMD (nonsense-mediated mRNA decay) lorsqu'un codon stop prématuré est présent sur l'ARNm (Hwang et al., 2010; Maquat et al., 2010a, 2010b). Cependant, il a été montré que le NMD n'est pas exclusivement dédié à la traduction CBC dépendante, puisqu'il peut également être déclenché en présence du facteur de liaison à la coiffe eIF4E (Durand and Lykke-Andersen, 2013; Rufener and Mühlemann, 2013) (Figure 39). Après le premier cycle de traduction de l'ARNm, le complexe CBC est remplacé par le facteur d'initiation eIF4E qui permet alors le

maintient de plusieurs cycles de traduction consécutifs (Chiu et al., 2004; Maquat et al., 2010b).

III.1.2 Polyadénylation de l'extrémité 3' des ARNm

La majorité des ARNm eucaryotes sont polyadénylés à leur extrémité 3' (pour revue récente, Shi and Manley, 2015). La queue poly(A) est essentielle à l'expression des ARNm puisqu'elle est impliquée dans la stabilité des ARNm, leur export cytoplasmique et leur traduction. La polyadénylation a lieu en deux étapes consistant en un clivage endonucléolytique au niveau d'un site spécifique (appelé site poly(A)), suivi de la synthèse d'une queue poly(A) (Chan et al., 2011; Colgan and Manley, 1997; Zhao et al., 1999). Ce processus requiert une machinerie bien plus complexe que celle du mécanisme d'addition de la coiffe et fait appel à plus d'une douzaine de facteurs (Shatkin and Manley, 2000) (Figure 37).

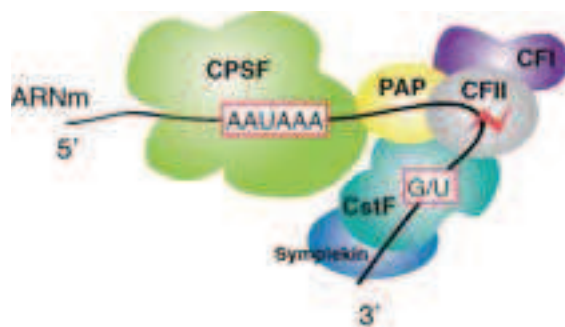


Figure 37. Le complexe de clivage et de polyadénylation des ARNm.

(d'après Shatkin and Manley, 2000)

Les signaux de polyadénylation en *cis* sont encadrés en rouge sur l'ARNm. La position de l'ensemble des facteurs nécessaires à ce processus est représentée. L'éclair rouge indique le site de clivage endonucléolytique.

Le site poly(A) chez les mammifères est riche en AU et sa séquence nucléotidique est hautement conservée (Shi, 2012). Ce site est composé de plusieurs éléments en *cis* incluant (i) une séquence AAUAAA (ou AUUAAA) localisée 10 à 30 nucléotides en amont du site de clivage, (ii) une séquence CA en 5' de la séquence précédente; (iii) une région riche en U/GU localisée environ 40 nucléotides en aval du site de clivage (nommée DSE) ; (iv) un élément auxiliaire riche en U (consensus UGUA) localisé en amont (appelé USE). Les facteurs de maturation de l'extrémité 3' des ARNm CPSF (cleavage and polyadenylation specificity factor) et CstF (Cleavage stimulation Factor) se lient de manière synergique à l'hexamère AAUAAA et à la région DSE respectivement tandis que le complexe CFI lie le motif UGUA (Hu et al., 2005) (Figure 37). CPSF, CstF et CFI forment un complexe et recrutent d'autres facteurs additionnels incluant les protéines CFII, Symplekin et la poly(A) polymérase (PAP), formant ainsi le complexe actif de maturation en 3' des ARNm (Chan et al., 2011; Shi

et al., 2009). Ce complexe permet le clivage de l'ARNm et l'ajout de la queue poly(A). L'assemblage de ces facteurs sur le site de polyadénylation se produit co-transcriptionnellement et est facilité par le domaine C-terminal de l'ARN polymérase II (Bentley, 2005; Hirose and Manley, 2000; Mandel et al., 2008). La queue poly(A) est reconnue par les PABP (poly(A) Binding Proteins), des protéines hautement conservées qui sont capables de lier de nombreux partenaires protéiques, participant ainsi à une diversité d'événements cellulaires (Derry et al., 2006; Eliseeva et al., 2013). Elles permettent en l'occurrence la circularisation des ARNm, prérequis pour leur traduction.

III.2 La traduction des ARNm

La traduction des ARNm est un processus complexe qui se divise en 4 étapes : l'initiation, l'élongation, la terminaison et le recyclage du ribosome. L'initiation est le processus d'assemblage du ribosome 80S pendant lequel les bases du codon initiateur sont appariées à celles de l'anticodon de l'ARNt initiateur (Met-ARNt^{Met}) dans le site P du ribosome. L'initiation canonique requiert l'intervention d'un minimum de 9 facteurs appelés facteurs d'initiation ou 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) (Figure 38). Durant l'étape d'élongation, la chaîne polypeptidique est allongée et les acides aminés sont liés successivement de manière covalente, chacun étant correctement positionné dans le ribosome par son ARNt, via l'interaction codon/anticodon avec l'ARNm. L'élongation requiert 3 facteurs d'élongation : eEF1A, qui apporte l'aminoacyl-ARNt au ribosome, le facteur d'échange eEF1B qui recycle le GTP fixé à eEF1 et le facteur eEF2, qui promeut la translocation du ribosome (Dever and Green, 2012). L'étape de terminaison prend place lorsque la fin de la région codante est atteinte par le ribosome et qu'un codon stop (UAA, UGA, or UAG) entre dans le site A. Cette étape est catalysée par deux facteurs de terminaisons appelés eRF1 et eRF3 (Blanchet et al., 2015; Hatin et al., 2009). La chaîne peptidique complète est alors libérée et le ribosome sera recyclé (Rodnina and Wintermeyer, 2009). La synthèse des protéines est principalement régulée à l'étape d'initiation de la traduction, permettant un contrôle rapide et réversible de l'expression des gènes. Nous nous focaliserons sur cette étape de la traduction.

III.2.1 Mécanisme d'initiation de la traduction canonique, coiffe dépendant

III.2.1.1 Recrutement de eIF4F au niveau de la coiffe m⁷G

Le mécanisme d'initiation canonique de la traduction est dépendant de la présence d'une coiffe monométhylée m⁷G en 5' de l'ARNm. Le processus d'initiation commence par la formation du complexe eIF4F au niveau de la coiffe m⁷G. Ce complexe est composé du facteur de liaison à la coiffe eIF4E (Altmann et al., 1985), de l'hélicase à motif "DEAD" eIF4A et de la protéine de structure eIF4G (Sonenberg et al., 1978) (Figure 38). En fixant la coiffe, le rôle de eIF4F est de déplier la structure secondaire de l'extrémité 5' de l'ARNm dans le but de préparer l'attachement du ribosome (Asano et al., 2001; Etchison et al., 1982). Ceci est réalisé par l'activité hélicase de eIF4A, stimulée par les facteurs eIF4G et le facteur auxiliaire eIF4B (Rozen et al., 1990). La coiffe est empiée entre deux résidus tryptophane de la surface concave de eIF4E et des contacts additionnels avec les nucléotides à proximité de la coiffe stabilisent la liaison de eIF4E à l'ARNm (von der Haar et al., 2004; Sonenberg, 2008). En plus de son rôle dans l'initiation de la traduction, eIF4E est également présent sur la coiffe d'ARNm soumis au mécanisme du NMD (Durand and Lykke-Andersen, 2013; Rufener and Mühlemann, 2013). De plus, eIF4E est retrouvé dans le noyau où il jouerait un rôle dans l'export cytoplasmique de certains ARNm qui contiennent un motif de liaison appelé 4E-SE (eIF4E Sensitive Element) et qui sont majoritairement impliqués dans la progression du cycle cellulaire (Culjkovic et al., 2005, 2007) (Figure 39). eIF4G a un rôle de stabilisation du complexe puisqu'elle interagit avec eIF4E, eIF4A, les protéines de liaison à la queue poly(A) en 3' de l'ARNm (PABP) et eIF3 (Hentze, 1997; Lamphear et al., 1995). eIF4G augmente l'affinité de eIF4E pour la coiffe (Gross et al., 2003) et, en interagissant avec les PABP, rapproche les extrémités 5' et 3' de l'ARNm, ce qui permet sa circularisation (Gingras et al., 1999; Wells et al., 1998) (Figure 39). Il est proposé que cette configuration de l'ARNm en « boucle fermée » stimule la traduction en augmentant le recyclage et l'assemblage du complexe 48S sur l'ARNm (Jackson et al., 2010; Sonenberg and Hinnebusch, 2009; Svitkin et al., 2009; Topisirovic et al., 2011). L'ARNm est alors prêt à recruter le complexe 43S.

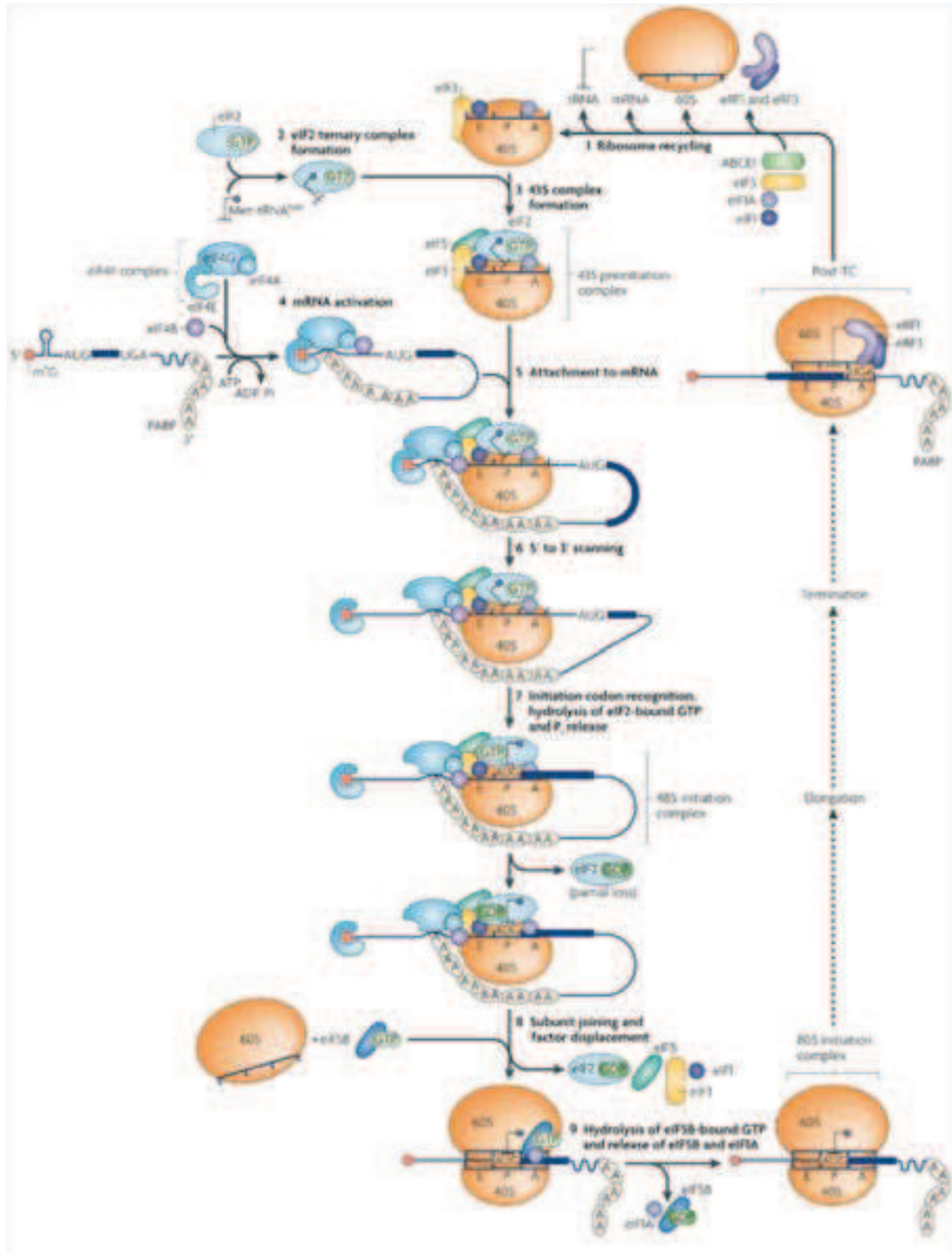


Figure 38. Modèle du mécanisme d'initiation canonique coiffe-dépendant chez les eucaryotes. (d'après Jackson et al., 2010) Le processus d'initiation coiffe-dépendant canonique peut être divisé en 8 étapes (2-9) qui commencent après le recyclage (1) des complexes traductionnels en phase de terminaison. Le complexe ternaire eIF2-GTP-Met-ARNT^{Met}_i (2) est assemblé dans le complexe de pré-initiation 43S (3). L'ARNm est activé : la coiffe lie le complexe eIF4F qui interagit avec les PABP en 3' induisant une circularisation de l'ARNm (4). Le complexe 43S est alors recruté au niveau de l'extrémité 5' de l'ARNm (5) et va procéder au mécanisme de scanning (6). La reconnaissance de l'AUG initiateur induit l'hydrolyse du GTP associé à eIF2 et la formation du complexe 48S (7). L'assemblage de la sous-unité 60S sur le complexe 48S provoque l'évacuation de la majorité des facteurs d'initiation (8). Le ribosome 80S est alors prêt à entrer en phase d'élongation.

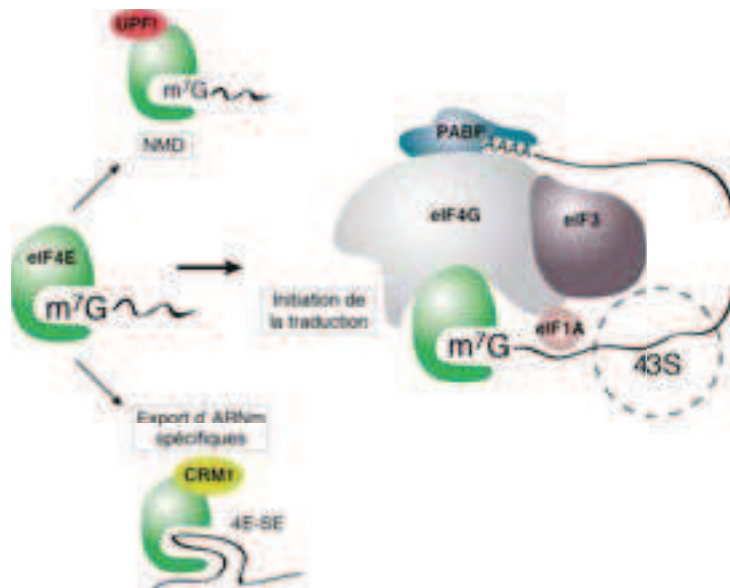


Figure 39. Rôles du facteur d'initiation de la traduction eIF4E.

eIF4E lie directement la coiffe m⁷G des ARNm et permet le recrutement de divers facteurs essentiels à l'initiation de la traduction. Les ARNm associés à eIF4E sont sensibles au mécanisme du NMD (Durand and Lykke-Andersen, 2013; Rufener and Mühlemann, 2013). eIF4E régule également l'export nucléaire CRM1-dépendant de plusieurs ARNm impliqués dans la progression du cycle cellulaire. Cette fonction implique l'interaction de eIF4E avec l'élément structural 4E-SE (eIF4E sensitive element) dans la région 3'UTR de ces ARNm (Culjkovic et al., 2005, 2006; Feoktistova et al., 2013).

III.2.1.2 Le complexe 43S

La traduction est un processus cyclique et les sous-unités ribosomiques qui participent à l'initiation proviennent du recyclage des complexes après terminaison qui fait intervenir la protéine ABCE1 (Skabkin et al., 2013). La sous-unité 40S issue de cette étape de recyclage, associée aux facteurs d'initiation eIF3, eIF1, eIF1A et probablement eIF5, recrute le complexe ternaire eIF2- GTP-Met-ARNt^{Met}_i pour former le complexe de pré-initiation 43S (Majumdar et al., 2003; Trachsel and Staehelin, 1978, 1979) (Figure 38). eIF3 est un complexe de 13 sous-unités qui stimule la fixation du complexe ternaire à la sous-unité 40S, promeut l'attachement du complexe 43S à l'ARNm et empêche l'association de la sous-unité 60S (Hershey, 2010; Jackson et al., 2010). Récemment, un modèle de structure du complexe 40S-eIF1-eIF3 de levure a été proposé à partir de données cristallographiques (Erzberger et al., 2014), qui montre que eIF3 fournit une plateforme pour le recrutement, l'assemblage et la régulation de la machinerie d'initiation de la traduction.

III.2.1.3 Formation du complexe 48S sur l'ARNm

Le complexe 43S est recruté à l'extrémité 5' de l'ARNm grâce à l'interaction entre le complexe eIF3 et le facteur eIF4G (Villa et al., 2013). Après son recrutement, le complexe 43S va glisser sur l'ARNm de 5' vers 3' selon un mécanisme appelé « scanning » jusqu'au codon initiateur AUG (Hinnebusch, 2014; Kozak, 1989) (Figure 38). Le processus de « scanning » est assuré par eIF1/eIF1A qui stabilise la sous-unité 40S dans une conformation favorable dite ouverte (Passmore et al., 2007) et est aidé par des hélicases (eIF4A/4B et DHX29) qui déplient la structure secondaire de la région 5'UTR de l'ARNm (Hashem et al., 2013; Hinnebusch, 2014; Parsyan et al., 2009). Au cours de ce processus, la fidélité de reconnaissance du codon d'initiation AUG 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; Pestova et al., 1998a; Reibarkh et al., 2008). Le codon AUG sélectionné doit se trouver dans un contexte de séquence optimal c'est à dire avec une purine en position -3 et un G en position +4 (relatif au A de l'AUG considéré comme le +1) (Kozak, 1991). La reconnaissance et le positionnement du codon initiateur par l'anticodon du Met-ARNt^{Met}_i au niveau du site P du ribosome induit un changement d'organisation du complexe de « scanning » qui adopte alors une conformation dite fermée. Ce changement provoque le déplacement d'eIF1, permet l'hydrolyse du GTP lié à eIF2 catalysé par eIF5 et, par conséquent, facilite l'ancrage du Met-ARNt^{Met}_i dans le site P du ribosome et son interaction avec le codon initiateur (Algire and Lorsch, 2006; Lee et al., 2002; Lorsch and Herschlag, 1999) (Figure 38). L'ensemble de ces éléments constitue le complexe 48S.

III.2.1.4 Assemblage du complexe 80S

L'association de la sous-unité 60S au niveau du complexe 48S s'accompagne de la dissociation des facteurs eIF1, eIF1A, eIF3 et eIF2-GDP (Pestova and Hellen, 2000) (Figure 38). Ceci est modulé par le facteur eIF5B, une GTPase ribosome-dépendante. L'hydrolyse du GTP par eIF5B est requise à sa propre dissociation et celle de eIF1A lorsque le complexe 80S est assemblé (Acker and Lorsch, 2008). Le ribosome est alors prêt à entrer en phase d'élongation et le processus d'initiation canonique est terminé.

III.2.2 Mécanismes d'initiation coiffe-indépendant induits par les IRES

Au cours de l'évolution, les virus ont mis en place des stratégies pour détourner la machinerie traductionnelle de la cellule hôte en faveur de la traduction de leur propres ARNm et au détriment de celle des ARNm cellulaires ; ceci dans le but de se multiplier rapidement et de se propager efficacement (Au and Jan, 2014). La machinerie traductionnelle est détournée à l'aide d'éléments structuraux appelés IRES (pour Internal Ribosome Entry Site) dont la fonction est de recruter directement les particules ribosomiques ou indirectement via les facteurs d'initiation (Jackson and Kaminski, 1995; Jackson et al., 1990; Pelletier and Sonenberg, 1988) (Figure 40).

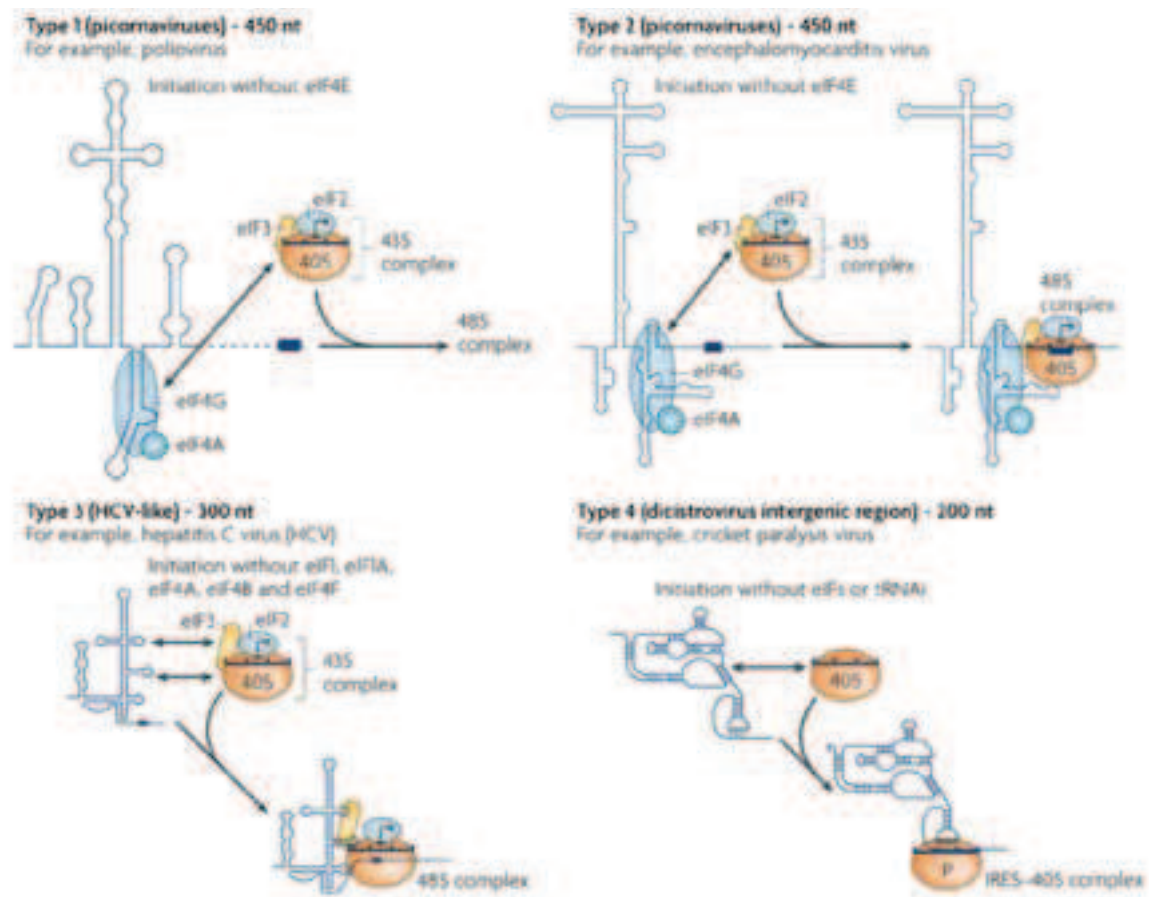


Figure 40. Mécanismes d'initiation de la traduction coiffe-indépendants induits pas les IRES. (d'après Jackson et al., 2010)

Quatre types d'IRES ont été caractérisés. L'initiation à partir des IRES de type 1 et 2 impliquent leur reconnaissance par les facteurs eIF4G/eIF4A (de Breyne et al., 2009). Les IRES de type 3 interagissent directement avec le facteur eIF3 et la sous-unité ribosomique 40S du complexe 43S (Pestova et al., 1998b). Les IRES de type 4 se lient directement à la sous-unité 40S sans aucun facteurs d'initiation (Deniz et al., 2009; Pisarev et al., 2005).

Il a été proposé que la traduction de certains ARNm cellulaires (3%) puissent également faire appel à un mécanisme d'initiation de type IRES (Macejak and Sarnow, 1991; Sharathchandra et al., 2014; Soo-Kyung et al., 1992). Ce concept d'IRES cellulaire reste cependant controversé (Jackson, 2013; Shatsky et al., 2014). A ce jour, 4 types d'IRES viraux ont été caractérisés (Jackson et al., 2010; Lozano and Martínez-Salas, 2015) (Figure 40). Ces différents types d'IRES diffèrent par leur architecture et par les facteurs d'initiation qu'ils recrutent.

III.2.2.1 IRES de type I

Les IRES de type I sont les premiers à avoir été identifiés et sont trouvés dans la famille des *Picornaviridae*, des virus à ARN de polarité positive, dont le génome est directement traduit par la machinerie cellulaire (Pelletier and Sonenberg, 1988; Sweeney et al., 2014). Dans le but de positionner le ribosome sur le codon AUG initiateur, les IRES de type I recrutent tous les facteurs d'initiation impliqués dans le modèle consensus d'initiation de la traduction eucaryotes à l'exception de eIF4E (Sonenberg and Meervitch, 1990). En effet, au début de l'infection virale, le poliovirus synthétise une protéase (2Apro) dont la fonction est de cliver eIF4G au niveau de son domaine de liaison à eIF4E afin d'inactiver l'initiation coiffe dépendante utilisée par la cellule hôte, permettant ainsi au virus de s'approprier la machinerie traductionnelle cellulaire (Davies et al., 1991; Gradi et al., 1998; Haghighat et al., 1996). Le complexe de pré-initiation 43S est recruté en amont du codon initiateur par l'intermédiaire de eIF4G tronqué qui interagit directement avec l'IRES (Figure 40). La suite du processus d'initiation est identique au modèle d'initiation canonique : le complexe 43S va glisser le long de l'ARNm (mécanisme de scanning) jusqu'au codon AUG initiateur où sera assemblé le complexe 80S (Belsham, 1992; Sonenberg and Meervitch, 1990). En plus des facteurs d'initiation canoniques, l'initiation de la traduction dépendante des IRES de type I fait intervenir d'autres facteurs cellulaires appelés ITAF (IRES Trans-Acting Factors) (pour revue, Sweeney et al., 2014) qui modulent et stabilisent la structure de l'IRES et facilitent le recrutement du 43S (Gamarnik and Andino, 1996; Hunt and Jackson, 1999; Meervitch et al., 1993).

III.2.2.2 IRES de type II

Les IRES de type II sont également retrouvés dans le génome des *Picornaviridae* (Jang et al., 1988) tel le virus de l'encéphalomyocardite virale (EMCV). Les facteurs requis pour l'initiation de la traduction dépendante de ce type d'IRES sont identiques à ceux des IRES de type I. Une des différences majeures entre ces deux types d'IRES concerne la stratégie utilisée pour s'approprier de manière exclusive la machinerie traductionnelle de l'hôte (pour revue, Chamond et al., 2014). Dans les deux cas, il s'agit d'une inhibition spécifique de la traduction coiffe dépendante ciblant la formation du complexe eIF4F. Dans le cas des IRES de type II, cette stratégie est basée sur l'activation d'un répresseur de la traduction cellulaire, 4E-BP1 qui se fixe sur eIF4E empêchant son interaction avec eIF4G et donc la formation de eIF4F (Gingras et al., 1999; Pause et al., 1994). L'autre différence entre les deux IRES est que l'IRES de type II recrute le complexe 43S directement sur le codon AUG initiateur et ne requiert donc pas d'étape de scanning (Kaminski et al., 1990) (Figure 40).

III.2.2.3 IRES de type III

Ces IRES sont retrouvés dans le génome des *Flaviviridae*, des virus ARN à polarité positive dont fait partie le virus de l'hépatite C (HCV) (Niepmann, 2013; Reynolds et al., 1995; Tsukiyama-Kohara et al., 1992). Comme dans le cas des IRES de type II, le complexe 43S est directement recruté au niveau de l'AUG initiateur sans passer par une étape de scanning. Les deux facteurs eIF2 et eIF3 sont nécessaires et suffisants pour assurer la formation du complexe 48S (Pestova et al., 1998b) (Figure 40). En effet, la sous-unité 40S et le facteur eIF3 se fixent directement et avec haute affinité à des éléments structuraux présents dans l'IRES III, ce qui permet de former un complexe stable (Ji et al., 2004; Kieft et al., 2001; Otto and Puglisi, 2004). Le complexe ternaire eIF2-GTP-Met-ARNt^{Met}_i est alors recruté et l'assemblage du complexe 80S fait intervenir les facteurs eIF5B-GTP et eIF5 comme dans le modèle d'initiation canonique (Locker et al., 2007; Pestova et al., 1998b).

III.2.2.4 IRES de type IV

Les IRES de type IV sont trouvés dans le génome des *Dicistroviridae* dont le virus de la paralysie du criquet (CrPV) est un des représentants (Wilson et al., 2000a). Ce type

d'IRES a la capacité exceptionnelle d'assembler un ribosome 80S fonctionnel sans l'aide d'aucun facteur d'initiation, ni même du Met-ARNt^{Met}_i (Deniz et al., 2009; Pisarev et al., 2005) (Figure 40). Ce processus d'initiation est entièrement assuré par des éléments structuraux de l'IRES (Wilson et al., 2000b) qui interagissent de manière directe avec les sous-unités ribosomiques. De manière intéressante, plusieurs études suggèrent que l'IRES du CrPV mime le Met-ARNt^{Met}_i dans le site P du ribosome (Costantino et al., 2008; Jan et al., 2003; Muhs et al., 2015a; Pestova and Hellen, 2003), expliquant l'autonomie de ce mécanisme d'initiation vis à vis du Met-ARNt^{Met}_i. Ce mécanisme atypique requiert donc une étape de translocation de l'IRES. Récemment, des données de cryo-microscopie électronique de l'IRES du CrPV sur le ribosome ont démontré la nécessité du facteur eEF2 pour cette première étape de translocation (Fernández et al., 2014; Muhs et al., 2015b). Cette étape est requise pour amener le premier codon de l'ARNm dans le site A, le rendre accessible à l'ARNt^{Ala} et permettre l'entrée du ribosome en élongation (Muhs et al., 2015b). Durant cet événement de translocation inhabituel, l'IRES subit des changements conformationnels importants.

III.2.3 Mécanisme d'initiation de la traduction coiffe-dépendant par “tethering”

Ce mécanisme permet de recruter le complexe de pré-initiation 43S directement au niveau du codon d'initiation AUG (“tethering”) via un élément de structure présent dans la région codante qui interagit avec le facteur eIF4E. Ce mécanisme d'initiation de la traduction non-canonique a été décrit dans notre laboratoire pour l'ARNm de l'histone H4 (Martin et al., 2011). Cet ARNm présente des caractéristiques atypiques : sa région 5'UTR, particulièrement courte, peut être constituée de 9 nucléotides seulement pour l'ARNm de souris (Piontkivska et al., 2002). De plus, les ARNm d'histones sont dépourvus de queue poly(A) et leur extrémité 3' se termine par une structure en tige-boucle (Jaeger et al., 2005; Martin et al., 1997). Notre laboratoire a montré que la région codante (CDS) de l'ARNm de l'histone H4 contient au moins deux éléments structuraux essentiels au processus d'initiation de la traduction (Figure 41). Le premier élément, localisé à proximité de l'extrémité 5', est replié sous forme d'une jonction triple-hélice et permet le positionnement du ribosome directement au niveau de l'AUG initiateur sans étape de scanning préalable (Martin et al., 2011). Le deuxième élément structural, plus proche de l'extrémité 3', adopte un repliement en

double tige-boucle similaire à l'élément 4E-SE (eIF4E Sensitive Element) retrouvé dans la région 3'UTR de l'ARNm de la cycline D1 (Culjkovic et al., 2006, 2007). Cet élément fixe le facteur eIF4E, qui, sous forme de complexe eIF4F, recrute la particule ribosomique 43S et la dépose au niveau du codon initiateur (Martin et al., 2011).

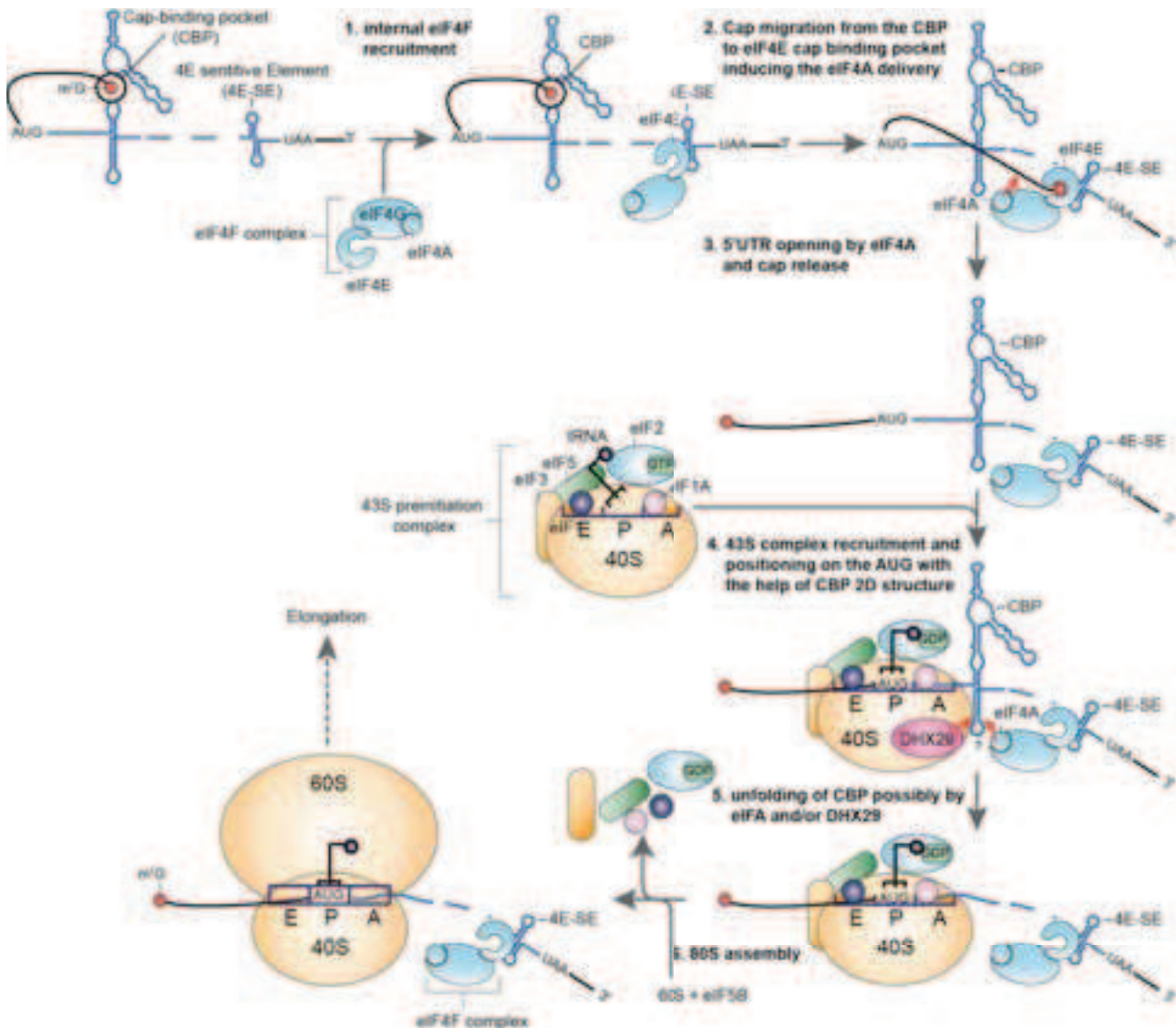


Figure 41. Modèle d'initiation de la traduction coiffe-dépendante par tethering de l'ARNm de l'histone H4

(d'après Martin et al., 2011 et HDR de F. Martin 2012)

La région codante de l'ARNm d'histone H4 contient deux éléments structuraux impliqués dans le processus d'initiation de la traduction : une jonction triple-hélice contenant une poche de séquestration de la coiffe, et une tige-boucle 4E-SE qui recrute le facteur d'initiation eIF4E de manière interne.

Contrairement au modèle consensus d'initiation de la traduction, le recrutement d'eIF4E sur l'ARNm de H4 se fait de manière indépendante de la coiffe, bien que celle-ci soit essentielle à l'entrée du ribosome en phase d'élongation (Figure 41). La

coiffe est dans un premier temps séquestrée dans une région de la jonction triple-hélice appelée CBP (cap-binding pocket) afin d'empêcher une compétition éventuelle avec l'élément 4E-SE pour la liaison à eIF4E. La fixation du complexe eIF4F dans la région codante de l'ARNm induit un changement conformationnel dans l'ARNm qui conduit à la libération de la coiffe (Martin et al., 2011). S'en suit l'interaction de eIF4E avec la coiffe qui entraîne le rapprochement de l'hélicase eIF4A de la jonction triple-hélice. Le dépliement de cette structure par eIF4A libère le ribosome qui entre alors en phase d'élongation (Figure 41). Ce mécanisme d'initiation, dit par « tethering » de la machinerie de traduction, présente des caractéristiques hybrides entre le modèle consensus d'initiation coiffe-dépendant et celui non-canonique induit par les IRES. Il est vraisemblable qu'il soit à l'origine de l'efficacité de traduction qui caractérise les ARNm d'histones lors de la phase S du cycle cellulaire.

III.2.4 Mécanismes d'initiation coiffe-indépendant chez les virus de plantes

Chez les plantes, la majorité des virus à ARN de polarité positive ne portent pas de coiffe à leur extrémité 5', ne sont pas polyadénylés en 3' et font donc appel à des mécanismes non-conventionnels pour recruter le ribosome en 5' de leur génome viral (pour revue, Newburn and White, 2015). Une des stratégies employées par ces virus (essentiellement *tombusviridae*, *luteovirus*, et *umbravirus*) implique un élément ARN en *cis* nommé « élément 3' d'activation de la traduction coiffe-indépendante » (3' cap-independent translation enhancer, 3'CITE), localisé dans la région 3'UTR du génome ARN (pour revue, Nicholson and White, 2011; Simon and Miller, 2013) (Figure 42A). Les 3'CITE permettent une traduction efficace des protéines virales codées dans la région 5' en recrutant soit des facteurs d'initiation, soit la sous-unité 60S vers l'ARN viral (Nicholson et al., 2010; Stupina et al., 2008; Treder et al., 2008) (Figure 42B). Ces éléments peuvent prendre différentes formes : en I, en Y, en T, ou en structure pseudonœud (Simon and Miller, 2013; Wang et al., 2009) (Figure 42B). La plupart des 3'CITE, tels que ceux des *luteovirus*, *tombusvirus* et *panicovirus*, fonctionnent en recrutant eIF4E et/ou eIF4G et en interagissant avec leur 5'UTR correspondante via des interactions ARN-ARN en « kissing loop » avec une structure en tige-boucle localisée à l'extrémité 5' de l'ARN viral (Nicholson et al., 2010, 2013; Treder et al., 2008). L'interaction des deux extrémités relocalise les facteurs d'initiation à l'extrémité 5' du génome qui est le site où est recruté le ribosome.

D'autres 3'CITE, tels que ceux du TCV (Turnip crinkle virus), se lient directement à la sous-unité ribosomique 60S (Stupina et al., 2008), qui va recruter et s'associer à la sous-unité 40S au niveau d'une séquence riche en pyrimidines dans la région 5'UTR (McCormack et al., 2008; Simon, 2015; Stupina et al., 2011). Les 3'CITE, par l'intermédiaire d'un mécanisme d'initiation de la traduction efficace, permettent aux ARN viraux d'entrer en compétition avec les ARNm cellulaires pour la machinerie traductionnelle de l'hôte.

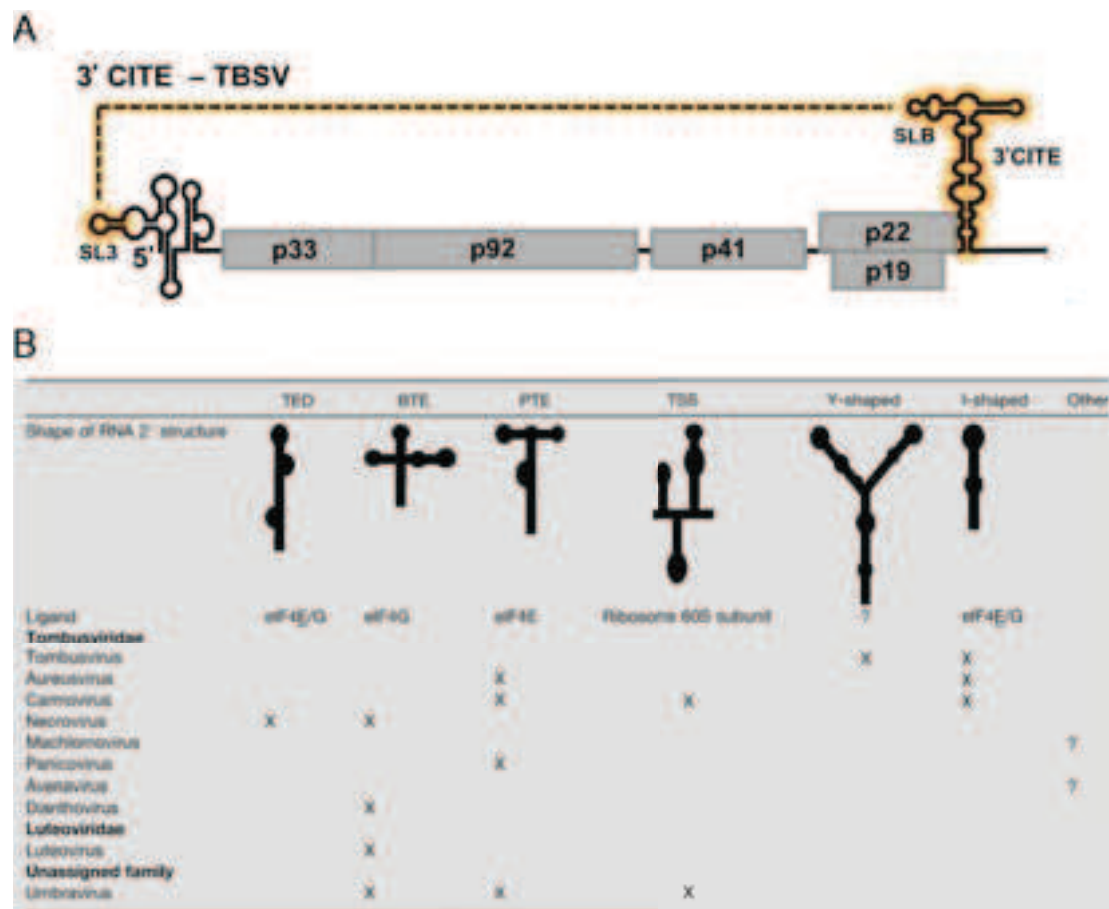


Figure 42. Initiation de la traduction coiffe-indépendante chez les virus de plantes.

(A) (Figure de Newburn and White, 2015) Représentation linéaire du génome du virus TBSV (*Tomato bushy stunt virus*). Les régions codantes sont représentées par des rectangles gris, les régions non-codantes par des traits noirs. Les structures ARN agissant en *cis* dans le mécanisme de traduction sont surlignées en jaune. L'interaction à distance du 3'CITE avec la région 5' est indiquée par une ligne en pointillé. (B) (Figure de Nicholson and White, 2011) Tableau récapitulatif des structures secondaires, ligands et distribution phylogénétique des éléments 3'CITE. La première ligne indique les différentes classes de 3'CITE. TED : translation enhancer domain (Danthinne et al., 1993), BTE : *Barley yellow dwarf virus* (BYDV)-like translation element (Guo et al., 2001), PTE : *Panicum mosaic virus* (PMV)-like translation element (Wang et al., 2011), TSS : T-shaped structure (Stupina et al., 2008). La structure secondaire des différentes classes de 3'CITE est représentée sur la 2^{ème} ligne. Les ligands reconnaissant les différents 3'CITE sont indiqués.

IV Objectifs et réalisations du travail de thèse

Les objectifs de ma thèse étaient d'une part de mieux comprendre le mécanisme de biogenèse des mRNP de sélénoprotéines et d'autre part d'élucider l'ensemble des étapes du mécanisme de traduction et de recodage des ARNm de sélénoprotéines. Mon travail de thèse a contribué à mettre en évidence l'existence d'un mécanisme particulier de maturation de la coiffe en 5' des ARNm de sélénoprotéines (Wurth et al., 2014). Nos résultats montrent que certains ARNm de sélénoprotéines possèdent une coiffe hyperméthylée $m_3^{2,2,7}G$ à la manière des sn et snoARN. Ces travaux seront présentés dans le chapitre I. Nous avons déterminé le mécanisme de biogenèse de cette coiffe, et avons montré qu'il fait appel à l'enzyme Tgs1 (Triméthyl-guanosine synthase). Tgs1 peut être recrutée vers les ARNm de sélénoprotéines via la protéine cœur Nop58 et le chaperon d'assemblage SMN. L'analyse du rôle du complexe SMN et de son cofacteur, le méthylosome dans la biogenèse des mRNP de sélénoprotéines est présentée dans le chapitre II. Nous avons également révélé l'existence d'un lien inattendu entre l'expression des sélénoprotéines et l'amyotrophie spinale (SMA), une pathologie provoquée par des défauts d'expression de la protéine SMN. L'impact de la coiffe m_3G , mais également le rôle du codon UGA^{Sec} et de l'élément SECIS sur la traduction des ARNm de sélénoprotéines ont été analysés *in vitro* et sont présentés dans le chapitre III. L'ensemble de mes résultats de thèse va dans le sens de l'existence d'un mécanisme d'initiation atypique de la traduction des ARNm de sélénoprotéines. Enfin, j'ai mis au point des méthodes de purification des complexes de traduction associés à un ARNm de sélénoprotéine. Notre but est de résoudre la structure tridimensionnelle du complexe de recodage par cryo-microscopie électronique et d'identifier les facteurs impliqués dans ce mécanisme. L'objectif final est d'élucider l'ensemble des étapes de la traduction et du recodage des ARNm de sélénoprotéines.

Résultats

I La coiffe des ARNm de sélénoprotéines est hyperméthylée chez les mammifères

I.1 Résumé de la publication

Notre laboratoire a montré précédemment que l'assemblage des mRNP de sélénoprotéines en 3'UTR, pré-requis pour leur traduction, est un processus particulier qui obéit aux mêmes règles que celles des snRNP et des snoRNP. SBP2, est capable de transiter entre le noyau et le cytoplasme et lie les composants d'une machinerie d'assemblage liée au chaperon protéique Hsp90 (Boulon et al., 2008). De plus, SBP2 possède des motifs de liaisons à l'ARN communs avec des protéines cœur des sn et des snoRNP (Allmang et al., 2002b). Ainsi, il est tout à fait envisageable que les mRNP de sélénoprotéines puissent également partager, avec les sn et les snoRNP, une voie commune de maturation de leur extrémité 5'.

Nos résultats montrent que certains ARNm de sélénoprotéines ne sont pas reconnus efficacement par le facteur canonique d'initiation de la traduction eIF4E car ils possèdent une coiffe hyperméthylée $m_3^{2,2,7}G$ à la manière de certains petits ARN non codants. Ce travail est résumé ci-dessous et est présenté dans la publication qui suit.

Une sous-classe d'ARNm de sélénoprotéines porte une coiffe hyperméthylée et montre une affinité réduite pour le facteur d'initiation eIF4E

Les ARNm eucaryotes se caractérisent généralement par la présence d'une coiffe monométhylée m^7G à leur extrémité 5' (Shatkin, 1976; Sonenberg, 2008). Celle-ci est reconnue par de multiples effecteurs protéiques et joue un rôle essentiel dans l'initiation de la traduction des ARNm mais également dans leur localisation et dégradation (Houseley and Tollervey, 2009; Köhler and Hurt, 2007; Mitchell and Tollervey, 2001). Les ARN non-codants tels que les snARN et snoARN acquièrent une modification supplémentaire de la coiffe qui devient triméthylée $m_3^{2,2,7}G$ (ou m_3G), ce qui est primordial pour leur localisation, assemblage et maturation (Huber et

al., 1998; Matera et al., 2007a; Mouaikel et al., 2002; Strasser et al., 2004). Puisque notre laboratoire a montré que les mRNP de sélénoprotéines subissent à leur extrémité 3' le mécanisme d'assemblage des snRNP et snoRNP, nous avons testé si les ARNm de sélénoprotéines étaient également soumis à un événement de maturation similaire au niveau de la coiffe située en 5' et sont reconnus par la machinerie de traduction canonique. Le facteur d'initiation eIF4E reconnaît de manière spécifique la coiffe m⁷G des ARNm et est indispensable à la traduction coiffe-dépendante des ARNm (Jackson et al., 2010; Sonenberg, 2008). Par des expériences de GST pull-down nous avons montré que la majorité des ARNm de sélénoprotéines sont moins bien reconnus par eIF4E que les ARNm non-sélénoprotéines. Par des expériences d'immunoprécipitation à l'aide d'anticorps reconnaissant spécifiquement la coiffe m₃G des ARN (IP m₃G), nous avons établi que la majorité des ARNm de sélénoprotéines porte une coiffe hyperméthylée (Figure 43). Ces expériences ont permis de révéler l'existence d'une corrélation inverse entre la liaison à eIF4E et la présence de coiffe m₃G. Elles nous ont permis d'identifier trois classes d'ARNm de sélénoprotéines : les ARNm majoritaires portent une coiffe hyperméthylée et ne sont pas reconnus par le facteur d'initiation eIF4E, d'autres sont coiffés de m⁷G et interagissent avec eIF4E alors que certains ARNm de sélénoprotéines semblent porter un mélange de l'une ou l'autre des coiffes m₃G ou m⁷G.

Tgs1 est l'enzyme responsable de l'hyperméthylation de la coiffe des ARNm des sélénoprotéines

Il a été précédemment montré que la Triméthylguanosine synthase 1 (Tgs1) est responsable de la conversion de la coiffe m⁷G en coiffe m₃G des snARN et snoARN (Matera et al., 2007a; Mouaikel et al., 2002). Cette étape de modification joue un rôle critique dans l'assemblage et la maturation des RNP correspondantes (Matera et al., 2007a). Elle est requise pour l'obtention de particules RNP matures et fonctionnelles. L'enzyme Tgs1 existe sous deux formes : la protéine entière (LF), majoritairement cytoplasmique, et une protéine isoforme plus courte (SF), essentiellement nucléaire (Girard et al., 2008). L'interaction de Tgs1 avec les ARNm de sélénoprotéines a été recherchée par analyse microarray des ARN co-immunoprécipités avec Tgs1 (Collaboration avec le laboratoire d'Edouard Bertrand, IGM, Montpellier). Pour discriminer les deux formes de Tgs1, des lignées de cellules HeLa exprimant

stablement l'une ou l'autre de ces deux formes avec une étiquette 3XFlag ont été utilisées et les immunoprécipitations ont été réalisées à l'aide d'anticorps anti-Flag. Nous avons montré que les deux formes de Tgs1 sont capables de s'associer aux ARNm de sélénoprotéines (Figure 43). L'affinité de Tgs1 LF pour les ARNm de sélénoprotéines apparaît légèrement plus forte que celle de Tgs1 SF, ce qui est très comparable au cas des snARN et différent des snoARN qui ne recrutent que la forme courte nucléaire de Tgs1. Nous avons ensuite recherché si Tgs1 pouvait être responsable de l'hyperméthylation de la coiffe des ARNm de sélénoprotéines. Ainsi, des tests d'inhibition de Tgs1 par ARN interférence (siRNA) combinés à une analyse de l'immunoprécipitation de la coiffe m⁷G ont été réalisés. L'inhibition de Tgs1 mène à une réduction globale du taux d'immunoprécipitation des ARNm des sélénoprotéines. Ceci indique que Tgs1 est bien responsable de l'hyperméthylation de la coiffe des ARNm de sélénoprotéines.

Tgs1 est recrutée sur les ARNm de sélénoprotéines via la protéine Nop58 et le chaperon d'assemblage SMN

Le mode de recrutement de Tgs1 est dépendant de la nature de sa cible ARN. Dans le cas des snARN, Tgs1 est recruté vers la coiffe m⁷G via un complexe de chaperons, le complexe SMN (Massenet et al., 2002; Mouaikel et al., 2003b; Paushkin et al., 2002) tandis que dans le cas des snoARN à boîte C/D ce sont des protéines constitutives de ces ARNs (Nop56 et Nop58) qui interagissent avec Tgs1 (Girard et al., 2008; Mouaikel et al., 2002; Verheggen et al., 2001). Pour déterminer si Tgs1 pouvait être recruté vers les ARNm de sélénoprotéines via l'une de ces deux voies, nous avons inactivé séparément les protéines Nop58 et SMN par siARN. Dans les deux cas, l'hyperméthylation des ARNm de sélénoprotéines a été très fortement affectée indiquant que les deux protéines participent au recrutement de Tgs1 vers les ARNm de sélénoprotéines et donc à l'hyperméthylation de la coiffe (Figure 43).

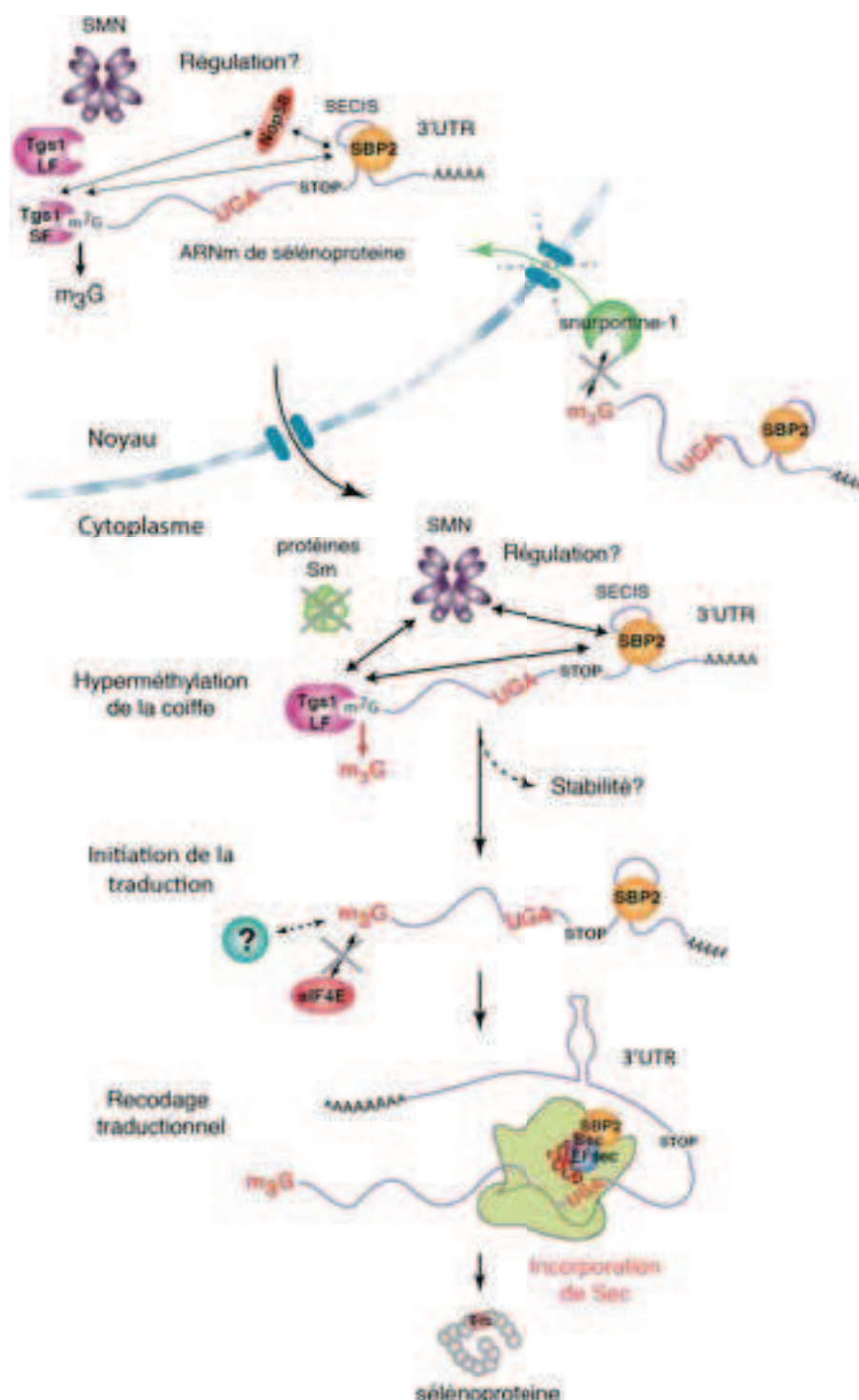


Figure 43. Modèle de l'hyperméthylation de la coiffe des ARNm de sélénoprotéines.

La coiffe des ARNm de sélénoprotéines est hyperméthylée par Tgs1. Les deux isoformes de Tgs1 : Tgs1 LF (nucéaire/cytoplasmique) et Tgs1 SF (exclusivement nucléaire) interagissent avec SBP2 et les ARNm de sélénoprotéines. Tgs1 est majoritairement recrutée par le complexe ternaire SBP2/SMN/Tgs1LF (flèches épaisses) mais également via SBP2/Nop58/Tgs1SF (flèches fines). L'hyperméthylation pourrait avoir lieu dans le noyau ou le cytoplasme et être soumise à des mécanismes de régulation. L'activité de Tgs1 est requise pour la stabilité des ARNm, l'incorporation de Sec ou l'initiation de la traduction. Les ARNm de sélénoprotéines coiffés de m₃G sont majoritairement localisés dans le cytoplasme (80%), ils n'interagissent ni avec la snurportine-1 ni avec les protéines Sm ; et ne semblent donc pas réimportés vers le noyau. Ils présentent une affinité réduite pour le eIF4E. Nos résultats suggèrent l'existence d'un mécanisme d'initiation de la traduction atypique faisant potentiellement intervenir d'autres facteurs.

Les ARNm de sélénoprotéines portant une coiffe m₃G sont localisés dans le cytoplasme et sont traduits

La découverte d'ARNm portant une coiffe m₃G mène à la question fondamentale de leur capacité à être traduits dans le cytoplasme. Par des expériences de fractionnement subcellulaire, nous avons montré que les ARNm de sélénoprotéines triméthylés sont plus abondants dans le cytoplasme que dans le noyau. Pour déterminer si ces derniers peuvent être traduits, nous avons évalué leur capacité à s'associer à des ribosomes actifs. Des expériences de fractionnement sur gradient de sucrose, combinées à des IP m₃G, ont montré que les ARNm triméthylés des sélénoprotéines sont retrouvés dans la fraction contenant les polysomes ce qui suggère qu'ils sont bien traduits (Figure 43).

Contrairement à eIF4E, l'activité de Tgs1 est requise pour la traduction de GPx1 in vivo

De par son activité antioxydante, GPx1 est l'une des sélénoprotéines les plus impliquées dans le cancer. Puisque son ARNm porte une coiffe hyperméthylée par Tgs1, nous l'avons choisi comme modèle pour déterminer l'importance fonctionnelle de la coiffe m₃G sur la synthèse des sélénoprotéines *in vivo*. J'ai construit deux lignées cellulaires humaines (T-Rex HEK293) capables d'exprimer de manière inductible des versions étiquetées (HA) de la sélénoprotéine GPx1 sauvage (HA-GPx1) soumise au recodage traductionnel, ou mutante (HA-GPx1Cys), qui suit un mécanisme de traduction canonique. L'utilisation de ces deux lignées nous a permis d'analyser sur la synthèse *de novo* de la sélénoprotéine GPx1, l'impact de l'inactivation de l'hyperméthylase Tgs1 et du facteur d'initiation de la traduction eIF4E, qui reconnaît moins bien les ARNm de sélénoprotéines coiffés de m₃G. L'inhibition de Tgs1 par siARN, affecte fortement l'expression de HA-GPx1 mais pas de HA-GPx1Cys ce qui démontre que Tgs1, probablement par son action sur l'hyperméthylation de la coiffe des ARNm, joue un rôle important dans la synthèse *in vivo* des sélénoprotéines, sans doute au niveau du mécanisme de recodage. Cependant, nos résultats ne nous permettent pas de différencier si l'activité de Tgs1 est requise pour l'étape d'incorporation de la sélénocystéine ou au niveau de l'initiation de la traduction. Il est possible que les deux mécanismes soient liés pour une traduction optimale des sélénoprotéines. En revanche, des résultats inverses sont

observés lorsque le facteur eIF4E est inhibé. Ces résultats suggèrent fortement qu'eIF4E n'est pas essentiel à la traduction des ARNm de sélénoprotéines portant une coiffe m₃G et que le mécanisme d'initiation de la traduction des sélénoprotéines est atypique et pourrait faire appel à d'autres facteurs encore non identifiés (Figure 43).

I.2 Publication

Hypermethylated-capped selenoprotein mRNAs in mammals

Laurence Wurth[†], Anne-Sophie Gribling-Burrer^{1†}, Céline Verheggen, Michael Leichter, Akiko Takeuchi, Stéphanie Baudrey, Franck Martin, Alain Krol, Edouard Bertrand et Christine Allmang

[†] contribution équivalente

Nucleic Acids Research, 2014, Vol. 42, No. 13 **8663–8677**

Hypermethylated-capped selenoprotein mRNAs in mammals

Laurence Wurth^{1,†}, Anne-Sophie Gribbling-Burrer^{1,†}, Céline Verheggen², Michael Leichter¹, Akiko Takeuchi¹, Stéphanie Baudrey¹, Franck Martin¹, Alain Krol¹, Edouard Bertrand² and Christine Allmang^{1,*}

¹Architecture et Réactivité de l'ARN, Université de Strasbourg, Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 67084 Strasbourg, France and ²Equipe labélisée Ligue contre le cancer, Institut de Génétique Moléculaire, Centre National de la Recherche Scientifique, UMR 5535, 34293 Montpellier, France

Received April 24, 2014; Revised June 05, 2014; Accepted June 17, 2014

ABSTRACT

Mammalian mRNAs are generated by complex and coordinated biogenesis pathways and acquire 5'-end m⁷G caps that play fundamental roles in processing and translation. Here we show that several selenoprotein mRNAs are not recognized efficiently by translation initiation factor eIF4E because they bear a hypermethylated cap. This cap modification is acquired via a 5'-end maturation pathway similar to that of the small nucleolar RNAs (sn- and snoRNAs). Our findings also establish that the trimethylguanosine synthase 1 (Tgs1) interacts with selenoprotein mRNAs for cap hypermethylation and that assembly chaperones and core proteins devoted to sn- and snoRNP maturation contribute to recruiting Tgs1 to selenoprotein mRNPs. We further demonstrate that the hypermethylated-capped selenoprotein mRNAs localize to the cytoplasm, are associated with polysomes and thus translated. Moreover, we found that the activity of Tgs1, but not of eIF4E, is required for the synthesis of the GPx1 selenoprotein *in vivo*.

INTRODUCTION

Eukaryotic mRNAs, synthesized by RNA polymerase II (polII), are characterized by the presence of a 7-methylguanosine (m⁷G) cap structure at their 5'-end (1). Nuclear RNA capping serves as an important quality control checkpoint and is an essential determinant of mRNA expression, stability and biogenesis (2). The m⁷G cap promotes splicing of cap-proximal introns, 3'-end formation, nucleocytoplasmic export, control of RNA stability and

translation (3–7). Two major cap-binding proteins mediate these processes: the cap-binding complex (CBC) in the nucleus, composed of the CBP20 and CBP80 subunits (3), and the translation initiation factor eIF4E in the cytoplasm (8–10). The binding of both CBC and eIF4E is highly specific for the m⁷G cap, and other modified caps are poorly recognized by these proteins (11–14). PolII small nuclear (snRNAs) and a few small nucleolar RNAs (snoRNAs) also acquire m⁷G caps in the nucleus but become further methylated at guanosine position 2 to generate the 2,2,7-trimethylguanosine cap (m₃G or TMG cap) (15). TMG modification of snRNAs is performed in the cytoplasm by the trimethylguanosine synthase 1 (Tgs1) which catalyzes the m⁷G to m₃G hypermethylation. After export to the cytoplasm, precursor snRNAs are assembled with Sm core proteins by the Survival of Motor Neuron (SMN) complex (16). Binding of SMN and Sm proteins is a prerequisite for further recruitment of Tgs1. The m₃G cap together with the Sm ring represent bipartite signals that promote nuclear import of snRNPs for final maturation and assembly into the active spliceosome (17,18). The transfer of TMG-capped snRNPs to the nucleus is governed by snurportin which interacts specifically with m₃G and Sm proteins, but not with m⁷G cap structures (19,20). Cap hypermethylation, therefore, represents a critical step in determining the fate of sn- and snoRNAs.

Unlike snRNAs, hypermethylation of m⁷G-capped snoRNAs takes place in the nucleus, and snoRNAs transit to the Cajal bodies (CB) in a PHAX (phosphorylated adaptor for RNA export)-dependent manner where they are hypermethylated by nuclear Tgs1 (21,22). In this case, Tgs1 is recruited to the snoRNA by the core proteins Nop56 and Nop58 for the box C/D snoRNAs, and by dyskerin for the H/ACA snoRNAs (23,24). Two Tgs1 isoforms have been identified; the long form (LF) is mainly

*To whom correspondence should be addressed. Tel: +33 3 88 41 70 80; Fax: +33 3 88 60 22 18; Email: c.allmang@unistra.fr

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Present address:

Laurence Wurth, Gene Regulation, Stem Cells and Cancer Programme, Centre for Genomic Regulation and UPF, 08003 Barcelona, Spain.

involved in snRNA hypermethylation and is present in both the nucleus and cytoplasm, whereas a short nuclear isoform (SF) is dedicated to snoRNA maturation (23). In the nucleoplasm, Tgs1 isoforms are present in a large multiprotein complex that contains ribonucleoprotein (RNP) maturation, transport and assembly factors (15,22,25).

Selenoprotein mRNAs constitute an interesting class of mRNAs. Indeed, because of the presence of an in-frame UGA codon, recoded as selenocysteine (Sec) and otherwise read as a stop codon, they are submitted to distinctive biogenesis and translation pathways (26,27). Several *cis*- and *trans*-acting factors participate in the co-translational recoding of UGA Sec codons. Pivotal in this process are the SelenoCysteine Insertion Sequence (SECIS), a stem-loop in the 3' untranslated region (3'UTR) of selenoprotein mRNAs, and the SECIS-binding protein 2 (SBP2) (28–31). We have previously shown that the assembly mechanism of selenoprotein mRNPs, in particular in the 3'UTR region, is partly common with that of snRNP, snoRNP particles and the telomerase (27). SBP2 which undergoes nucleocytoplasmic shuttling (32,33), binds to components of this conserved RNP assembly machinery linked to the protein chaperone Hsp90 (27). Besides, SBP2 shares functional RNA binding motifs with primary core proteins of sn- and snoRNPs (27,34), suggesting that selenoprotein mRNPs, sn- and snoRNPs could also share a common 5' maturation pathway.

In this work, we demonstrate that several mammalian selenoprotein mRNAs, in contrast to other cellular mRNAs, undergo a 5'-end maturation pathway similar to that of sn- and snoRNPs and bear hypermethylated caps.

MATERIALS AND METHODS

Cell culture

HEK293FT, HeLa and U2OS cells were cultured at 37°C in 5% CO₂ in Dulbecco's modified Eagle's media (DMEM) containing 10% Fetal Calf Serum (FCS), 1% penicillin–streptomycin and 10 nM sodium selenite. One percent geneticine was added to HEK293FT cells. Cells were transfected using Turbofect (Fermentas) and siRNA inactivation was done for 48 h using 30 or 100 nM of siRNA and Lipofectamine 2000 (Invitrogen) following the manufacturer's conditions. siTGS1, siSMN, siSBP2 and siEIF4E RNAs and sicontrol were ON-TARGET plus smart pools of four different siRNAs (Thermo Scientific Dharmacon), siNop58 RNAs were from Santa Cruz. For rescue experiments, cells were transfected for 24 h after siRNA inactivation with pcGFP-Tgs1 (21,24), non-tagged plenti6/V5-SBP2 or pcDNA5-eIF4E expression plasmids and grown for an additional 24 h treatment. Cells were extracted with HNTG buffer (20 mM HEPES–NaOH pH 7.9, 150 mM NaCl, 1% Triton, 10% glycerol, 1 mM MgCl₂, 1 mM EGTA (Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid), 1 mM PMSF (phenylmethylsulfonyl fluoride), anti-protease cocktail from Roche). Subcellular fractionation was performed according to (35).

Generation of stable cell lines and induced protein expression

Flp-In™ T-Rex™ 293 system (Invitrogen) was used according to the manufacturer's instructions to generate stable cell lines with regulated expression of 3xHA-GPx1 (HA-GPx1) and 3xHA-GPx1mutCys (HA-Gpx1Cys). Stable clones containing the required open reading frame under the control of the cytomegalovirus/tetO₂ hybrid promoter were selected by culturing in selective medium containing 250 µg/ml of hygromycin (InvivoGen) and 15 µg/ml of blasticidin (InvivoGen). Resistant colonies were expanded and tested for doxycycline-regulated protein expression. Protein expression was induced by addition of 0.5–1 µg/ml of doxycycline in culture medium and in the presence of 10 nM of sodium selenite.

Stable isogenic HeLa cells expressing the Flag-tagged Tgs1 proteins were obtained with HeLa Flp-In cells, by cotransfection of the parental cells with pcDNA5-3X Flag-Tgs1 (SF or LF) and a Flippase expression vector with Lipofectamine and Plus reagent (Invitrogen). Stable clones were then selected with 50 mM hygromycin B (Calbiochem), picked individually, expanded and characterized by western blots. Individual clones usually expressed similar levels of the tagged protein. DNA cloning was performed with the Gateway system (Invitrogen).

Immunopurification, co-immunoprecipitation and western blotting

Immunopurification of endogenous SBP2 complexes was performed as described in (27). For co-immunoprecipitation experiments, cell extracts (100–300 µl) were incubated with 50 µl of protein A-sepharose (Sigma) coupled to the indicated antibody for 2 h at 4°C in NT2 buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 1 mM MgCl₂, 0.05% NP40, 1 mM DTT (Dithiothreitol), 400 µM VRC (Vanadyl ribonucleoside complexes), 100 U RNasin/ml, anti-protease cocktail) or with anti-GFP magnetic beads (Milenyi) in IPP150 (10 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.1% NP-40, 0.5 mM PMSF, 2 mM VRC). Beads alone were used as a control. Beads were washed with NT2 or IPP150, respectively, resuspended in Laemmli buffer, analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by western blot. Antibodies used were: anti-His (H3) mouse monoclonal (Santa Cruz), anti-SBP2 rabbit polyclonal (27), anti-Tgs1 mouse monoclonal (24), anti-Nop58 (C-20) polyclonal goat antibodies (Santa Cruz), anti-SMN (2B1) mouse monoclonal (Santa Cruz), anti-rpS21 goat polyclonal (Santa Cruz) and anti-AspRS rabbit polyclonal (36). For RNA IP, extracts were first clarified on protein A-sepharose to eliminate unspecific RNA binders. For microarray experiment, co-immunoprecipitation was performed according to (37).

TMG RIP and qRT-PCR analysis

Trimethyl-capped RNAs were immunoprecipitated using a strictly specific rabbit polyclonal anti-m³G (TMG) serum (Synaptic Systems) also referred to as R1131 (38,39). Ten microliters of serum was coupled to 50 µl of protein A-Sepharose beads saturated with 20 µg of both purified bovine serum albumin (BSA) and total yeast tRNA in NT2

buffer for 18 h at 4°C. The immobilized antibody was incubated with 200 µg of pre-cleared total RNAs prepared from HEK293FT cells (Tri-reagent, Euromedex) in a total volume of 1 ml for 2 h at 4°C. Beads were washed six times in NT2 buffer, the bound RNA was extracted by phenol/chloroform and precipitated. After DNase treatment, RNAs were reverse transcribed using AMV-RT (Q-Biogen) and cDNAs were amplified by classical or quantitative RT-PCR (qRT-PCR). Level of mRNAs after siRNA treatment was measured directly after RNA extraction using the same qRT-PCR method. qRT-PCR was performed on MX3005P (Stratagene) using the Maxima SYBR Green PCR kit (Fermentas). Oligonucleotides used for qRT-PCR are listed in Supplementary Table S1.

Anti-FLAG immunoprecipitation and microarray analysis

Anti-Flag Tgs1 LF and Tgs1 SF immunoprecipitations were done with M2 antibody-coated beads (Sigma-Aldrich). HeLa cells stably expressing Tgs1 LF, Tgs1 SF or parental cells as control, were extracted with lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton, antiprotease cocktail) for 30 min at 4°C and centrifuged for 10 min at 10 000 rpm at 4°C. Extracts were incubated with M2 beads for 2 h at 4°C. Beads were washed twice in HNTG and three times in phosphate buffered saline (PBS), RNAs from beads were homogenized in Trizol and purified with RNeasy mini kit (Qiagen). cDNA were prepared and labeled as recommended by the manufacturer (Affymetrix). RNAs were analyzed using human gene 1.0ST microarray covering all annotated transcripts (Affymetrix). Files were analyzed and normalized with expression console software (Affymetrix). The 3XFlag-Tgs1 IPs were performed in duplicates, and the control IP (non-tagged parental cell line) was done in triplicate. Microarray data are accessible at GEO GSE57625.

Recombinant proteins and GST pull-down assays

Radiolabeled proteins were synthesized in the presence of ³⁵S-Met in rabbit reticulocyte lysate (TNT; Promega). Binding was performed with 5 µg GST-tagged protein in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM MgCl₂, 0.1 mM EDTA and 0.1% NP-40. Washing was performed with the same buffer. GST- and His-tagged proteins were produced by standard procedure and purified using glutathione sepharose (GE Healthcare) and Ni-NTA agarose (Qiagen), respectively. GST-4E_{K119A} protein was expressed and purified as described by (40). For GST pull-down experiments purified GST-4E_{K119A} or GST (200 µg) was bound to 200 µl of MagneGST beads (Promega) and incubated with 100 µg of HEK293 total RNA in RNP buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 0.1 mM EDTA, 1 mM DTT, 10% glycerol, 100 U RNasin/ml, anti-protease cocktail) for 2 h at 4°C. Beads were washed five times with RNP buffer. The RNAs present in the flow-through or on the beads were extracted by phenol/chloroform and precipitated.

Yeast-two hybrid interaction tests

For yeast-two hybrid interaction (Y2H) assays, appropriate pGBKT7 (DB) or pGADT7 (AD) plasmids were co-

transformed into AH109 (Clontech Laboratories, Inc.) and plated on triple selective media (-Leu -Trp -Ade).

Polysome analysis

Polysome profiles were analyzed on sucrose gradients. HEK293 cells were cultivated up to 80% confluence. 2 × 10⁶ cells were collected, washed in 3 volumes PBS and incubated in the presence of 150 µg/ml cycloheximide for 20 min on ice. Lysis of the cells was performed in polysome buffer (PB) containing 10 mM HEPES-NaOH pH 7.5, 50 mM KCl, 5 mM MgCl₂, 1 mM DTT, 100 units RNasin/ml, 400 µM VRC, protease inhibitors and 150 µg/ml cycloheximide or in polysome denaturing buffer (PDB) containing 10 mM HEPES-NaOH pH 7.5, 10 mM potassium acetate, 0.5 mM magnesium acetate, 5 mM DTT. Cytoplasmic extracts were loaded onto a linear 7–47% sucrose gradient prepared in PB or PDB. Polysomes were separated by 2.5 h centrifugation at 37 000 rpm using a Beckmann SW41 rotor. Gradients were monitored by following absorbance at 254 nm. RNAs were extracted and analyzed by RT-qPCR as described previously.

RESULTS

A subclass of selenoprotein mRNAs harbor hypermethylated caps and show reduced affinity for eIF4E

Having previously established that selenoprotein mRNPs and sn-, snoRNPs share a common 3' assembly pathway (27), we asked whether selenoprotein mRNAs could consequently undergo similar 5' maturation events. Indeed, unlike mRNAs, PolII sn- and snoRNAs that are modified co-transcriptionally with a 5' m⁷G cap, subsequently acquire a trimethylguanosine cap m^{2,2,7} (TMG) structure (15). We asked whether selenoprotein mRNAs could also bear similar cap modifications and could be recognized by the canonical translational machinery. A prerequisite to cap-dependent translation is the recognition of the eukaryotic mRNA cap structure by the translation initiation complex eIF4F (composed of the three subunits eIF4E, eIF4A and eIF4G). eIF4E specifically recognizes the m⁷G moiety of the 5' m⁷GpppN cap of eukaryotic mRNAs and was shown to poorly recognize TMG caps (41). Therefore, we first examined the ability of selenoprotein mRNAs to be recognized by eIF4E.

We performed glutathione-S-transferase (GST) pull-down experiments using total RNAs from HEK293FT cells and a high-affinity mutant of eIF4E, GST-eIF4E_{K119A} (40). This mutant was developed for specific isolation of 5' m⁷G-capped mRNAs and showed strict specificity but 10-fold higher affinity for the m⁷G cap. The RNA content of eIF4E bound and unbound fractions was determined by qRT-PCR analysis. For the detection of selenoprotein mRNAs, we used primers complementary to 12 out of the 25 selenoprotein mRNAs characterized in mammals (Figure 1A). β-actin, HPRT (hypoxanthine guanine phosphoribosyltransferase) and LDHA (lactate dehydrogenase A) mRNAs were used as the m⁷G-capped controls, U3 snoRNA and U2 snRNA for TMG-capped controls. Coherently, an average of 75% of the canonical β-actin, HPRT and LDHA mRNAs were recovered in the eIF4E bound fraction, whereas

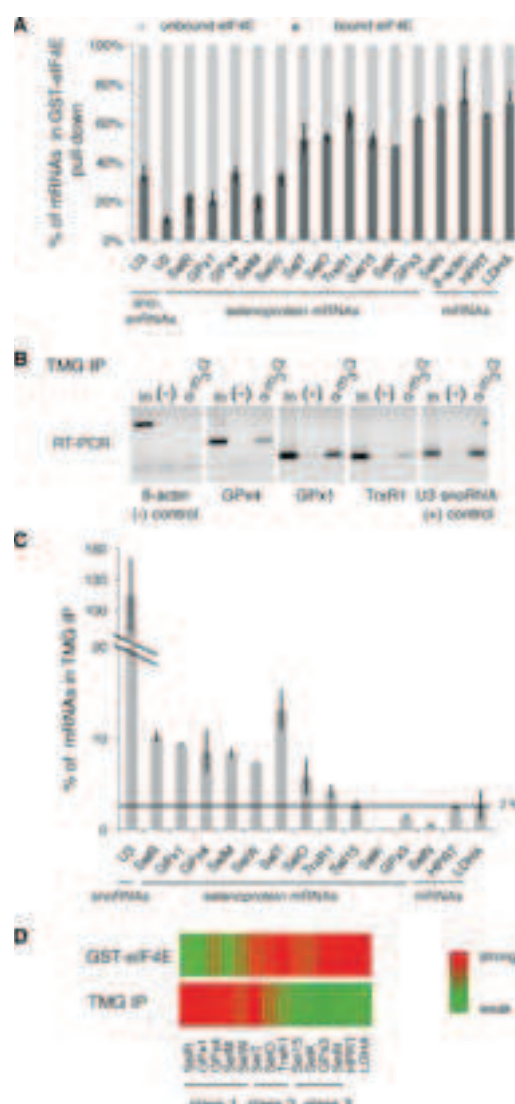


Figure 1. The cap of selenoprotein mRNAs is hypermethylated and poorly recognized by eIF4E. (A) Binding of recombinant GST-eIF4E_{K119A} to HEK293 total RNA. GST-eIF4E_{K119A} (or GST alone) was bound to glutathione beads and incubated with the extract. The percentage of mRNAs present in the bound and unbound fractions were determined separately by qRT-PCR and normalized to 100%. Asymmetric error bars represent the minimum and maximum observations for three independent biological replicates, reflecting intrinsic variability. *SelR*, *GPx1*, *GPx4*, *SelM*, *SelW*, *SelT*, *SelO*, *TrxR1*, *Sel15*, *SelK*, *GPx3* and *SelN* are selenoprotein mRNAs. *U3 snoRNA* was used as a positive control, *β-actin*, *HPRT* and *LDHA* are housekeeping mRNAs used as negative controls. (B and C) Total RNA extracted from HEK293FT cells was immunoprecipitated with anti-TMG serum (α -m³G). Bound RNA was analyzed by (B) RT-PCR or (C) qRT-PCR. (–) Control without antibodies. In: input 10%. The graph represents the percentage of mRNAs in IP compared with the input RNA. Error bars represent standard deviation of an average of three independent experiments. The horizontal line represents the level of housekeeping mRNA binding (1–2% in average). See also Supplementary Figure S1 for specificity controls. (D) Heatmap representation of mRNA binding in GST-eIF4E pull-down and TMG IP experiments. The binding scale is represented to the right, maximum binding values in each set of experiments are represented in red and minimal binding values in green. Heat maps were generated with the MeV software. The three classes of selenoprotein mRNAs are indicated.

74% of sn-, snoRNAs were found in the unbound fraction (Figure 1A). The results revealed that selenoprotein mRNAs showed differential binding patterns to eIF4E. The selenoprotein mRNAs of *SelR*, glutathione peroxidases 1 and 4 (*GPx1*, *GPx4*), *SelM* and *SelW* showed a distribution pattern similar to that of sn-, snoRNAs with only 20–35% of the mRNA recovered in the eIF4E bound fraction (Figure 1A). *SelT*, *SelO*, thioredoxin reductase 1 (*TrxR1*), *Sel15*, *SelK* selenoprotein mRNAs showed an intermediate pattern with over 50% of the mRNAs in the bound fraction (Figure 1A), whereas selenoprotein mRNAs comprising glutathione peroxidase 3 (*GPx3*) and selenoprotein N (*SelN*) were enriched up to 70% in the eIF4E bound fraction with patterns similar to non-selenoprotein mRNAs (Figure 1A).

To explain the differential binding to eIF4E, we asked whether some of the selenoprotein mRNAs could bear 5' modifications similar to sn- and snoRNAs as a trimethylguanosine cap m^{2,2,7} (TMG) structure. To answer this question, RNAs extracted from HEK293FT cells were immunoprecipitated with the highly specific anti-TMG cap R1131 serum that was demonstrated not to recognize monomethylated caps (38,39) (see also Supplementary Figure S1 for specificity validation of the antibody). The immunoprecipitated RNAs were extracted and analyzed by RT-PCR. Results showed that anti-TMG antibodies recognized *GPx1*, *GPx4* and *TrxR1* selenoprotein mRNAs as well as *U3 snoRNA* used as the control but not the m⁷G-capped *β-actin* mRNA, indicating that the cap of these endogenous selenoprotein mRNAs is hypermethylated (Figure 1B). Quantitative RT-PCR analysis revealed that seven selenoprotein mRNAs, namely *SelR*, *GPx1*, *GPx4*, *SelM*, *SelW*, *SelT* and *SelO*, were specifically immunoprecipitated with anti-TMG antibody (Figure 1C) while *TrxR1* was only weakly recognized. These results correlate with the GST-eIF4E pull-down experiments, as these TMG-capped selenoprotein mRNAs correspond to those poorly recognized by eIF4E. *TrxR1*, *SelT* and *SelO*, that show an intermediate binding pattern of eIF4E, are consistently recognized in the TMG-IP (Figure 1A, C). Four selenoprotein mRNAs were not pulled-down over the background (*Sel15*, *GPx3*, *SelN* and *SelK*) in TMG immunoprecipitations (Figure 1C); they were consistently bound by eIF4E (Figure 1A). As expected, the *U3 snoRNA* positive control was fully retained but not the control *HPRT* or *LDHA* housekeeping mRNAs (Figure 1C). A significant fraction (5–15%) of the TMG-capped selenoprotein mRNAs was recovered in the anti-TMG immunoprecipitation. This precipitation is weaker compared with *U3 snoRNA* (100%) but is nevertheless consistent with previously reported levels for snRNAs (39).

Determinants such as the larger size, under-representation of selenoprotein mRNAs versus sno-, snRNAs as well as the shorter half-life of selenoprotein mRNAs compared to sn(o)RNAs may contribute to the lower recovery of selenoprotein mRNAs in the immunoprecipitation. It is also envisageable that different mRNA folded structures in the 5'UTR could lead to differential recognition of individual mRNAs by the antibody.

TMG-IP experiments, in combination to the GST-eIF4E pull-down results, suggest that a substantial fraction of se-

lenoprotein mRNAs possesses a trimethylated guanosine cap. Altogether, our results reveal an inverse relationship between eIF4E binding and TMG capping (Figure 1D) and suggest the existence of three classes of selenoprotein mRNAs. In the first class, <50% of the selenoprotein mRNAs are bound to eIF4E and the TMG-IP efficiency is >5%. These mRNAs include SelR, GPx1, GPx4, SelM and SelW (Figure 1D); they harbor a hypermethylated cap and are not recognized efficiently by the translation factor eIF4E. A second class of selenoprotein mRNAs showed an intermediate pattern with >50% of the mRNAs bound by eIF4E and >5% TMG-IP efficiency; this class includes selenoprotein mRNAs such as SelT, SelO and TrxR1 that can possibly bear both types of caps (Figure 1D). Finally, the third class represents selenoprotein mRNAs—SelI5, SelK, GPx3 and SelN—(Figure 1D) bound with <5% in the TMG-IP, that are m⁷G-capped and recognized by eIF4E. Selenoprotein mRNAs thus appear to be subjected to differential 5' processing events.

Tgs1 hypermethylates the cap of selenoprotein mRNAs

The trimethylguanosine synthase 1 (Tgs1) converts the m⁷G cap of sn- and snoRNA precursors into the functional TMG cap (21,24). This modification step plays a critical role in targeting sn- and snoRNA precursors to their final maturation and functional site (15). Tgs1 is present in a large U3 snoRNA processing complex that also contains components of the conserved RNP assembly complex linked to Hsp90 (27). Because selenoprotein mRNAs are assembled into mRNPs by the same molecular machinery (27), we asked whether Tgs1 was linked to selenoprotein mRNP biogenesis. Tgs1 exists as two molecular species: a full-length, mostly cytoplasmic (Tgs1 LF) and a shorter nuclear isoform (Tgs1 SF) (23). First, we examined if Tgs1 could be associated to selenoprotein mRNPs *in vivo* and performed immunoprecipitation of RNA–protein complexes from HEK293FT cells with an anti-Tgs1 antibody that recognizes both isoforms (Figure 2A). RNAs associated with endogenous Tgs1 were detected by RT-PCR. Several selenoprotein mRNAs, such as GPx1 and 4 as well as TrxR1 were specifically associated with Tgs1 *in vivo*, whereas this was not the case for the β -actin mRNA; U3 snoRNA was the positive interaction control that binds both Tgs1 isoforms (Figure 2A). Whether the interaction of Tgs1 could be generalized to all selenoprotein mRNAs was investigated by analyzing the RNA content of the Tgs1 immunoprecipitation reaction (IP) by microarrays. Because we did not have antibodies specific for either form of Tgs1, we used two HeLa cell lines that stably express either Tgs1 LF or Tgs1 SF with a 3XFlag tag, and performed the IPs using anti-Flag beads, using the parental cell line to perform control IPs (37). On average, selenoprotein mRNAs were significantly enriched in the IPs versus non-selenoprotein mRNAs that were not retained (Table 1). We found that both Tgs1 LF and Tgs1 SF associated to selenoprotein mRNAs (Table 1 and Figure 2B); likewise and as expected from previous studies (23), Tgs1 LF preferentially recognized snRNAs whereas Tgs1 SF bound more strongly to snoRNAs in our experiment (Table 1 and Figure 2B, (37)). To date, 25 selenoprotein genes have been identified in humans (42). Among

the 21 mRNAs expressed in our experiment, 14 were bound by Tgs1 LF and/or SF (Figure 2B), and 7 were not detected: TrxR1, SelH, SelI, SPS2, DI2, TrxR3 and GPx1-2. Binding efficiencies were variable, the strongest signals being observed for SelT, SelI5, GPx1-1 and SelM selenoprotein mRNAs for which the enrichment was similar to that seen for snRNAs and snoRNAs (Figure 2B). TrxR1 results show differences with those obtained in the endogenous Tgs1 RNA IP experiments (Figure 2A) most likely because qRT-PCR is more sensitive and specific than microarrays for the detection of the endogenous population of TrxR1 mRNAs. Binding of Tgs1 does not appear to strictly correlate with TMG-capping, suggesting that recruitment of the enzyme does not necessarily lead to cap modification. This is similar to the case of intronic snoRNAs which bind Tgs1 without subsequent cap modification. Altogether our results show that selenoprotein mRNAs interact with the two isoforms of Tgs1. The affinity of Tgs1 LF for selenoprotein mRNAs appears only slightly higher than for Tgs1 SF and resembles that of snRNAs.

Having shown the interaction of Tgs1 with selenoprotein mRNAs, we further investigated whether this enzyme was indeed responsible for hypermethylation of the selenoprotein mRNA cap. To this end, Tgs1 was knocked-down by RNA interference (RNAi) in HEK293FT cells (Figure 2C). Tgs1 mRNA was decreased to 36% (Figure 2D), which resulted in 35% of residual Tgs1 protein (Figure 2C). The drop of Tgs1 did not affect selenoprotein mRNA steady state levels (Figure 2D) but led to a global reduction of TMG-IP efficiency in correlation with the level of Tgs1 knock-down (Figure 2E and Supplementary Figure S2). Individual selenoprotein mRNAs responded differentially to Tgs1 inactivation. TMG-IP efficiency dropped down to 45% in the case of SelR (Figure 2E), comparably to the U3 snoRNA positive control. Milder effects were observed in the case of GPx1, GPx4, SelW and TrxR1 as the TMG-IP efficiency was only reduced to 80%. This can be attributed to residual Tgs1 and suggests that its depletion does not affect all selenoprotein mRNAs to the same extent. Hierarchy of selenoprotein expression and variations in selenoprotein mRNA stability have been reported in numerous studies, particularly with regard to the glutathione peroxidase family (43–46). Individual selenoprotein mRNAs respond in a unique fashion to various imbalances, including Tgs1 depletion. Altogether, these results indicate that Tgs1 is involved in hypermethylation of the cap of the majority of selenoprotein mRNAs. Our data therefore conclude that the cap hypermethylation activity of Tgs1 is thus not only restricted to that of sn- and snoRNA substrates.

Tgs1 is recruited to selenoprotein mRNAs predominantly via the assembly chaperone SMN but also by the core protein Nop58

The recruitment mode of Tgs1 is dependent on the nature of its RNA targets. In the case of sn- and snoRNAs, the best characterized Tgs1 targets, two different strategies allow Tgs1 isoforms to be recruited to the m⁷G cap. Recruitment of Tgs1 LF to snRNAs is dependent on both a major assembly factor, the Survival of Motor Neuron (SMN) complex, and the core protein SmB, whereas the core pro-

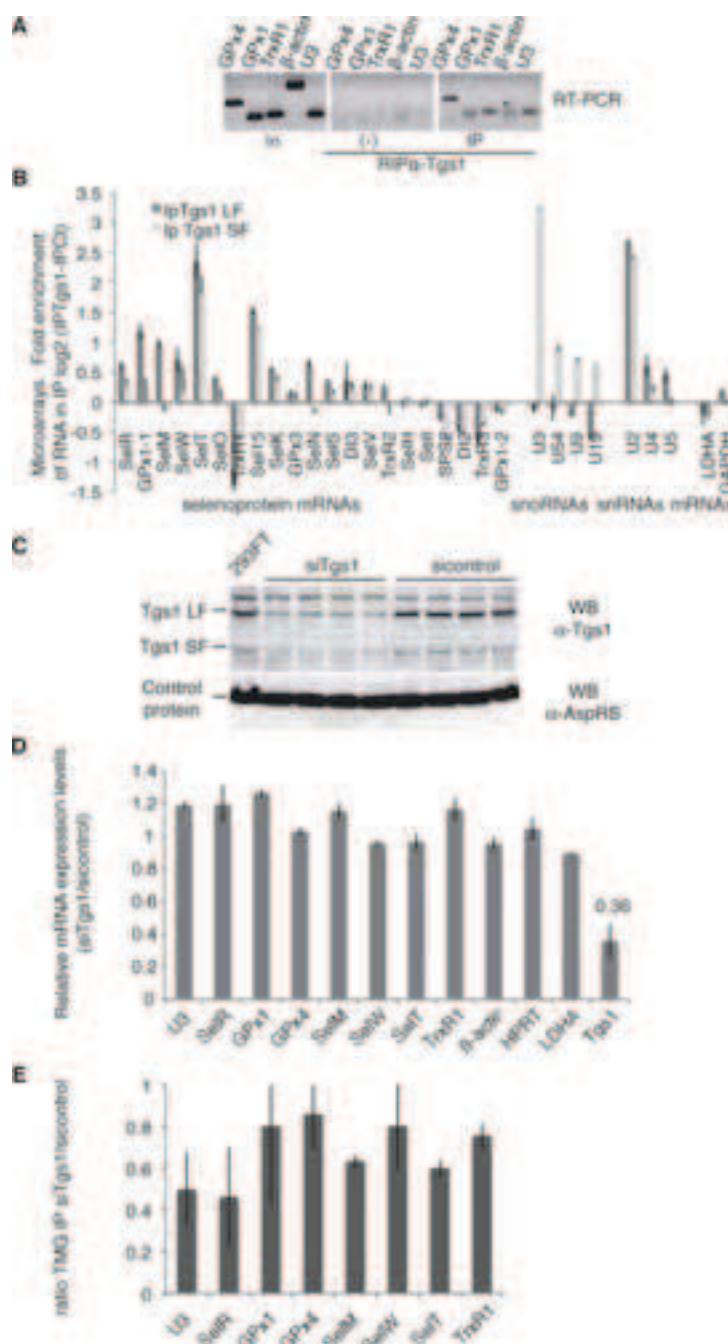


Figure 2. Tgs1 is involved in selenoprotein mRNA cap hypermethylation. (A) Tgs1 interacts with selenoprotein mRNAs *in vivo*. HEK293FT cells transfected with SBP2 were immunoprecipitated using anti-Tgs1 antibodies. Bound RNAs were detected by RT-PCR. In: input 15%; (–) no antibodies; U3: positive control; β -actin: negative control. (B) Tgs1 LF and SF associate with selenoprotein mRNAs. Isogenic HeLa cells stably expressing 3XFlag tagged Tgs1 LF and SF were used for anti-FLAG IPs, the RNA content was analyzed on microarrays. The control was the parental cell line that did not express any tagged protein. The graph represents the fold RNA enrichment in the IP in a log₂ scale. Dark bars: Tgs1 LF IP; gray bars: Tgs1 SF IP. Data of all the expressed selenoprotein mRNAs are represented followed by examples of snoRNAs (U3 to U19), snRNAs (U2, U4, U5) and housekeeping mRNAs (LDHA and GAPDH). GPx1-1 and GPx1-2 are splice variants of GPx1. (C) Tgs1 inactivation by siRNA. Tgs1 mRNA was decreased to 36% which resulted in 35% of residual Tgs1 protein. Western blot detection of Tgs1 was performed using anti-Tgs1 antibodies; anti-AspRS antibodies were used as a control. siRNAs directed against firefly luciferase were used as controls (sicontrol). (D) siTgs1 did not have any effect on the steady state level of selenoprotein mRNAs. qRT-PCR was used to determine relative expression levels by the $\Delta\Delta C_t$ method. (E) siTgs1 reduces hypermethylation efficiency. RNA-IP using anti-TMG serum was performed as described in Figure 1 under siTgs1 and sicontrol conditions. IP ratios between siTgs1 and sicontrol experiments are represented by the histogram and deduced from Supplementary Figure S2. Error bars represent standard deviation of an average of three independent experiments.

Table 1. Microarray analysis of RNA immunoprecipitations using Tgs1 SF and LF isoforms

log ₂ (IP Tgs1-IPCT)	Tgs1-LF-Cont	Tgs1-SF-Cont
Seleno mRNA (20 genes)	0.32 (0.0038)	0.16 (0.0001)
snRNA (9 genes)	0.82 (0.0031)	0.25 (0.0031)
C/D snoRNA (112 genes)	−0.34 (0.0033)	1.2 (0.0019)
H/ACA snoRNA (108 genes)	−0.11 (0.0037)	0.61 (0.0001)
mRNA (18685 genes)	−0.022 (0.2612)	−0.027 (0.1910)

Average of fold enrichment for each RNA family in IP (Tgs1 SF or LF) versus control measured in duplicate microarray analysis for Tgs1SF and Tgs1 LF RNA IPs. *P*-values were obtained from Mann–Whitney tests and are indicated in brackets. The average expression level of the entire set of selenoprotein mRNAs expressed in the HeLa cell line is indicated (20 expressed mRNAs among the 25 selenoprotein mRNA genes). The average expression level of the entire set of non-selenoprotein mRNAs is also indicated.

teins Nop56 and Nop58 recruit Tgs1 SF to box C/D snoRNAs (23,47,48). To determine if Tgs1 could be recruited to selenoprotein mRNAs via one of these two pathways, we separately inactivated expression of the Nop58 and SMN proteins by RNAi. This resulted in residual levels of 2 and 8% of Nop58 and SMN proteins, respectively, but did not affect expression of Tgs1 LF, SF and of the selenoprotein biosynthesis factor SBP2 (Figure 3 and Supplementary Figures S3A and S4A). Hypermethylation of the selenoprotein mRNA cap was affected in each case (Figure 3A and B). Cap hypermethylation was decreased in Nop58 depleted-cells by 60–80% (Figure 3A), but was totally abolished after siSMN treatment (Figure 3B). The steady state levels of selenoprotein mRNAs, Tgs1 and SBP2 mRNAs, as measured by qRT-PCR were not significantly affected in both cases (Figure 3C and D). As expected, Nop58 depletion strongly affected the stability of U3 snoRNA (Figure 3C) while SMN knock-down deregulated expression of U2 (Figure 3D). This is consistent with earlier reports showing that SMN deficiency or knock-down altered the stoichiometry of snRNAs in mammalian cells and could even increase their level, particularly in the case of U2 snRNA (49). We conclude that both Nop58 and SMN participate in hypermethylation of the selenoprotein mRNA cap. The strong effect of SMN inactivation suggests that recruitment of Tgs1 LF likely constitutes the major pathway.

SBP2 interacts with Nop58 and SMN *in vivo* and recruits Tgs1

Next, we asked how Tgs1 isoforms, SMN and Nop58 could be recruited to selenoprotein mRNAs. As SBP2 plays central roles in selenoprotein biosynthesis by binding to the selenoprotein mRNA SECIS element (26), we first tested if SBP2 could interact with Tgs1. Endogenous protein complexes associated with SBP2 were immunoprecipitated from HeLa cell extracts using antibodies against the N-terminal region of SBP2. Western blotting using anti-Tgs1 antibody revealed the association of SBP2 with endogenous Tgs1 LF, and little, if any, with Tgs1 SF (Figure 4A). No association was seen with the control protein Hsp70 (Figure 4A). To confirm this finding, we co-transfected SBP2 with either GFP-Tgs1 LF or GFP-Tgs1 SF and immunoprecipitated the total cell lysates with anti-GFP antibodies. As shown in Figure 4B, SBP2 associated with Tgs1 LF *in vivo*, while binding to Tgs1 SF was not detected. Y2H tests confirmed protein–protein interactions between SBP2 and both Tgs1 LF and Tgs1 SF (Figure 4C). Finally, *in vitro* binding assays

between ³⁵S-labeled SBP2 proteins expressed in micrococcal nuclease treated rabbit reticulocyte lysate (RRL) and the recombinant His-Tgs1 LF protein produced in *Escherichia coli*, showed quantitative binding of SBP2 to Tgs1 LF. However, little interaction was observed between His-SBP2 and ³⁵S-Tgs1 SF (Figure 4D). Similar results were obtained using ³⁵S-labeled proteins translated in RNase-treated bacterial S30 lysates (Figure 4D). Recombinant SBP2 protein preparations, produced from bacterial heterologous systems, do not contain snRNAs, snoRNAs and corresponding RNP proteins, nor components dedicated to synthesis of mammalian selenoprotein. This excludes the possibility that the protein–protein interactions shown are RNA dependent or mediated via protein components of the sn-, sno- and selenoprotein mRNP assembly machinery. Thus, Tgs1 LF and SBP2 associate *in vivo* and *in vitro* and this association is RNA independent. Because SMN and Nop58 interact with Tgs1 and appear to be required for selenoprotein mRNA cap-hypermethylation *in vivo* (Figure 3A and B), we next analyzed whether Nop58 and SMN also interacted with SBP2. SBP2 was co-transfected with Nop58-YFP in HEK293FT cells and we immunoprecipitated the total cell lysates with anti-GFP beads. As shown in Figure 4E, SBP2 interacted with Nop58 *in vivo*. We also verified the interaction *in vitro* by GST pull-down experiments (Figure 4F) and found that (³⁵S-Met)-SBP2 produced in RRL or bacterial S30 extracts bound strongly to GST-Nop58 independently of RNA. In addition, RNA-IPs in Nop58-YFP transfected cells revealed that GPx1 and GPx4 mRNAs were specifically associated with Nop58 *in vivo* (Supplementary Figure S5). To verify the link between SMN and selenoprotein mRNPs, we co-transfected GFP-SBP2 or GFP-SMN and SBP2 in HEK293FT cells for co-IP analysis. We found that GFP-SBP2 was able to interact with endogenous SMN and conversely that GFP-SMN interacted with transfected SBP2 *in vivo* (Figure 4G); these interactions were resistant to RNase A treatment and are therefore RNA independent (Figure 4G). GST pull-down experiments confirmed the interaction between (³⁵S-Met)-SBP2 and GST-SMN *in vitro* (Figure 4H). We conclude that SBP2 plays a central role by interacting with both SMN, Nop58 and Tgs1. The recruitment of Tgs1 is likely to be dependent on the formation of the ternary complexes between SBP2/SMN/Tgs1 LF on one side, and SBP2/Nop58/Tgs1 SF on the other.

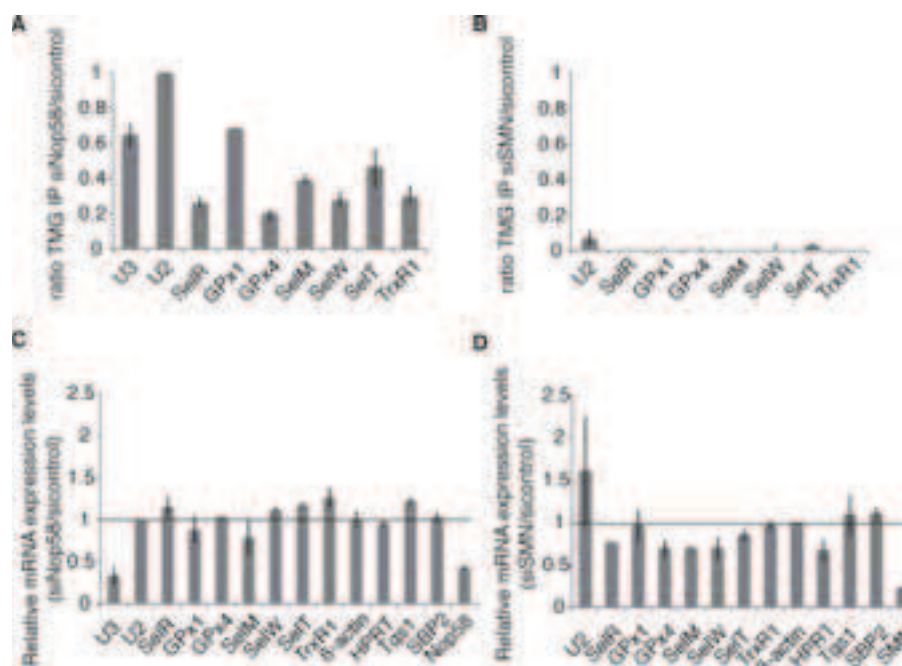


Figure 3. Nop58 and SMN contribute to selenoprotein mRNA hypermethylation. (A and B) Nop58 and SMN inactivation. RNA-IP using anti-TMG serum was performed as described in Figure 1 under siNop58 or siSMN and sicontrol conditions. IP ratios between siNop58, siSMN and sicontrol conditions are represented by the histogram bars and deduced from Supplementary Figures S3B and S4B, respectively. (C and D) siNop58 and siSMN have no effect on the steady state level of selenoprotein mRNAs nor on Tgs1 or SelP2 mRNAs. Relative expression levels were determined by qRT-PCR. U2 snRNA and β -actin were used as normalizers, respectively.

Hypermethylated-capped selenoprotein mRNAs localize to the cytoplasm and are polysome-associated

The discovery of hypermethylated-capped selenoprotein mRNAs raises the fundamental question of their ability to be present and translated in the cytoplasm. Indeed, since the TMG cap is a part of the nuclear localization signal for snRNAs, it could also be envisaged that hypermethylation leads to sequestration of selenoprotein mRNAs in the nucleus. We thus performed subcellular fractionation of HEK293 cells (Figure 5A) followed by TMG-IP experiments and determined the percentage of each TMG-capped mRNA in the cytoplasm compared with the nucleus (Figure 5B). To assess the quality of the nuclear-cytoplasmic fractions, we have performed western blot analysis using antibodies directed against the transcription factor ZNF143 (a strictly nuclear protein (50)) and the cytoplasmic ribosomal protein rpS21 (Figure 5A). Results showed that globally selenoprotein mRNAs are more abundant in the cytoplasmic than the nuclear compartment; indeed 70–84% of TMG-capped selenoprotein mRNAs are found in the cytoplasm (Figure 5B). As expected, both U3 snoRNA and U6 snRNA were predominantly immunoprecipitated from the nuclear fraction. U6 was used as a control because it does not exit the nucleus during biogenesis; although it is not TMG-capped, it is always recovered in TMG-IPs because of its interaction with U4 snRNA (51). Non-selenoprotein mRNA controls were not recognized by the anti-TMG antibody. These results strongly suggest that hypermethylated-capped selenoprotein mRNAs are present in the cytoplasm,

a prerequisite to their translation. To evaluate the ability of selenoprotein mRNAs to associate with actively translating ribosomes, we analyzed the polysome distribution of endogenous selenoprotein mRNAs. Cytoplasmic extracts of cycloheximide-treated HEK293 cells (blocking translation elongation) were fractionated on linear 7–47% sucrose gradients and the abundance of individual mRNAs in each fraction was measured by qRT-PCR (Figure 6A–C). The profiles revealed that all the endogenous selenoprotein mRNAs tested, except TrxR1, sedimented to fractions of lower molecular weight than non-selenoprotein mRNAs. The peak of selenoprotein mRNA population was found in fractions 10–26 in close proximity to the 80S monosomes (Figure 6B), whereas β -actin, HPRT and LDHA mRNAs sedimented in the heavier fractions 5–16 containing the polysomes (Figure 6C). As expected, U3 snoRNA was not found in polysomes (Figure 6C). These results indicate that fewer ribosomes are loaded on selenoprotein mRNAs, consistent with elongation pausing at the UGA Sec codon as previously reported for GPx4 (52). TrxR1 selenoprotein mRNA is the only mRNA found in heavier polysome fractions (see Figure 6B). This correlates with the fact that TrxR1 is the sole mRNA tested for which the UGA Sec codon is at the antepenultimate position before the stop codon; in all the other selenoprotein mRNAs of our test panel, the UGA Sec codon resides between 39 (SelW) and 285 nucleotides (SelR) downstream of the start codon. Worth of note, the proportion of free versus ribosome bound mRNAs is higher for selenoprotein than non-selenoprotein mRNAs. We next tested whether selenopro-

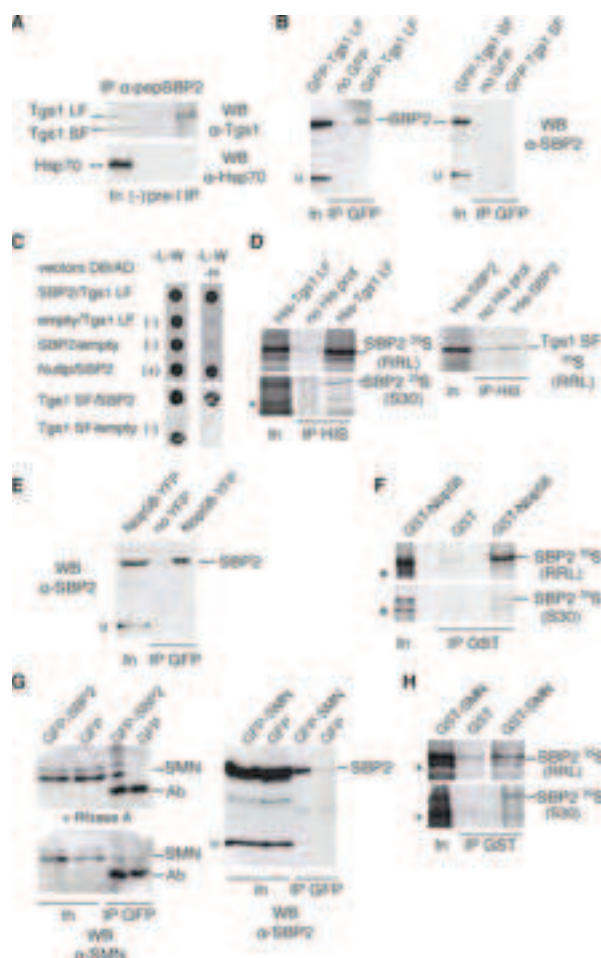


Figure 4. Tgs1 is recruited to selenoprotein mRNAs via interactions with SBP2. (A) Immunoprecipitation of endogenous SBP2 from HeLa cytoplasmic extracts using anti-peptide antibodies (α -pepSBP2) against amino acids 380–852. In: input 4%; (–) beads without antibodies; PreI: beads with preimmune serum; IP: beads with antibodies. (B) Co-immunoprecipitations using anti-GFP beads and HEK293FT cells transfected with SBP2 and GFP-Tgs1 or SBP2 alone (no GFP). (C) Y2H interaction tests performed in *S. cerevisiae* between SBP2 and Tgs1 LF or SF. (–) Controls with empty DNA binding (DB) or activation domain (AD) fusion vectors; (+) positive interaction control between SBP2 and Nup1. (D) Binding of recombinant His-Tgs1 to *in vitro* translated (^{35} S-Met)-SBP2 using rabbit reticulocyte lysate (RRL) and bacterial S30 extracts (S30). His-Tgs1 was bound to protein A-Sepharose using anti-His antibodies. (E) Co-immunoprecipitations using anti-GFP beads and HEK293FT cells transfected by SBP2 and Nop58-YFP. (F) Binding of recombinant GST-Nop58 to *in vitro* translated (^{35} S-Met)-SBP2 using RRL and S30 extracts. GST-Nop58 (or GST alone) was bound to glutathione beads and incubated with the extract. (G) Co-immunoprecipitations using anti-GFP beads and HEK293FT cells transfected by GFP-SBP2 or GFP-SMN and SBP2 without tag. (+ RNase A) co-immunoprecipitation performed in the presence of RNase A. (H) Binding of recombinant GST-SMN (or GST alone) to *in vitro* translated (^{35} S-Met)-SBP2 using RRL and S30 extracts. Asterisks denote frequently encountered hSBP2 proteolysis fragments. u represents non-specific protein signal revealed by the anti-SBP2 antibodies and that serves as an internal control. Ab: cross-reactivity signal from antibodies bound to the beads.

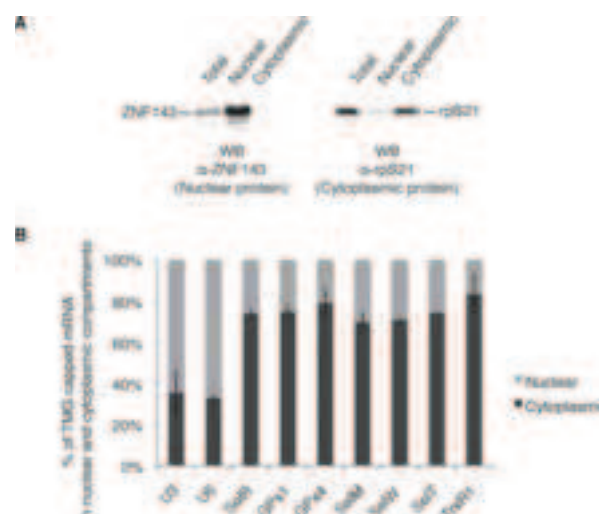


Figure 5. Hypermethylated selenoprotein mRNAs are more abundant in the cytoplasm. (A) Western blot analysis of the nuclear and cytoplasmic extracts using antibodies directed against the nuclear restricted ZNF143 transcription factor and the cytoplasmic rpS21. (B) RNA was extracted from total, nuclear and cytoplasmic extracts of HEK293FT cells; RNA-IPs using anti-TMG antibodies were performed and analyzed as described in Figure 1. The relative percentage of TMG-capped mRNA in both compartments is represented. Error bars represent standard deviation of an average of two independent experiments. U3 snRNA and U6 snRNA were used as controls for the quality of the fractionation.

tein mRNAs present in polysomes bear hypermethylated caps. TMG-IP experiments were performed on pooled fractions 6–26 (polysomes) or 28–41 (RNP) that contain free or non-polysome associated RNAs. We reasonably considered that the two pools contained 100% of the RNAs. Immunoprecipitation yields dropped importantly after fractionation of the RNAs on sucrose gradients; nevertheless, the results clearly showed the presence of the hypermethylated selenoprotein mRNAs in the polysome pool (Figure 6D). Indeed, between 50% (SelT) and 80% (SelR) of the recovered TMG-capped mRNAs was present in the polysome fraction, the rest of the TMG-capped selenoprotein mRNAs being in the RNP fractions (Figure 6D). As expected, U3 snoRNA and U2 snRNA were predominantly recovered from the RNP pool (Figure 6D). The non-selenoprotein mRNA controls were not recognized by anti-TMG antibodies (see also Figure 1C). When RNA fractionation was performed under low magnesium concentration leading to ribosome dissociation into subunits (Figure 6A), the signal of selenoprotein mRNA TMG-IP was shifted to the RNP fractions (Figure 6E). Altogether these results show that hypermethylated selenoprotein mRNAs are found in polysomes and are therefore translated.

Tgs1 activity, but not eIF4E, is required *in vivo* for translation of GPx1

In order to determine the functional importance of the TMG-cap modification by Tgs1 on selenoprotein synthesis *in vivo*, we constructed two stable cell lines capable of expressing HA tagged wild-type (wt) and mutant GPx1 se-

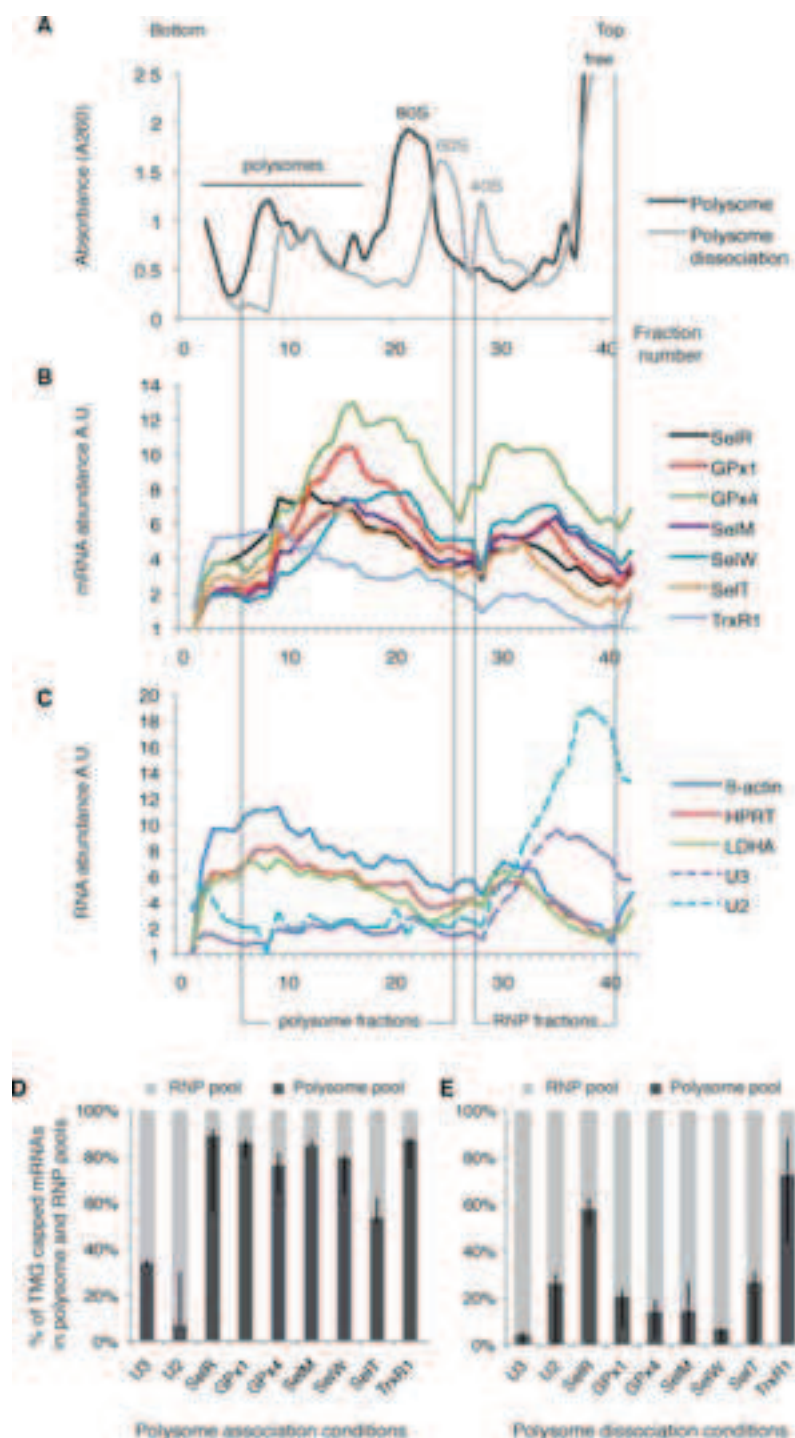


Figure 6. Polysome distribution of endogenous HEK293 selenoprotein mRNAs. (A) Cytoplasmic extracts from HEK293 cell were fractionated onto 7–47% (w/v) linear sucrose gradient and collected in 40 fractions. Typical absorbance profiles are shown and the positions of the polysomes, 80S, 60S, 40S ribosomal subunits as well as free RNA are indicated. Fractionation was performed in polysome association (black profile) or low magnesium dissociation conditions (gray profile). (B and C) The RNA content of each fraction was analyzed by qRT-PCR and the relative mRNA abundance was represented in arbitrary units A.U. Vertical bars indicate the position of the polysome and RNP fractions that were pooled and analyzed in (D). (B) Sedimentation profiles of selenoprotein mRNAs; (C) non-selenoprotein mRNAs, U3 snoRNA and U2 snRNA are represented. (D and E) Hypermethylated selenoprotein mRNAs co-fractionate with polysomes. RNA-IP using anti-TMG antibodies and qRT-PCR analysis was performed as described in Figure 1 in polysome association (D) and dissociation conditions (E). The amount of RNA immunoprecipitated from the polysome and RNP pool were determined separately by qRT-PCR and normalized to 100%. Error bars represent standard deviation of an average of two independent experiments.

lenoproteins. GPx1 was chosen as a model of class 1 selenoprotein mRNAs which bear a TMG cap and are poorly recognized by eIF4E. The mRNAs contained the complete natural 5' and 3'UTRs and were under the control of a tetracycline inducible promoter (HA-GPx1 and HA-GPx1Cys, respectively). In HA-GPx1 cell lines, synthesis of GPx1 was selenium dependent (Supplementary Figure S6) and relied on translational recoding. In the HA-GPx1Cys mutant, the Sec codon was mutated to Cys. Expression of this construct is thus independent of the recoding mechanism. Translation of GPx1Cys was indeed unaffected by the selenium level (Supplementary Figure S6); it was used as a non-selenoprotein control. This system allowed us to analyze the effects of the depletion of biogenesis or translation factors on *de novo* synthesis of the GPx1 selenoprotein. We first inhibited expression of Tgs1 by siRNA using pools of four different, non-overlapping, siRNAs and two different siRNA conditions. After 48 h of siRNA treatment, the level of Tgs1 mRNA was reduced down to 52% (siRNA2) and the corresponding Tgs1 protein levels to 62% (Figure 7A). At this stage, expression of HA-GPx1 was selectively induced for an additional 12 h and we monitored the impact of Tgs1 depletion on *de novo* selenoprotein synthesis (Figure 7B). Partial Tgs1 depletion resulted in a statistically significant 47% drop of HA-GPx1 selenoprotein synthesis compared to the endogenous control protein AspRS that remained unaffected (Figure 7B). With a similar efficacy of siRNA treatments, the drop in protein synthesis was 66% in the case of SBP2 knock-down, a protein factor essential to the Sec incorporation process *in vivo* and no effect was observed in samples treated with control siRNAs (Figure 7B). Both the siTgs1 and siSBP2 effects could be rescued by transfection of SBP2 and Tgs1 expression plasmids 24 h after knock-down. Restoration of Tgs1 and SBP2 wt levels (Figure 7A) re-established 89 and 100% of HA-GPx1 expression, respectively (Figure 7B). Interestingly, depletion of neither Tgs1 nor SBP2 affected translation of the HA-GPx1Cys mutant (Figure 7B). In this case, synthesis of the HA-GPx1Cys protein does not rely on translational recoding events and becomes independent of the presence of SBP2, in agreement with depletion of SBP2 having no effect on protein production arising from this construct (Figure 7B). Altogether these results suggest that Tgs1 does play a functional role in selenoprotein synthesis *in vivo*, most likely by its direct action on selenoprotein mRNA cap hypermethylation.

Because we showed that hypermethylation affected the recognition by eIF4E, we also examined the effects of eIF4E knock-down, using pools of four different siRNAs, on the *in vivo* expression of GPx1 using the same HA-GPx1 and HA-GPx1Cys inducible stable cell lines (Figure 7). eIF4E mRNAs were reduced to 58%. In this case, a 40% depletion of the eIF4E protein (Figure 7A) hardly affected HA-GPx1 selenoprotein synthesis, in agreement with our data showing that eIF4E only poorly recognized the GPx1 mRNA. In contrast, translation of the HA-GPx1Cys mutant was more affected and reduced between 23% (siRNA1) and 30% (siRNA2) (statistically significant, $P = 0.0002$ and $P = 0.02$ for siRNA1 and siRNA2, respectively) (Figure 7B). In this case, protein synthesis is independent of UGA Sec recoding and relies on canonical translation mechanisms. The results of eIF4E inactivation show opposite effects compared

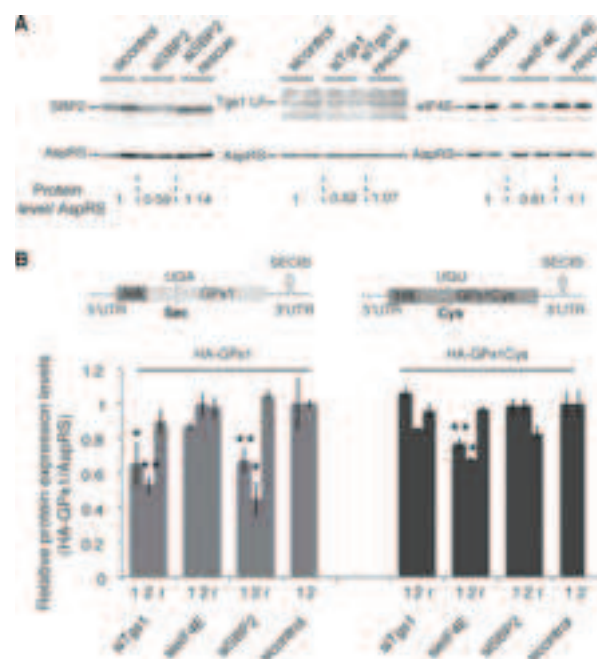


Figure 7. Tgs1, but not eIF4E, is required for *de novo* synthesis of GPx1. Stable cell lines expressing HA-GPx1 and HA-GPx1Cys mutant proteins under the control of an inducible promoter were used to monitor *de novo* selenoprotein synthesis following 48 h of Tgs1, eIF4E and SBP2 knock-down (siTgs1, siEIF4E and siSBP2). Two different siRNA conditions using 30 and 100 nM siRNA (marked 1 and 2 below the histogram) were analyzed for each gene. r indicates rescue experiments of siRNA2 conditions by transfection of a plasmid 24 h after siRNA. (A) Western blot analysis of siTgs1, siEIF4E and siSBP2 efficiency using antibodies directed against the targeted proteins. Levels of protein compared with the control are indicated below the panels and normalized against AspRS. (B) Quantification results of *de novo* HA-GPx1 protein expression levels by western blot analysis using anti-HA antibodies. Normalization was performed using AspRS antibodies. Gray bars: HA-GPx1; dark bars: HA-GPx1Cys. Error bars represent standard deviation of an average of four independent experiments. Asterisks indicate statistically significant differences between siTgs1, siEIF4E, siSBP2, rescues and corresponding siRNA control conditions. * $P < 0.05$ and ** $P < 0.005$ based on Student's *t*-test. Expression constructs HA-GPx1 and HA-GPx1Cys, integrated in the stable cell lines, flanked by their natural GPx1 5' and 3'UTRs as well as SECIS RNA, are represented above the histograms.

to the siTgs1 and siSBP2 control conditions that solely impacted UGA recoding (Figure 7B). The mild effect observed is consistent with the existence of strong homeostatic control mechanisms of eIF4E activity; indeed knockdowns have been reported to cause only minor reduction in translation (53). Rescue experiments of siEIF4E (Figure 7A) restored 98% of HA-GPx1Cys expression (Figure 7B). These results suggest that eIF4E is not essential for the translation of hypermethylated-capped selenoprotein mRNAs.

DISCUSSION

Cellular mRNAs are known to be m⁷G-capped (1). Here we showed that mammalian mRNAs encoding selenoproteins bear a hypermethylated cap. Some mRNAs in *Caenorhabditis elegans*, *Ascaris lumbricoides*, *Ciona intestinalis* and *Euglena gracilis* acquire TMG caps but by spliced-leader

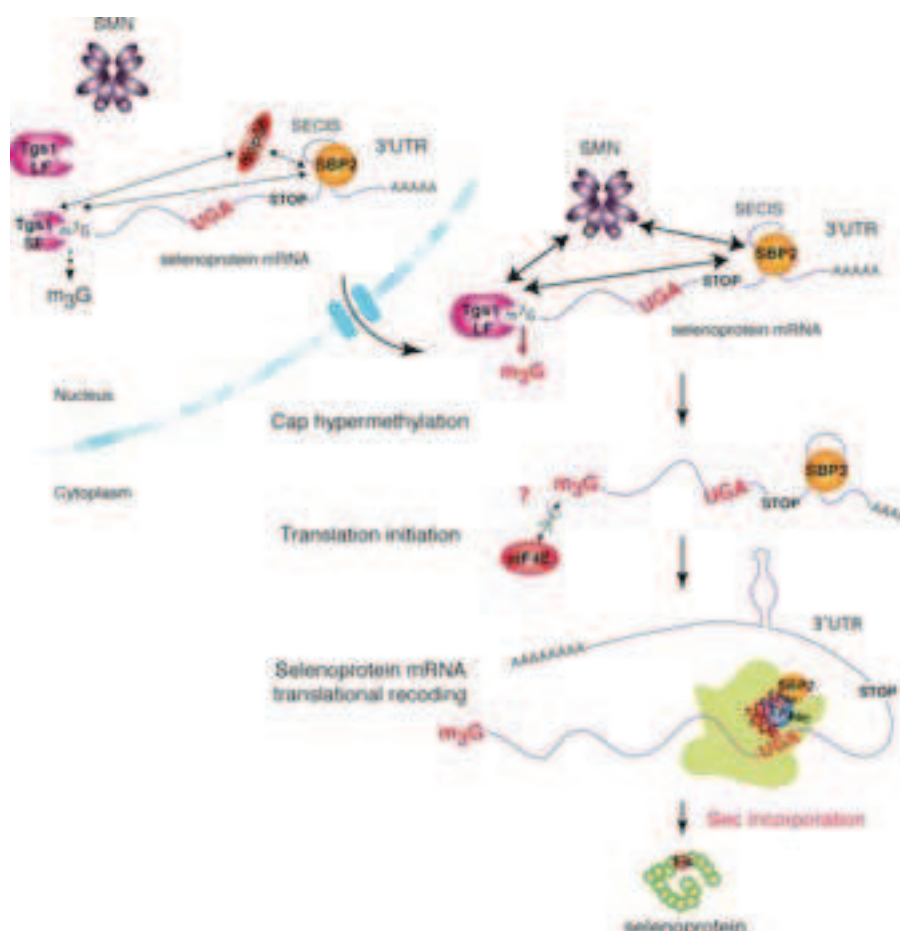


Figure 8. Model for cap hypermethylation and translation of selenoprotein mRNPs. The caps of selenoprotein mRNAs are hypermethylated by Tgs1. Tgs1 isoforms (Tgs1 SF and Tgs1 LF) interact with SBP2 and selenoprotein mRNAs and are recruited predominantly via the ternary complex SBP2/SMN/Tgs1 LF (thick arrows) but also via SBP2/Nop58/Tgs1 SF (thin arrows). Hypermethylated-capped selenoprotein mRNAs are localized to the cytoplasm and are polysome-associated but show reduced affinity for eIF4E. Tgs1 activity is required for Sec incorporation or translation initiation.

(SL) trans-splicing, thus a radically different pathway than described here (54–57). Some Rev/RRE-dependent HIV-1 RNAs can also be TMG-capped by Tgs1, TMG capping representing in this case a regulation mechanism for selective expression (58). To assess the impact of TMG capping on selenoprotein mRNA export, stability and translation is challenging. Hypermethylation of selenoprotein mRNA cap may serve different purposes: dedicated assembly pathway for selenoprotein mRNPs and/or control of selenoprotein expression at the level of translation initiation or UGA Sec recoding. Our experiments reveal that selenoprotein mRNAs are subjected to differential 5' processing and translational initiation events. Although some are substrates for Tgs1, harbor hypermethylated caps and are not recognized efficiently by the translation factor eIF4E, others are m⁷G-capped and recognized by eIF4E. These differences between selenoprotein mRNAs are not surprising as differential expression and stability regulation mechanisms have been reported for various subsets of selenoprotein mRNAs under conditions as different as selenium levels, type of SBP2–SECIS interactions, and sensitivity to

NMD (nonsense mediated decay) (44,59–63). Control of gene-specific selenoprotein expression and differential regulation of UGA Sec recoding of a subset of selenoprotein mRNAs were recently shown to be selenium-dependent and mediated by the degree of tRNA^{[Ser]Sec} Um34 methylation (43). Likewise, it is conceivable that cap hypermethylation could also be part of a regulation process of selenoprotein expression. As suggested by our results, hypermethylation does not necessarily occur equally on all selenoprotein mRNAs, possibly regulating translation initiation in a way that could contribute to the mechanisms of hierarchy of selenoprotein expression. Mechanisms and factors influencing the fate of selenoprotein mRNAs are far from being understood and so far no unified mechanism could be proposed for all selenoprotein mRNAs, at any stage of their mRNA life cycle. Our findings point to the existence of an additional level of complexity in the mechanism of selenoprotein synthesis residing at the stage of mRNP assembly, processing and translation initiation (Figure 8). This process is governed by the interactions between SBP2, Tgs1 and selenoprotein mRNAs. We established that the hypermethylase

Tgs1 is recruited to selenoprotein mRNAs predominantly via SMN but also Nop58, and therefore probably belongs to two possible ternary complexes: SBP2/Nop58/Tgs1 or SBP2/SMN/Tgs1 (Figure 8). Interestingly, oxidative stress was identified as a novel regulation mechanism for selective selenoprotein synthesis (64). It was shown to induce nuclear shuttling of SBP2 and selenocysteine incorporation factors, and to promote assembly of selenoprotein mRNPs (32,33). It is conceivable that oxidative stress could also regulate hypermethylation of selenoprotein mRNAs by stimulating recruitment of the nuclear isoform of Tgs1.

SBP2 is a member of the L7Ae family of RNA-binding proteins (65) that include the archaeal L7Ae and eukaryotic 15.5 kDa core proteins of the box C/D sRNPs and snoRNPs, respectively. We have previously shown that the structural motifs of the SECIS and box C/D RNAs as well as binding strategies governing the interaction of SBP2, L7Ae and 15.5 kDa to their cognate RNAs are similar and are controlled by a common conserved RNA chaperone complex (27,34,66). Interestingly, it is the L7Ae and 15.5 kDa proteins that play a key role in box C/D RNP assembly in Archaea and Eukarya, respectively, by specifically recognizing their cognate RNA motif. Moreover, in the case of U3 snoRNA, assembly and maturation was shown to be dependent on a large multiprotein complex that contains, in addition to the core proteins, components of the RNA chaperone complex but also RNA processing factors including Tgs1 (22,25). Likewise, our present data suggest that SBP2 could trigger the assembly of the cap modification complex by recruiting Nop58 or SMN in a similar manner, and this would ultimately lead to the binding of Tgs1 methylase for modification. Tgs1 SF is strictly nuclear but Tgs1 LF is present both in the nucleus and the cytoplasm. This suggests that m⁷G hypermethylation could occur in either of the two cellular compartments via two alternative pathways and could obey a fast and potentially reversible regulation process (Figure 8). Altogether, our results demonstrate that selenoprotein mRNAs indeed undergo a sophisticated assembly and 5' end maturation pathway similar to that of sn- and snoRNAs, and which involve common maturation factors and core proteins.

However, unlike TMG-capped sn- or snoRNAs, our results also show that hypermethylated-capped selenoprotein mRNAs can be found in the cytoplasm and in actively translating ribosomes, excluding nuclear retention as a consequence of cap hypermethylation. Hypermethylation by Tgs1 rather seems to be required for translational recoding events, as its knock-down reduced expression of GPx1. Our current data do not allow us to differentiate whether Tgs1 activity is required for Sec incorporation or translation initiation (Figure 8). It is conceivable that both mechanisms could be linked for optimal selenoprotein translation. Specialized translation initiation factors or alternative mechanisms are therefore likely to be involved in hypermethylated mRNA recognition. One possibility would be an IRES-dependent translation. Consistently, earlier reports showed that the cricket paralysis virus internal ribosome entry site is able to support Sec incorporation in a luciferase reporter construct, although the incorporation efficiency was decreased (67). Our results also showed a reduced affinity of the translation initiation factor eIF4E for

selenoprotein mRNA cap structures. However, eIF4E variants that can accommodate the TMG cap were identified in the genome of *C. elegans* (68). Whether such a situation prevails for some selenoprotein mRNAs will be the aim of future investigations.

ACCESSION NUMBER

NCBI Gene Expression Omnibus accession number: GSE57625.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGMENTS

We thank R. Lührmann for the gift of HeLa cell extracts and antibodies, C.H. Hagedorn for the pGST-eIF4E plasmid, A. Schweigert for technical assistance, E. Westhof for invaluable input and constant support, P. Carbon and G. Eriani for helpful discussions and critical reading of the manuscript and J. Steitz for useful comments and critical reading of the manuscript.

Authors contributions: L.W., A.S.G.B., C.V., A.K., E.B. and C.A. designed research. A.S.G.B., L.W., C.V., M.L., A.T., S.B., F.M. and C.A. performed research. L.W., A.S.G.B., C.V., F.M., A.K., E.B. and C.A. analyzed data. C.A. wrote the paper.

FUNDING

Centre national de la recherche scientifique (CNRS); Agence Nationale de la Recherche [BLAN06-2.134682, ANR-2011-svse8 02501 to F.M.]. French Ministry of Research and the National Research Fund, Luxembourg [to L.W.]; Contract Doctoral of the University of Strasbourg [A.S.G.]. Funding for open access charge: CNRS.

Conflict of interest statement. None declared.

REFERENCES

- Shatkin, A.J. (1976) Capping of eucaryotic mRNAs. *Cell*, **9**, 645–653.
- Topisirovic, I., Svitkin, Y.V., Sonenberg, N. and Shatkin, A.J. (2011) Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev. RNA*, **2**, 277–298.
- Izaurrealde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C. and Mattaj, I.W. (1995) A cap-binding protein complex mediating U snRNA export. *Nature*, **376**, 709–712.
- Kohler, A. and Hurt, E. (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell. Biol.*, **8**, 761–773.
- Mitchell, P. and Tollervey, D. (2001) mRNA turnover. *Curr. Opin. Cell Biol.*, **13**, 320–325.
- Wilusz, C.J., Wormington, M. and Peltz, S.W. (2001) The cap-to-tail guide to mRNA turnover. *Nat. Rev. Mol. Cell. Biol.*, **2**, 237–246.
- Houseley, J. and Tollervey, D. (2009) The many pathways of RNA degradation. *Cell*, **136**, 763–776.
- Fortes, P., Inada, T., Preiss, T., Hentze, M.W., Mattaj, I.W. and Sachs, A.B. (2000) The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. *Mol. Cell*, **6**, 191–196.
- McKendrick, L., Thompson, E., Ferreira, J., Morley, S.J. and Lewis, J.D. (2001) Interaction of eukaryotic translation initiation factor 4G with the nuclear cap-binding complex provides a link between nuclear and cytoplasmic functions of the m(7) guanosine cap. *Mol. Cell. Biol.*, **21**, 3632–3641.

10. Sonenberg, N. (2008) eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. *Biochem. Cell Biol.*, **86**, 178–183.
11. Calero, G., Wilson, K.F., Ly, T., Rios-Steiner, J.L., Clardy, J.C. and Cerione, R.A. (2002) Structural basis of m⁷GpppG binding to the nuclear cap-binding protein complex. *Nat. Struct. Biol.*, **9**, 912–917.
12. 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.
13. Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E. and Mattaj, J.W. (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell*, **78**, 657–668.
14. Darzynkiewicz, E., Stepinski, J., Ekiel, I., Jin, Y., Haber, D., Sijuwade, T. and Tahara, S.M. (1988) Beta-globin mRNAs capped with m⁷G, m².7(2)G or m².2.7(3)G differ in intrinsic translation efficiency. *Nucleic Acids Res.*, **16**, 8953–8962.
15. Matera, A.G., Terns, R.M. and Terns, M.P. (2007) Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell. Biol.*, **8**, 209–220.
16. Chari, A., Paknia, E. and Fischer, U. (2009) The role of RNP biogenesis in spinal muscular atrophy. *Curr. Opin. Cell Biol.*, **21**, 387–393.
17. Fischer, U. and Luhrmann, R. (1990) An essential signaling role for the m³G cap in the transport of U1 snRNP to the nucleus. *Science*, **249**, 786–790.
18. Fischer, U., Sumpster, V., Sekine, M., Satoh, T. and Luhrmann, R. (1993) Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m³G cap. *EMBO J.*, **12**, 573–583.
19. Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M. and Luhrmann, R. (1998) Snurportin1, an m³G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J.*, **17**, 4114–4126.
20. Hamm, J., Darzynkiewicz, E., Tahara, S.M. and Mattaj, J.W. (1990) The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. *Cell*, **62**, 569–577.
21. Verheggen, C., Lafontaine, D.L., Samarsky, D., Mouaikel, J., Blanchard, J.M., Bordonne, R. and Bertrand, E. (2002) Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *EMBO J.*, **21**, 2736–2745.
22. Boulon, S., Verheggen, C., Jady, B.E., Girard, C., Pesca, C., Paul, C., Ospina, J.K., Kiss, T., Matera, A.G., Bordonne, R. et al. (2004) PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Mol. Cell*, **16**, 777–787.
23. Girard, C., Verheggen, C., Neel, H., Cammas, A., Vagner, S., Soret, J., Bertrand, E. and Bordonne, R. (2008) Characterization of a short isoform of human Tgs1 hypermethylase associating with small nucleolar ribonucleoprotein core proteins and produced by limited proteolytic processing. *J. Biol. Chem.*, **283**, 2060–2069.
24. Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J. and Bordonne, R. (2002) Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell*, **9**, 891–901.
25. Watkins, N.J., Lemm, I., Ingelfinger, D., Schneider, C., Hobach, M., Urlaub, H. and Luhrmann, R. (2004) Assembly and maturation of the U3 snRNP in the nucleoplasm in a large dynamic multiprotein complex. *Mol. Cell*, **16**, 789–798.
26. Allmang, C., Wurth, L. and Krol, A. (2009) The selenium to selenoprotein pathway in eukaryotes: more molecular partners than anticipated. *Biochim. Biophys. Acta*, **1790**, 1415–1423.
27. Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jady, B.E., Rothe, B., Pesca, C., Robert, M.C., Kiss, T. et al. (2008) The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J. Cell Biol.*, **180**, 579–595.
28. Copeland, P.R., Fletcher, J.E., Carlson, B.A., Hatfield, D.L. and Driscoll, D.M. (2000) A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J.*, **19**, 306–314.
29. Lescure, A., Allmang, C., Yamada, K., Carbon, P. and Krol, A. (2002) cDNA cloning, expression pattern and RNA binding analysis of human selenocysteine insertion sequence (SECIS) binding protein 2. *Gene*, **291**, 279–285.
30. Takeuchi, A., Schmitt, D., Chapple, C., Babaylova, E., Karpova, G., Guigo, R., Krol, A. and Allmang, C. (2009) A short motif in Drosophila SECIS Binding Protein 2 provides differential binding affinity to SECIS RNA hairpins. *Nucleic Acids Res.*, **37**, 2126–2141.
31. Kinzy, S.A., Caban, K. and Copeland, P.R. (2005) Characterization of the SECIS binding protein 2 complex required for the co-translational insertion of selenocysteine in mammals. *Nucleic Acids Res.*, **33**, 5172–5180.
32. Papp, L.V., Lu, J., Striebel, F., Kennedy, D., Holmgren, A. and Khanna, K.K. (2006) The redox state of SECIS binding protein 2 controls its localization and selenocysteine incorporation function. *Mol. Cell. Biol.*, **26**, 4895–4910.
33. de Jesus, L.A., Hoffmann, P.R., Michaud, T., Forry, E.P., Small-Howard, A., Stillwell, R.J., Morozova, N., Harney, J.W. and Berry, M.J. (2006) Nuclear assembly of UGA decoding complexes on selenoprotein mRNAs: a mechanism for eluding nonsense-mediated decay? *Mol. Cell. Biol.*, **26**, 1795–1805.
34. Clery, A., Bourguignon-Igel, V., Allmang, C., Krol, A. and Branlant, C. (2007) An improved definition of the RNA-binding specificity of SECIS-binding protein 2, an essential component of the selenocysteine incorporation machinery. *Nucleic Acids Res.*, **35**, 1868–1884.
35. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.*, **11**, 1475–1489.
36. Frugier, M., Ryckelynck, M. and Giege, R. (2005) tRNA-balanced expression of a eukaryal aminoacyl-tRNA synthetase by an mRNA-mediated pathway. *EMBO Rep.*, **6**, 860–865.
37. Pradet-Balade, B., Girard, C., Boulon, S., Paul, C., Azzag, K., Bordonne, R., Bertrand, E. and Verheggen, C. (2011) CRM1 controls the composition of nucleoplasmic pre-snoRNA complexes to licence them for nucleolar transport. *EMBO J.*, **30**, 2205–2218.
38. Luhrmann, R., Appel, B., Bringmann, P., Rinke, J., Reuter, R., Rothe, S. and Bald, R. (1982) Isolation and characterization of rabbit anti-m³ 2,2,7G antibodies. *Nucleic Acids Res.*, **10**, 7103–7113.
39. Tycowski, K.T., Aab, A. and Steitz, J.A. (2004) Guide RNAs with 5' caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa. *Curr. Biol.*, **14**, 1985–1995.
40. Choi, Y.H. and Hagedorn, C.H. (2003) Purifying mRNAs with a high-affinity eIF4E mutant identifies the short 3' poly(A) end phenotype. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 7033–7038.
41. Marcotrigiano, J., Gingras, A.C., Sonenberg, N. and Burley, S.K. (1997) X-ray studies of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Nucleic Acids Symp. Ser.*, **8**, 1–11.
42. Kryukov, G.V., Castellano, S., Novoselov, S.V., Lobanov, A.V., Zehab, O., Guigo, R. and Gladyshev, V.N. (2003) Characterization of mammalian selenoproteomes. *Science*, **300**, 1439–1443.
43. Howard, M.T., Carlson, B.A., Anderson, C.B. and Hatfield, D.L. (2013) Translational redefinition of UGA codons is regulated by selenium availability. *J. Biol. Chem.*, **288**, 19401–19413.
44. Sun, X., Li, X., Moriarty, P.M., Henics, T., LaDuca, J.P. and Maquat, L.E. (2001) Nonsense-mediated decay of mRNA for the selenoprotein phospholipid hydroperoxide glutathione peroxidase is detectable in cultured cells but masked or inhibited in rat tissues. *Mol. Biol. Cell*, **12**, 1009–1017.
45. Weiss Sachdev, S. and Sunde, R.A. (2001) Selenium regulation of transcript abundance and translational efficiency of glutathione peroxidase-1 and -4 in rat liver. *Biochem. J.*, **357**, 851–858.
46. Moriarty, P.M., Reddy, C.C. and Maquat, L.E. (1998) Selenium deficiency reduces the abundance of mRNA for Se-dependent glutathione peroxidase 1 by a UGA-dependent mechanism likely to be nonsense codon-mediated decay of cytoplasmic mRNA. *Mol. Cell. Biol.*, **18**, 2932–2939.
47. Paushkin, S., Gubitz, A.K., Massenot, S. and Dreyfuss, G. (2002) The SMN complex, an assemblyosome of ribonucleoproteins. *Curr. Opin. Cell Biol.*, **14**, 305–312.
48. Mouaikel, J., Narayanan, U., Verheggen, C., Matera, A.G., Bertrand, E., Tazi, J. and Bordonne, R. (2003) Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Rep.*, **4**, 616–622.
49. Zhang, Z., Lotti, F., Dittmar, K., Younis, I., Wan, L., Kasim, M. and Dreyfuss, G. (2008) SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell*, **133**, 585–600.

50. Myslinski, E., Krol, A. and Carbon, P. (1998) ZNF76 and ZNF143 are two human homologs of the transcriptional activator Staf. *J. Biol. Chem.*, **273**, 21998–22006.
51. Mottram, J., Perry, K.L., Lizardi, P.M., Luhrmann, R., Agabian, N. and Nelson, R.G. (1989) Isolation and sequence of four small nuclear U RNA genes of *Trypanosoma brucei* subsp. *brucei*: identification of the U2, U4, and U6 RNA analogs. *Mol. Cell. Biol.*, **9**, 1212–1223.
52. Fletcher, J.E., Copeland, P.R. and Driscoll, D.M. (2000) Polysome distribution of phospholipid hydroperoxide glutathione peroxidase mRNA: evidence for a block in elongation at the UGA/selenocysteine codon. *RNA*, **6**, 1573–1584.
53. Yanagiya, A., Suyama, E., Adachi, H., Svitkin, Y.V., Aza-Blanc, P., Imataka, H., Mikami, S., Martineau, Y., Ronai, Z.A. and Sonenberg, N. (2012) Translational homeostasis via the mRNA cap-binding protein, eIF4E. *Mol. Cell*, **46**, 847–858.
54. Liou, R.F. and Blumenthal, T. (1990) trans-spliced *Caenorhabditis elegans* mRNAs retain trimethylguanosine caps. *Mol. Cell. Biol.*, **10**, 1764–1768.
55. Van Doren, K. and Hirsh, D. (1990) mRNAs that mature through trans-splicing in *Caenorhabditis elegans* have a trimethylguanosine cap at their 5' termini. *Mol. Cell. Biol.*, **10**, 1769–1772.
56. Maroney, P.A., Denker, J.A., Darzynkiewicz, E., Laneve, R. and Nilsen, T.W. (1995) Most mRNAs in the nematode *Ascaris lumbricoides* are trans-spliced: a role for spliced leader addition in translational efficiency. *RNA*, **1**, 714–723.
57. Hastings, K.E. (2005) SL trans-splicing: easy come or easy go? *Trends Genet.*, **21**, 240–247.
58. Yedavalli, V.S. and Jeang, K.T. (2010) Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 14787–14792.
59. Lei, X.G., Evenson, J.K., Thompson, K.M. and Sunde, R.A. (1995) Glutathione peroxidase and phospholipid hydroperoxide glutathione peroxidase are differentially regulated in rats by dietary selenium. *J. Nutr.*, **125**, 1438–1446.
60. Low, S.C., Grundner-Culemann, E., Harney, J.W. and Berry, M.J. (2000) SECIS-SBP2 interactions dictate selenocysteine incorporation efficiency and selenoprotein hierarchy. *EMBO J.*, **19**, 6882–6890.
61. Squires, J.E., Stoytchev, I., Forry, E.P. and Berry, M.J. (2007) SBP2 binding affinity is a major determinant in differential selenoprotein mRNA translation and sensitivity to nonsense-mediated decay. *Mol. Cell. Biol.*, **27**, 7848–7855.
62. Budiman, M.E., Bubenik, J.L., Miniard, A.C., Middleton, L.M., Gerber, C.A., Cash, A. and Driscoll, D.M. (2009) Eukaryotic initiation factor 4a3 is a selenium-regulated RNA-binding protein that selectively inhibits selenocysteine incorporation. *Mol. Cell*, **35**, 479–489.
63. Latreche, L., Duhieu, S., Touat-Hamici, Z., Jean-Jean, O. and Chavatte, L. (2012) The differential expression of glutathione peroxidase 1 and 4 depends on the nature of the SECIS element. *RNA Biol.*, **9**, 681–690.
64. Touat-Hamici, Z., Legrain, Y., Bulteau, A.L. and Chavatte, L. (2014) Selective up-regulation of human selenoproteins in response to oxidative stress. *J. Biol. Chem.*, **289**, 14750–14761.
65. Koonin, E.V., Bork, P. and Sander, C. (1994) A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme? *Nucleic Acids Res.*, **22**, 2166–2167.
66. Allmang, C., Carbon, P. and Krol, A. (2002) The SBP2 and 15.5 kD/Snu13p proteins share the same RNA binding domain: identification of SBP2 amino acids important to SECIS RNA binding. *RNA*, **8**, 1308–1318.
67. Donovan, J. and Copeland, P.R. (2010) The efficiency of selenocysteine incorporation is regulated by translation initiation factors. *J. Mol. Biol.*, **400**, 659–664.
68. Keiper, B.D., Lamphear, B.J., Deshpande, A.M., Jankowska-Anyszka, M., Aamodt, E.J., Blumenthal, T. and Rhoads, R.E. (2000) Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans*. *J. Biol. Chem.*, **275**, 10590–10596.

Figures supplémentaires

Nucleic Acids Research, 2014, Vol. 42, No. 13

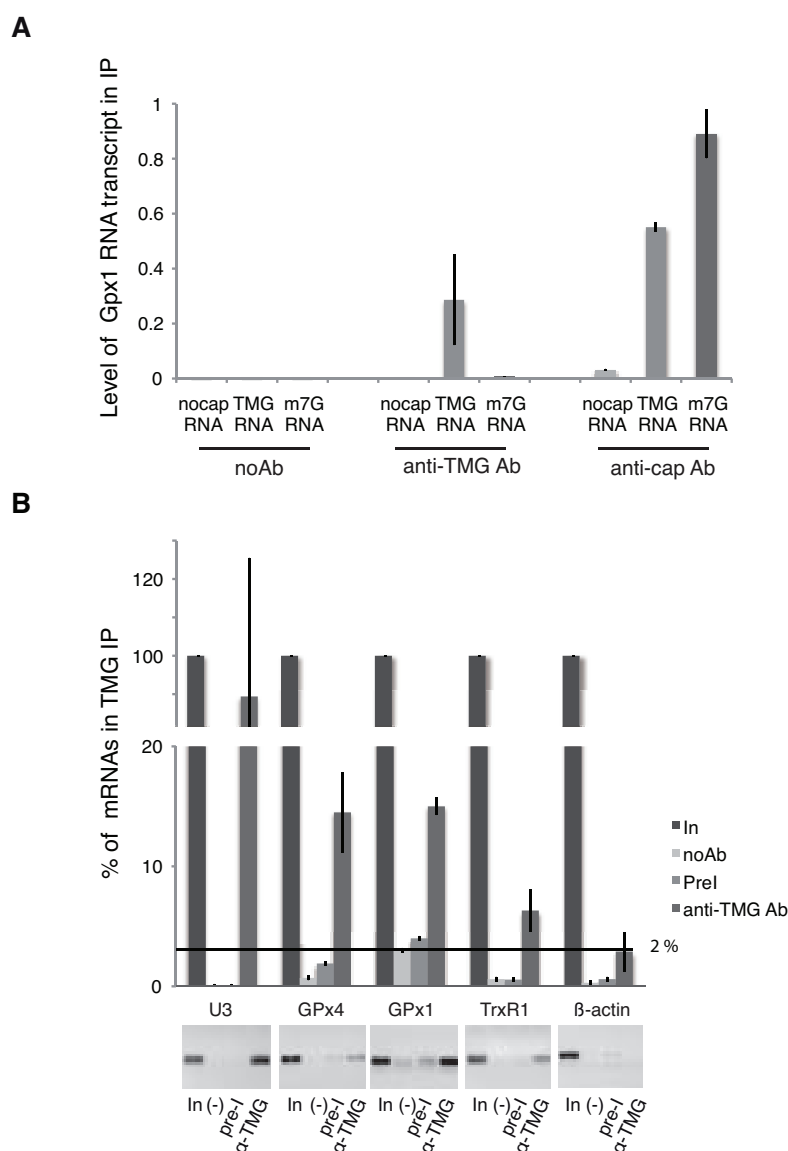


Figure S1: Validation of the anti-TMG serum specificity for TMG-capped RNAs. (A) In vitro transcribed GPx1 RNAs, capped co-transcriptionally with either m3G (TMG RNA) or m7G (m7G RNA) cap analogs were used to test the specificity of TMG-IPs. Non-capped RNAs (nocap RNA) were used as negative controls. Immunoprecipitations were performed with no antibodies, specific anti-TMG serum (anti-m3G R1131 serum, Synaptic Systems) and anti-cap antibodies (anti-m7(m3)G antibodies (Synaptic Systems)). Bound GPx1 RNA transcripts were analyzed by qRT-PCR. Maximum RNA binding levels were set to 1. RNA IPs reveal that the anti-TMG serum specifically recognizes TMG-capped RNA but not m7G-capped RNA transcripts while anti-cap antibodies recognize both RNAs. (B) Example total RNA TMG-IP quality control. Total RNA extracted from HEK293FT cells was immunoprecipitated with anti-TMG serum (anti-TMG Ab), no antibodies (noAb or (-)) or unrelated rabbit pre-immune serum (PreI). Bound RNA was analyzed by qRT-PCR. The graphs represents the % of mRNAs in IP compared to the input RNA. The horizontal line represents the level of housekeeping mRNA binding and background (1-2% in average). Background values due to beads or pre-immune serum are retrieved from specific binding values in all subsequent representations. (A, B) Error bars represent standard deviation of an average of 3 independent experiments. Panels below the graph show the RT-PCR products obtained after 35 PCR cycles. In: input 10 %.

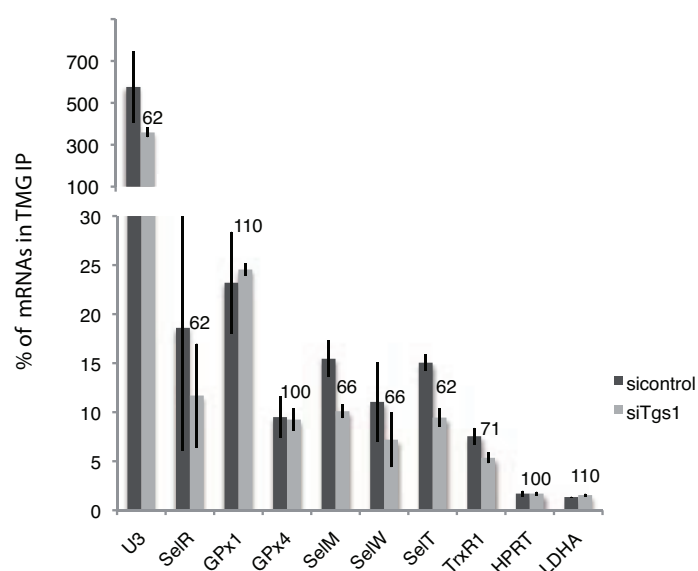


Figure S2: Tgs1 is required for selenoprotein mRNA cap hypermethylation. RNA-IP using anti-TMG serum was performed as described in Fig. 1. The housekeeping HPRT and LDHA mRNAs are negative controls for the TMG-IP.

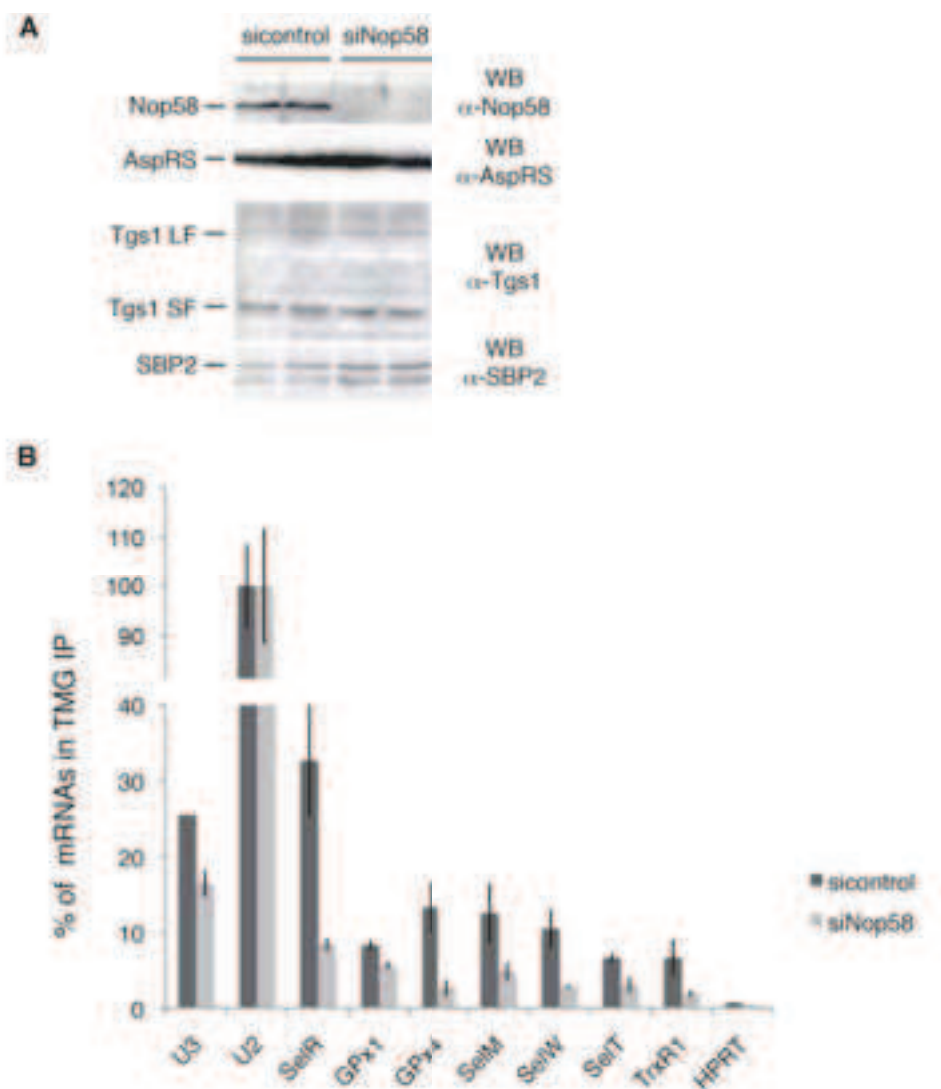


Figure S3: Nop58 is required for cap hypermethylation. (A) Western blot analysis of siNop58 efficiency. Levels of Nop58 protein were reduced to 2% compared to the control, Tgs1 and SBP2 were unaffected. (B) Nop58 contributes to selenoprotein mRNA hypermethylation. RNA-IP using anti-TMG serum was performed as described in Fig. 1.

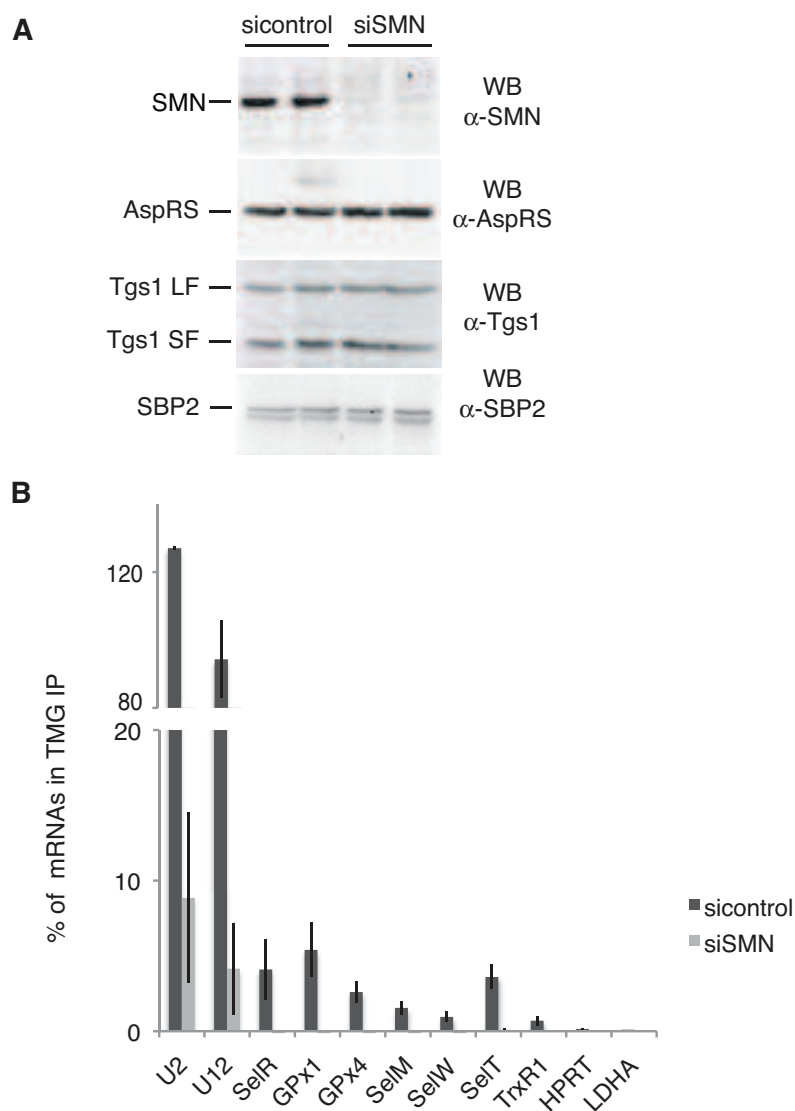


Figure S4: SMN is required for cap hypermethylation. (A) Western blot analysis of siSMN efficiency. Levels of SMN protein were reduced to 8% compared to the control, Tgs1 and SBP2 were unaffected. (B) SMN is required for selenoprotein mRNA hypermethylation. RNA-IP using anti-TMG serum in siSMN conditions were performed as described in Fig. 1.

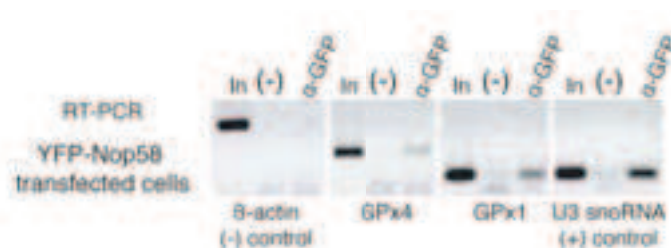


Figure S5: Nop58 interacts with selenoprotein mRNAs in vivo. (A) HEK293FT cells transfected with SBP2 and Nop58-YFP were immunoprecipitated using anti-GFP antibodies. Detection of bound RNAs was done by RT-PCR. In: input 15%; (-) no antibody; U3: positive control; β-actin: negative control.

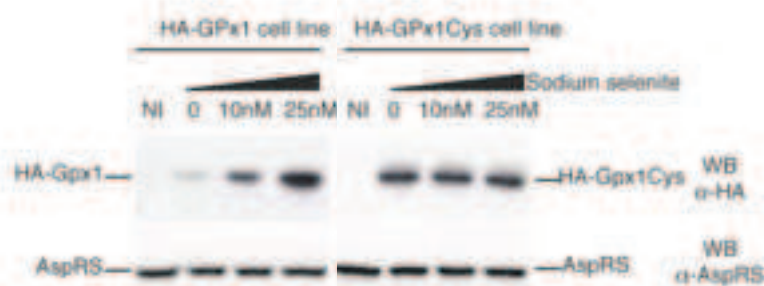


Figure S6: Experimental validation of the stable cell lines HA-GPx1 and HA-GPx1Cys. Western blot analysis of the expression of HA-GPx1 and HA-GPx1Cys proteins using anti-HA antibodies, after doxycycline induction. NI: non induced. The levels of HA-GPx1 protein expression, but not of HA-GPx1Cys, are selenium dependant. Sodium selenite concentrations are indicated above the panels. Western blot using anti-AspRS antibodies were used as control.

I.3 Résultats supplémentaires

I.3.1 Les ARNm de sélénoprotéines n'interagissent pas avec la protéine snurportine-1

Dans le cas des snARN l'hyperméthylation de la coiffe par Tgs1 a lieu dans le cytoplasme. La coiffe m₃G et la présence de protéines Sm assemblées sur le snARN forment un signal de localisation nucléaire bipartite (Fischer and Lührmann, 1990) reconnu par la protéine snurportine-1 (Huber et al., 1998; Narayanan et al., 2002; Strasser et al., 2004) qui permet le réimport des particules dans le noyau (pour une revue récente voir Matera et al. 2014). Puisque nous avons montré que les ARNm de sélénoprotéines portent une coiffe hyperméthylée et que leur assemblage présente de fortes similitudes avec celui des snRNP, nous avons analysé leur capacité à interagir avec la snurportine-1 *in vivo*. Pour cela, la snurportine-1 a été surexprimée en fusion avec une étiquette GST dans les cellules HEK293FT (Figure 44A) et des expériences de GST pull-down ont été réalisées. Les ARN associés à la GST-snurportine-1 ont été analysés par RT-qPCR et la souche parentale a été utilisée comme contrôle (Figure 44B). Il apparaît que, contrairement aux snARN U1 et U2 qui sont enrichis dans les précipités et associés à la GST-snurportine-1, les ARNm de sélénoprotéines ne le sont pas. Ces résultats, corrélés avec le fait que ces ARNm sont plus abondant dans le cytoplasme que dans le noyau (Wurth et al., 2014), indiquent que les ARNm de sélénoprotéines, après hyperméthylation de leur coiffe, ne sont vraisemblablement pas réimportés dans le noyau par la snurportine-1.

Il serait cependant possible que l'interaction entre la snurportine-1 et la coiffe m₃G des ARNm de sélénoprotéines soit trop transitoire pour être observée dans nos conditions expérimentales ou ne soit pas suffisante. Dans ce cas, nous pourrions envisager de réaliser un pontage aux UV ou chimique, liant ainsi de manière covalente les macromolécules en interaction et facilitant leur détection. Cette méthode a été utilisée pour étudier les sites de contact entre la snurportine-1 et le snARN U1 et a révélé que la snurportine-1 établissait également des contacts avec une tige boucle du snARN U1 (Kühn-Hölsken et al., 2010).

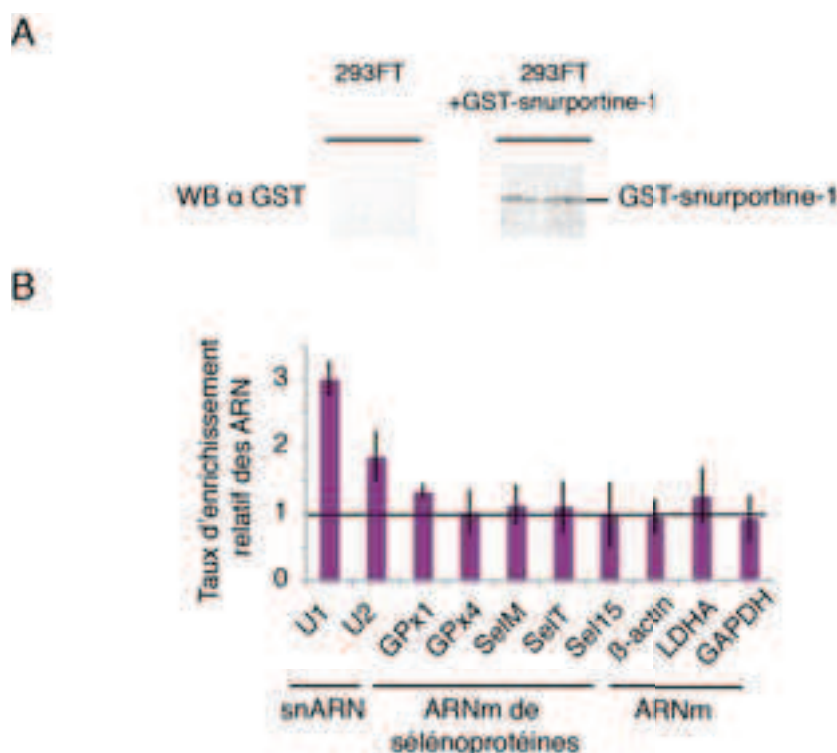


Figure 44. Les ARNm de sélénoprotéines n'interagissent pas avec la protéine snurportine-1.

(A) Des cellules HEK293FT ont été transfectées par un plasmide codant pour la GST-snurportine-1 et la surexpression de la protéine a été vérifiée par western blot. (B) Expériences de GST pull-down. Les ARN précipités par la GST-snurportine-1 à partir des extraits cellulaires ont été analysés par RT-qPCR. Les résultats sont représentés sous forme de taux d'enrichissement des ARN après immunoprécipitation, relatifs à celui de l'ARNm de HPRT et normalisés par rapport à l'expérience réalisée sur les extraits de cellules non transfectées. U1 et U2 sont les snARN utilisés comme contrôles positifs. Les ARN β-actine, LDHA et GAPDH sont des ARN non-sélénoprotéines utilisés comme contrôles négatifs.

1.3.2 Les protéines Sm ne sont pas assemblées sur les ARNm de sélénoprotéines

L'assemblage cytoplasmique des snRNP fait intervenir le complexe SMN et le méthylosome pour associer les protéines Sm aux niveau des sites Sm (sites riches en uridines, flanqués de purines) des snARN (Chari et al., 2008; Meister et al., 2002). Les protéines Sm et le complexe SMN jouent ensuite un rôle majeur dans l'import nucléaire des snARN portant une coiffe m₃G. A ce jour, aucune analyse bioinformatique systématique n'a permis de définir l'existence de sites de type Sm au sein des ARNm de sélénoprotéines. Nous avons cependant testé expérimentalement la présence de protéines Sm sur les ARNm de sélénoprotéines. Pour cela, nous avons immunoprécipité les complexes ARN/protéines associés aux protéines Sm à partir d'extraits de cellules HEK293FT et analysé les ARN co-immunoprécipités par RT-PCR (Figure 45A) et qRT-PCR (Figure 45B). Contrairement aux snARN U1 et U2

qui sont fortement enrichis par le biais de protéines Sm, les ARNm de sélénoprotéines ne semblent pas associés aux protéines Sm (Figure 45).

Ainsi, contrairement aux snRNP qui sont réimportés dans le noyau suite à l'hyperméthylation de leur coiffe, il semblerait que les mRNP de sélénoprotéines ne le soient pas car elles n'interagissent ni avec les protéines Sm, ni avec la snurportine-1, et sont majoritairement cytoplasmiques (Wurth et al., 2014). Malgré des similitudes importantes entre les mécanismes d'assemblage des snRNP et des mRNP de sélénoprotéines, des différences majeures liées au devenir et à la localisation des mRNP ne sont pas surprenantes.

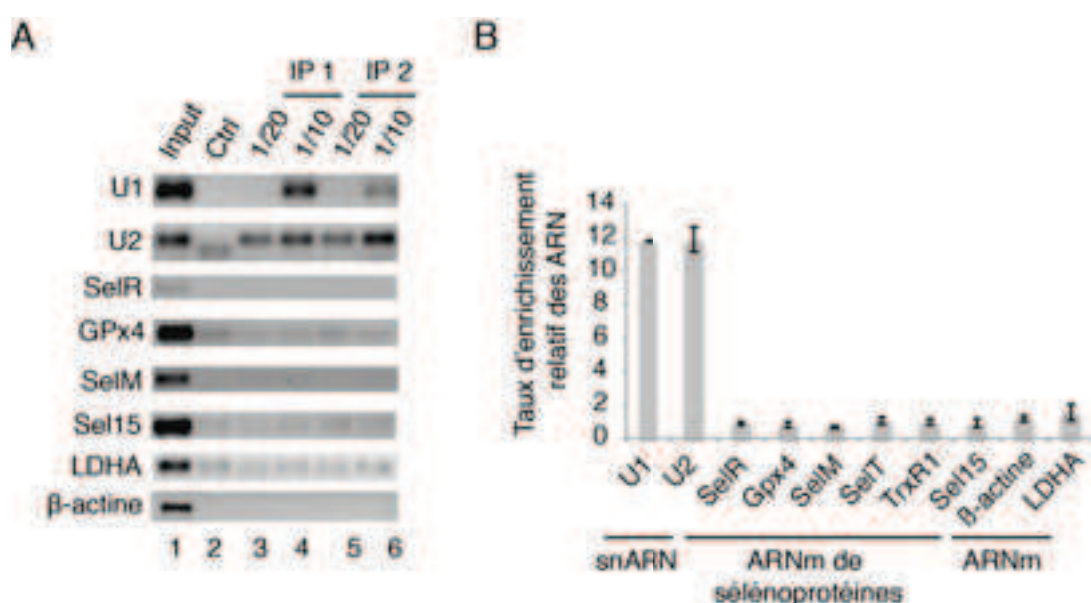


Figure 45. Les protéines Sm ne s'associent pas aux ARNm de sélénoprotéines.

Analyse par RT-PCR (A) et RT-qPCR (B) des ARN co-immunoprécipités par les anticorps dirigés contre les protéines Sm B/B' et D. IP 1 et IP 2 représentent les IP issues de deux réplicats biologiques. U1 et U2 sont les snARN utilisés comme contrôles positifs. Les ARNm de la β -actine et de LDHA sont les contrôles négatifs. Les résultats sont représentés sous forme de taux d'enrichissement des ARN après IP, relatifs à celui de l'ARNm de HPRT et normalisés par rapport à l'IP contrôle réalisée sans anticorps.

1.3.3 Impact des facteurs de liaison à la coiffe sur la synthèse de la sélénoprotéine HA-GPx1

1.3.3.1 Le complexe de liaison à la coiffe CBC

La coiffe m⁷G des ARNm eucaryotes est un déterminant essentiel du devenir des ARNm. Les deux complexes protéiques majeurs de liaison à la coiffe sont le facteur d'initiation eIF4E dans le cytoplasme et le complexe CBC (Cap Binding Complex)

dans le noyau. Le complexe CBC est le premier complexe recruté sur la coiffe m⁷G lorsque celle-ci est ajoutée de manière post-transcriptionnelle en 5' des ARN. CBC est un complexe multifonctionnel, essentiel pour l'expression des gènes et impliqué dans l'assemblage, la maturation, l'export nucléaire et la traduction des ARN (Gonatopoulos-Pournatzis and Cowling, 2014). CBC est composé des protéines CBP20 et CBP80 qui lient la coiffe m⁷G de manière synergique. L'affinité du complexe pour la coiffe monométhylée m⁷G est spécifique et diminue de plus de 1000 fois pour une coiffe triméthylée m₃G (Izaurre et al., 1994; Worch et al., 2005). Puisque les ARNm de sélénoprotéines portent une coiffe m₃G, nous avons analysé l'impact de l'inhibition des facteurs de liaison à la coiffe par siARN sur l'expression des protéines HA-GPx1 (Figure 46). Ces expériences ont été réalisées dans les lignées cellulaires HA-GPx1 et HA-GPx1Cys par pulse à la méthionine ³⁵S (Figure 46B) ou par Western blot (Figure 46C) (les méthodes sont discutées dans le chapitre Résultats II.2.). L'inhibition de CBP20 par siARN, affecte plus fortement l'expression de HA-GPx1Cys que de HA-GPx1 (Figure 46B). En effet une réduction de 60% de l'ARNm de CBP20 aboutit à une diminution de 50% de la traduction de HA-GPx1Cys alors que la traduction de HA-GPx1 n'est réduite que de 35%. Ces effets sont comparables à ceux obtenus pour siEIF4E (diminution de 65% pour HA-GPx1Cys contre 20% seulement pour HA-GPx1) (Figure 46B et article) et opposés à ceux de siTgs1, utilisés comme contrôles. Ces résultats suggèrent que la traduction de l'ARNm HA-GPx1 est peu dépendante de la liaison de CBP20 à la coiffe, Ceci est en accord avec la présence d'une coiffe m₃G au niveau de l'ARNm HA-GPx1. L'inactivation de CBP80 en revanche diminue à la fois le niveau d'expression HA-GPx1 et de HA-GPx1Cys ; dans les deux cas le niveau d'expression des protéines varie respectivement entre 20 et 30% (Figure 46B). Selon la méthode de détection et de normalisation utilisée le niveau d'expression a également été vu augmenté jusqu'à 1,5 fois (Figure 46C, méthode de Western blot). Ceci résulte clairement d'une dérégulation de l'expression des ARNm de sélénoprotéines qu'il est pour l'instant difficile d'interpréter mais qui pourrait être liée au fait que CBP80, contrairement à CBP20, n'établit pas de contact direct avec la coiffe mais sert de plateforme d'interaction avec d'autres facteurs (Mazza et al., 2001). L'un des rôles du complexe CBC est d'assister le premier cycle de traduction de l'ARNm (« pioneer round of translation ») avant d'être remplacé par le facteur eIF4E (Chiu et al., 2004; Maquat et

al., 2010b). Il a été proposé qu'une des fonctions du premier cycle de traduction CBC-dépendant était de contrôler la qualité de l'expression des gènes *via* le mécanisme du NMD (nonsense-mediated mRNA decay) (Hwang et al., 2010; Maquat et al., 2010a, 2010b). De façon intéressante, plusieurs ARNm de sélénoprotéines sont sensibles au NMD (Squires et al., 2007b), voire régulés par le mécanisme du NMD en conditions de sélénium limitantes (Seyedali et al., 2014) ; il est donc envisageable que l'absence d'interaction du complexe CBC avec la coiffe m₃G des ARNm de sélénoprotéines permette d'éviter leur dégradation par ce mécanisme. Il sera donc intéressant de déterminer le rôle de CBC dans l'expression des sélénoprotéines.

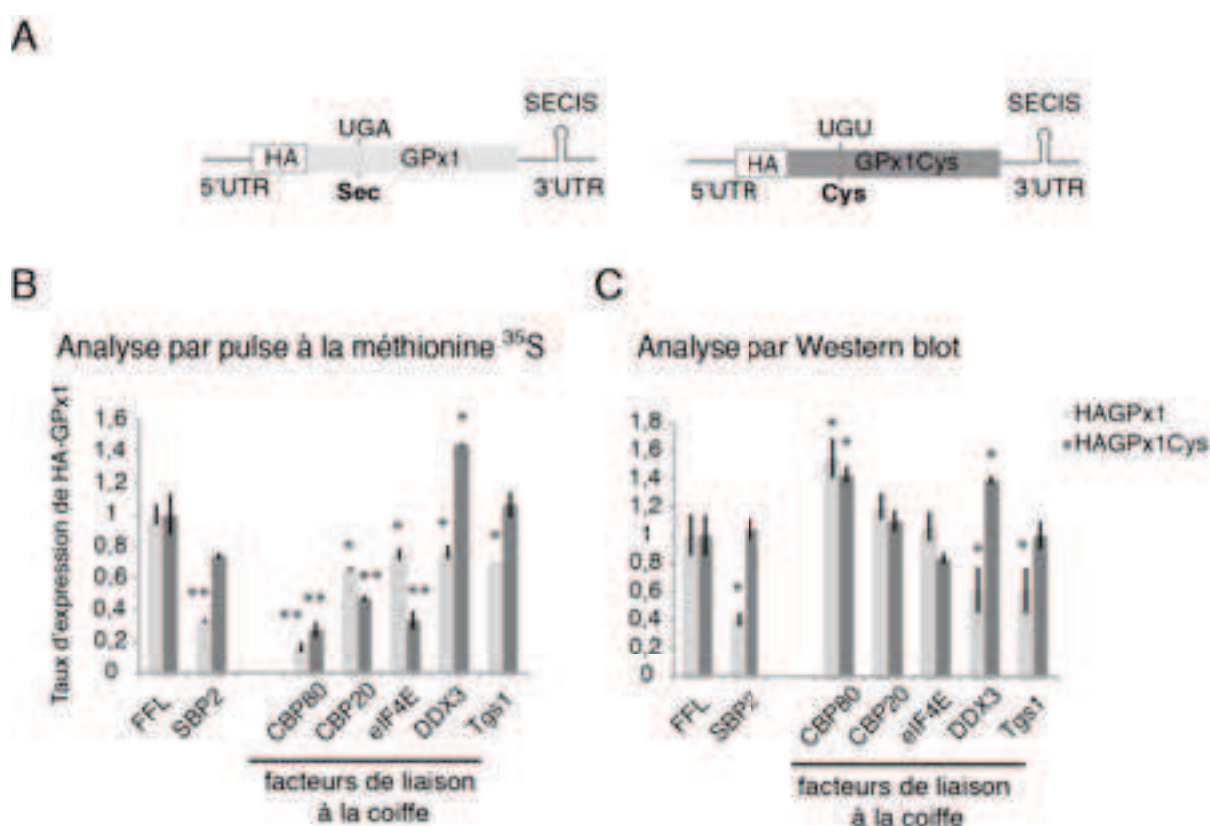


Figure 46. Impact de l'inhibition de facteurs de liaison à la coiffe sur la synthèse de la sélénoprotéine HA-GPx1 *in vivo*.

(A) Représentation des constructions exprimées par les cellules HA-GPx1 et HA-GPx1Cys. Les traits représentent les UTR de l'ARNm GPx1 et les rectangles gris les phases codantes de GPx1 et GPx1Cys. Le codon UGA^{Sec} ou UGU (mutant cystéine), sont représentés ainsi que l'élément SECIS et l'étiquette HA. L'expression de différents facteurs de liaison à la coiffe a été inhibée dans les cellules HA-GPx1 et HA-GPx1Cys par siARN et l'impact sur la synthèse de HA-GPx1 a été quantifié par pulse à la méthionine ³⁵S (B) ou par Western blot (C).

1.3.3.2 L'hélicase DDX3

L'hélicase DDX3 a récemment été proposée comme protéine de liaison à la coiffe. En effet, DDX3 est capable de se substituer au facteur d'initiation eIF4E dans la liaison à la coiffe m⁷G de l'ARNg du virus VIH-1 où il génère la formation d'un complexe trimérique DDX3/PABP/eIF4G (Soto-Rifo et al., 2013). Ce complexe de pré-initiation ne contient pas les protéines de liaisons à la coiffe CBP20/CBP80 ou eIF4E et est requis pour l'initiation de la traduction de l'ARNg (Soto-Rifo et al., 2012, 2013). De manière intéressante, il a été montré que la coiffe de certains ARN du VIH-1, dont l'ARNg, était également hyperméthylée par Tgs1 et que cette modification module leur expression (Yedavalli and Jeang, 2010). Cette similarité avec notre système nous a conduit à analyser l'impact de l'inactivation par siARN de DDX3 sur l'expression de la sélénoprotéine HA-GPx1. Etonnamment, l'inhibition de DDX3 entraîne une forte diminution de l'expression de la sélénoprotéine HA-GPx1 mais pas du mutant HA-GPx1Cys (Figure 46B et C). En Effet, pour une inactivation de DDX3 entraînant une réduction de 45% de son ARNm, la traduction de HA-GPx1 est diminuée de 30-40%, alors que celle de HA-GPx1Cys est même augmentée de 40% (Figure 46B et C). Ces résultats indiquent que DDX3 pourrait jouer un rôle dans la synthèse des sélénoprotéines. DDX3 est une DEAD-box ARN hélicase impliquée dans tous les aspects du métabolisme des ARN (Ariumi, 2014). L'hélicase semble moduler la traduction des protéines cellulaires. En condition de stress, elle est associée à des complexes traductionnels inactif dans les granules de stress, sites de répression de la traduction (Lai et al., 2008; Shih et al., 2011), et réprime la traduction coiffe-dépendante en piégeant eIF4E (Shih et al., 2008). Paradoxalement, DDX3 stimulerai la traduction d'un pool de transcrits très spécifiques contenant des régions 5'UTR structurées et notamment lorsque qu'il y a présence d'IRES (Internal Ribosome Entry Site) (Lai et al., 2008; Shih et al., 2011; Soto-Rifo et al., 2012). Cette dualité de fonction se reflète dans nos résultats. Puisque l'inactivation de DDX3 par siARN inhibe l'expression de HA-GPx1, l'hélicase semble jouer un rôle positif sur la synthèse de la sélénoprotéine. Au contraire, elle semble réprimer l'expression de HA-GPx1Cys et cette inhibition est levée dans les conditions d'inactivation de DDX3 entraînant une surexpression de la protéine. La région 5'UTR des ARNm de sélénoprotéines est courte et peu structurée, il serait cependant envisageable que DDX3 puisse être recrutée au niveau du SECIS. La protéine DDX3 pourrait avoir une

fonction dans le mécanisme d'initiation de la traduction des ARNm de sélénoprotéines, similaire à celle de observée pour les ARNg du VIH-1 (Soto-Rifo et al., 2012, 2013). Ainsi, DDX3 constitue une piste intéressante dans l'étude du mécanisme d'initiation de la traduction non-canonique des ARNm de sélénoprotéines.

I.4 Conclusion

Nous avons montré que les ARNm de sélénoprotéines sont sujets à une maturation différentielle de leur extrémité 5' : certains portent une coiffe m₃G, d'autres, une coiffe monométhylée m⁷G et d'autres semblent avoir les deux. Ainsi, il est tout à fait envisageable que l'hyperméthylation de la coiffe fasse partie d'un processus de régulation de l'expression hiérarchique des sélénoprotéines. En effet, la régulation de la stabilité des ARNm de sélénoprotéines et de leur traduction varie pour chaque ARNm de sélénoprotéine et dépend de nombreux facteurs (Budiman et al., 2009a; Latrèche et al., 2012; Lei et al., 1995; Low et al., 2000a; Squires et al., 2007b; Sun et al., 2001c). Ce mécanisme de régulation est loin d'être compris et, jusqu'à présent, aucun mécanisme unificateur n'a pu être proposé. Il a été montré que la coiffe m₃G pouvait être impliquée dans un mécanisme de régulation différentielle. En effet, l'hyperméthylation de la coiffe de certains ARN du virus du VIH-1 contenant un motif RRE (Rev Responsive Element) représente un mécanisme de régulation sélective de leur expression (Yedavalli and Jeang, 2010). La coiffe m₃G pourrait avoir une fonction similaire dans le cas des ARNm de sélénoprotéines.

L'hyperméthylase Tgs1 semble recrutée vers les ARNm de sélénoprotéines via deux réseaux d'interaction SBP2/Nop58/Tgs1 ou SBP2/SMN/Tgs1, respectivement dans le noyau ou le cytoplasme. Cela suggère que l'hyperméthylation puisse avoir lieu dans différents compartiments cellulaires, ce qui pourrait constituer un moyen de régulation de l'hyperméthylation. Il a été montré que le stress oxydant, provoque l'accumulation de SBP2 dans le noyau et stimule l'assemblage des mRNP de sélénoprotéines (de Jesus et al., 2006; Papp et al., 2006; Touat-Hamici et al., 2014). Ainsi, le stress oxydant pourrait réguler l'hyperméthylation de la coiffe des ARNm de sélénoprotéines en stimulant le recrutement de l'isoforme nucléaire de Tgs1 via Nop58. Il serait donc intéressant de comprendre le mode de régulation de l'hyperméthylation de la coiffe des ARNm de sélénoprotéines.

La coiffe m₃G des sn et des snoARN joue un rôle important dans leur stabilité (Reddy et al., 1992; Speckmann et al., 2000). Puisque certains ARNm de sélénoprotéines sont sensibles au mécanisme du NMD (Squires et al., 2007b), un des rôles potentiels de la coiffe m₃G de ces ARNm pourrait être d'augmenter leur stabilité en leur permettant d'échapper à ce mécanisme de contrôle. Il est envisageable que l'hyperméthylation de la coiffe m₃G permette d'échapper au premier cycle de traduction CBC-dépendant, conduisant à une régulation de l'expression des protéines par le NMD, comme cela a été observé lorsque le sélénium est limitant (Seyedali et al., 2014).

La découverte d'ARNm portant une coiffe m₃G mène à la question fondamentale de leur capacité à être traduits. Il a été montré que la traduction *in vitro* de l'ARNm de la β -globine portant une coiffe m₃G est moins efficace que lorsque l'ARNm est coiffé par du m⁷G (Darzynkiewicz et al., 1988). Nous avons cependant montré la présence des ARNm de sélénoprotéines coiffés de m₃G dans les polysomes ce qui suggère qu'ils sont traduits (Wurth et al., 2014). De manière intéressante, les ARNm de 70% des gènes de *Caenorhabditis elegans* portent une coiffe triméthylée m₃G et sont traduits (Liou and Blumenthal, 1990). Cette coiffe est acquise par un mécanisme très différent de ce que nous avons décrit jusqu'à présent, appelé trans-épissage et qui fait appel à un snARN SL (spliced-leader) (Blumenthal, 2012; Van Doren and Hirsh, 1990; Hastings, 2005). Elle est retrouvée dans plusieurs espèces d'eucaryotes primitifs (Pettitt et al., 2010). Durant ce phénomène, les 22 nucléotides en 5' du snARN SL sont ajoutés aux extrémités 5' des pré-ARNm. Dans le cas des opérons, la fonction du trans-épissage permet de générer des transcrits traductibles à partir des régions en aval du premier cistron (Blumenthal and Gleason, 2003). La présence de la coiffe m₃G en 5' de la séquence SL, stimule la traduction (Lall et al., 2004; Maroney et al., 1995; Wallace et al., 2010). Chez *Ciona intestinalis* une corrélation entre la fréquence de trans-épissage de l'ARNm et la fonction du gène a été montrée (Matsumoto et al., 2010). Ainsi, le trans-épissage pourrait être un mode de régulation de l'expression des gènes chez ces organismes. Il existe cinq isoformes du facteur d'initiation de la traduction eIF4E chez *C. elegans* dont trois sont capables de reconnaître la coiffe m₃G (Keiper et al., 2000). Jusqu'à présent, aucun équivalent de ces trois isoformes n'a été identifié chez les mammifères (Joshi et al., 2005). Cependant il est possible que des facteurs, encore non identifiés, soient capables de

lier la coiffe hyperméthylée des ARNm de sélénoprotéines), à la place de eIF4E (voir Figure 43). Nous avons tenté d'identifier de tels facteurs à l'aide d'une méthode de purification de complexes associés *in vivo* à ces ARNm (Résultats III). L'ensemble de nos résultats suggère l'existence d'un mécanisme d'initiation de la traduction non-canonique lié à la présence de la coiffe m₃G (Figure 43). L'analyse détaillée de l'impact de la coiffe m₃G sur les mécanismes d'initiation de la traduction des ARNm de sélénoprotéines fait l'objet du chapitre III de la partie Résultats.

II Le chaperon d'assemblage SMN et le méthylosome sont impliqués dans la synthèse des sélénoprotéines

II.1 Objectif

Nous avons montré que la protéine SMN participe au recrutement de l'enzyme Tgs1 vers les ARNm de sélénoprotéines et donc à l'hyperméthylation de la coiffe (voir Résultats I). Dans l'assemblage des snRNP, le complexe SMN fonctionne en association avec le méthylosome. Notre objectif est de comprendre l'implication de ces deux complexes et de l'ensemble des protéines qui les composent dans le mécanisme d'assemblage des mRNP de sélénoprotéines. SBP2 joue un rôle clé dans cet assemblage en stimulant l'interaction entre des facteurs de biogenèse, des protéines cœurs et l'ARN SECIS. Nous avons, dans un premier temps, testé l'interaction de SBP2 avec les différentes sous-unités des complexes SMN et du méthylosome. Une partie de ce travail a été initiée par Michael Leichter, post-doctorant dans notre équipe. Nous avons ensuite analysé l'impact de l'inhibition de ces facteurs sur la synthèse *in vivo* de la sélénoprotéine HA-GPx1. Le rôle fonctionnel des deux complexes sur l'assemblage des mRNP de sélénoprotéines a été recherché.

II.2 Résultats

II.2.1 Le complexe SMN et le méthylosome interagissent avec la protéine SBP2

L'assemblage des snRNP requière la présence de la protéine SMN sous forme d'un complexe de haut poids moléculaire (Meister et al., 2001b) composé des protéines Gemin 2 à 8 et de la protéine Unrip (Charroux et al., 1999; Fischer et al., 1997; Otter et al., 2007), pour revue voir (Gubitz et al., 2004; Paushkin et al., 2002; Pellizzoni, 2007). Afin de comprendre si l'ensemble du complexe SMN est requis pour l'assemblage des sélénoprotéines, nous avons testé l'interaction des différentes protéines qui le composent avec SBP2. Nous avons montré que les protéines SMN, Gemin 2 et Gemin 3 co-immunoprécipitent avec la protéine GFP-SBP2 surexprimée dans les cellules HEK293FT, indiquant qu'au moins trois protéines du complexe

SMN, seraient impliquées dans l'assemblage des mRNP de sélénoprotéines (Figure 47A).

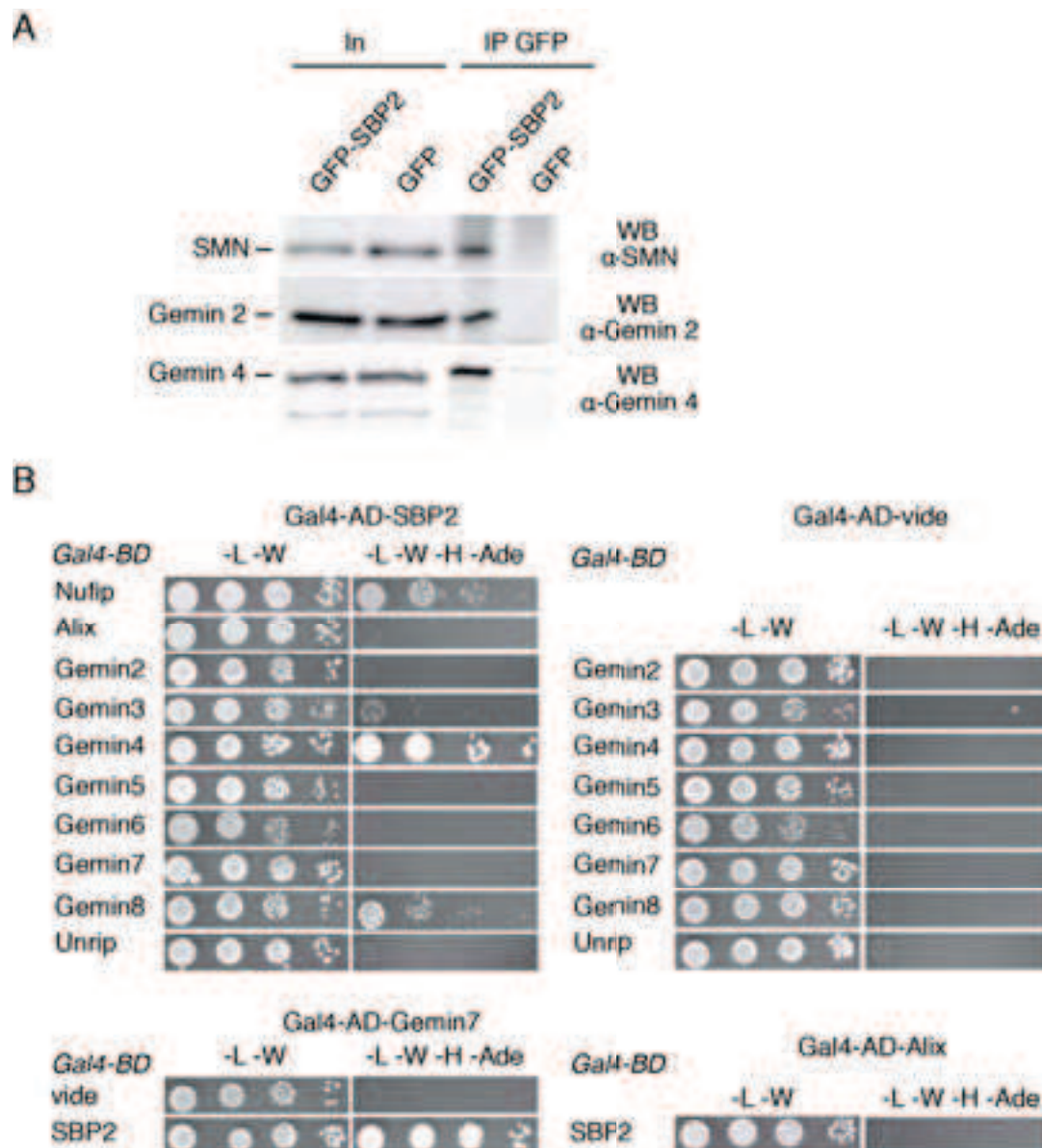


Figure 47. L'ensemble du complexe SMN interagit avec la protéine SBP2

(A) Immunoprécipitations à partir d'extraits de cellules HEK293FT transfectées par GFP-SBP2 ou GFP seul, à l'aide de billes anti-GFP. Les protéines co-immunoprécipitées ont été analysées par Western-blot à l'aide d'anticorps dirigés contre SMN, Gemin2 et Gemin4. (B) Tests d'interaction double-hybride réalisés dans la levure *S. Cerevisiae* entre les protéines du complexe SMN et protéine SBP2. Des constructions exprimant les différentes protéines en fusion avec le domaine de liaison à l'ADN Gal4-(BD), ou le domaine activateur Gal4-(AD), ont été utilisées. La capacité des cellules à pousser sur un milieu dépourvu d'histidine et d'adénine (-H -Ade) indique l'existence d'une interaction entre les protéines testées. Les résultats obtenus pour SBP2-BD ne sont représentés que s'ils diffèrent des résultats obtenus avec SBP2-AD. La protéine Nufip a été choisie comme contrôle positif et Alix comme contrôle négatif. (Résultats de M. Leichter)

Des tests d'interaction double-hybride dans la levure ont été réalisés pour tester l'interaction de l'ensemble des protéines du complexe SMN avec SBP2 (Figure 47B, expériences réalisées par M. Leichter). La protéine Nufip a été choisie comme contrôle positif (Boulon et al., 2008). Des constructions permettant d'exprimer les différentes protéines en fusion avec le domaine de liaison à l'ADN (BD), ou le domaine activateur (AD), du facteur de transcription Gal4 ont été utilisées. Lorsqu'il y a interaction entre deux protéines testées, les domaines BD et AD sont à proximité, la fonction du facteur de transcription Gal4 est rétablie et permet la croissance des cellules sur un milieu dépourvu d'histidine et d'adénine. Les tests d'interaction double-hybride révèlent que les protéines Gemin 3, Gemin 4, Gemin 7 et Gemin 8 interagissent avec la protéine SBP2 (Figure 47B et Tableau 2). La protéine SMN n'a pas pu être testée puisqu'elle est capable à elle seule d'activer la transcription du promoteur Gal4 (Tableau 2 et données non incluses). Ces interactions sont vraisemblablement directes car SMN, SBP2 ainsi que tous les composants requis pour la synthèse des sélénoprotéines sont absents dans la levure *S. Cerevisae*. Par ailleurs nous avons montré que l'interaction de SBP2 avec SMN est indépendante de l'ARN, puisqu'il est possible de co-immunoprécipier les deux protéines en présence de RNaseA (voir Chapitre I, Figure 4G de l'article (Wurth et al., 2014)), et que l'interaction a également lieu lorsque les deux protéines sont traduites *in vitro* dans des RRL et des extraits d'*Escherichia coli* S30 (Wurth et al., 2014, Figure 4H). Ainsi, ces résultats montrent que l'ensemble du complexe SMN interagit avec la protéine SBP2 et serait donc impliqué dans le recrutement de Tgs1 vers les ARNm de sélénoprotéines. Le bilan des interactions caractérisées est résumé Figure 51 et Tableau 2.

La biogenèse des snRNP fait également intervenir le complexe du méthylosome, qui assiste le complexe SMN dans l'assemblage des protéines Sm sur les snARN (Chari et al., 2008; Neuenkirchen et al., 2008). Nous avons testé si le méthylosome était également impliqué dans l'assemblage des ARNm de sélénoprotéines. Le cœur du méthylosome composé de l'enzyme PRMT5 et de la protéine MEP50, adopte une structure hétéro-octamérique (Friesen et al., 2001; Meister et al., 2001b; Stopa et al., 2015). Ce cœur peut être soit associé à pICln, soit à Riok1 (Guderian et al., 2011) aboutissant à deux complexes de spécificités de substrat différentes. Ainsi, le

méthylosome associé à pICln est impliqué dans la méthylation des protéines Sm et la biogenèse des snRNP, tandis qu'il méthyle la nucléoline lorsqu'il contient RioK1 (Stopa et al., 2015). De manière intéressante, des expériences d'immunopurification des complexes endogènes associés à SBP2 dans des cellules HeLa nous ont permis de détecter l'existence d'interactions entre PRMT5, MEP50 et SBP2 (Figure 48A, d'après la thèse de Laurence Wurth, 2009).

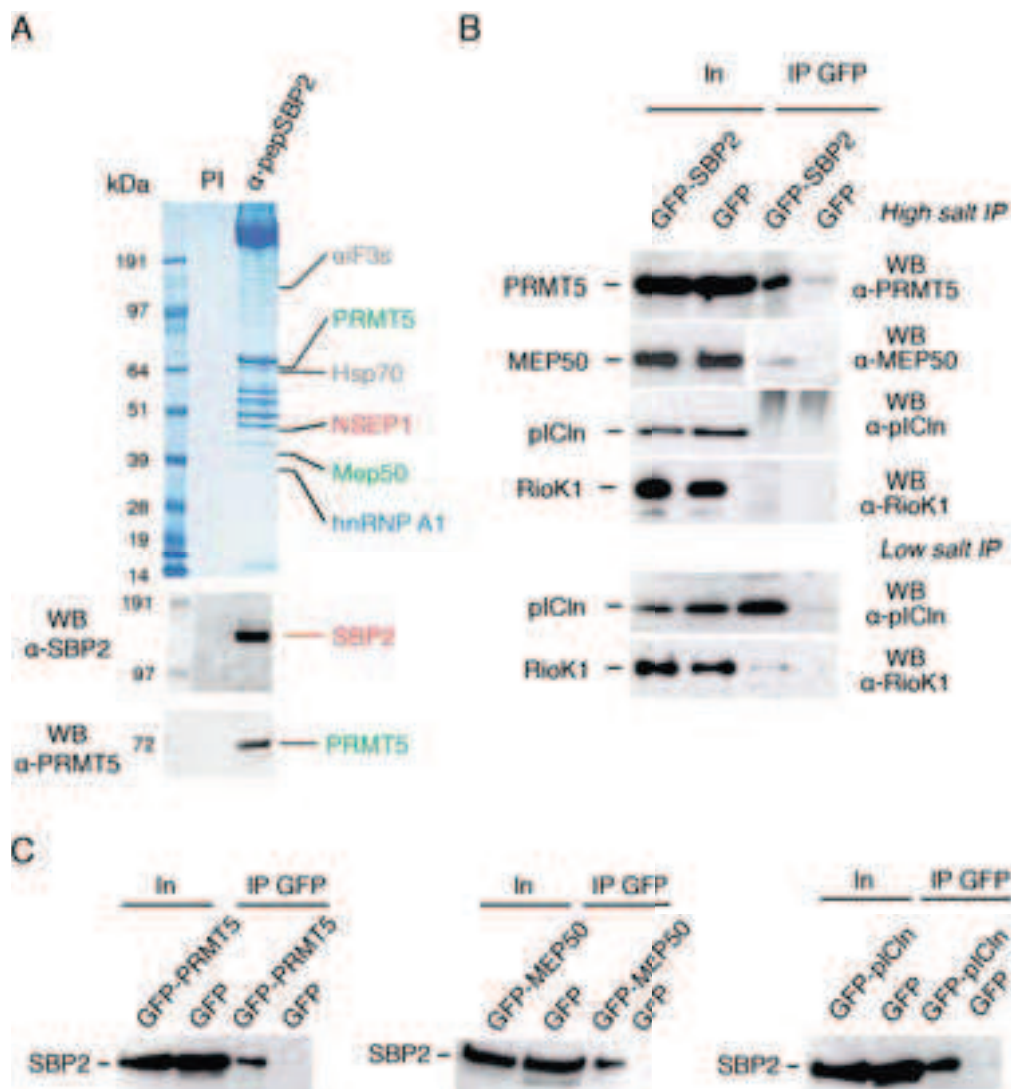


Figure 48. Le méthylosome interagit avec SBP2 dans les cellules HeLa

(A) Immunoprécipitation de complexes endogènes associés à SBP2 (d'après la thèse de Laurence Wurth, 2009). Expériences réalisées à partir d'extraits cytoplasmiques de cellules HeLa à l'aide d'anticorps anti-peptide-SBP2 (α-pepSBP2). PI : billes avec sérum pré-immun, α-pepSBP2 : billes avec anticorps anti-peptide-SBP2, kDa : kiloDaltons. Les protéines présentes dans les complexes ont été détectées par spectrométrie de masse. Les protéines SBP2 et PRMT5 ont été révélées par Western blot. (B) Immunoprécipitations anti-GFP à partir d'extraits de cellules HEK293FT transfectée par GFP-SBP2 ou GFP seul. Les IP ont été réalisées en présence de 100mM de NaCl (Low salt IP) ou 300 mM NaCl (High salt IP). Les protéines co-immunoprécipitées ont été analysées par Western-blot à l'aide d'anticorps dirigés contre les protéines du méthylosome PRMT5, MEP50, pICln et RioK1. (C) Co-immunoprécipitation à l'aide d'anticorps anti-GFP et d'extraits cellulaires HEK293FT transfectés par GFP-PRMT5, GFP-MEP50 ou GFP-pICln et analyse des résultats par Western-blot. (AS-GB avec M. Leichter et C. Allmang)

Les deux protéines PRMT5 et MEP50 co-immunoprécipitent également avec la protéine GFP-SBP2 surexprimée dans les cellules HEK293, indiquant que les protéines cœur du méthylosome interagissent avec SBP2 (Figure 48B). Ni pICln, ni Riok1 ne sont co-immunoprécipitées lorsque les lavages sont réalisés dans des conditions stringentes (300 mM NaCl). Cependant, pICln interagit avec GFP-SBP2 en conditions de faible concentration saline (150 mM NaCl), indiquant que pICln est présent au sein du complexe mais que son interaction est peu stable. L'interaction de SBP2 avec les différentes sous-unités du méthylosome a également été vérifiée par des tests d'interaction double-hybrides dans la levure (Figure 49).



Figure 49. Révélation des interactions SBP2-méthylosome par des expériences de double-hybride dans la levure.

Tests d'interaction double-hybride dans la levure *S. Cerevisiae* entre les protéines du complexe du méthylosome et SBP2. Les résultats sont représentés comme dans la Figure 47B. (Résultats de M. Leichter)

Un signal d'interaction positif est observé pour les protéines PRMT5, MEP50 et pICln (Figure 49, Figure 51 et Tableau 2). Ainsi, il semblerait que ce soit le complexe du méthylosome comprenant la protéine pICln qui soit retrouvé en interaction avec SBP2 et non le complexe lié à Riok1. Ceci révèle une similitude de plus entre le mécanisme d'assemblage des mRNP de sélénoprotéines et celui des snRNP. Afin de vérifier si ces interactions protéiques sont directes, nous avons purifié le complexe SBP2/méthylosome exprimé dans des cellules d'insectes grâce au système d'expression baculovirus (Figure 50).

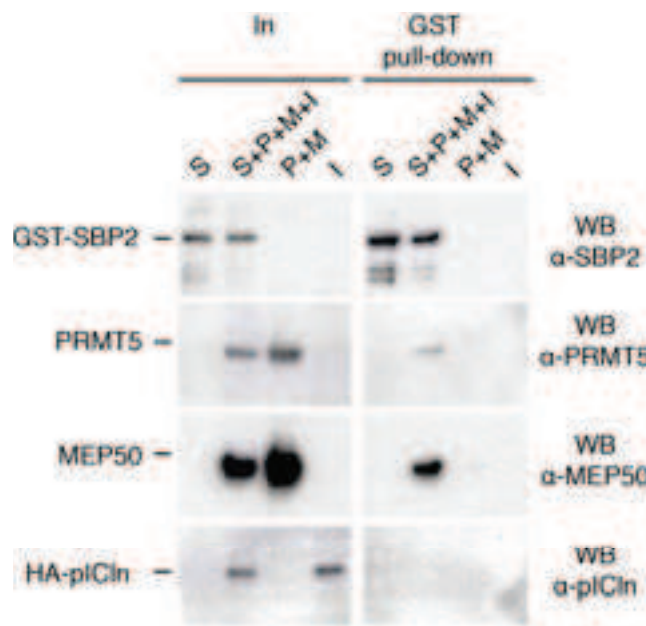


Figure 50. Expression et purification du complexe méthylosome/SBP2 dans des cellules d'insectes infectées par baculovirus.

Des baculovirus recombinants permettant l'expression de GST-hSBP2 (S), hPRMT5 (P), hMEP50 (M) et HA-hpICln (I) ont été utilisés pour infecter ou co-infecter des cellules d'insecte Sf9 (*Spodoptera frugiperda*). La purification des complexes protéiques associés à SBP2 a été réalisée par GST pull-down et analysée par Western blot à l'aide d'anticorps anti-SBP2, anti-PRMT5, anti-MEP50 ou anti-pICln. In : extrait cellulaire (Input).

Ce travail a été réalisé en collaboration avec l'équipe de J. Cavarelli (IGBMC, Strasbourg). Des expériences de GST pull-down montrent que PRMT5 et MEP50 sont co-purifiés avec GST-SBP2 (Figure 50), confirmant que ces protéines font partie du même complexe. En revanche pICln n'est pas retrouvé au sein du complexe (Figure 50), ceci est en accord avec son caractère labile (Figure 51). Ainsi, par son interaction directe avec SBP2, le méthylosome, pourrait être impliqué dans l'assemblage des ARNm de sélénoprotéines et participer à leur mécanisme de synthèse.

	SMN	Gem2	Gem3	Gem4	Gem5	Gem6	Gem7	Gem8	Unrip	PRMT5	MEP50	pICln	Riok1
Y2H, AD-SBP2	/	-	+	+	-	-	-	+	-	+	+	+	nd
Y2H, BD-SBP2	/	-	-	-	-	-	+	-	-	-	+	-	nd
co-IP	+	+	nd	+	nd	nd	nd	nd	nd	+	+	+	-
co-purif	nd	nd	nd	nd	nd	nd	nd	nd	nd	+	+	-	nd

Tableau 2 : Récapitulatif des interactions détectées entre les protéines des complexes SMN, du méthylosome et SBP2. Résultats obtenus par l'approche double-hybride (Y2H), par co-immunoprécipitations à partir d'extraits de cellules HEK293FT (co-IP) et par co-expression et purification à partir de cellules d'insectes infectées par baculovirus. + : interaction, - : pas d'interaction, nd: interaction non-testée (indéterminable dans le cas de SMN en Y2H). Pour les expériences de double-hybride, les résultats des tests dans les deux sens de fusion de AD et BD sont représentés.

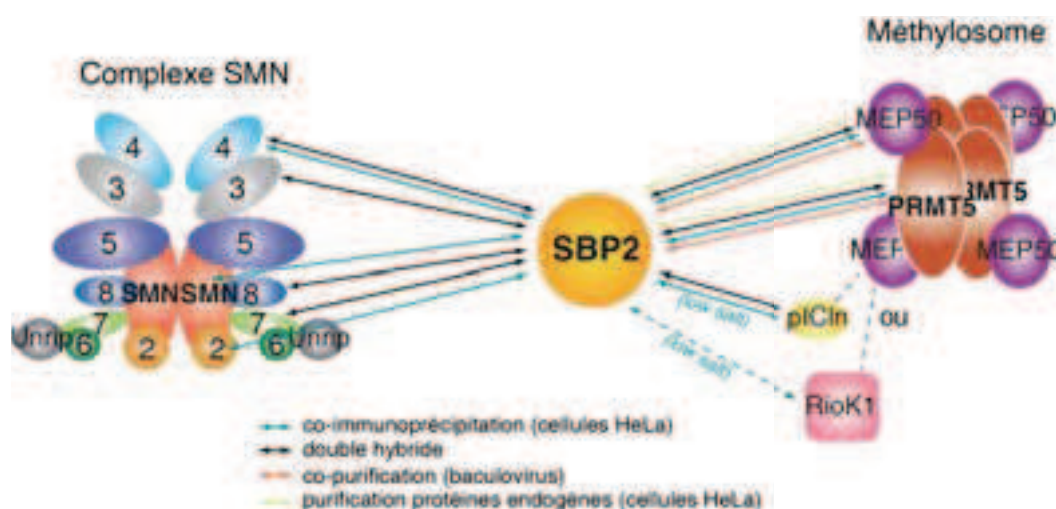


Figure 51. Représentation schématique du réseau d'interaction protéine-protéine entre SBP2 et les complexes SMN et méthylosome.

Etabli expérimentalement par l'approche double-hybride (flèches noires), par co-immunoprécipitations à partir d'extraits de cellules HEK293FT transfectées (flèches bleues) et par co-purification de SBP2 à partir de cellules d'insectes infectées par baculovirus (flèches rouges) ou par co-immunoprécipitation de protéines endogènes dans les cellules HeLa (flèches jaunes). Les chiffres dans le complexe SMN correspondent aux protéines Gemin 2-8. Pour simplifier, nous avons représenté les complexes SMN et méthylosome de manière non-stoechiométrique (Pellizzoni, 2007; Stopa et al., 2015). MEP50 et PRMT5 sont les protéines Coeur du méthylosome; pICln et RioK1 sont deux protéines associées à des complexes du méthylosome distincts (Guderian et al., 2011; Stopa et al., 2015; Yong et al., 2004).

II.2.2 L'inhibition des complexes SMN et du méthylosome a un impact fonctionnel sur la synthèse de la sélénoprotéine HA-GPx1 *in vivo*

Durant ma thèse j'ai créé deux lignées cellulaires humaines capables d'exprimer de manière inductible les protéines HA-GPx1 et HA-GPx1Cys (représentées Figure 52A) et mis au point les conditions qui me permettent d'analyser l'impact de l'inactivation de facteurs protéiques sur la synthèse *de novo* des deux protéines (voir résultats I.1). Ce système m'a permis de tester de façon systématique l'importance fonctionnelle de tous les facteurs supposés intervenir dans la synthèse des sélénoprotéines (voir Tableau M4 de la partie Matériel et Méthodes). Pour cela, j'ai utilisé une librairie de siARN « à façon » dirigés contre 20 cibles correspondant à tous les facteurs de biogenèse des sélénoprotéines caractérisés à ce jour ainsi que l'ensemble des facteurs d'assemblage que nous avons identifié dans le cadre de ma thèse. Pour chaque gène inhibé, un mélange de 4 siARN dirigés contre des séquences différentes de l'ARN cible est utilisé pour une meilleure efficacité et spécificité. J'ai utilisé cette méthode pour déterminer le rôle des protéines des complexes SMN et du méthylosome sur la synthèse de la sélénoprotéine HA-GPx1 (Figure 52).

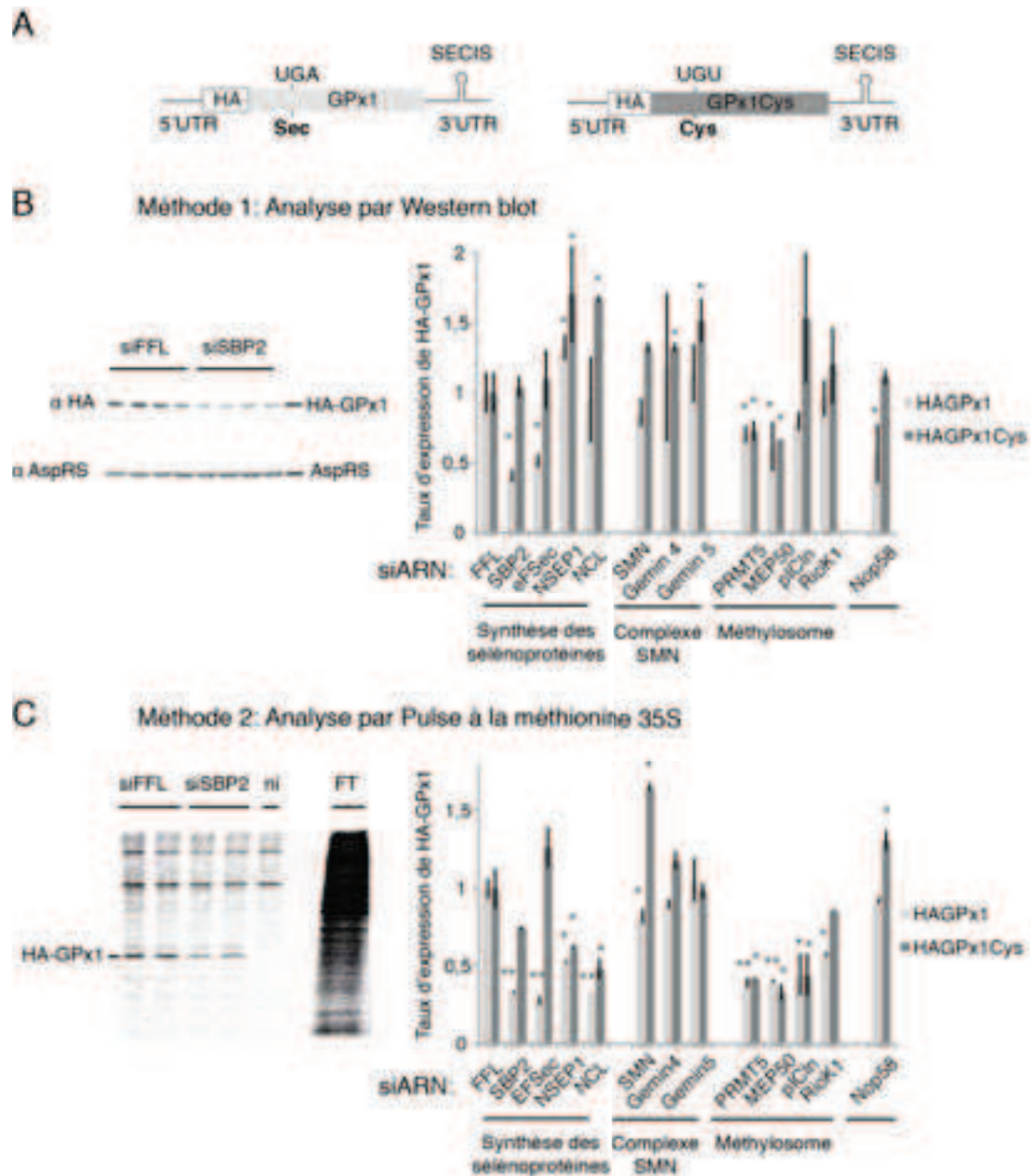


Figure 52. Impact de l'inhibition de composants des complexes SMN et méthylosome sur la synthèse de la sélénoprotéine HA-GPx1 *in vivo*.

(A) Représentation des constructions exprimées par les cellules HA-GPx1 et HA-GPx1Cys. Les traits représentent les UTR de l'ARNm GPx1 et les rectangles gris les phases codantes de GPx1 et GPx1Cys. Le codon UGA^{Sec} ou UGU (mutant cystéine), sont représentés ainsi que l'élément SECIS et l'étiquette HA. L'expression de différents facteurs de recodage et des composants du complexe SMN et du méthylosome a été inhibée par siARN dans les cellules pendant 24h. L'expression des protéines HA-GPx1 et HA-GPx1Cys est alors induite à la doxycycline pendant 7h et analysée par deux méthodes différentes. (B) Analyse et quantification par Western blot anti-HA ; le taux d'expression de HA-GPx1 est normalisé par rapport à celui de l'AspRS et au contrôle siFFL. (C) Analyse par pulse à la méthionine ³⁵S. Après 3h d'induction, le milieu est remplacé par un milieu contenant de la méthionine ³⁵S pendant 1h. Les protéines HA-GPx1 et HA-GPx1Cys issues des extraits cellulaires correspondants sont immunoprécipitées à l'aide de billes magnétiques μ MACs anti-HA (Miltenyi) et analysées sur un gel SDS-PAGE 15%. Après migration le gel est séché, exposé 48h et révélé par autoradiographie. Le taux d'expression révélé par autoradiographie est quantifié et normalisé par rapport à une protéine non-spécifique et au contrôle siFFL. Les résultats sont modulés en fonction de l'efficacité d'inhibition des siARN. Les astérisques représentent des différences significatives par rapport au contrôle siFFL. * P<0,05 et ** P<0,005 selon le test de Student.

Deux méthodes d'analyse de l'expression de HA-GPx1 ont été utilisées : la détection par Western blot ou par pulse à la méthionine ^{35}S . Dans le cas de l'analyse Western blot, l'expression de HA-GPx1 et HA-GPx1Cys est normalisée par rapport à celle de l'Asparaginyl-ARNt synthétase (AspRS) (Figure 52B). Les expériences de pulse à la méthionine ^{35}S sont suivies de l'immunoprécipitation de la protéine à l'aide d'un anticorps anti-HA (Figure 52C). Dans ce cas l'expression de HA-GPx1 et HA-GPx1Cys a été normalisée par rapport à celle obtenue dans la lignée HA-GPx1Cys Δ SECIS, pour laquelle l'expression de HA-GPx1 est indépendante du mécanisme de recodage des sélénoprotéines. De manière générale, les deux méthodes d'analyse donnent des tendances identiques. La méthode de pulse à la méthionine ^{35}S a permis d'obtenir des résultats statistiquement plus significatifs. L'inhibition des protéines SBP2 et eEFSec, facteurs clés du mécanisme de recodage des sélénoprotéines, utilisés comme contrôles, affecte fortement l'expression de HA-GPx1 en la réduisant de plus de 50% mais pas de HA-GPx1Cys, validant nos conditions expérimentales (Figure 52A et B).

Les protéines SMN, Gemin4 et Gemin5 du complexe SMN et l'ensemble des composants du méthylosome ont été inactivés. Une inhibition de l'ARNm de SMN de 50% a un effet négatif sur l'expression de HA-GPx1 qui diminue de 25% mais augmente l'expression du mutant HA-GPx1Cys de 70% (Figure 52A et B). L'inactivation de Nop58, que nous avons également montré impliqué dans le recrutement de Tgs1 pour l'hyperméthylation de la coiffe (Wurth et al., 2014), conduit à des résultats similaires à savoir une diminution de 40% de l'expression de HA-GPx1 et une augmentation l'expression du mutant HA-GPx1Cys de 20% (Figure 52A et B). Ces résultats indiquent que leur rôle dans le recrutement de Tgs1 et l'hyperméthylation de la coiffe des ARNm de sélénoprotéines (Wurth et al., 2014) est important pour leur traduction. L'inhibition de l'expression des protéines Gemin4 et Gemin5 du complexe SMN par siARN n'a, en revanche, aucun effet sur la synthèse de HA-GPx1 (Figure 52A et B). Gemin5 est une protéine de liaison à l'ARN située en périphérie du complexe SMN (Gubitz et al., 2002). La protéine reconnaît et délivre les snARN au complexe SMN, permettant ainsi l'assemblage des protéines Sm (Battle et al., 2006). Nous avons montré précédemment que les protéines Sm ne sont pas

associées aux ARNm de sélénoprotéines (cf Résultats I.3), il n'est donc pas étonnant que son inactivation n'affecte pas la synthèse de la sélénoprotéine HA-GPx1.

L'interaction du complexe SMN avec les protéines Sm est nécessaire et suffisant pour l'assemblage des snRNP *in vitro* (Meister and Fischer, 2002). Cependant, cet assemblage requiert la coopération du complexe SMN avec le méthylosome. L'inhibition des trois protéines du méthylosome (PRMT5, MEP50 et pICln) par siARN (taux d'inhibition des ARNm cibles variant entre 40 et 60%), affecte fortement et de manière égale l'expression des deux protéines HA-GPx1 et HA-GPx1Cys (Figure 52B et C), qui chute de 60% (Figure 52C) ou de 40% (Figure 52B) selon la méthode d'analyse. Ainsi, ces protéines semblent jouer un rôle dans la synthèse des sélénoprotéines mais à une étape différente du recodage. De manière intéressante, cette inhibition n'affecte pas la synthèse de HA-GPx1Cys Δ SECIS (Figure 52C), indiquant que le mode d'action des protéines du méthylosome est lié à la présence de l'élément SECIS. L'un des complexes du méthylosome caractérisé comprend la protéine RioK1 (Guderian et al., 2011) et est impliqué dans la méthylation d'une protéine de liaison à l'ARN, la nucléoline. Cette dernière se lie à l'ARN SECIS et est impliquée dans l'expression des sélénoprotéines (Miniard et al., 2010; Wu et al., 2000) mais nous n'avons pas observé d'interaction entre Riok1 et SBP2 dans nos expériences de co-immunoprécipitation (Figure 48B). Nous avons testé l'effet de l'inhibition de Riok1 et de la nucléoline sur la synthèse de HA-GPx1. De manière intéressante, l'effet de l'inactivation de ces deux protéines est très similaire et induit une inhibition légère mais significative, d'environ 20%, de l'expression de HA-GPx1 par rapport à HA-GPx1Cys (Figure 52C). Riok1 pourrait potentiellement être impliquée dans le recrutement de la nucléoline vers les ARNm de sélénoprotéines. Ainsi, nous ne pouvons pas exclure que les deux formes du complexe du méthylosome puissent être impliquées dans l'expression des sélénoprotéines : le complexe contenant Riok1 favoriserait le recodage tandis que celui contenant pICln aurait un rôle plus global sur la traduction de l'ARNm. Il est également envisageable que les protéines RioK1 et pICln fassent partie du même complexe et ne s'excluent pas mutuellement dans tous les cas.

Comme la nucléoline, NSEP1 est une protéine de liaison à l'ARN SECIS impliquée dans la régulation de l'expression des ARNm de sélénoprotéines (Fagegaltier et al., 2000d; Qichang Shen, 2006). Cependant, l'effet de l'inhibition de NSEP1 n'affecte pas significativement l'expression de HA-GPx1 dans notre système puisque les deux méthodes d'analyses n'ont pas la même tendance (Comparer Figure 52B avec 52C). L'ensemble de ces résultats nous permet de conclure que la protéine SMN ainsi que les deux complexes du méthylosome ont un impact sur la synthèse de la sélénoprotéine HA-GPx1.

II.2.3 Rôle du complexe SMN dans la synthèse des sélénoprotéines - origine du stress oxydant observé dans la pathologie SMA

Des défauts d'expression de la protéine SMN sont à l'origine de l'atrophie musculaire spinale (SMA), caractérisée par une dégénérescence des motoneurones et menant progressivement à l'atrophie musculaire (Jablonka et al., 2000; Li et al., 2014b). Une étude a montré que les motoneurones de patients SMA sont soumis à un stress oxydant important pouvant être impliqué dans leur dégénérescence (Hayashi et al., 2012). Nous avons analysé si des défauts d'expression de SMN d'origine pathologique sont susceptibles d'affecter la synthèse des sélénoprotéines, expliquant les stress oxydants observés dans les cellules des patients SMA. Grâce à une collaboration avec Séverine Massenet (IMoPa, Ingénierie Moléculaire et Physiopathologie Articulaire, Nancy) nous avons accès à des cDNA provenant de cerveau et de moelle épinière de 3 souris modèles de la pathologie SMA, et de 3 souris contrôles (voir figure 53) sacrifiées à 10 jours. Les souris modèles SMA que nous avons utilisées ont été créées par l'équipe de M. Barkats, au Génomique (Evry) et possèdent un génotype *Smn*^{-/-} ; *SMN2*^{+/+} ; *SMN1* Δ Ex7^{+/+}. Le gène *Smn* de souris sauvage a été délété et les souris transgéniques expriment le gène de SMN1 humain dans lequel l'exon 7 a été délété (*SMN1* Δ Ex7) ainsi que le gène *SMN2*, de façon à reproduire les mutations et les symptômes observés dans la pathologie SMA humaine (voir Figure 53 et introduction).

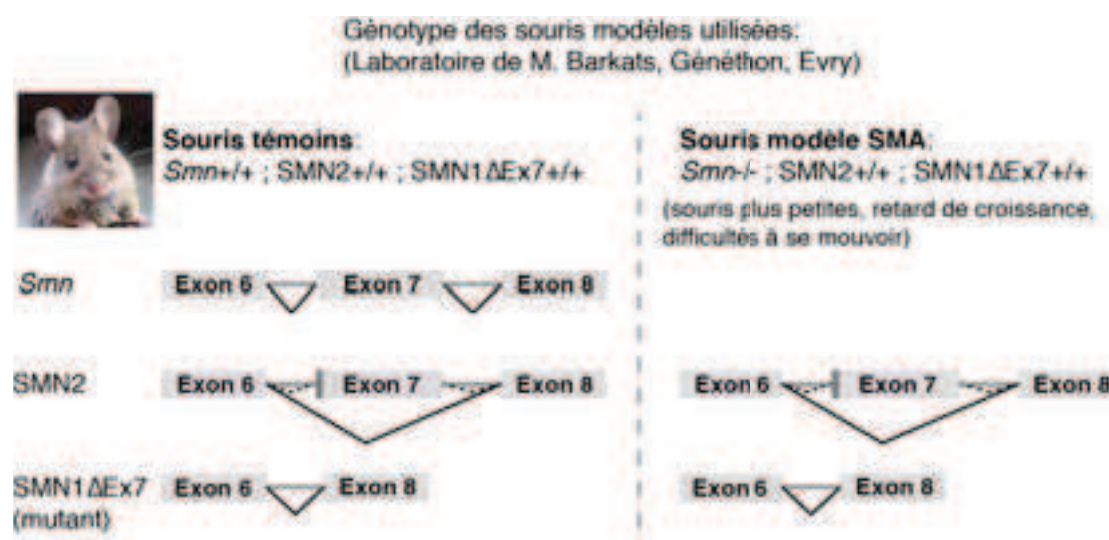


Figure 53. Représentation schématique des génotypes de souris modèle SMA utilisées.

Les rectangles gris représentent des exons et les traits gris foncés des introns. Les schémas d'épissage alternatifs sont représentés : les voies majoritaires sont représentées d'un trait noir et les voies minoritaires sont en pointillés. *Smn* : gène *Smn1* de la souris sauvage ; *SMN1* et *SMN2*, gènes humains ; *SMN1*Δ*Ex7*, gène mutant de *SMN1* humain dans lequel l'exon 7 a été délété. Seule la région comprenant les Exons 6 à 8 a été représentée puisque les différents gènes sont identiques dans le reste de la séquence. Le rectangle noir sur l'exon 7 du gène *SMN2* représente la mutation à l'origine de l'épissage alternatif de cet exon.

L'analyse des échantillons par RT-qPCR révèle que dans les cellules de cerveau (Figure 54A), aucune différence dans le taux d'ARNm de sélénoprotéines n'est visible entre les cellules de souris modèles SMA et les cellules de souris "contrôles". Ainsi, il semble que l'inactivation de *SMN* n'a pas d'effet sur les ARNm de sélénoprotéines dans ces cellules. De manière surprenante, dans les cellules de moelle épinière (Figure 54B), l'expression de la majorité des ARNm de sélénoprotéines testés est diminuée dans le cas de souris SMA (Figure 54B). Ceci est intéressant puisque ce sont spécifiquement les motoneurones, situés dans la moelle épinière, qui sont atteints dans la pathologie. Ainsi, il est possible que la diminution du complexe *SMN* entraîne des défauts de synthèse des sélénoprotéines dans les motoneurones, conduisant à des dégâts oxydants tels que ceux observés dans la pathologie SMA et pouvant contribuer aux symptômes de la maladie.

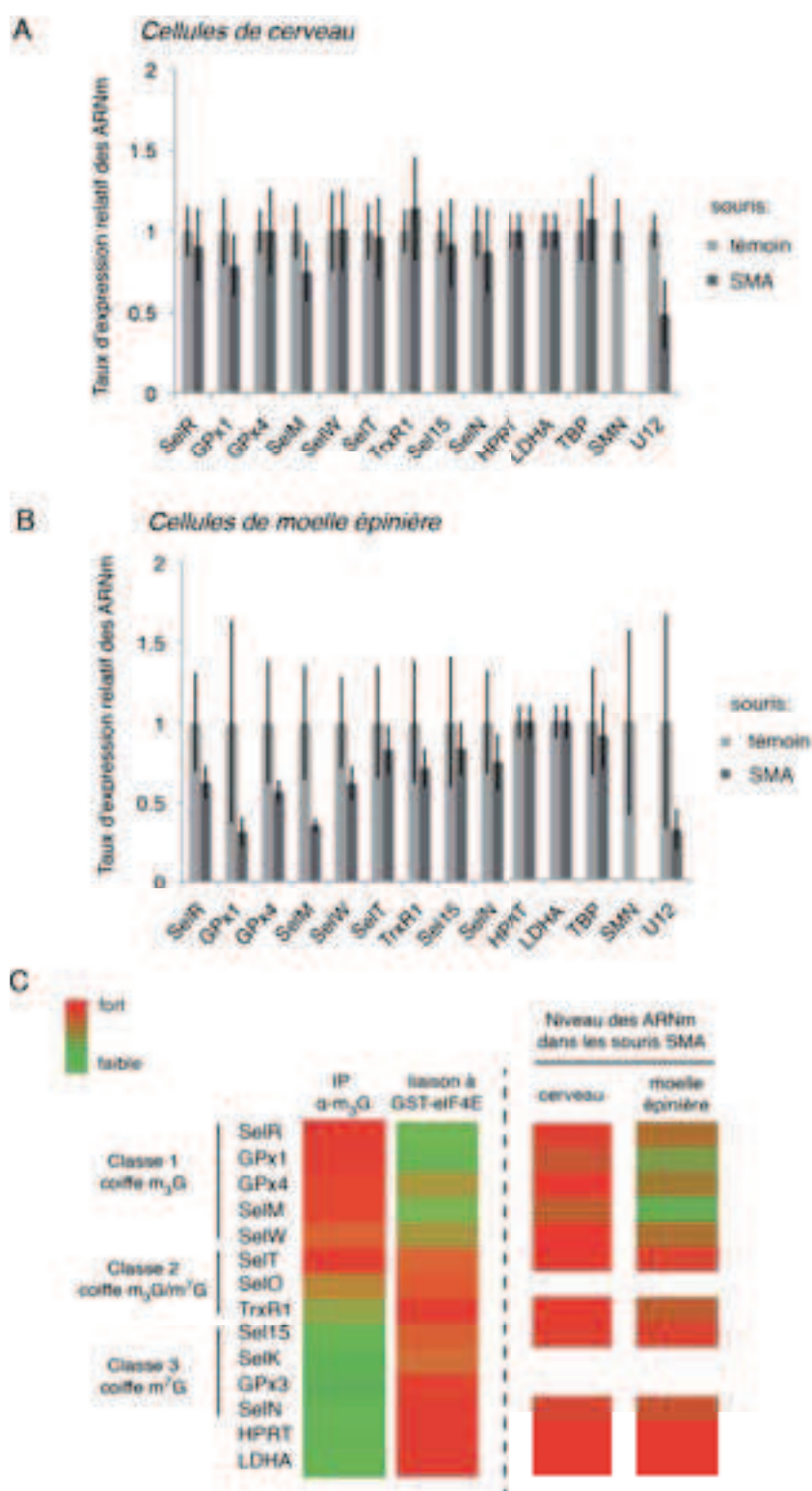


Figure 54. Le taux d'expression des ARNm de sélénoprotéines est inhibé dans les cellules de moelle épinière de souris modèles SMA

L'expression des ARNm de sélénoprotéines a été analysée dans des cellules de cerveau (A) et de la moelle épinière (B) de 3 souris témoins ou de 3 souris modèles de la pathologie SMA par RT-qPCR. Les taux d'expression des ARNm ont été normalisés par ceux de HPRT et de LDHA. (C) Comparaison des taux d'expression des ARNm de sélénoprotéines dans les souris SMA avec les résultats d'IP m_3G et de liaison à GST-eIF4E obtenus dans les cellules HEK293FT (voir chapitre I). La représentation « heat map » a été générée par le logiciel MeV. L'échelle est représentée sur la gauche. Rouge : fort signal d'interaction ou fort taux d'expression. Vert : faible signal d'interaction ou faible taux d'expression.

De manière intéressante, les ARNm de sélénoprotéines les plus sensibles à l'absence de SMN correspondent à ceux qui portent une coiffe hyperméthylée m₃G (Wurth et al., 2014), comme on le voit dans l'analyse comparative des taux d'expression des ARNm de sélénoprotéines dans les souris SMA avec les résultats d'IP m₃G et de liaison à GST-eIF4E obtenus dans les cellules HEK293FT (Figure 54C). Une excellente corrélation est ainsi observée entre les ARNm de la classe 1, c'est à dire ceux qui portent une coiffe m₃G et les ARNm sous-exprimés dans les cellules de moelle épinière des souris SMA. Ces résultats indiquent que SMN semble influencer sur le niveau des ARNm de sélénoprotéines dans les cellules de moelle épinière de souris, sans doute du fait de son rôle dans le recrutement de l'enzyme Tgs1 et l'hyperméthylation de leur coiffe.

II.2.4 A la recherche du rôle du Méthylosome

Le méthylosome catalyse la méthylation des arginines de protéines Sm, augmentant leur affinité pour le complexe SMN et facilitant leur assemblage sur les snARN. Nous avons précédemment montré que les protéines Sm ne sont pas associées aux ARNm de sélénoprotéines (Résultats I.I.3). Ceci indique que le rôle du méthylosome serait différent dans le cas de l'assemblage des mRNP de sélénoprotéines, de celui des snRNP. PRMT5 est responsable de la méthylation de diverses protéines (Branscombe et al., 2001), impliquant l'enzyme dans une variété de processus cellulaires (Karkhanis et al., 2011) incluant la biogénèse des ribosomes (Ren et al., 2010), l'assemblage de l'appareil de Golgi (Zhou et al., 2010) ou encore la différenciation cellulaire (Ancelin et al., 2006; Dacwag et al., 2007, 2009). SBP2 ne contient pas de répétition de séquence RG comme les protéines Sm mais possède 3 sites potentiels de méthylation (RG) en positions 120, 176 et 494 (représentés sur la Figure 55A). Nous avons donc testé si SBP2 pouvait être la cible de PRMT5 *in vitro*. Pour cela, nous avons incubé les protéines hSBP2 et hSBP2 515-854 en présence de PRMT5, MEP50 et de SAM (S-Adénosyl-Méthionine) tritiée. Le dimère d'histone H3/H4 a été utilisé comme contrôle positif puisqu'il est substrat de PRMT5. Nos résultats révèlent que SBP2 n'est pas méthylée par PRMT5/MEP50 *in vitro* (Figure 55B). Puisque les protéines pICln ou Riok1, protéines adaptatrices du méthylosome, pourraient être requises pour la méthylation de SBP2, il serait intéressant de tester la méthylation *in vitro* de SBP2 en présence de ces deux protéines. La nucléoline pourrait être la cible

du méthylosome dans notre système. La protéine hnRNP A1, que nous avons retrouvé associée à SBP2 et aux mRNP de sélénoprotéines (voir résultats II.1 et III.2) pourrait également constituer une cible du méthylosome.

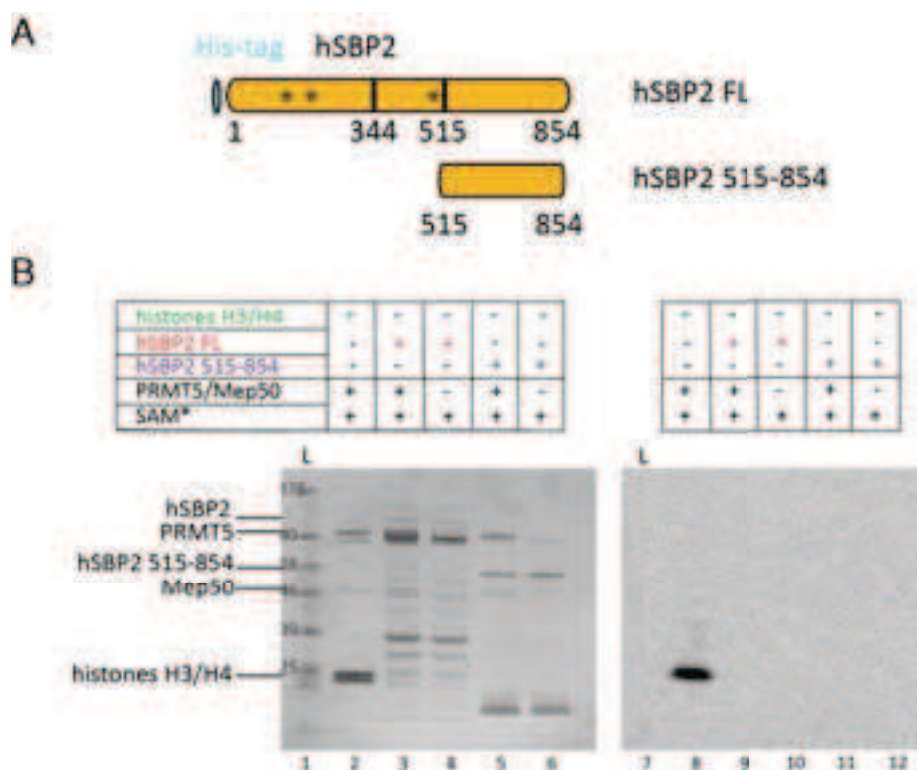


Figure 55. Test de méthylation in vitro de hSBP2 et hSBP2 515-854 par la méthyltransférase PRMT5.

(A) Représentation schématique des deux formes de hSBP2 produites dans la souche BL21 (DE3) RIL utilisées pour les tests de méthylation. Les protéines ont été purifiées sur colonne Ni-NTA. Les étoiles (*) représentent les 3 séquences RG, sites potentiels de méthylation. (B) Les protéines sont incubées 30 min à 37°C en présence de PRMT5/Mep50 et de SAM ^{14}C . Les protéines sont analysées sur gel SDS-PAGE, coloré au bleu de Coomassie (panneau de gauche). Le signal radioactif est révélé au Bioimager (panneaux de droite). Le dimère d'histones H3/H4 est le contrôle positif.

Nous avons montré que les complexes SMN et du méthylosome interagissent avec la protéine SBP2 et l'inhibition de leur composants protéiques a un impact sur la synthèse des sélénoprotéines. Ils sont donc vraisemblablement impliqués dans l'assemblage des mRNP de sélénoprotéines. Le complexe SMN permet le recrutement de l'enzyme Tgs1 vers les ARNm et donc l'hyperméthylation de leur coiffe, qui stimulerait leur traduction. Le méthylosome quant à lui pourrait être impliqué dans la méthylation d'une protéine constitutive de la RNP que nous devons identifier.

III Etude du mécanisme de traduction des ARNm de sélénoprotéines

III.1 Objectifs

De nombreuses études ont contribué à l'élucidation du mécanisme d'incorporation de la sélénocystéine dans les sélénoprotéines et à la compréhension du rôle de facteurs co-traductionnels ou d'assemblage spécifiques impliqués dans ce processus. Les ARNm de sélénoprotéines partagent des éléments caractéristiques bien spécifiques, le codon UGA^{Sec} et l'élément SECIS dans la région 3'UTR (Allmang et al., 2009; Cléry et al., 2007; Fagegaltier et al., 2000a; Tujebajeva et al., 2000), dont la fonction dans le mécanisme de recodage est connu. Cependant divers aspects du mécanisme de traduction des ARNm de sélénoprotéines restent encore incompris et le déficit majeur est de comprendre comment les éléments *cis* et *trans* essentiels au mécanisme de recodage s'inscrivent dans le contexte général de la traduction. En effet, l'impact de ces éléments sur le mécanisme d'initiation de la traduction n'a pas été étudié.

Au cours de ma thèse nous avons montré que la présence d'une coiffe hyperméthylée m₃G (Wurth et al., 2014) empêche les ARNm de sélénoprotéines d'être reconnus par le facteur d'initiation de la traduction eIF4E. Ainsi, il semblerait que ces ARNm présentent un mécanisme d'initiation de la traduction non-canonique. Notre objectif est de comprendre ce mécanisme.

Dans ce chapitre, nous nous sommes focalisé sur deux aspects : l'impact fonctionnel de la coiffe m₃G sur l'initiation de la traduction des ARNm de sélénoprotéines et l'analyse du mécanisme d'assemblage des complexes de traduction sur les ARNm de sélénoprotéines. Nous avons choisi comme modèles les ARNm de la sélénoprotéine M (SelM) et de la glutathion peroxydase (GPx1), auxquels nous avons pu spécifiquement ajouter une coiffe m⁷G ou m₃G.

L'étude de l'assemblage de complexes traductionnels sur ces ARNm a été réalisée *in vitro*, en présence ou non d'inhibiteurs de la traduction, afin d'analyser l'étape d'initiation de la traduction. Nous avons également tenté de localiser le ribosome à l'étape d'initiation et durant le recodage sur l'ARNm par des expériences de toeprint. L'ensemble de nos résultats montre que les ARNm de sélénoprotéines sont soumis à un mécanisme d'initiation de la traduction atypique, qui pourrait faire appel à d'autres

facteurs d'initiation. Afin d'identifier de nouveaux facteurs potentiels impliqués dans ce mécanisme, j'ai mis au point une méthode permettant de purifier les complexes de traduction assemblés *in vivo* sur l'ARNm de HA-GPx1. L'analyse par spectrométrie de masse de ces complexes a permis d'obtenir des candidats potentiels. Enfin, nous avons initié un projet d'analyse structurale par cryo-microscopie électronique de complexes de traduction assemblés sur l'ARNm de SelM.

Notre but est de comprendre le mécanisme de traduction des ARNm de sélénoprotéines dans sa globalité et d'y intégrer le rôle de la coiffe m₃G, du codon UGA^{Sec} et de l'élément SECIS.

III.2 Résultats

III.2.1 L'initiation de la traduction des sélénoprotéines est atypique

III.2.1.1 Analyse du rôle de la coiffe m₃G sur la traduction des ARNm de sélénoprotéines *in vitro*

III.2.1.1.a La coiffe m₃G inhibe le recrutement des ribosomes sur l'ARNm de SelM *in vitro*

Nous avons précédemment montré que l'inactivation de Tgs1 par siARN inhibe la synthèse *in vivo* de la sélénoprotéine HA-GPx1 (Wurth et al., 2014). Ceci suggère que Tgs1 joue un rôle fonctionnel dans la synthèse des sélénoprotéines *in vivo*, vraisemblablement de part son activité d'hyperméthylation de la coiffe. De plus, les ARNm de sélénoprotéines coiffés par du m₃G sont enrichis dans les fractions contenant les polysomes, indiquant qu'ils sont traduits (Wurth et al., 2014). Nous avons donc voulu analyser l'impact de la coiffe m₃G sur la traduction des sélénoprotéines de manière plus directe, *in vitro*. La majorité des études du mécanisme de traduction des sélénoprotéines sont basées soit sur l'utilisation de systèmes rapporteurs, comme la luciférase comprenant un codon UGA^{Sec} dans la phase codante et un élément SECIS dans la région 3'UTR (Gupta et al., 2013; Taylor et al., 2015), soit sur l'utilisation d'ARNm de sélénoprotéines dépourvus de la région 5'UTR (Takeuchi et al., 2009). Nous avons cloné les ARNm de GPx1 et SelM entiers incluant les régions 5' et 3'UTR naturelles afin de n'exclure aucun élément de séquence ou de structure de l'ARNm, potentiellement important pour leur traduction (Figure 56A). Les différents ARN utilisés lors des expériences d'assemblage de

complexes de traduction sont représentés sur la Figure 56. La synthèse des sélénoprotéines SelM et GPx1 entières requiert le recodage traductionnel du codon UGA^{Sec}. En effet, le recrutement de SBP2 et de la machinerie de recodage sur l'élément SECIS en 3'UTR des ARNm de SelM et GPx1 permet d'incorporer la sélénocystéine en réponse au codon UGA^{Sec} et empêche la terminaison prématurée de la traduction au niveau de ce codon. L'expression des mutants SelMCys et GPx1Cys, dans lesquels le codon UGA^{Sec} a été remplacé par un codon UGU^{Cys} codant pour la cystéine, ne nécessite pas de mécanisme de recodage. Enfin, les mutants SelMΔSECIS et GPx1ΔSECIS, délétés de l'élément SECIS, ne permettent pas le recrutement de la machinerie de recodage. Leur traduction se termine au niveau du codon UGA^{Sec} et il n'y a pas synthèse de la protéine entière (Figure 56).

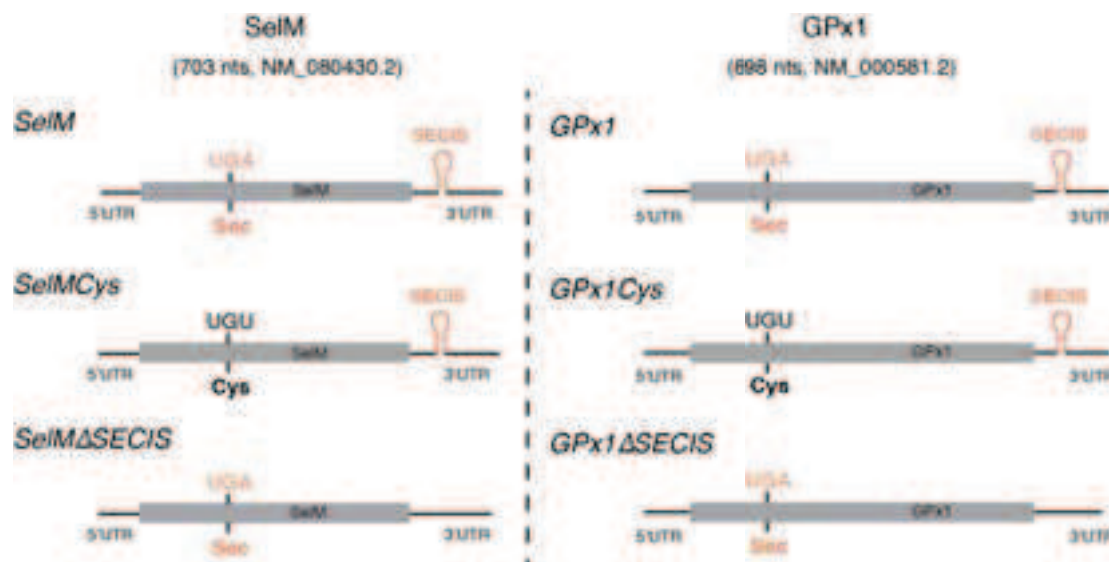


Figure 56. ARNm de SelM et de GPx1 utilisés comme modèles pour l'étude des mécanismes de traduction *in vitro*

Représentation des différents transcrits d'ARNm SelM, GPx1, SelMCys, GPx1Cys, SelMΔSECIS et GPx1ΔSECIS utilisés dans les expériences *in vitro*. Les traits représentent les UTR des ARNm, les rectangles gris les phases codantes de SelM ou GPx1, le codon UGA^{Sec} ou son mutant cystéine sont indiqués ainsi que la position de l'élément SECIS.

La traduction de ces ARNm synthétisés *in vitro* et coiffés par du m⁷G ou du m₃G de manière co-transcriptionnelle a été testée dans des extraits de réticulocytes de lapins (RRL) en présence de méthionine ³⁵S afin de détecter les protéines néosynthétisées. Ces extraits contiennent tous les composants nécessaires à la traduction protéique, ils sont cependant limitants en protéine SBP2, essentielle à la synthèse des sélénoprotéines et qui doit y être ajoutée (Copeland et al., 2000; Low et al., 2000b; Takeuchi et al., 2009).

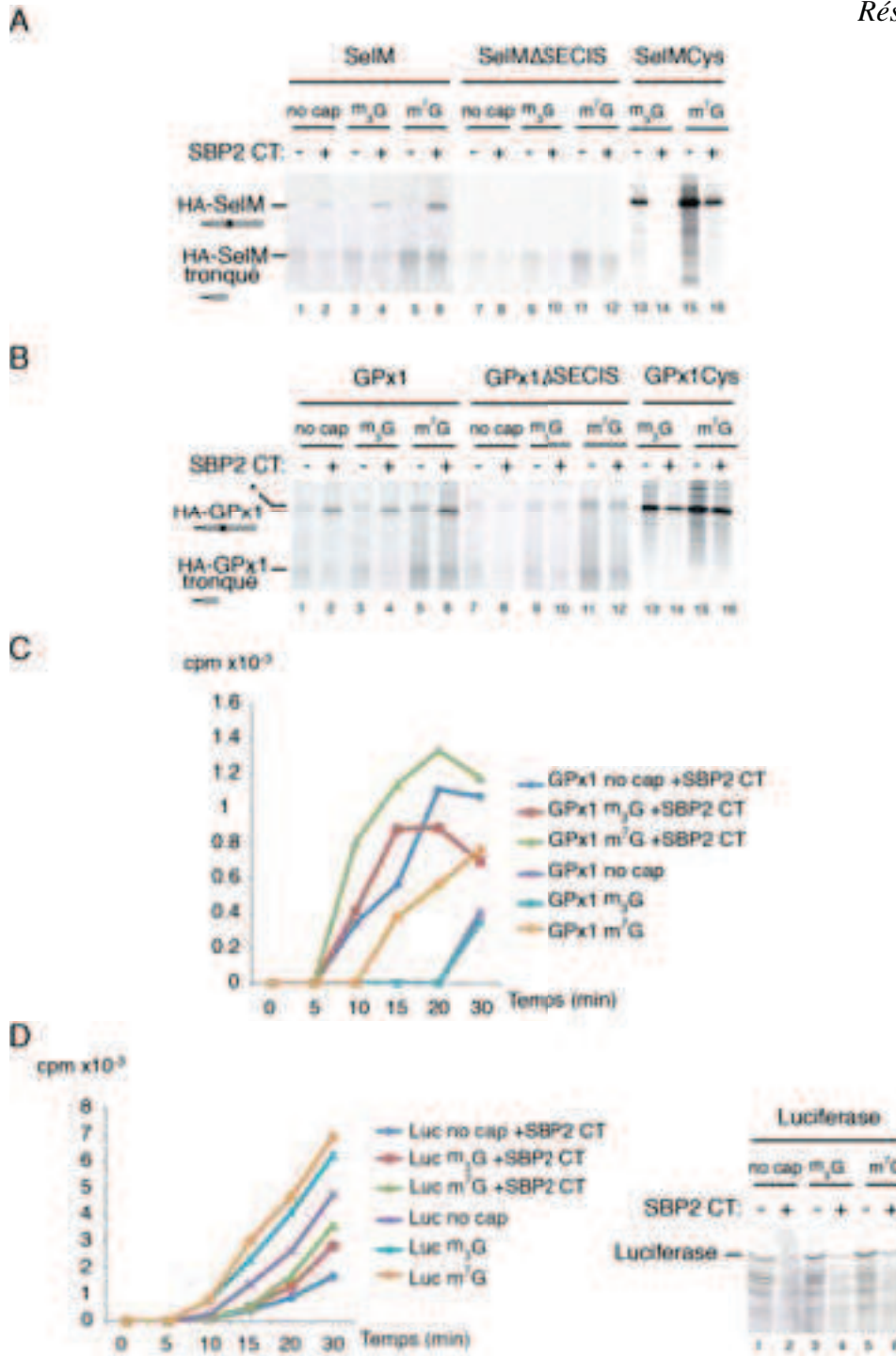


Figure 57. La traduction *in vitro* des ARNm de SelM et de GPx1 coiffés par du m₃G est SBP2-dépendante mais peu efficace.

(A) Traduction *in vitro* des ARNm HA-SelM, HA-SelMCys et HA-SelMΔSECIS dans les RRL ou (B) des ARNm HA-GPx1, HA-GPx1Cys et HA-GPx1ΔSECIS non-coiffés (no cap), coiffés par du m⁷G ou du m₃G avec ajout (+) ou non (-) de la protéine SBP2 CT dans les RRL, en présence de méthionine ³⁵S. Les produits de traduction ont été purifiés à l'aide de billes couplées à un anticorps anti-HA, puis déposés sur un gel SDS-PAGE 15%. Les sélénoprotéines entières ou tronquées sont indiquées.

* représente un produit de traduction aberrant d'origine inconnue, il pourrait résulter d'un changement de cadre de lecture au cours de la traduction. (C) Cinétique de traduction des ARNm HA-GPx1 avec ou sans coiffe, en présence ou non de SBP2 CT. Des aliquots de milieu traductionnel ont été prélevés et précipités au TCA aux différents temps indiqués. La radioactivité incorporée dans les protéines a été mesurée à l'aide d'un compteur à scintillation liquide et exprimée en nombre de coups par minute (cpm) x 10⁻³. (D) Traduction *in vitro* et cinétique de traduction de l'ARNm de la Luciferase avec ou sans coiffe, en présence ou non de SBP2 CT.

Dans toutes les expériences, nous avons utilisé la protéine SBP2 344-854, une version tronquée de SBP2 dont la production est plus efficace dans *E. coli* que celle de la protéine entière (1-854). SBP2 344-854 lie le SECIS avec la même affinité que SBP2 entière et permet la même efficacité de recodage du codon UGA^{Sec}, nous l'appellerons SBP2 CT pour simplification (Lescure et al., 2002b). Pour pouvoir être visualisés les produits de traduction, et notamment ceux soumis au recodage, sont obtenus à partir d'ARNm codant pour une protéine de fusion présentant une étiquette HA en N-terminal. Les protéines marquées à la méthionine ³⁵S dont la traduction est très inefficace sont ainsi purifiées à l'aide d'anticorps anti-HA (Takeuchi et al., 2009). Les mutants SelMCys et GPx1Cys, non soumis au recodage sont traduits indépendamment de la présence de SBP2, qu'ils soient coiffés par du m⁷G (Figure 57A et B, pistes 15-16) ou du m₃G (Figure 57A et B, pistes 13-14). La délétion de l'élément SECIS (SelMΔSECIS et GPx1ΔSECIS) ne permet pas la synthèse de sélénoprotéines entières (Figure 57A et B, pistes 7 à 12) ce qui est en accord avec les données de la littérature (Berry et al., 1993). Comme attendu, la traduction de l'ARNm de SelM et de GPx1 est strictement dépendante de la présence de la protéine SBP2 CT (Figure 57A et B, pistes 2, 4 et 6). Ceci est également le cas en absence de coiffe (Figure 57A et B, pistes 2) ou en présence d'une coiffe m₃G (Figure 57A et B, pistes 4), le taux de traduction est cependant plus faible qu'en présence d'une coiffe m⁷G. Une analyse cinétique de l'activité de traduction des ARNm de GPx1 a été réalisée afin d'étudier l'impact de la nature de la coiffe sur l'efficacité de traduction (Figure 57C et D). Pour cela, l'ARNm de GPx1 sans coiffe ou portant une coiffe m⁷G ou m₃G, a été traduit dans les RRL en présence ou non de SBP2 CT. Toutes les 5 min une fraction du milieu réactionnel a été prélevée, précipitée sur filtre et la radioactivité liée à l'incorporation de méthionine ³⁵S a été mesurée. L'ARNm de la Luciférase a été utilisé comme contrôle non-sélénoprotéine (Figure 57D). Dans la phase initiale, la traduction des ARNm de GPx1 et de la Luciférase est plus efficace lorsque les ARNm sont coiffés par du m⁷G, plutôt que du m₃G ou en absence de coiffe. En effet la traduction des ARNm de GPx1 m₃G et non coiffés est négligeable lors des 20 premières minutes de la réaction et les premiers produits de traduction n'apparaissent qu'au bout de 30 minutes. Ceci est en accord avec des résultats obtenus par la même méthode pour la β-globine (Darzynkiewicz et al., 1988). Cependant, l'ajout de la protéine SBP2 CT augmente très significativement

l'efficacité initiale de traduction des ARNm de GPx1, qu'ils soient coiffés par du m⁷G, du m₃G ou non-coiffés (Figure 57C). Après 15 minutes en présence de SBP2 CT, la traduction de l'ARNm GPx1 m⁷G est augmentée de 3X. Pour les ARNm de GPx1 m₃G et non coiffés, ce taux est largement surpassé, puisque d'aucun produit de traduction en absence de SBP2, on arrive à une quantité de protéine GPx1 synthétisée comparable à celle observée pour les ARNm coiffés par du m⁷G. Ceci n'est pas le cas pour les ARNm de la Luciférase, au contraire, la présence de SBP2 CT diminue leur taux de traduction en phase initiale de 6, 5 ou 3X pour les ARNm luciférase coiffés par du m⁷G, m₃G et non coiffés respectivement (Figure 57D). Ceci est peut-être lié à la capacité de SBP2 à se fixer aux ribosomes (Caban et al., 2007; Copeland et al., 2001; Kossinova et al., 2014; Takeuchi et al., 2009) les rendant potentiellement indisponibles pour la traduction de l'ARNm de la luciférase. Il est tentant de proposer que la présence d'une coiffe m₃G sur les ARNm de sélénoprotéines participe à un mécanisme de régulation de leur traduction. En absence de SBP2, la traduction de l'ARNm de sélénoprotéine serait complètement inhibée par la coiffe m₃G. La présence de SBP2, leverait cette inhibition permettant une traduction spécifique par des ribosomes directement compétents pour le mécanisme de recodage du codon UGA^{Sec}.

Pour conclure, la coiffe m₃G permet la traduction des ARNm de sélénoprotéines GPx1 et SelM dans les lysats de réticulocytes, mais avec une efficacité plus faible que la coiffe m⁷G. Deux hypothèses se présentent quant à cette différence d'efficacité : la coiffe m₃G inhibe le recrutement du ribosome au niveau de l'ARNm, ou elle ralentit le processus d'élongation de la traduction. Pour répondre à cette question, nous avons analysé la capacité des ARNm de SelM coiffés par du m₃G ou du m⁷G à s'associer aux ribosomes. Pour cette étude, notre équipe a mis au point des conditions permettant d'hyperméthyliser et de marquer au ³²P la coiffe des ARNm de sélénoprotéines *in vitro* de manière quantitative à l'aide de l'enzyme Tgs1 purifiée (voir chapitre Matériel et méthodes et Figure 58A). Ces ARNm ont été incubés avec des RRL, et la formation des complexes de traduction a été analysée par fractionnement sur gradient de sucrose (Figure 58B). Nous nous focaliserons principalement sur les résultats obtenus avec l'ARNm de SelM, certains résultats obtenus pour l'ARNm de GPx1 sont représentés dans les figures 61 et 68. Il apparaît

que les ARNm de SelM coiffés par du m_3G sont très faiblement pris en charge par les ribosomes comparés à ceux coiffés par du m^7G . La coiffe m_3G ne permettrait donc pas le recrutement du ribosome vers l'ARNm. L'ajout de SBP2 CT n'a aucun effet sur le profil de polysome (Figure 58). Ceci est en contradiction avec nos résultats précédents (Figure 58) qui indiquent qu'en présence de SBP2 CT, les ARNm qui portent une coiffe m_3G peuvent être traduits dans les lysats de réticulocytes, donc pris en charge par les ribosomes. Ceci est également en désaccord avec des résultats de Fletcher et al., 2000 qui montrent que l'ajout de SBP2 modifie légèrement le profil de polysome dans le cas de l'ARNm de GPx4 en permettant la fixation d'un ribosome supplémentaire sur l'ARNm. Les différences majeures entre ces expériences sont les temps d'incubation de l'ARNm avec les RRL qui sont respectivement de 5 minutes, 25 minutes (Fletcher et al., 2000) ou 2 h (traduction *in vitro*). Cependant, au bout de 25 minutes, les ribosomes sont majoritairement dissociés de l'ARNm dans nos conditions expérimentales et ne permettent plus l'analyse de l'effet de l'ajout de SBP2 (données non montrées).

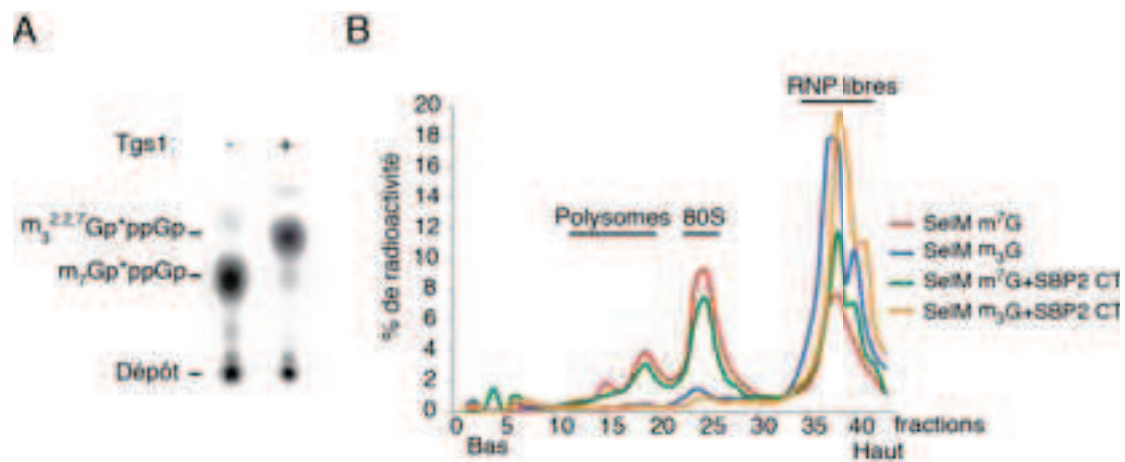


Figure 58. Les ARNm de SelM coiffés par du m_3G recrutent peu de ribosomes *in vitro* dans les RRL

(A) Hyperméthylation de la coiffe de l'ARNm de SelM par Tgs1. L'ARNm de SelM marqué à l' $[\alpha^{32}P]$ -GTP au niveau de sa coiffe a été incubé ou non en présence de l'enzyme Tgs1. Afin de vérifier l'efficacité de la modification par Tgs1, les ARNm ont été digérés par la nucléase S1, les dinucléotides résultants ont été séparés par CCM sur une plaque PEI-cellulose. p* représente le phosphate marqué.

(B) Assemblage de complexes de traduction sur les ARNm SelM coiffés par du m^7G ou du m_3G dans les RRL en présence ou absence de SBP2 CT. Les extraits traductionnels ont été fractionnés sur un gradient linéaire de sucrose 7-47%. La radioactivité présente dans chaque fraction a été mesurée par comptage Cerenkov et exprimée en pourcentage de nombre de coups total. La position des fractions contenant les polysomes, les complexes 80S et les RNP libres est indiquée sur le graphique.

Par ailleurs, il est possible que d'autres facteurs nécessaires à la synthèse ou la régulation de l'expression des sélénoprotéines soient également limitants dans les

RRL. L'assemblage de complexes de traduction sur les ARNm de SelM et SelMCys coiffés par du m₃G ou m⁷G a donc été testé dans des extraits de traduction réalisés à partir de cellules HeLa. Quelque soit la nature de la coiffe, les ARNm de SelM et SelMCys ne sont pas associés aux ribosomes dans les extraits de cellules HeLa (données non-présentées). L'expression des sélénoprotéines varie selon le type cellulaire et est tissu-spécifique (Behne et al., 1988; Schomburg and Schweizer, 2009). Notre objectif sera, à plus long terme, de tenter d'optimiser les extraits de traduction cellulaire par l'ajout, ou la surexpression préalable dans les cellules, de SBP2 ou d'autres composants de la machinerie de recodage traductionnel. Nous pourrions également rechercher un type cellulaire plus adapté à l'expression des sélénoprotéines. Différents types cellulaires, incluant des lignées cellulaires d'hépatoblastome (HEPG2) (Dreher et al., 1997) et de mésothéliome (MSTO-211H) (Curcio et al., 2001), riches en sélénoprotéines, seront testés pour leur capacité à générer des extraits de traduction compatibles avec l'expression des sélénoprotéines. Il a été suggéré que l'assemblage des mRNP de sélénoprotéines se déroulerait dans le noyau (Allmang et al., 2009; de Jesus et al., 2006; Papp et al., 2006; Small-Howard et al., 2006). Il est donc également possible qu'un ARN transcrit *in vitro* ne soit pas mûré ou assemblé de la même manière qu'il ne le serait *in vivo*, ce qui pourrait empêcher sa prise en charge par la machinerie traductionnelle. Une meilleure compréhension du mécanisme d'assemblage et de maturation des ARNm de sélénoprotéines est requise pour la mise en place d'un meilleur système d'étude de la coiffe m₃G *in vitro*. En conclusion nous avons montré que la coiffe m₃G inhibe le recrutement du ribosome sur les ARNm de sélénoprotéines, mais que la présence de SBP2 stimule préférentiellement la traduction des ARNm contenant un codon UGA^{Sec} dans les lysats de réticulocytes de lapins quelque soit la nature de leur coiffe.

III.2.1.1.b La coiffe m⁷G des ARNm en cours de traduction est accessible et son hyperméthylation n'induit pas la dissociation des ribosomes

Nous avons montré précédemment que les ARNm de sélénoprotéines coiffés par du m₃G *in vivo* sont présents dans le cytoplasme et enrichis dans les fractions contenant les polysomes, indiquant qu'ils sont traduits (Wurth et al., 2014). Cependant, pour l'instant, aucune donnée ne nous a permis de définir à quelle étape avait réellement

lieu l'hyperméthylation de la coiffe de ces ARNm. Ainsi, il est envisageable que ce processus survienne sur des ARNm déjà en cours de traduction, après l'étape d'initiation par exemple. Ajoutée à ce stade, la coiffe m_3G pourrait éventuellement jouer un rôle dans le mécanisme de recodage du codon UGA^{Sec} . Puisque les ARNm de SelM coiffés par du m_3G ne semblent pas initier la traduction *in vitro*, nous avons analysé l'impact de l'hyperméthylation de la coiffe sur des ARNm déjà assemblés sous forme de complexes de traduction (Figure 59).

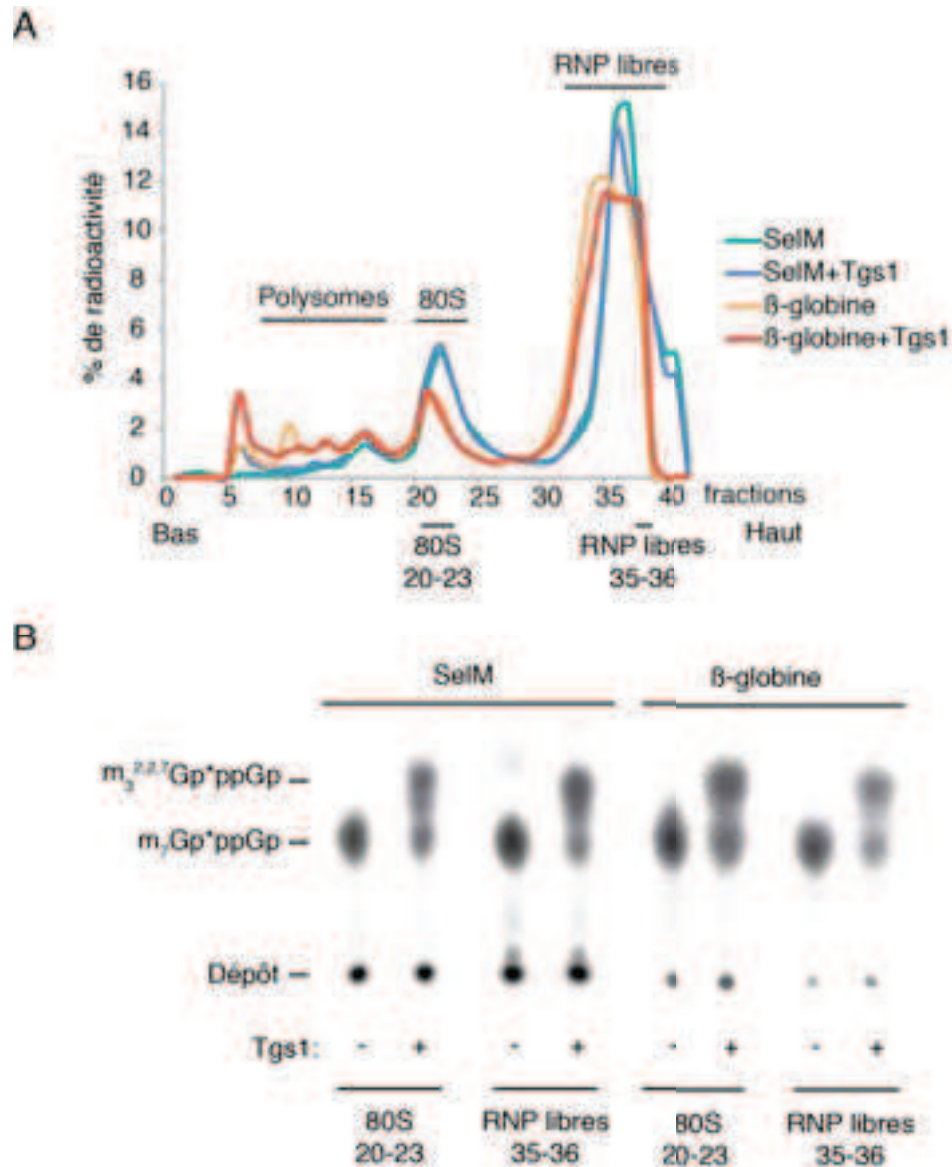


Figure 59. La coiffe m_7G des ARNm en cours de traduction est accessible et son hyperméthylation n'induit pas la dissociation des ribosomes.

Les ARNm de SelM et de la β -globine (contrôle non-sélénoprotéine), ont été marqués radioactivement au niveau de la coiffe m_7G et incubés 5 minutes avec des extraits RRL afin d'y assembler les complexes de traduction. Tgs1 a été, ajoutée (+) ou non (-) et l'incubation a été poursuivie pendant 10 minutes supplémentaires. (A) Les réactions d'assemblage ont été analysées comme décrit dans la Figure 58. (B) Analyse de la nature de la coiffe des ARNm libres ou associés aux ribosomes après action de Tgs1. Les ARNm extraits des fractions 35-36 (RNP libres) et des fractions 20-23 (complexe 80S) ont été digérés par la nucléase S1 et analysés par CCM, comme décrit Figure 58.

Ainsi, les ARNm de SelM et de la β -globine, choisi comme contrôle non-sélénoprotéine, ont été coiffés par du m⁷G et assemblés en présence d'extraits RRL. Les complexes ainsi formés ont été incubés en présence de Tgs1 pendant 10 minutes. Dans les deux cas, l'ajout de l'enzyme ne modifie pas le profil des polysomes (Figure 59A). Afin de vérifier l'impact de Tgs1 sur le statut de méthylation de la coiffe, les ARNm présents au sein des particules 80S (fractions 20-23) ou des RNP libres (fractions 35-36) ont été rassemblés et digérés à la nucléase S1. Les dinucléotides résultants ont été analysés par chromatographie sur couche mince (CCM). Il apparaît que la coiffe des ARNm libres mais également celle des ARNm des fractions 80S a été hyperméthylée par Tgs1 (Figure 59B). Ainsi, l'hyperméthylation de la coiffe est encore possible après la formation des complexes de traduction, ce qui indique que la coiffe reste accessible. Ceci est également le cas pour l'ARNm de la β -globine. La coiffe de tous les ARNm de la cellule pourrait être accessible durant leur traduction. En revanche leur hyperméthylation dépendrait de la capacité des ARNm à recruter ou non l'enzyme Tgs1. Il est possible que nos conditions *in vitro* ne garantissent pas cette spécificité. Cette expérience devra être réalisée avec des quantités plus faibles de Tgs1, mimant au mieux les conditions *in vivo*.

Puisque l'hyperméthylation de la coiffe des ARNm, après assemblage des complexes de traduction, n'induit pas la dissociation des ribosomes, on pourrait imaginer que Tgs1 soit recrutée vers la coiffe des ARNm de sélénoprotéines en cours de traduction. La coiffe m₃G ainsi formée, pourrait avoir un effet positif sur l'étape de recodage de ces ARNm, expliquant l'impact fonctionnel de Tgs1 dans la synthèse *in vivo* des sélénoprotéines (Wurth et al., 2014). D'après nos résultats, la coiffe m₃G ne permet pas l'initiation de la traduction dans les lysats de réticulocytes de lapin, mais pourrait être ajoutée ultérieurement et jouer un rôle dans le mécanisme de recodage.

III.2.1.2 Analyse du mécanisme d'initiation et de recodage de la traduction des ARNm de sélénoprotéines *in vitro*

Puisque nous avons montré qu'*in vivo* certains ARNm de sélénoprotéines peuvent porter les deux types de coiffes (Wurth et al., 2014), nous avons, en première approche, poursuivi l'étude détaillée de l'assemblage des complexes de traduction *in*

vitro sur les ARNm coiffés par du m⁷G. Nos objectifs sont de comprendre les mécanismes d'initiation et de recodage de la traduction des ARNm de sélénoprotéines.

III.2.1.2.a La présence du codon UGA^{Sec} induit l'accumulation de ribosomes 80S sur l'ARNm de SelM

Dans le but d'observer l'impact du codon UGA^{Sec} et de l'élément SECIS sur l'assemblage des complexes de traduction *in vitro*, les ARNm radiomarqués de SelM et des mutants SelMCys et SelMΔSECIS ont été incubés avec des extraits RRL et fractionnés sur gradient de sucrose (Figure 60). Il apparaît que le profil de polysome de l'ARNm SelM semble majoritairement bloqué sous forme de particule 80S et n'assemble que des monosomes ou des disomes, par comparaison au mutant SelMCys qui forme d'avantage de polysomes (Figure 60B), ce qui est tout à fait cohérent avec nos observations *in vivo* (Wurth et al., 2014). En effet, nous avons montré que la majorité des ARNm de sélénoprotéines présents dans les cellules HEK293FT sédimentent dans des fractions de poids moléculaire plus bas que des ARNm non-sélénoprotéines, indiquant qu'ils sont associés à un plus faible nombre de ribosomes. Ceci est vraisemblablement lié à une pause ou un blocage du ribosome au niveau du codon UGA^{Sec}. Il est également concevable que les ribosomes soient bloqués à l'initiation. En effet, la distance entre le codon UGA^{Sec} et le codon AUG initiateur (141nts) devrait permettre l'assemblage de plus d'un ou deux ribosomes. Des résultats similaires ont déjà été observés pour l'ARNm de la sélénoprotéine PHGPx (Fletcher et al., 2000). Dans ce cas, l'ajout de SBP2 a permis d'augmenter le nombre de ribosomes assemblés sur l'ARNm. Les auteurs ont suggéré que SBP2 augmente l'efficacité de la traduction en améliorant le décodage du codon UGA^{Sec}. L'ajout de la protéine SBP2 CT, nécessaire à la traduction de SelM dans les extraits de réticulocytes, ne modifie cependant pas le profil des polysomes dans nos conditions (Figure 58). De même, l'ajout de sélénium n'a eu aucun effet (données non présentées). L'ARNm du mutant SelMΔSECIS assemble moins de complexes que l'ARNm de SelM, ce qui indique que l'élément SECIS est non seulement essentiel au mécanisme de recodage (Copeland et al., 2000; Lescure et al., 2002d; Takeuchi et al., 2009), mais est également important pour le recrutement des ribosomes sur l'ARNm (Figure 60B).

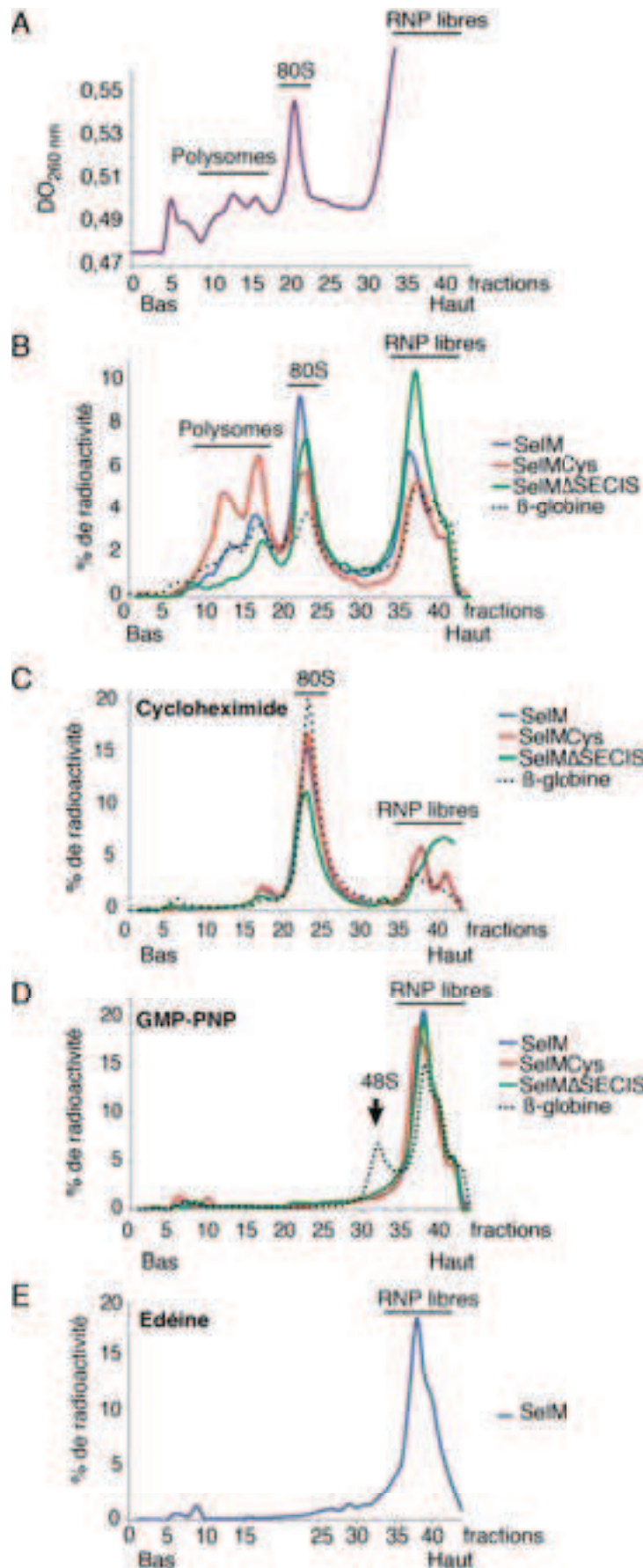


Figure 60. Analyse des complexes de traduction assemblés sur les ARNm de SelM coiffés par du m⁷G.

Les ARNm de SelM, SelMCys, SelMΔSECIS et de la β-globine (contrôle non-sélenoprotéine) marqués radioactivement au niveau de la coiffe m⁷G, sont incubés avec des extraits RRL afin d'y assembler les complexes de traduction. Les réactions d'assemblage ont été analysées de la même manière que décrit Figure 58. **(A)** Profil d'absorbance permettant de déterminer la position des polysomes, du complexe 80S et des RNP libres. **(B-E)** La radioactivité présente dans chaque fraction est quantifiée par comptage Cerenkov et exprimée en pourcentage de radioactivité. **(B)** Les complexes ont été formés en absence d'inhibiteurs de l'initiation de la traduction, **(C)** de cycloheximide, **(D)** de GMP-PNP, **(E)** d'édéine.

L'ensemble de ces observations a également été constaté pour l'ARNm de GPx1 (voir III.3, Figures 61A et B). Il est par ailleurs frappant de constater qu'un seul nucléotide de différence entre le codon UGA^{Sec} et le codon UGU^{Cys} présent au sein du mutant SelMCys ait un tel impact sur l'assemblage des polysomes. Il est également envisageable que le codon UGA^{Sec} constitue un signal essentiel au recrutement efficace de ribosomes actifs pour le recodage.

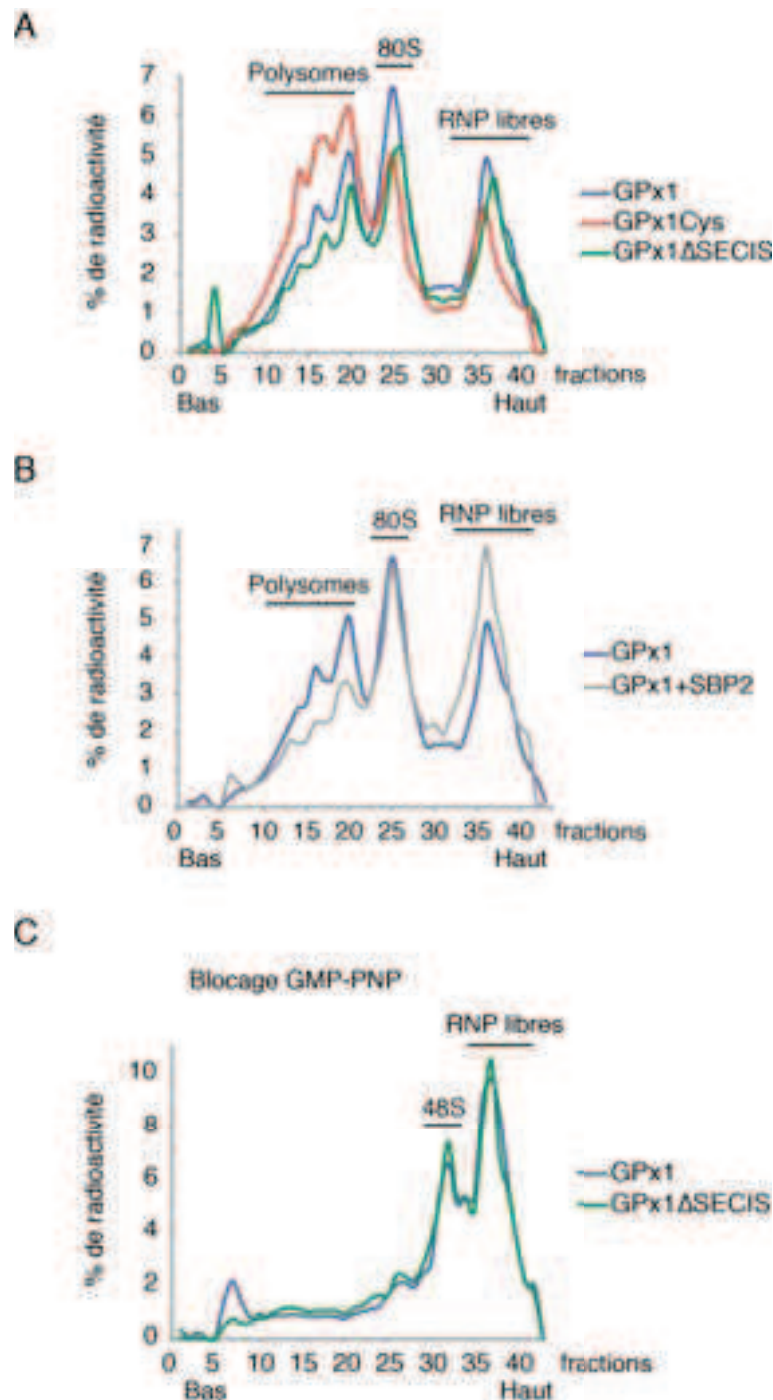


Figure 61. Analyse des complexes de traduction assemblés sur les ARNm de GPx1 coiffés par du m⁷G *in vitro*.

Les ARNm de GPx1, GPx1Cys et GPx1ΔSECIS, marqués radioactivement au niveau de la coiffe m⁷G, sont incubés avec des extraits de traduction afin d'y assembler les complexes traductionnels. Les réactions d'assemblage ont été analysées de la même manière que décrite Figure 58. (A), (B), (C) Assemblage dans les extraits RRL, (A) en absence ou (B) en présence de la protéine SBP2 CT, ou (C) en présence d'inhibiteur GMP-PNP.

III.2.1.2.b Un mécanisme atypique d'initiation de la traduction pour l'ARNm de SelM

Dans le but d'analyser le mécanisme d'initiation de la traduction des ARNm de sélénoprotéines, nous avons analysé la composition des complexes de traduction assemblés sur l'ARNm de SelM, SelMCys, SelMΔSECIS et de la β-globine en présence de cycloheximide, de GMP-PNP et d'édéine ; des inhibiteurs ciblant spécifiquement différentes étapes de l'initiation de la traduction. La cycloheximide bloque le ribosome dans l'étape de translocation, l'empêchant de passer en phase d'élongation. Lorsqu'elle est ajoutée avant l'étape d'assemblage des complexes de traduction, cela résulte en une accumulation de complexes 80S bloqués en phase d'initiation. Cet antibiotique lie le site E du ribosome et y bloque l'ARNt, empêchant les étapes de translocation ultérieures (de Loubresse et al., 2014; Schneider-Poetsch et al., 2010). Le GMP-PNP (5'Guanylyl-imidodiphosphate) entre en compétition avec le GTP pour la liaison au facteur d'initiation eIF5B. L'hydrolyse du GTP par eIF5B est nécessaire à l'association de la sous-unité ribosomique 60S. Ainsi, l'ajout du GMP-PNP avant l'étape d'assemblage de complexes de traduction abolit complètement la formation de complexes 80S et induit l'accumulation de complexes 48S (Hiremath et al., 1989; Locker et al., 2006; Wilson et al., 2000b). Enfin, l'édéine empêche l'interaction entre le codon de l'ARNm et l'anticodon de l'aminoacyl-ARNt (Dinos et al., 2004), rendant la formation de complexes ribosomes-ARNm impossible.

En présence de cycloheximide, comme attendu, les quatre ARNm ne sont plus assemblés sous forme de polysomes et il apparaît une accumulation de complexes ARNm-80S (Figure 60C). Il est intéressant de noter que le mutant SelMΔSECIS présente un léger défaut d'assemblage également en présence de cycloheximide, confirmant le rôle de l'élément SECIS dans le recrutement des ribosomes vers l'ARNm. Comme attendu, la présence de GMP-PNP empêche la formation des complexes 80S (Figure 60D). En revanche, le GMP-PNP empêche également la formation de complexe 48S pour les ARNm de SelM et des mutants SelMCys et SelMΔSECIS, contrairement à l'ARNm de la β-globine. Ces résultats sont révélateurs d'un mécanisme d'initiation non-canonique pour l'ARNm de SelM. Le GMP-PNP est un analogue du GTP et un inhibiteur compétitif des enzymes utilisant le GTP (les GTPases). Dans le mécanisme d'initiation de la traduction canonique, la formation du

complexe 48S requiert l'hydrolyse de GTP par eIF2. La molécule de GTP liée à eIF2, contrairement à celle associée à eIF5B, ne s'échange pas avec le GMP-PNP (Pestova et al., 2000). L'absence de complexes 48S sur les ARNm de SelM en présence de GMP-PNP pourrait indiquer, dans ce cas précis, que le mécanisme d'initiation requière l'activité d'une GTPase supplémentaire, c'est à dire différente de eIF2 ou eIF5B (Wilson et al., 2000b). Enfin, comme attendu, l'édéine empêche la formation de tout complexe de traduction sur les quatre ARNm, confirmant que l'interaction codon-anti-codon est nécessaire à l'initiation de la traduction de ces ARNm (Figure 60E). L'ensemble de ces résultats indique que le recrutement du ribosome au niveau de l'ARNm de SelM fait intervenir un mécanisme atypique.

Des résultats différents ont été observés pour l'ARNm de GPx1 (Figure 61C) qui est capable d'assembler des complexes 48S en présence de GMP-PNP. Ainsi cette singularité dans le mécanisme d'initiation observée pour l'ARNm de SelM pourrait être une spécificité unique à cet ARNm. Il est également envisageable que différents mécanismes d'initiation puissent exister pour les ARNm de sélénoprotéines. Ainsi, l'étape de formation du complexe 48S devra être analysée pour l'ensemble des ARNm de sélénoprotéines.

III.2.1.2.c Localisation du ribosome sur l'ARNm

Nous avons précédemment montré que les ARNm de sélénoprotéines endogènes sédimentent dans les fractions de poids moléculaires plus faibles que les ARNm classiques (Wurth et al., 2014), indiquant la présence d'un plus faible nombre de ribosomes associés à ces ARNm. De même, nous avons observé l'accumulation de ribosomes 80S sur l'ARNm de SelM alors que l'ARNm du mutant SelMCys est trouvé dans des polysomes. Ceci nous conduit à émettre l'hypothèse d'une pause du ribosome en cours d'élongation au niveau du codon UGA^{Sec}. Il a été proposé que cette pause soit une caractéristique cinétique de l'incorporation de sélénocystéine et que le mécanisme de décodage du codon UGA^{Sec} serait un processus lent (Fletcher et al., 2000; Stoytcheva et al., 2006b). Une étude de « ribosome profiling » réalisée chez la souris a montré qu'une pause du ribosome est présente plusieurs nucléotides en amont du codon UGA^{Sec} pour les ARNm de GPx1, SelW, SelX et SelH ; un supplément de

sélénium dans l'alimentation augmente considérablement la densité de ribosomes en amont du codon UGA^{Sec} (Howard et al., 2013).

Pour vérifier d'une part l'hypothèse d'une pause naturelle du ribosome sur l'ARNm de SelM lors de l'élongation et d'autre part le mode de recrutement du ribosome lors de l'initiation nous avons réalisé des expériences de toeprint. Celles-ci ont été réalisées sur l'ARNm de SelM coiffé par du m⁷G en présence d'extraits de réticulocytes de lapin, et l'arrêt de la reverse transcriptase (RT) correspondant au ribosome a été recherché dans les séquences avoisinant le codon initiateur AUG et le codon UGA^{Sec}. Les complexes de traduction, de haut poids moléculaire, sont enrichis dans les culots après ultracentrifugation, ce qui permet de les différencier des protéines de liaison à l'ARNm, qui sont à la fois présentes dans les culots et les surnageants. Afin de bloquer les ribosomes à l'initiation, nous avons utilisé des inhibiteurs de la traduction, la cycloheximide (voir ci-dessus) et l'hygromycine-B qui bloquent toutes deux l'étape de translocation du ribosome, empêchant ainsi l'entrée en élongation. L'hygromycine-B augmente légèrement l'affinité de l'ARNt pour le site A du ribosome, bloquant ainsi l'étape de translocation (Cabañas et al., 1978; Peske et al., 2004). Contrairement aux ribosomes, l'association des protéines de liaison à l'ARNm n'est pas sensible aux antibiotiques utilisés. On peut ainsi observer un certain nombre de pauses de la reverse transcriptase qui pourraient correspondre à la position de protéines provenant des extraits de réticulocytes indiqués par des étoiles.

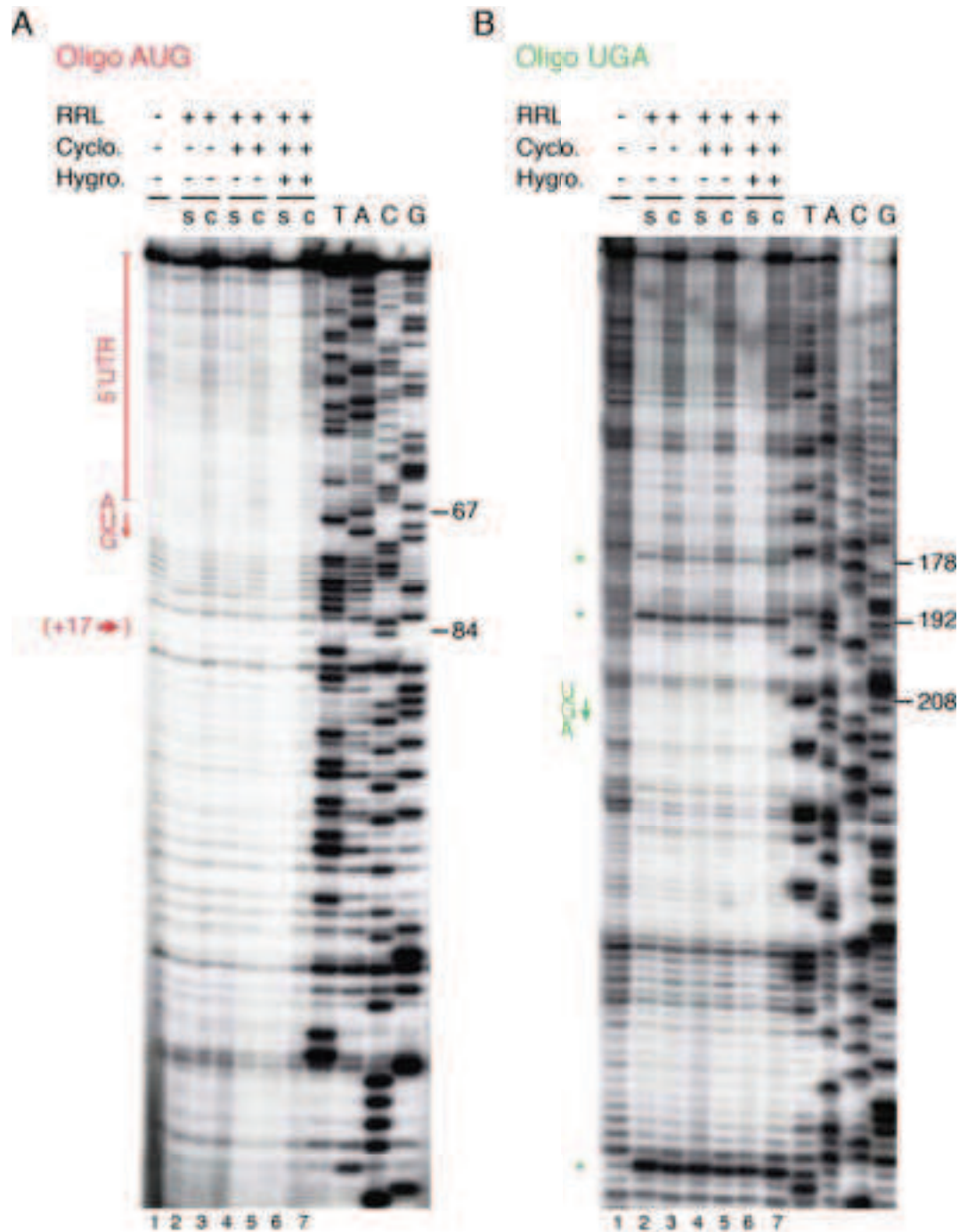


Figure 62. Localisation du ribosome sur l'ARNm de SelM.

Des expériences de toeprint ont été réalisées sur l'ARNm de SelM coiffé par du m^7G . Les complexes de traduction ont été assemblés dans des extraits RRL, en présence ou en absence de cycloheximide et d'hygromycine et fractionnés par ultracentrifugation en culot (C) et surnageant (S) afin de différencier les arrêts de la RT liés à la liaison de protéines provenant des RRL à l'ARN (marqués par des étoiles), des arrêts liés à la présence du ribosome sur l'ARN. Des oligonucléotides complémentaires à des régions en aval du codon initiateur AUG ou Oligo AUG (A) et du codon UGA^{Sec} ou Oligo UGA (B) ont été utilisés pour l'étape de RT. Les positions des nucléotides de l'ARN de SelM sont indiquées en noir, la position de la région 5'UTR, du codon d'initiation et du codon UGA^{Sec} sont indiqués en couleur. La position +17 correspond à la position habituelle d'un toeprint lorsque le codon initiateur AUG est localisé dans le site P du ribosome. TACG, représentent la réaction de séquençage de l'ARN.

Les expériences de toeprint n'ont révélé aucune pause du ribosome aux alentours du codon AUG, notamment en +17, site de pause habituel de la reverse transcriptase lorsque le codon AUG initiateur est dans le site P du ribosome (Figure 62A). Par ailleurs aucune pause de la RT liée à la présence du ribosome n'a été observée au niveau du codon UGA^{Sec} (Figure 62B). Ces résultats indiquent que l'accumulation de 80S, observée lors des expériences d'assemblage de complexes de traduction sur l'ARNm de SelM (Figure 60B), ne se ferait ni au niveau du codon initiateur, ni du codon UGA^{Sec}. Des résultats identiques ont été obtenus lors des expériences de toeprint sur l'ARNm de GPx1 (Figure 68A). Il est également possible que le complexe de traduction formé ne soit pas suffisamment stable pour pouvoir être analysé par toeprint, ou qu'il soit assemblé au niveau d'une structure de l'ARNm, induisant une pause de la RT que le ribosome soit présent ou absent. L'absence de pause du ribosome au niveau du codon initiateur AUG de l'ARNm, même en présence de cycloheximide ou d'hygromycine-B, pourrait également indiquer l'existence d'un mécanisme d'initiation de la traduction différent. Le ribosome pourrait être recruté ailleurs sur les ARNm de sélénoprotéines.

III.2.1.2.d La région 3'UTR de l'ARNm de SelM est impliquée dans le recrutement du ribosome

Plusieurs études ont montré que la région 3'UTR des ARNm peut être impliquée dans la régulation de leur traduction (De Moor et al., 2005; Nicholson and White, 2011). Ainsi, nous avons analysé le rôle de la région 3'UTR entière de l'ARNm de SelM sur l'assemblage des complexes de traduction. Tandis que la délétion de l'élément SECIS n'induit qu'une faible modification du profil de polysomes, la délétion de la totalité de la région 3'UTR abolit totalement la formation de 80S et tout l'ARNm est retrouvé dans les fractions de RNP libres (Figure 63A). Ces résultats indiquent que d'autres éléments présents dans la région 3'UTR, en dehors du SECIS, sont impliqués dans le recrutement des ribosomes vers l'ARNm de SelM.

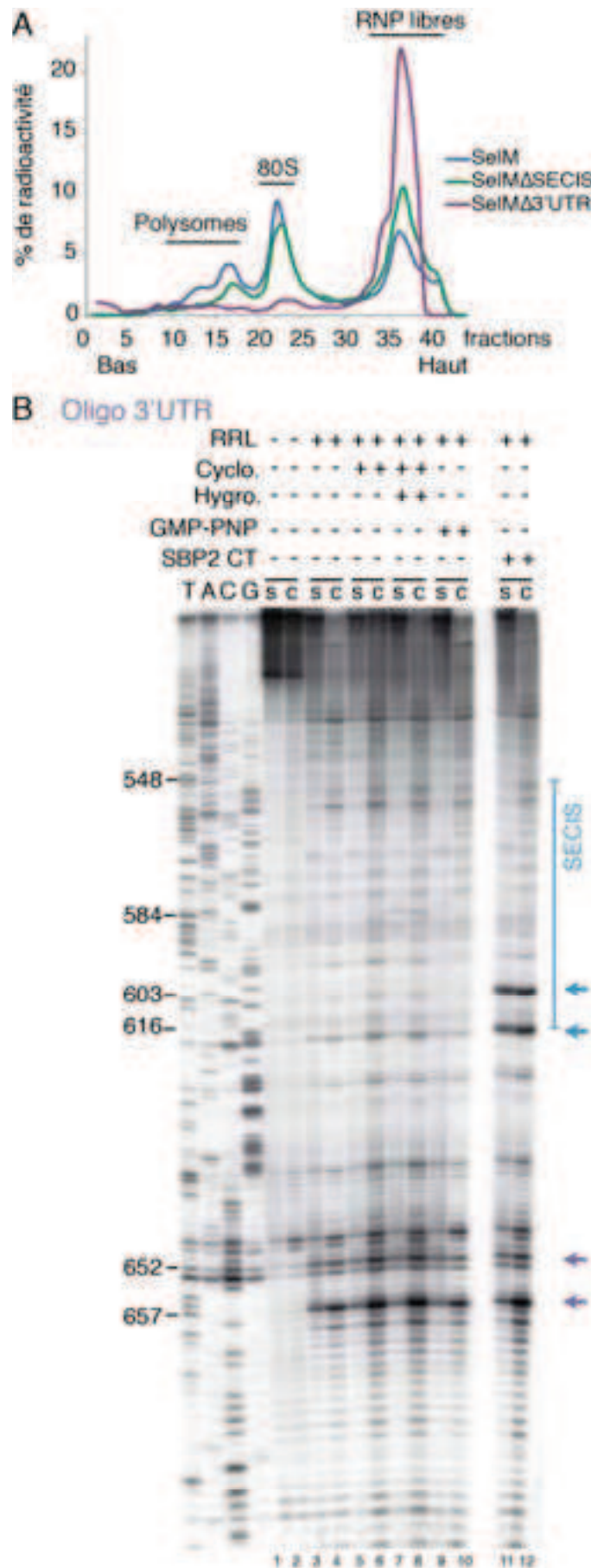


Figure 63. La région 3'UTR est requise pour l'assemblage des complexes de traduction sur l'ARNm de SelM.

(A) Assemblage de complexes de traduction sur les ARNm de SelM, de SelM Δ SECIS et de SelM Δ 3'UTR dans les RRL. Les extraits traductionnels ont été fractionnés comme décrit dans la Figure 58. (B) Expériences de toeprint sur l'ARNm de SelM coiffé par du m⁷G en absence ou en présence de cycloheximide, d'hygromycine, de GMP-PNP ou de SBP2 CT ; l'analyse a été réalisée au niveau de la région 3'UTR de l'ARNm en utilisant un oligonucléotide complémentaire à l'extrémité 3' de l'ARNm (Oligo 3'UTR) durant l'étape de RT. Les positions de pause de la RT liées à la présence de SBP2 CT sur l'ARNm de SelM sont indiquées par des flèches bleues et celles liées aux RRL, par des flèches violettes. Les gels sont annotés comme décrit Figure 62.

L'analyse de la région 3'UTR par des expériences de toeprint, révèle la présence de fortes pauses de la RT en aval de l'élément SECIS lorsque l'ARNm est assemblé dans les RRL (Figure 63B, pistes 3 à 10). Ces pauses sont plus intenses dans les culots qui contiennent généralement les ribosomes, que dans les surnageants (Figure 63B, comparer les pistes 4, 6, 8, 10 avec 3, 5, 7, 9) indiquant qu'il pourrait s'agir de complexes de haut poids moléculaires interagissant avec l'ARN. L'ajout d'antibiotiques destinés à inhiber les diverses étapes de la traduction ne modifie pas le profil des pauses (Figure 63B, pistes 5 à 10); il en va de même pour l'ajout de SBP2 CT (Figure 63B, pistes 11,12), indiquant qu'il ne s'agit pas de ribosomes. Comme attendu, en présence de SBP2 CT, des pauses de la RT sont observées au niveau de l'élément SECIS (Figure 63B, pistes 11-12, positions 603 et 616), indiquant que notre méthode d'extension d'amorce permet de détecter des interaction protéines-ARN.

Nos résultats indiquent que la région 3'UTR de l'ARNm de SelM joue un rôle important dans la traduction, probablement via l'interaction avec des facteurs ou des complexes, ou du fait de sa structure. De plus, l'élément SECIS ne serait pas le seul élément de la région 3'UTR dont la fonction est importante pour la synthèse des sélénoprotéines. Récemment, il a été montré que des structures en tige-boucle à proximité de l'élément SECIS dans la région 3'UTR de l'ARNm de SelS régulent l'incorporation de sélénocystéine au cours de la synthèse de SelS (Bubenik et al., 2013). Des éléments régulateurs ont également été retrouvés dans la phase codante de certains ARNm de sélénoprotéines, sous la forme de structures en tige-boucle situées quelques nucléotides en aval du codon UGA^{Sec} appelé SRE (pour Selenocysteine Redefinition Element) et seraient impliqués dans l'efficacité du mécanisme de recodage (Howard et al., 2005a, 2007b). La capacité de divers éléments des ARNm à influencer l'incorporation de sélénocystéine et la traduction des sélénoprotéines, souligne l'importance du contexte de séquence et de structure de l'ARN lors de l'étude des mécanismes traductionnels. Le logiciel RNA fold (<http://rna.tbi.univie.ac.at/>) prédit une structure en tige-boucle dans la région 3'UTR, en aval de l'élément SECIS, représentée Figure 64. Les pauses de la RT observées en présence de RRL sont localisées au pied de cette structure en tige-boucle, indiquant que celle-ci pourrait être impliquée dans une interaction avec des protéines ou complexes, éventuellement importants pour la traduction de l'ARNm. Des

expériences futures viseront à analyser l'impact de cette tige-boucle sur la traduction de l'ARNm de SelM.



Figure 64. Prédiction de structure dans la région 3'UTR.

L'élément SECIS et la structure en tige-boucle prédite dans la région 3'UTR de l'ARNm de SelM par le logiciel RNA fold sont représentés. Les positions des pauses de la RT (observées figure 63B) liées à la présence de SBP2 CT sont indiquées par des flèches bleues et celles induites par les extraits RRL, par des flèches violettes.

III.2.1.2.e La protéine SBP2 se lie en amont du codon UGA^{Sec} de l'ARNm de SelM

Puisque SBP2 est essentielle au recodage traductionnel de l'ARNm de SelM, des expériences d'élongation d'amorce ont été réalisées en présence de la protéine. Une pause de la RT en présence de SBP2 CT a étonnamment été observée 48 nucléotides (position 160) en amont du codon UGA^{Sec} (Figure 65, comparer pistes 7-8 avec 5-6 et Figure 66B, comparer pistes 9-10 avec 7-8). Cette pause est directement induite par SBP2 (Figures 65 et 66B, pistes 3-4), elle est absente dans les extraits de réticulocytes seuls (Figure 65, pistes 5-6 et Figure 66B, pistes 8-9). La même pause d'élongation en position 160 est observée sur les ARNm SelMCys et SelMΔSECIS (Figure 65, pistes 11-12, 15-16 et 19-20 et 23-24), indiquant qu'elle ne dépend ni de la présence du codon UGA^{Sec}, ni de l'élément SECIS. Une protéine tronquée SBP2 545-854, qui n'est plus capable de lier l'ARN SECIS (Takeuchi et al., 2009) a été utilisée comme contrôle (Figure 66A, pistes 5-6). La pause d'élongation située 48 nucléotides en amont du codon UGA^{Sec} est aussi observée avec la protéine SBP2 545-854 (Figure 66B, pistes 7-8 et 13-14). Ces résultats suggèrent que la protéine SBP2 est capable de se lier à l'ARNm de SelM dans une autre région que le SECIS et que cette interaction

ne nécessite pas la présence d'un domaine de liaison à l'ARN SECIS fonctionnel. Il serait intéressant de tester si le domaine L7AE, qui reste présent dans la protéine SBP2 545-854, pourrait être suffisant pour cette interaction.

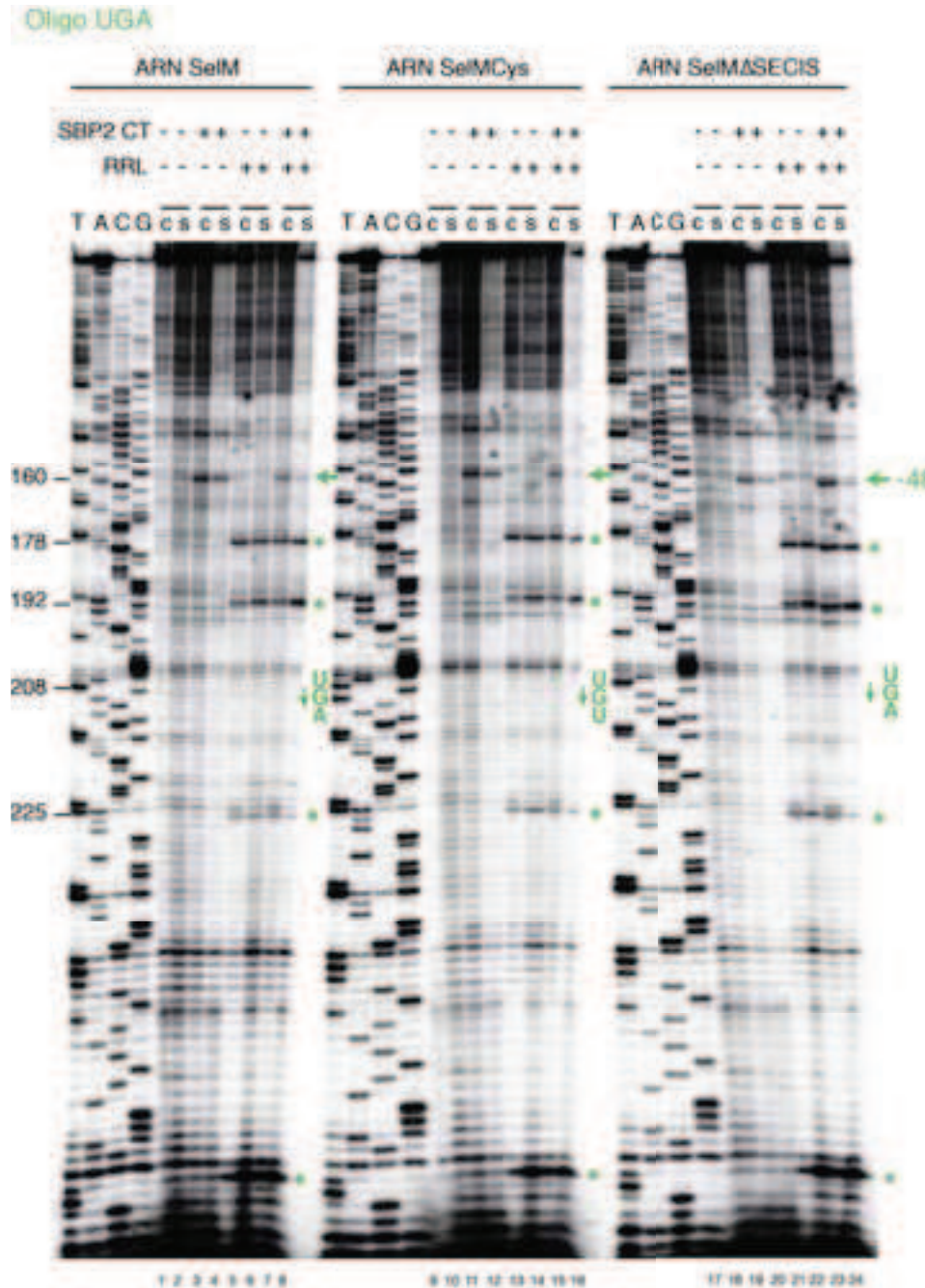


Figure 65. Empreinte de la protéine SBP2 en amont du codon UGA^{Sec} de l'ARNm de SelM.

Expérience de toeprint réalisée sur les ARNm de SelM, SelMCys et SelMΔSECIS coiffés par du m⁷G. Les complexes de traduction ont été assemblés en présence ou en absence d'extraits RRL et de la protéine SBP2 CT. La position de SBP2 CT a été recherchée dans la région entourant le codon UGA^{Sec}. Les annotations sont comme indiquées Figure 62. Les pauses de RT dues à SBP2 CT sont pointées d'une flèche verte. Les étoiles représentent la position des pauses de RT dues aux extraits RRL.

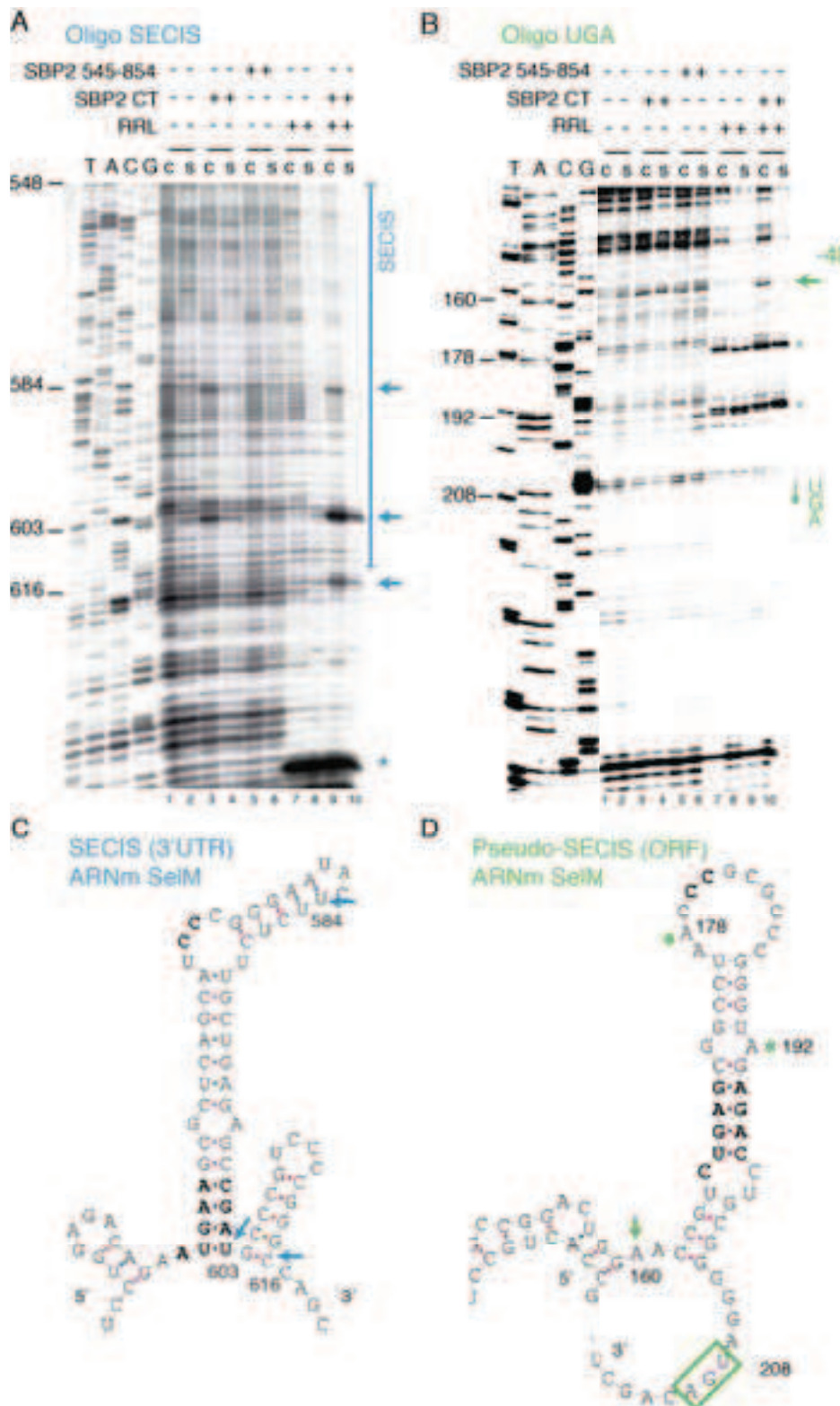


Figure 66. Empreinte de la protéine SBP2 dans les régions UGA^{Sec} et SECIS de l'ARNm de SelM et prédiction de structure

La position de SBP2 CT sur l'ARNm de SelM coiffé par du m⁷G a été recherchée par analyse toeprint dans la région SECIS en 3'UTR (A) et dans la région entourant le codon UGA^{Sec} (B). Les complexes de traduction ont été assemblés en présence ou en absence d'extraits RRL, de SBP2 344-854 (SBP2 CT) ou de SBP2 545-854. Les annotations sont expliquées Figure 62. (C) et (D) Prédiction de structure par le logiciel SECISaln de la région SECIS en 3'UTR (C) et de la région située en amont du codon UGA^{Sec} (D). La position des pauses de la RT liées à la présence de SBP2 sur l'ARNm de SelM est indiquée par des flèches. Les étoiles représentent la position des pauses de RT dues aux extraits RRL.

Nous avons recherché des éléments de structure secondaire dans la région entourant le codon UGA^{Sec} à l'aide des logiciels M-Fold (Zuker, 2003) (<http://mfold.rna.albany.edu/>), un outil de prédiction de la structure secondaire des ARN, et SECISaln (Chapple et al., 2009a) (<http://genome.crg.es/software/secisaln/>), qui recherche la présence d'éléments SECIS eucaryotiques dans une séquence nucléotidique. Le logiciel SECISaln a permis de prédire la présence d'une structure de type SECIS en amont du codon UGA^{Sec} de l'ARNm de SelM, que nous appellerons "pseudo-SECIS". Cette prédiction est représentée Figure 66D. La prédiction du SECIS en 3'UTR est représentée Figure 66C et diffère légèrement de la structure SECIS publiée (Korotkov et al., 2002b), représentée Figure 64) par une petite structure en tige-boucle à la place de la boucle-interne. La position des pauses de RT induites par la protéine SBP2 CT dans les expériences d'élongation d'amorce y sont reportées. Les pauses observées au niveau de l'ARN SECIS sont localisées au niveau du quartet de paires de bases non Watson-Crick et de la boucle apicale en accord avec les modèles d'interaction proposés précédemment (Allmang et al., 2002a; Cléry et al., 2007; Fletcher et al., 2001; Korotkov et al., 2002a; Takeuchi et al., 2009). La pause d'élongation due à SBP2 CT présente en amont du codon UGA^{Sec}, est localisée en 5' de la structure "pseudo-SECIS" prédite ce qui pourrait indiquer un mode d'interaction différent de celui décrit pour l'élément SECIS. En vue de confirmer la présence de cet élément "pseudo-SECIS" en amont du codon UGA^{Sec}, il sera nécessaire de déterminer la structure secondaire de l'ARNm de SelM par des méthodes d'analyses structurales en solution telles le sondage chimique et enzymatique (Jaeger et al., 2006) ou la méthode de SHAPE (Selective 2'-Hydroxyl Acylation analyzed by Primer Extension) (Merino et al., 2005). Dans le but de valider l'interaction de SBP2 avec l'élément "pseudo-SECIS" de l'ARNm SelM, des expériences de retard sur gel ont été réalisées à l'aide d'un transcrit ARN correspondant aux nucléotides 113-203 situé en amont du codon UGA^{Sec} (Figure 67). Ce transcrit contenant l'élément "pseudo-SECIS" (représenté Figure 67A), a été marqué au ³²P et incubé en présence de quantités croissantes de protéine SBP2 CT (Figure 67C). Alors que SBP2 CT est capable de se lier à l'ARN SECIS (K_d= 1,6 μM), elle ne se lie que faiblement à l'ARN "pseudo-SECIS". Le déplacement observé, bien que faible, est significatif et pourrait correspondre à une conformation minoritaire capable d'interagir avec la protéine

SBP2 CT. En effet, cet ARN adopte plusieurs conformations sur gel natif (comparer les Figures 67B et C) et l'une d'elle est déplacée en présence de SBP2 CT.

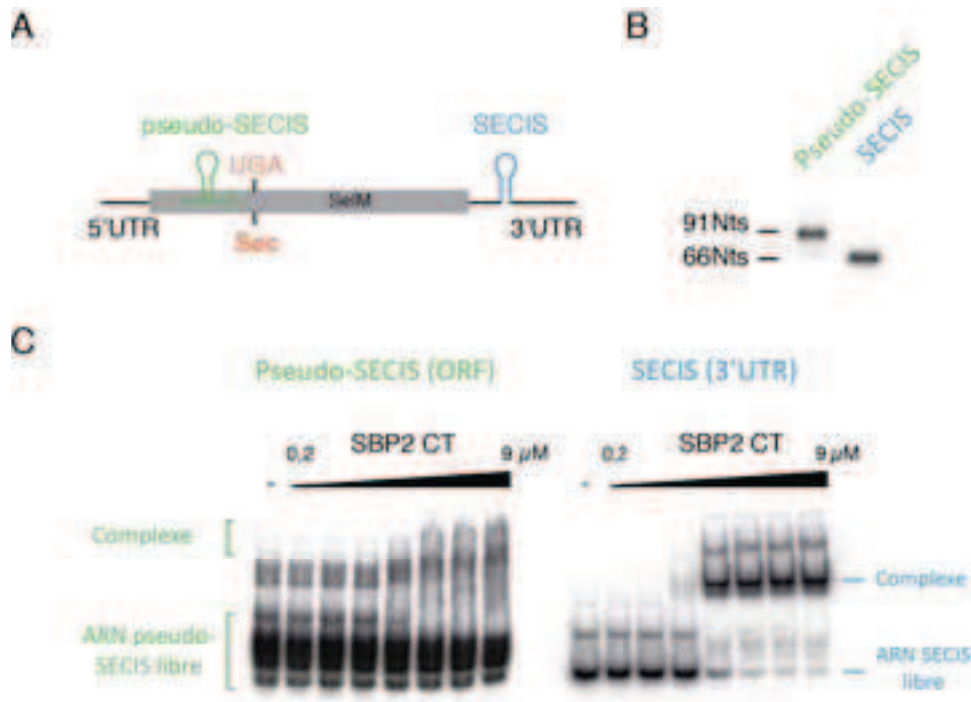


Figure 67. SBP2 interagit faiblement avec la région entourant le codon UGA^{Sec}.
Expériences de retard sur gel montrant l'interaction de SBP2 avec les transcrits d'ARN de SelM SECIS (3'UTR) et pseudo-SECIS (ORF). (A) Représentation schématique de l'ARNm de SelM. Les transcrits utilisés pour les expériences de retard sur gel, contenant l'élément SECIS ou le "pseudo-SECIS", sont représentés en bleu et en vert respectivement. (B) Analyse des transcrits marqués en interne à l'aide d'[α³²P]-ATP, sur gel de polyacrylamide dénaturant 8%. (C) Les transcrits contenant le "pseudo-SECIS" ou l'élément SECIS ont été incubés en présence d'une quantité croissante de la SBP2 CT et déposés sur gel de polyacrylamide 8% en conditions natives.

Des informations structurales nous permettront de mieux déterminer la taille et les bornes du transcrit d'ARNm de SelM dans la région du codon UGA^{Sec} en tenant compte de son repliement. Si cette interaction est confirmée, il sera nécessaire de la caractériser biochimiquement et fonctionnellement. Pour cela, des mutations dans la séquence du "pseudo-SECIS" seront réalisées et leur impact sur l'interaction de cet élément avec SBP2 et sur le mécanisme de recodage du codon UGA^{Sec} de l'ARNm SelM seront analysés. Les appariements non-canoniques seront mutés afin de déterminer si les mêmes principes de reconnaissance (Walczak et al., 1998) gouvernent l'interaction SBP2-"pseudo-SECIS" et SBP2-SECIS. Des résultats d'expériences par élongation d'amorces similaires ont été obtenus avec l'ARNm GPx1 (Figure 68) et une structure "pseudo-SECIS" a également pu être prédite en amont du codon UGA^{Sec}, l'ARNm de SelM ne semble donc pas être un cas particulier.

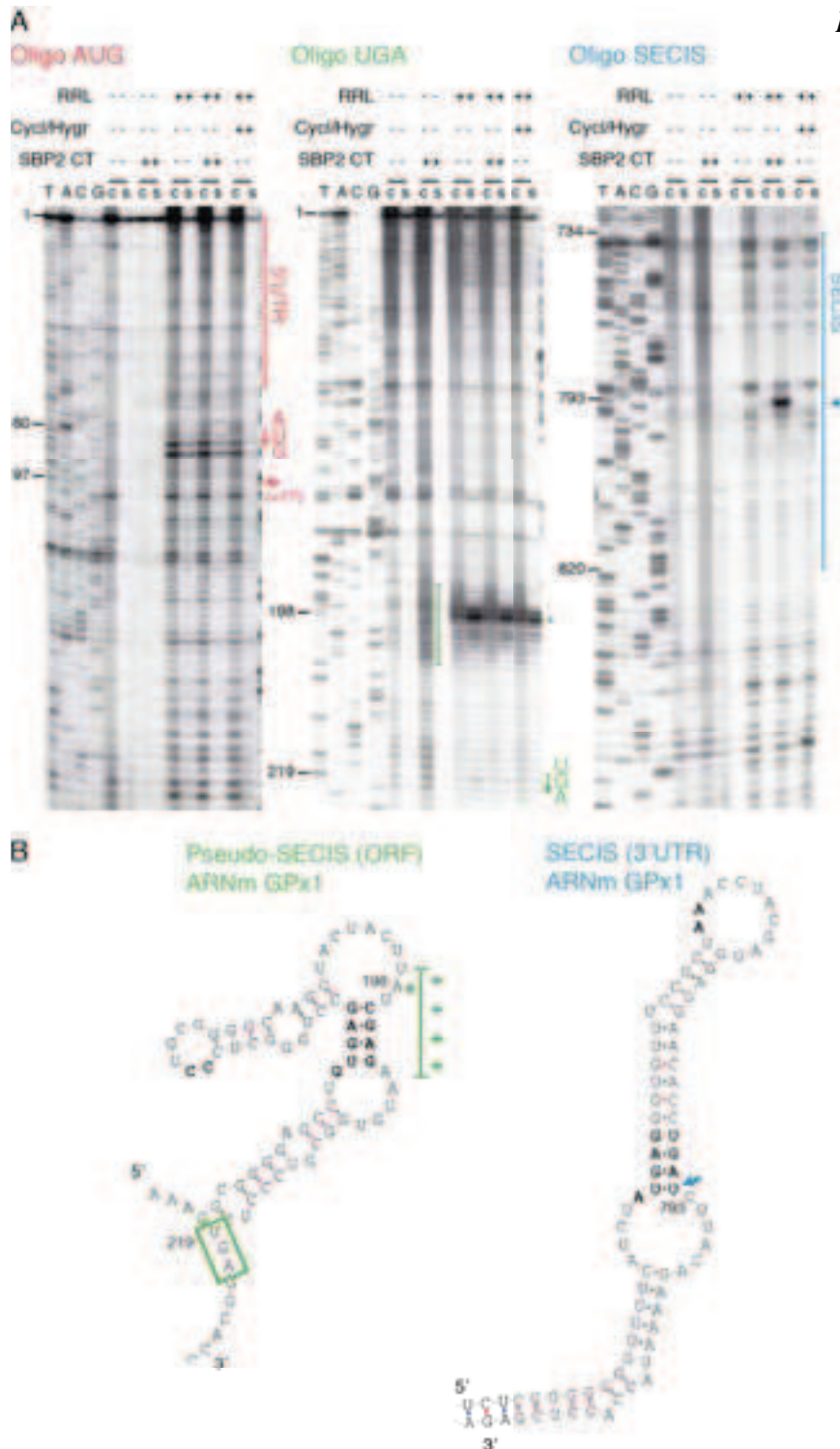


Figure 68. Localisation du ribosome sur l'ARNm de GPx1

(A) Des expériences de toeprint ont été réalisées sur l'ARNm de GPx1 coiffé par du m⁷G. Les complexes de traduction ont été assemblés dans des extraits RRL, en présence ou en absence de cycloheximide/hygromycine ou de SBP2 CT. Afin de différencier les arrêts de la RT liés à la liaison du ribosome sur l'ARN de celles provenant des protéines (marqués par des étoiles), les complexes ont été fractionnés par ultracentrifugation en culots (C) et surnageants (S). Des oligonucléotides complémentaires à des régions en aval du codon initiateur AUG (Oligo AUG), du codon UGA^{Sec} (Oligo UGA) ou de l'élément SECIS (Oligo SECIS) ont été utilisés pour l'étape de RT. Les positions des nucléotides de l'ARN de GPx1 sont indiquées en noir et la région 5'UTR, le codon d'initiation, le codon UGA^{Sec} et l'élément SECIS sont indiqués en couleur. (B) Prédiction de structure secondaire de la région SECIS (SECIS 3'UTR) et de la région située en amont du codon UGA^{Sec} (Pseudo-SECIS ORF) par le logiciel SECISal. Les positions de pause de la RT liées à la présence de SBP2 CT sur l'ARNm de SelM sont marquées par des flèches. Les étoiles représentent la position des pauses de RT dues aux extraits RRL.

Pour les deux ARNm SelM et GPx1, le quartet de paires de bases non Watson-Crick du “pseudo-SECIS” est située à une distance du codon UGA^{Sec} presque identique. De plus, en présence de RRL, une pause de la RT au même endroit dans le “pseudo-SECIS” est observée pour les ARNm de SelM (position 192) et GPx1 (position 198), et pourrait indiquer la liaison d’une protéine autre que SBP2 (Figures 65, 66D et 68). Il sera nécessaire de vérifier si ce modèle est généralisable à l’ensemble des ARNm de sélénoprotéines. L’existence d’une telle interaction pourrait fournir un début d’explication quant à la manière dont se fait la communication entre le SECIS en 3’UTR de l’ARNm et le codon UGA^{Sec} au cours du mécanisme de recodage des sélénoprotéines. En effet, SBP2, en interagissant avec le “pseudo-SECIS”, pourrait permettre d’apporter les facteurs d’incorporation de la sélénocystéine à proximité du codon UGA^{Sec}, rendant l’étape de recodage plus efficace.

Pour conclure, nos résultats indiquent que le mécanisme d’initiation de la traduction des ARNm de sélénoprotéines est atypique. Ce mécanisme requiert la présence d’une GTPase différente de celles impliquées dans le processus d’initiation canonique. Nous avons montré que la région 3’UTR est essentielle au recrutement des ribosomes sur l’ARNm de SelM. Par ailleurs, nous avons identifié un site de liaison de SBP2 dans la phase codante de l’ARNm à proximité du codon UGA^{Sec}; une structure de type « SECIS-like » peut être prédite dans cette région. La seule présence du codon UGA^{Sec} induit l’accumulation de ribosomes 80S sur l’ARNm de SelM. Ces éléments *in cis* pourraient constituer des signaux essentiels pour le recrutement des ribosomes sur les ARNm de sélénoprotéines.

III.2.2 Vers l’identification de nouveaux facteurs des complexes de traduction des ARNm de sélénoprotéines

Afin d’identifier de nouveaux facteurs potentiellement impliqués dans le mécanisme de traduction des sélénoprotéines, j’ai mis au point une méthode permettant de purifier les complexes ribonucléiques assemblés *in vivo* sur les ARNm de la sélénoprotéine HA-GPx1, basée sur l’utilisation d’oligonucléotides complémentaires biotinylés. Le schéma de purification est décrit dans la Figure 69 et les mises au point expérimentales sont présentées dans la partie Matériel et Méthodes.

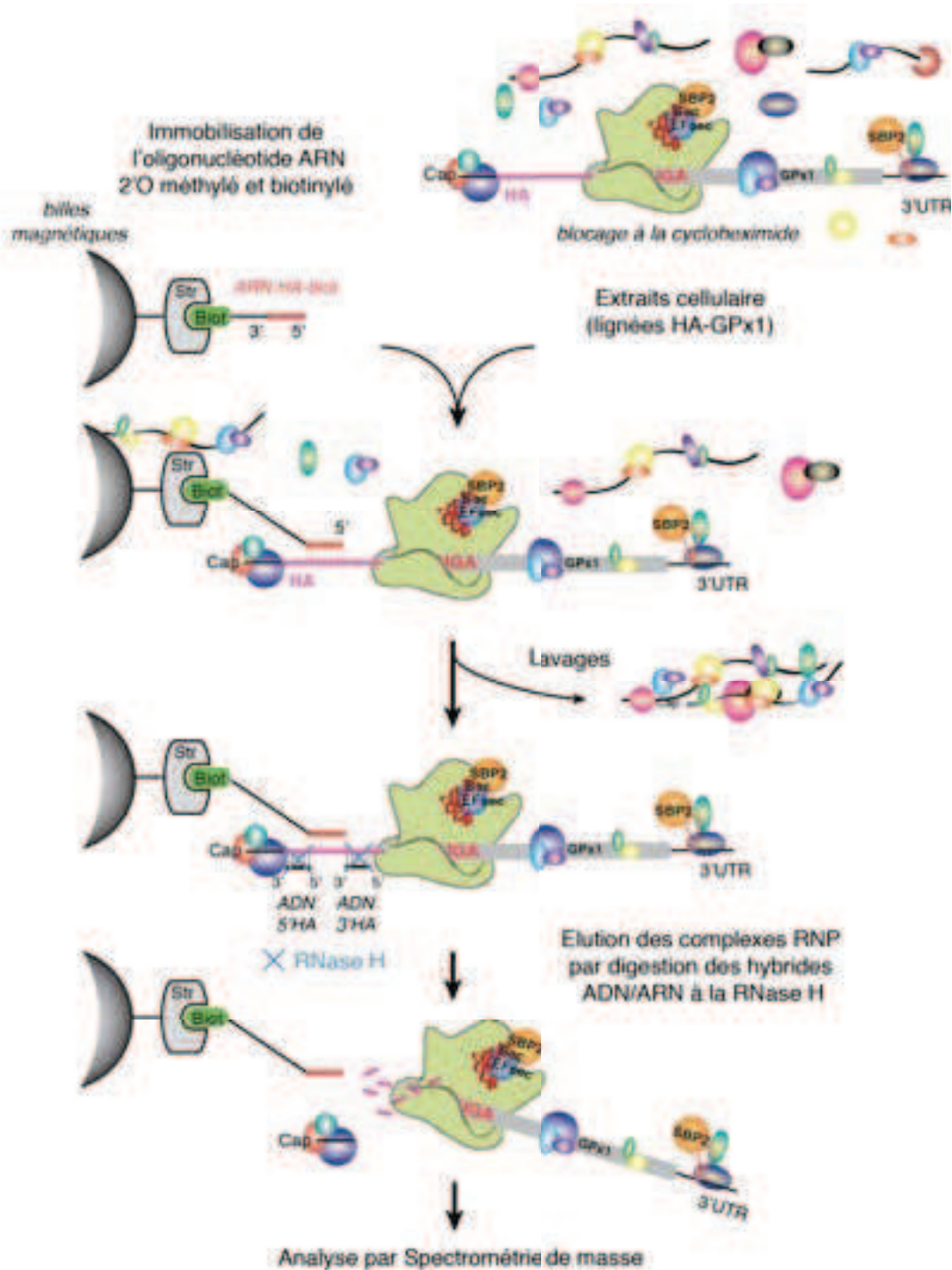


Figure 69. Schéma expérimental de la purification des complexes associés à l'ARNm de HA-GPx1 *in vivo*.

Un oligonucléotide ARN 2'O méthylé complémentaire à l'étiquette HA (ARN HA-Biot) est immobilisé sur des billes magnétiques couplées à la streptavidine (Str) et incubé en présence d'extrait de cellules HA-GPx1 dans lesquels les complexes de traductions ont été préalablement bloqués à la cycloheximide. Seule la rétention d'un ARNm ayant lié le ribosome est représentée sur la figure, les ARNm HA-GPx1 libres de ribosomes et associés sous forme de RNP sont également retenus. Après lavages, des oligonucléotides ADN (ADN 5'HA et ADN 3'HA) sont hybridés spécifiquement sur l'étiquette HA de l'ARNm HA-GPx1 en 5' et en 3' du site de l'oligonucléotide ARN HA-Biot. La digestion des hybrides ADN/ARN par la RNase H permet d'éluer spécifiquement les complexes associés à l'ARNm HA-GPx1, dont la composition protéique sera ensuite analysée par spectrométrie de masse.

Des extraits de cellules exprimant l'ARNm de HA-GPx1 de manière conditionnelle (voir Résultats 1.1), ont été purifiés par l'intermédiaire d'un oligonucléotide ARN 2'O-méthylé et biotinylé complémentaire à l'étiquette HA. Les complexes protéiques assemblés sur les ARNm HA-GPx1 ont été élués par clivage à la RNase H, à l'aide de 2 oligonucléotides ADN s'hybridant de part et d'autre de l'oligonucléotide ARN biotinylé. Le suivi de la purification par RT-qPCR est représenté dans la Figure 70A.

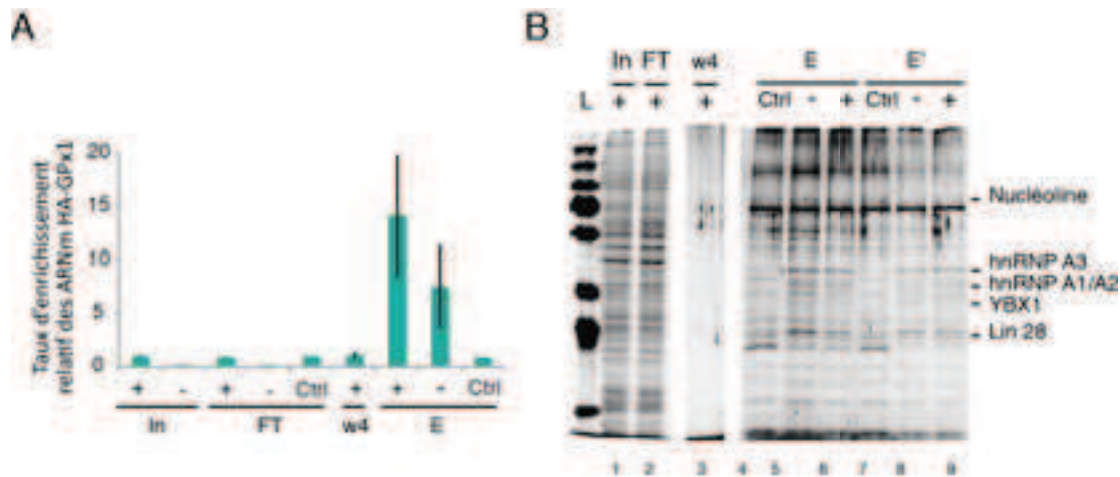


Figure 70. Purification de complexes associés à l'ARNm de HA-GPx1 *in vivo*.

Les complexes associés à l'ARNm HA-GPx1 ont été purifiés à partir d'extraits de cellules HA-GPx1 induites à la doxycycline (+) ou non-induites (-) d'après la méthode présentée Figure 69. Un contrôle de purification sans oligonucléotides ARN HA-Biot (Ctrl) a été réalisé. In : extrait de départ (input), FT : surnageant (flow-through), w4 = 4^{ème} lavage (wash), E : élution. (A) Analyse du taux d'enrichissement relatif des ARNm HA-GPx1 par RT-qPCR, les résultats sont normalisés par rapport à HPRT et GAPDH. (B) Analyse des protéines associées à l'ARNm HA-GPx1 par électrophorèse sur gel SDS-PAGE et coloration au nitrate d'argent.

L'ARNm HA-GPx1 purifié à partir d'extraits de cellules induites est enrichi de 15 fois (colonne +) par rapport au contrôle de purification réalisé sans oligonucléotides biotinylés (colonne Ctrl). En absence d'induction (colonne -) une quantité non-négligeable d'ARNm HA-GPx1 est retenue sur les billes. Ceci est expliqué par l'existence d'une expression basale de l'ARNm de HA-GPx1 à partir du promoteur CMV en absence d'induction. Par ailleurs, une différence significative dans la composition protéique est visible en présence d'oligonucléotide biotinylé comparé au contrôle sans oligonucléotide (Figure 70B, pistes 3 à 8). Cela pourrait indiquer la présence de protéines qui sont retenues par le biais de l'ARNm de HA-GPx1. Peu de différences sont cependant observées entre les éluats provenant des extraits induits (colonne +) et non-induits (colonne -), ce qui est vraisemblablement dû à l'expression basale de l'ARNm de HA-GPx1 (Figure 70A). L'analyse par spectrométrie de masse

des complexes ainsi purifiés a permis d'identifier 35 protéines retrouvées spécifiquement en présence d'oligonucléotide biotinylé (sur un ensemble de 63 protéines retenues au total). Parmi celles-ci 7 protéines spécifiques sont également retrouvées en absence d'induction de HA-GPx1. Les résultats sont présentés dans le Tableau 3.

ACCESSION No	Description	Non induit (-)		Induit (+)		GeneID	Symbol
		Score	#Spectra	Score	#Spectra		
XRN2_HUMAN	5'-3' exonuclease 2			154.5	9	22803	XRN2
THOC4_HUMAN	THO complex subunit 4			52.5	3	10189	ALYREF
RT28_HUMAN	28S ribosomal protein S28, mitochondrial			52.1	2	28957	MRPS28
RBBP5_HUMAN	Retinoblastoma-binding protein 5			105.8	6	5829	RBBP5
IGF2BP1_HUMAN	Insulin-like growth factor 2 mRNA-binding protein 1			144.8	4	10642	IGF2BP1
RS16_HUMAN	40S ribosomal protein S16			46.5	3	6217	RPS16
RS30_HUMAN	40S ribosomal protein S30			56.1	2	—	FAU
WBP11_HUMAN	WW domain-binding protein 11			64.4	2	51729	WBP11
WDR82_HUMAN	WD repeat-containing protein 82			39.2	2	80335	WDR82
MFAP1_HUMAN	Microfibrillar-associated protein 1			75.8	2	4230	MFAP1
CPSF3_HUMAN	Cleavage and polyadenylation specificity factor subunit 3			40.6	1	51692	CPSF3
TCOF_HUMAN	Treacle protein			57.2	1	6949	TCOF1
STRAP_HUMAN	Serine-threonine kinase receptor-associated protein			40.1	1	11171	STRAP
HNRPO_HUMAN	Heterogeneous nuclear ribonucleoprotein D0			234.1	11	3184	HNRNPD
EIF4A3_HUMAN	Eukaryotic initiation factor 4A-III			126.7	5	8775	EIF4A3
DBPA_HUMAN	DNA-binding protein A			103.7	3	—	—
EI28E_HUMAN	Translation initiation factor eIF-28 subunit e			50.6	2	8893	EIF285
MCM2_HUMAN	DNA replication licensing factor MCM2			88.9	1	4171	MCM2
MCM7_HUMAN	DNA replication licensing factor MCM7			74.8	3	4170	MCM7
ERH_HUMAN	Enhancer of rudimentary homolog			69.7	3	2079	ERH
PDCD2L_HUMAN	Programmed cell death protein 2-like			73.8	2	84306	PDCD2L
MCM4_HUMAN	DNA replication licensing factor MCM4			48.6	1	4173	MCM4
RSS8_HUMAN	40S ribosomal protein S8			41.1	1	6202	RPS8
USP1_HUMAN	116 kDa U5 small nuclear ribonucleoprotein component			79.2	1	9343	EFTUD2
RUXE_HUMAN	Small nuclear ribonucleoprotein E			39.6	1	6635	SNRPE
1433G_HUMAN	14-3-3 protein gamma			50.1	1	7532	YWHAQ
1433T_HUMAN	14-3-3 protein theta			49	1	10971	YWHAQ
K067_HUMAN	Antigen K0-67			86	2	4288	MO67
ROA2_HUMAN	Heterogeneous nuclear ribonucleoproteins A2/B1	207	5	612.2	36	3181	HNRNPA2B1
LN28E_HUMAN	Protein lin-28 homolog B	395.6	8	691.9	37	4E+05	LN28E
NUCL_HUMAN	Nucleolin	126.8	3	670.8	32	4691	NCL
ROA0_HUMAN	Heterogeneous nuclear ribonucleoprotein A0	87.2	6	270.6	19	10949	HNRNPA0
ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3	154.9	8	314.6	19	2E+05	HNRNPA3
YBOX1_HUMAN	Nucleic-acid sensitive element-binding protein 1	88.2	2	323.1	15	4604	YBX1
ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1	287.9	17	522.7	40	3178	HNRNPA1

épissage
polyadénylation
facteur d'initiation
transcription
localisation / export nucléaire
NMD
protéine ribosomique
cycle cellulaire
régulation des protéines
en interaction avec tBP2 dans d'autres expériences

Tableau 3. Analyse de la composition protéique de complexes purifiés associés à l'ARNm de HA-GPx1 *in vivo*, par spectrométrie de masse.

Au total, 36 protéines ont été retrouvées dans l'éluat contrôle (Ctrl), 13 dans l'éluat (-) et 63 dans l'éluat (+). Seules les protéines présentes dans les éluats (+) et (-) et absentes de l'éluat (Ctrl) sont représentées dans le tableau. Des codes couleurs ont été attribués aux protéines selon leur fonction. Spectres= Nombre de spectres fragmentés, Score= Score total Mascot. Ces valeurs sont indicatives de l'abondance des protéines dans l'éluat.

Les protéines identifiées sont impliquées dans différentes étapes du métabolisme de l'ARNm telles que la transcription, la polyadénylation, l'épissage, la localisation et la traduction. Plusieurs facteurs connus comme étant impliqués dans la régulation de

l'expression des sélénoprotéines sont retrouvés : eIF4A3, la nucléoline et YBX1, ce qui semble valider notre approche expérimentale (Tableau 3). Malheureusement, les protéines SBP2 et eEFSec, essentielles au mécanisme de recodage ne sont pas retrouvées. Il a été montré que l'interaction de eIF4A3 avec certains ARN SECIS, dont celui de GPx1, empêche le recrutement de SBP2 et inhibe le mécanisme de recodage des sélénoprotéines (Budiman et al., 2009a, 2011). Ceci est compatible avec nos données de purification et pourrait expliquer l'absence de SBP2. La nucléoline quant à elle se lie à l'élément SECIS de certains ARNm de sélénoprotéines et stimulerait leur traduction (Miniard et al., 2010; Wu et al., 2000). Enfin, la protéine YBX1 (ou NSEP1), interagit avec l'élément SECIS (Fagegaltier et al., 2000d; Shen et al., 1998) et son inhibition par siARN inhibe la synthèse des sélénoprotéines, ce qui laisse suggérer qu'elle joue un rôle dans le mécanisme de recodage (Qichang Shen, 2006). Deux des protéines que nous avons purifiées avec l'ARNm de HA-GPx1, hnRNP A1 et U5-116kD (U5S1 ou EFTuD2), avaient précédemment été retrouvées en interaction avec la protéine SBP2 par immunoprécipitation et crible double-hybride dans notre laboratoire (voir Figure 48A). U5-116kD est un constituant de la snRNP U5 et est impliquée dans l'épissage des ARNm (Fabrizio et al., 1997). Cette protéine est un homologue du facteur d'élongation EF-2 (translocase du ribosome) et possède une activité GTPase. L'une de nos hypothèses étant que l'ARNm de SelM requière une GTPase différente de eIF2 pour l'assemblage du complexe 48S (Figure 60D), la protéine U5-116kD pourrait constituer un candidat potentiel. hnRNP A1 est une des protéines les plus abondantes dans le noyau et est exprimée de manière ubiquitaire. Elle a des fonctions multiples dans la régulation de l'expression des gènes en régulant les étapes majeures de la maturation et de l'assemblage des transcrits naissants (Jean-Philippe et al., 2013). La protéine contient deux domaines RRM de liaison à l'ARN ainsi qu'une région C-terminale riche en glycines et arginines qui contient un domaine de liaison à l'ARN tri-RGG (Jean-Philippe et al., 2013; Thandapani et al., 2013). Ce domaine est la cible de Protéines Arginyl-MéthylTransférases (PRMT) responsables de la modulation de la liaison des hnRNP à l'ARN (Friend et al., 2013). En plus de la nucléoline, hnRNP A1 pourrait être une cible potentielle de la protéine PRMT5, recrutée vers les ARNm de sélénoprotéine via son interaction avec la protéine SBP2 (voir Résultats II). Le rôle de U5-116kD et de hnRNPA1 dans le mécanisme de synthèse des sélénoprotéines sera étudié. Dans un

premier temps, l'effet de l'inactivation de U5-116kD et de hnRNPA1 par siARN sur la synthèse de HA-GPx1 sera évalué. Leur interaction avec les différents facteurs de la machinerie de recodage, dont SBP2, sera analysée. La méthylation de hnRNPA1 par le méthylosome sera testée *in vitro*, afin de déterminer si hnRNPA1 peut être le substrat de ce complexe. D'autres protéines impliquées dans la régulation du cycle cellulaire ont été co-purifiées avec l'ARNm de HA-GPx1. Si ces interactions sont impliquées dans l'expression des sélénoprotéines, cela indique que leur synthèse, en plus d'être régulée par le taux de sélénium, pourrait dépendre du cycle cellulaire. Enfin, RBB5 ou WBP11 n'ont pas de fonction connue.

Au vu des différentes protéines identifiées (Tableau 3), il apparaît que notre méthode a bien permis de purifier des complexes ribonucléoprotéiques associés à l'ARNm de HA-GPx1 *in vivo*. Cependant, certaines améliorations pourront être apportées. En effet, l'absence de SBP2 et des facteurs de recodage et la faible abondance de protéines ribosomiques et de facteurs de traduction, indique que peu de complexes de traduction ont été purifiés. En revanche, la présence de XRN2 et de eIF4A3, composant de l'EJC (Exon Junction Complex) dans les complexes indique que ces ARNm seraient sujets au mécanisme du NMD. Il semblerait donc qu'au lieu de purifier des complexes de traduction, nous ayons majoritairement purifié des complexes de régulation négative de la traduction ou soumis au NMD. Il est envisageable de surexprimer SBP2 dans les cellules afin d'optimiser l'expression des sélénoprotéines et de modifier potentiellement la nature et la composition des complexes purifiés. Ceci permettra sans doute d'obtenir plus de complexes de traduction et donc de faciliter l'identification de facteurs impliqués dans ces mécanismes. Il a été montré que l'expression de eIF4A3 est dépendante du taux de sélénium (Budiman et al., 2009a). En effet, lorsque la concentration en sélénium est élevée (3 fois supérieure à nos conditions), l'expression de eIF4A3 est inhibée et celle de GPx1 est stimulée (Budiman et al., 2009a). Ainsi, dans les expériences futures, nous optimiserons les conditions en sélénium afin d'augmenter la traduction de HA-GPx1 et de lever la répression due à eIF4A3. Il est également envisageable que nos conditions expérimentales soient trop stringentes pour maintenir certaines interactions. La concentration saline pourra être diminuée et une étape de pontage pourra être envisagée pour faciliter la détection des interactions.

Après amélioration, la même méthode sera utilisée pour purifier les complexes de traduction associés aux ARNm de HA-GPx1, HA-GPx1Cys et HA-GPx1Cys Δ SECIS provenant des trois autres lignées cellulaires décrites précédemment (voir paragraphe II.2.2 de la partie résultats). La comparaison de la composition protéique des différents éluats permettra de différencier les protéines liées spécifiquement à la présence du codon UGA^{Sec} et de l'élément SECIS et intervenant peut-être dans l'étape de recodage des sélénoprotéines. Le rôle de ces protéines dans la synthèse des sélénoprotéines devra ensuite être validé et caractérisé. Leur impact sur le recodage par exemple pourra être évalué à l'aide de rapporteurs codant pour la luciférase et comprenant un codon UGA^{Sec} dans la phase codante et un élément SECIS en 3'UTR. Durant ma thèse, j'ai également construit une lignée cellulaire exprimant la protéine HA-GPx1 à partir de la séquence contenant l'intron naturel du gène GPx1 (HA-GPx1+Intron). Il semblerait que la présence de l'intron dans la séquence permette une meilleure expression de la protéine (Figure 71).



Figure 71. La protéine HA-GPx1 est mieux exprimée lorsque sa séquence ADN contient son intron naturel.

Après inhibition de SBP2 par siARN, la synthèse de HA-GPx1 a été induite dans les cellules HA-GPx1 et HA-GPx1+Intron pendant 24h. Des siARN dirigés contre la luciférase FFL sont utilisés comme contrôle. L'expression de HA-GPx1 a été analysée par Western blot à l'aide d'un anticorps anti-HA et comparée à celle de l'AspRS.

Lors de l'épissage, certains facteurs impliqués dans la traduction ou la stabilité des ARNm de sélénoprotéines pourraient être assemblés sur l'ARNm. Ces facteurs pourraient être identifiés en comparant la composition protéique des complexes associés à l'ARNm HA-GPx1 et HA-GPx1+Intron.

III.2.3 Vers une analyse structurale des complexes de traduction des ARNm de sélénoprotéines

Nous avons précédemment montré que la présence du codon UGA^{Sec} induit un blocage naturel du ribosome sur l'ARNm de SelM dans les extraits de réticulocytes de lapins (voir résultats III.2.1). Nous avons donc choisi de tirer avantage de cette propriété pour purifier des complexes de traduction, assemblés *in vitro* dans des lysats de réticulocytes de lapin en suivant une méthode développée dans notre laboratoire (Chicher et al., 2015; Prongidi-Fix et al., 2013). L'ARNm de SelM a été ligué à un oligo-désoxynucléotide biotinylé créant ainsi une molécule chimère ARNm/ADN biotinylé. Cette molécule a été immobilisée sur des billes magnétiques de Streptavidine et incubée avec les RRL afin de former les complexes de traduction. Après plusieurs étapes de lavage, les complexes sont élués spécifiquement par digestion de la partie ADN de la molécule chimère à la DNase I, libérant les complexes de traduction dans des conditions natives. Le contenu en ARN des différentes étapes de la purification a été analysé sur gel d'agarose (Figure 72A). On peut ainsi observer la présence de l'ARNm de SelM et des ARN ribosomiques 28S et 18S dans l'éluat final (Figure 72A, piste élution), confirmant la purification de ribosomes associés à l'ARNm de SelM. L'analyse des complexes par spectrométrie de masse a, par ailleurs, permis d'identifier la présence des 32 protéines composant la sous-unité 40S et de 45 protéines sur les 46 qui composent la sous-unité 60S, ainsi que de nombreux facteurs de traduction. Il est intéressant de remarquer que les facteurs eEF1A et eEF2 sont majoritaires ; les complexes purifiés semblent donc essentiellement correspondre à des complexes d'élongation. L'analyse structurale des complexes par microscopie électronique a été réalisée en collaboration avec Yaser Hashem (IBMC, Strasbourg). Dans un premier temps, la coloration négative par l'acétate d'uranyle et leur visualisation par microscopie électronique a confirmé que les complexes étaient homogènes, formaient peu d'agrégats, et étaient en quantité suffisante pour une analyse structurale à haute résolution (Figure 72B). Une première série d'images a ainsi pu être collectée par cryo-microscopie électronique à l'aide du microscope Tecnai F30 Polara (Centre de Biologie Intégrative, IGBMC, Illkirch). Des clichés d'environ 50.000 particules ont pu être collectés (Figure 72C). Le traitement des images a été effectué par A. Simonetti, à l'aide du logiciel RELION (Scheres, 2012).

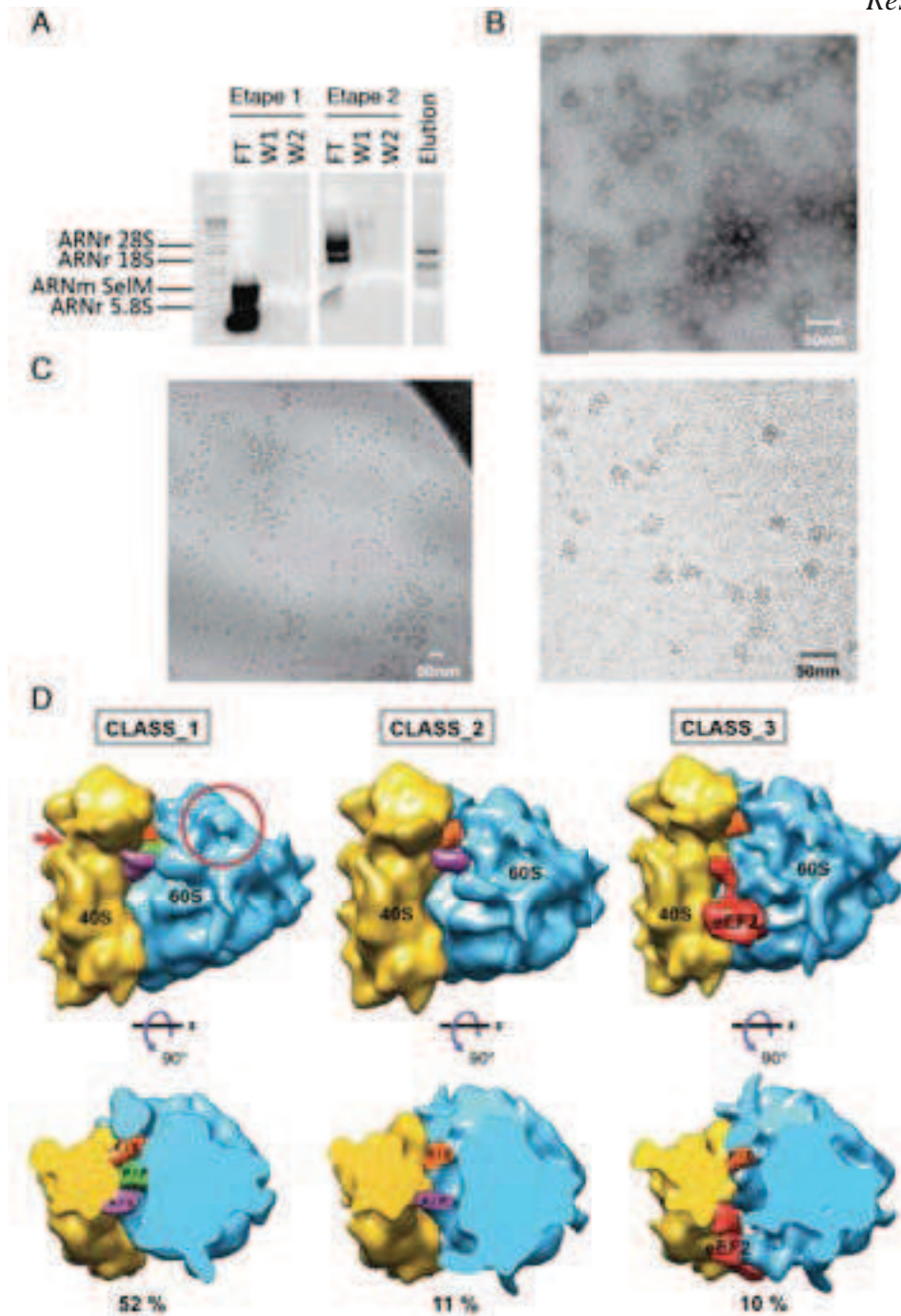


Figure 72. Analyse structurale des complexes traductionnels assemblés sur l'ARNm de SelM par cryo-microscopie électronique.

En collaboration avec le laboratoire de Y. Hashem (IBMC, Strasbourg)

(A) Suivi de purification des complexes. Etape 1 : l'ARNm de SelM est ligué à un oligodésoxynucléotide biotinylé puis immobilisé sur des billes magnétiques-streptavidine. Etape 2 : les billes couplées à l'ARNm de SelM sont incubées avec les RRL afin de former les complexes de traduction. Elution : Les complexes sont élués spécifiquement par digestion de l'extrémité ADN à la DNase I. FT : surnageant (Flow-Through), W : lavages (Wash). (B et C) Clichés de microscopie électronique des complexes de traduction associés à l'ARNm de SelM, par la méthode de coloration négative à l'acétate d'uranyle (B) ou par cryo-microscopie électronique (C). (D) Reconstitution 3D des structures des 3 classes de complexes sélectionnées (réalisé par A. Simonetti) grâce au programme UCSF Chimera. La sous-unité ribosomique 40S est représentée en jaune, la sous-unité 60S en bleu, l'ARNt dans le site A est en rose, dans le site P en vert et dans le site E en orange. Le site de liaison putatif de SBP2 au sein du ribosome (hélice ES7L-E de l'ARNr 28S) est indiqué par cercle rouge. La flèche rouge indique la zone d'interaction entre l'hélice H16 de l'ARNr 18S et la protéine ribosomique S3, indicatif de la présence d'ARNm.

Le tri des particules a été réalisé par classification 3D (RELION) à partir des 36.000 particules sélectionnées après classification 2D, et a conduit à l'obtention de 5 classes de complexes. Seules les trois classes les plus homogènes ont été analysées. La classe majoritaire (classe 1) est composée de 18.000 particules et a permis d'obtenir, par reconstitution 3D, une structure à 12Å de résolution (Figure 72D). Elle correspond au ribosome 80S en phase d'élongation avec 3 ARNt canoniques localisés respectivement dans les sites A/A, P/P, E/E. Un ARNt a été déposé dans le site A du ribosome par le facteur eEF1A, qui s'est ensuite dissocié, et la translocation n'a pas encore eu lieu. Au sein du complexe, on observe que l'hélice 16 de l'ARNr 18S interagit avec la protéine ribosomique S3, indiquant la présence d'un ARNm (flèche rouge, Figure 72D). Dans la classe 2 (4217 particules), le ribosome adopte une conformation différente et a subi une légère rotation appelée "mouvement de rochet" durant l'étape de transpeptidation (Figure 72D) (Frank and Agrawal, 2000). En effet, 2 ARNt sont visibles, dans des états hybrides intermédiaires au niveau des sites A/P et P/E (Moazed and Noller, 1989). Des états de ce type ont été observés dans d'autres structures obtenues par cryo-microscopie électronique (Budkevich et al., 2011), notamment lorsque le ribosome 80S est bloqué sur un ARNm par des structures de type pseudonœuds (Flanagan IV et al., 2010; Moran et al., 2008). Les anticodons sont restés dans leur position initiale par rapport à la petite sous-unité. Cet état précède l'étape finale de translocation qui est assurée par eEF2 (Noller et al., 2002; Wilden et al., 2006). Enfin, dans la classe 3 (3618 particules), l'ARNm semble être absent car l'interaction entre l'hélice 16 de l'ARNr 18S et la protéine S3 ne sont pas visibles. Dans cette classe, le ribosome adopte également une conformation de pré-translocation et on peut observer la présence du facteur eEF2 à l'entrée du site A. Curieusement, un seul ARNt est présent dans le site P/E, indiquant que eEF2 ne pourrait procéder à l'étape de translocation et que le ribosome serait bloqué dans cette configuration. Ce ribosome s'est peut-être dissocié de l'ARNm.

Le site putatif de liaison de SBP2 au ribosome, c'est à dire l'hélice ES7L-E localisée dans le segment d'extension 7 de l'ARNr 28S (Kossinova et al., 2014) est indiqué par un cercle rouge, on n'y remarque pas de densité supplémentaire (Figure 72D). Ceci est cohérent avec le fait que l'association de SBP2 au ribosome ne se ferait que lors de l'étape de recodage après accommodation du Sec-ARNt^{[Ser]Sec} dans le site A du

ribosome ; pendant les autres étapes de la traduction SBP2 reste liée à l'ARNm (Kossinova et al., 2013). Par ailleurs, ni la protéine SBP2 ni le facteur d'élongation eEFSec n'ont été retrouvés par spectrométrie de masse au sein des complexes que nous avons purifiés. L'ensemble de ces résultats indique que nous n'avons pas purifié de complexe de recodage. De même, aucun complexe de terminaison n'a été obtenu. Puisque les complexes de traduction sont naturellement bloqués sur l'ARNm contenant un codon UGA^{Sec} (voir Figure 60B), il est intrigant que nous n'ayons pas observé de signature structurale pouvant expliquer cette particularité. En effet, si le ribosome était bloqué au niveau du UGA^{Sec} nous devrions observer la présence d'un facteur empêchant le mécanisme de terminaison, ou une interaction de l'ARN SECIS au niveau du ribosome. Cependant, aucune densité additionnelle n'a été observée sur nos structures. L'ARNt^{[Ser]Sec} est le plus grand des ARNt et possède des caractéristiques structurales qui le distingue facilement des ARNt canoniques (voir chapitre I.2.1.1 de l'Introduction). Avec des bras A et D plus longs et un bras variable possédant une orientation particulière, l'ARNt^{[Ser]Sec} présente une conformation unique. eEFSec, est similaire au facteur d'élongation canonique eEF1A, qui est impliqué dans l'incorporation des 20 autres acides aminés. Il présente cependant un domaine additionnel en C-terminal absent dans eEF1A (voir chapitre I.2.2.2.a de l'introduction). Ainsi, si l'ARNt^{[Ser]Sec} et le facteur eEFSec avaient été présents dans les complexes purifiés, leurs caractéristiques structurales nous auraient permis de les distinguer de la machinerie de traduction canonique.

En conclusion, cette approche nous a permis de purifier pour la première fois des complexes de traduction associés à un ARNm de sélénoprotéine dans des conditions compatibles à une analyse structurale par cryo-microscopie électronique. Contrairement à ce que nous attendions, les complexes que nous avons purifié ne sont cependant pas des complexes de recodage, mais des complexes d'élongation. Des expériences futures viseront donc à compléter les extraits de RRL par les différents facteurs essentiels et spécifiques du mécanisme de recodage des sélénoprotéines tels que SBP2, eEFSec et le Sec-ARNt^{[Ser]Sec} afin d'augmenter les chances de purifier un complexe de recodage.

Discussion générale et perspectives

Les sélénoprotéines sont les effecteurs majeurs de la défense antioxydante des cellules. Grâce à leur fort potentiel réducteur, elles inactivent les espèces réactives oxygénées, qui, en excès, provoquent des stress oxydants pouvant être à l'origine de cancers ou de pathologies (Rayman, 2012; Zhuo and Diamond, 2009b). A ce jour, 25 gènes de sélénoprotéines ont été identifiés chez l'homme (Lobanov et al., 2009). En plus de leur rôle dans le maintien de l'homéostasie redox des cellules (Steinbrenner and Sies, 2009), les sélénoprotéines sont impliquées dans des mécanismes variés tels que la fertilité masculine (Guerriero et al., 2014; Imai et al., 2009), le métabolisme thyroïdien (Beckett and Arthur, 2005; Köhrle, 2013), la réponse immunitaire (Steinbrenner et al., 2015), le développement des muscles (Castets et al., 2012; Rederstorff et al., 2006) et vraisemblablement la protection des télomères (Squires et al., 2009). La synthèse des sélénoprotéines fait appel à un mécanisme de recodage très complexe qui, à ce jour, n'est pas encore entièrement compris. L'acide aminé sélénocystéine est incorporé au cœur du site actif des sélénoprotéines en réponse à un codon UGA^{Sec}, ce qui requiert l'intervention d'une machinerie spécialisée. Chez les mammifères, ce processus est conditionné par l'assemblage de facteurs spécifiques dans la région 3' non codante des ARNm des sélénoprotéines. Notre équipe a montré que le processus d'assemblage des mRNP de sélénoprotéines présente des similarités avec celui de l'assemblage des complexes de petits ARN non codants tels que les sn et les snoRNP (Boulon et al., 2008; Cléry et al., 2007). Mon travail de thèse a contribué à mettre en évidence l'existence d'un mécanisme particulier de maturation de la coiffe en 5' des ARNm de sélénoprotéines. Nos résultats montrent que certains ARNm de sélénoprotéines possèdent une coiffe hyperméthylée m₃^{2,2,7}G à la manière de certains petits ARN non codants et ne sont pas reconnus efficacement par le facteur canonique d'initiation de la traduction eIF4E. Nous avons déterminé le mécanisme de biogenèse de cette coiffe, et avons montré qu'il fait appel à l'enzyme Tgs1 (Triméthyl-guanosine synthase). Le complexe SMN et son cofacteur, le méthylosome, semblent essentiels à l'assemblage des mRNP de sélénoprotéines. Durant ma thèse, nous avons également analysé la fonction de la coiffe m₃G sur la

traduction des ARNm de sélénoprotéines *in vitro*. L'ensemble de mes résultats de thèse va dans le sens de l'existence d'un mécanisme d'initiation atypique de la traduction des ARNm de sélénoprotéines.

La coiffe des ARNm de sélénoprotéines est hyperméthylée

Nous avons identifié trois classes d'ARNm de sélénoprotéines. Dans la classe majoritaire les ARNm portent une coiffe hyperméthylée m_3G et ne sont pas reconnus par le facteur d'initiation eIF4E, d'autres sont coiffés de m^7G et interagissent avec eIF4E alors que certains ARNm de sélénoprotéines semblent porter les deux types de coiffes m_3G ou m^7G . Puisque l'hyperméthylation de la coiffe ne se produit pas de manière égale pour tous les ARNm de sélénoprotéines, il est possible qu'elle puisse contribuer à la régulation de l'expression différentielle des sélénoprotéines. Plusieurs études ont établi que la régulation du niveau d'expression et de la stabilité des ARNm de sélénoprotéines sont différents pour chaque ARNm et dépendent de nombreux facteurs (Budiman et al., 2009a; Latrèche et al., 2012; Lei et al., 1995; Low et al., 2000a; Squires et al., 2007b; Sun et al., 2001c) tels que le niveau de sélénium, le type de SECIS présent en 3'UTR et la sensibilité au mécanisme du NMD. Le Tableau 4 résume les différents classements proposés dans la littérature pour les ARNm de sélénoprotéines.

Nature de la coiffe	SECIS		Recodage / Incorporation de la sélénocystéine			Stabilité des ARNm		
	Type de SECIS	Influence du SECIS sur le recodage	Isoforme de l'ARNt ^{[Ser]Sec}	Niveau de sélénium et isoforme de l'ARNt ^{[Ser]Sec}	Stress oxydants	NMD	Niveau de sélénium	Inactivation de SBP2
Wurth et al., 2014	Chapple et al., 2009	Latreche et al., 2009	Carlson et al., 2005	Howard et al., 2013	Touat-Hamici et al., 2014	Seyedali et al., 2014	Sunde and Raines, 2011	Squires et al., 2007
Coiffe m ₃ G		Recodage efficace	Isoforme mcm5Um	<i>Sélénoprotéines sensibles au stress</i>				
SelR	2	SelR	SelR	SelR	GPx1	GPx1	GPx1	GPx1
GPx1	1	GPx1	GPx1	GPx1	GPx4	GPx4	SelM	SelT
GPx4	2	GPx4	SelW	SelW	TrxR1	SelM	SelW	TrxR1
SelM	2	Sel15	SelT		SelK	SelW	SelT	SelN
SelW	2	SelN				GPx3	TrxR1	
						SelN	SelK	
							GPx3	
Coiffe m ₃ G/m ⁷ G		Recodage peu efficace	Isoforme mcm5U	<i>Sélénoprotéines non-sensibles au stress</i>				
SelT	2	SelM	GPx4	GPx4	SelM	SelR	GPx4	SelR
SelO	2	SelW	TrxR1	SelT		SelT		GPx4
TrxR1	2	SelT	Sel15	SelO		SelO		SelW
		SelO		TrxR1		TrxR1		SelO
		TrxR1		Sel15		SelK		Sel15
		SelK		SelK				SelK
		GPx3		GPx3				
Coiffe m ⁷ G								
Sel15	2							
SelK	2							
GPx3	2							
SelN	1							

Tableau 4. Classements des sélénoprotéines proposés par différentes études.

Les ARNm de sélénoprotéines sont classés en fonction de la nature de leur coiffe (Wurth et al., 2014). Les ARNm portant une coiffe m₃G sont surlignés en vert, les ARNm à coiffe m⁷G en rouge et ceux portant les deux types de coiffe en jaune. Le type de SECIS présent sur chaque ARNm (type 1 ou type 2) ainsi que son influence sur l'efficacité de recodage sont indiqués respectivement dans les colonnes 2 et 3. Différentes études analysant l'expression des sélénoprotéines en fonction de la nature de l'ARNt^{[Ser]Sec}, du niveau de sélénium, du stress oxydant, de la sensibilité au NMD et de l'inactivation de SBP2 sont présentées. Les sélénoprotéines sont distribuées en deux groupes : les sélénoprotéines sensibles au stress et les sélénoprotéines non-sensibles au stress.

Deux éléments SECIS de structures légèrement différentes sont retrouvés chez les eucaryotes : la forme 1 et la forme 2, qui se distinguent au niveau de la boucle apicale (Chapple et al., 2009b; Fagegaltier et al., 2000a; Grundner-Culemann et al., 1999). Aucune corrélation évidente ne peut être faite entre la nature de la coiffe et la nature

de l'élément SECIS (type 1 ou type 2) présent sur l'ARNm de sélénoprotéines ni avec son influence sur l'efficacité du recodage (Latreche et al., 2009), voir Tableau 4. La corrélation entre la nature de la coiffe et la stabilité des ARNm (Seyedali et al., 2014; Squires et al., 2007b; Sunde and Raines, 2011) est également peu évidente.

De manière globale, il apparaît cependant que les ARNm de sélénoprotéines portant une coiffe m₃G sont majoritairement retrouvés dans les classes de sélénoprotéines sensibles au stress (Tableau 4). Notre classement présente le plus de similitudes avec celui proposé par Howard et al. (2013). En effet, il semblerait que les ARNm de sélénoprotéines qui portent une coiffe m₃G (Wurth et al., 2014) sont également ceux dont le mécanisme d'incorporation de la sélénocystéine, régulé par le niveau de sélénium, fait appel à l'isoforme mcm5Um de l'ARNt^{[Ser]Sec} (Carlson et al., 2005 ; Howard et al., 2013), voir Tableau 4. Deux isoformes de l'ARNt^{[Ser]Sec}, possédant soit la modification mcm⁵U, soit la modification mcm⁵Um en position 34, ont été caractérisés. Ces deux isoformes sont impliqués dans la traduction d'ARNm de sélénoprotéines différents, ce qui constituerait un moyen de régulation de l'expression des sélénoprotéines (Carlson et al., 2005, 2007). Howard et al. (2013) ont montré que l'expression sélénium-dépendante de certains ARNm de sélénoprotéines serait en partie modulée par la méthylation en position 34 du Sec-ARNt^{[Ser]Sec} et que l'augmentation du niveau de sélénium entraînerait une augmentation de la densité des ribosomes en aval du codon UGA^{Sec}, dans des expériences de ribosome profiling, indiquant une augmentation du taux de recodage.

Si les ARNm de sélénoprotéines qui portent une coiffe m₃G sont également ceux dont l'expression est sensible au stress, il est possible que l'hyperméthylation de la coiffe soit régulée par le niveau de sélénium ou de stress oxydant. Pour tester cette hypothèse, nous pourrions analyser si la nature de la coiffe des ARNm de sélénoprotéines varie en réponse au taux de sélénium ou en condition de stress oxydant. Les cellules seraient traitées par des quantités variables de sélénium ou de H₂O₂, et le taux d'ARNm de sélénoprotéines portant une coiffe m₃G serait analysé par immunoprécipitation à l'aide d'anticorps anti-m₃G spécifiques. Nous testerions également si l'expression de l'hyperméthylase Tgs1 varie dans ces conditions. Enfin, nous pourrions également analyser l'impact de l'inactivation par siARN de SECp43, l'enzyme responsable de la méthylation en position U34 de l'ARNt^{[Ser]Sec}, sur la

synthèse des sélénoprotéines. Ainsi, si l'isoforme mcm5Um de l'ARNt^{[Ser]Sec} est impliqué dans le mécanisme de recodage des ARNm de sélénoprotéines portant une coiffe m₃G, nous devrions observer une corrélation entre les sélénoprotéines affectées par l'inactivation de SECp43 et celles dont l'ARNm porte une coiffe m₃G. L'expression des sélénoprotéines est sans doute un processus multifactoriel complexe qui permet d'aboutir à une régulation fine et tissu-spécifique des différentes sélénoprotéines. La reconnaissance de la coiffe m₃G par des facteurs sensibles au taux de sélénium ou au stress oxydant par exemple pourrait permettre un mode de régulation fin et contrôlé de l'expression des sélénoprotéines en réponse aux besoins de la cellule et contribuerait à leur expression hiérarchique. Ces facteurs pourraient être impliqués dans divers aspects du métabolisme des ARNm de sélénoprotéines.

Biogenèse de la coiffe m₃G des mRNP de sélénoprotéines

Nous avons montré que Tgs1 est associée aux ARNm de sélénoprotéines et à SBP2 *in vivo* (Wurth et al., 2014). De plus, son inactivation par siARN entraîne une diminution du taux d'hyperméthylation de la coiffe des ARNm de sélénoprotéines et une diminution de l'expression de la sélénoprotéine HA-GPx1 *in vivo*. Ainsi, Tgs1 est responsable de la modification de la coiffe des ARNm de sélénoprotéines et stimule leur expression. Tgs1 existe sous deux formes : une protéine entière Tgs1 (LF), majoritairement cytoplasmique et une forme plus courte Tgs1(SF), essentiellement nucléaire (Girard et al., 2008). Deux voies de recrutement de l'enzyme Tgs1 vers les RNP ont été caractérisées. Tgs1(SF) est recrutée vers les snoRNP par le biais d'interactions avec les protéines cœur Nop58/Nop56 dans le noyau et Tgs1 (LF) est recrutée vers les snRNP via le complexe de chaperons SMN dans le cytoplasme (Matera et al., 2007a; Mouaikel et al., 2002). Nous avons montré que l'inactivation de Nop58 et de SMN par siARN abolit l'hyperméthylation de la coiffe des ARNm de sélénoprotéines et que ces deux protéines, ainsi que Tgs1, interagissent avec SBP2 (Wurth et al., 2014). Ainsi, l'enzyme Tgs1 pourrait être recrutée vers les ARNm de sélénoprotéines via deux complexes ternaires possibles : SBP2/Nop58/Tgs1 ou SBP2/SMN/Tgs1 (Figure 73). Cela suggère que l'hyperméthylation puisse avoir lieu soit dans le noyau soit dans le cytoplasme. La voie majoritaire semble impliquer SMN

et la forme longue de Tgs1, indiquant plus de similitudes avec le mécanisme de biogenèse des snRNP.

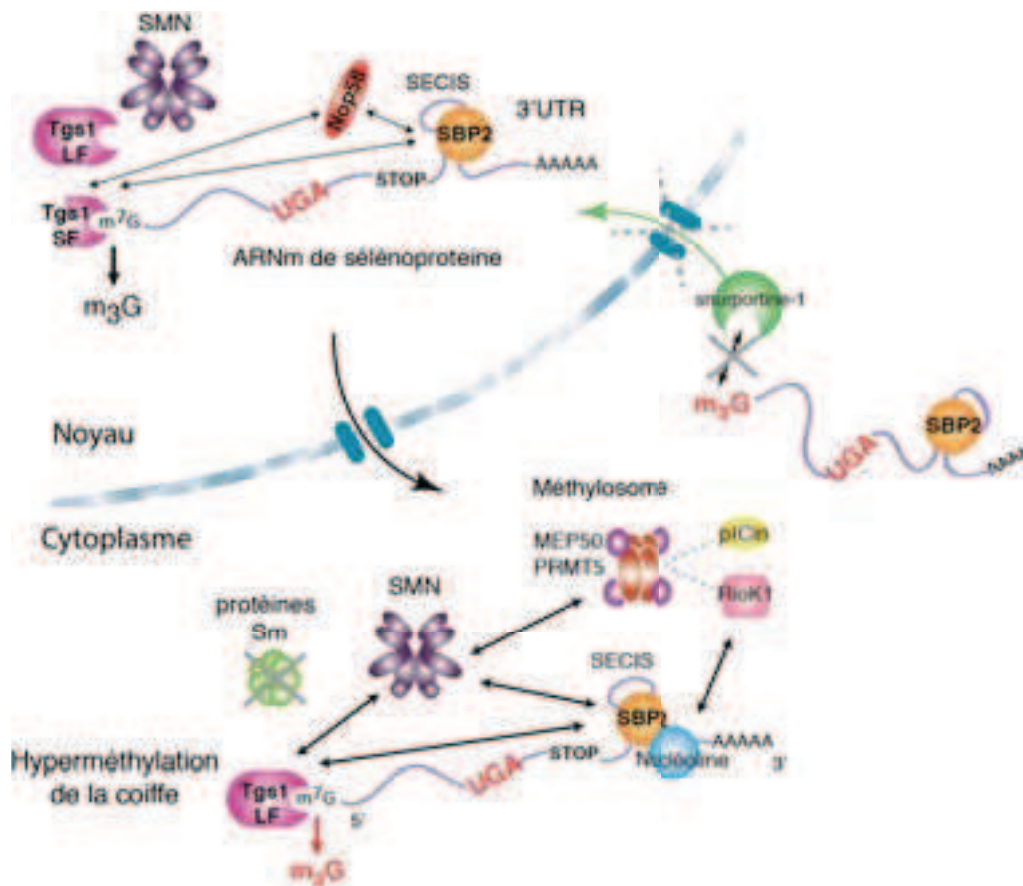


Figure 73. Mécanisme de biogenèse de la coiffe m₃G.

La coiffe des ARNm de sélénoprotéines est hyperméthylée par Tgs1. Les deux isoformes de Tgs1 : Tgs1 LF (nucléaire/cytoplasmique) et Tgs1 SF (exclusivement nucléaire) interagissent avec SBP2 et les ARNm de sélénoprotéines. Tgs1 est majoritairement recrutée sous la forme du complexe ternaire SBP2/SMN/Tgs1LF (flèches épaisses) mais également de SBP2/Nop58/Tgs1SF (flèches fines). L'hyperméthylation pourrait avoir lieu à la fois dans le noyau et le cytoplasme et être soumise à des mécanismes de régulation. Les ARNm de sélénoprotéines coiffés de m_3G sont majoritairement localisés dans le cytoplasme (80%), ils n'interagissent ni avec la snurportine-1 ni avec les protéines Sm et ne semblent donc pas réimportés vers le noyau. Un réseau d'interactions multiples a été identifié entre les sous-unités du complexe SMN, du métylosome et SBP2. Ces complexes permettraient le recrutement de l'enzyme Tgs1 et l'assemblage d'autres protéines cœurs de la mRNP, telles que la nucléoline ou hnRNPA1 par exemple.

L'existence d'un mécanisme hybride faisant intervenir à la fois la protéine Nop58 et le complexe SMN pourrait également être envisagée. Pour déterminer si l'ensemble de ces protéines font partie d'un même complexe, nous tenterons d'assembler et de co-purifier les protéines Tgs1, SBP2, SMN et Nop58 recombinantes *in vitro* en présence ou non d'ARNm de sélénoprotéine. Des expériences d'immunolocalisation pourraient permettre de déterminer le ou les compartiments dans lesquels ont lieu les

interactions entre Tgs1, SBP2, SMN et Nop58 et, indirectement, où l'hyperméthylation de la coiffe des ARNm de sélénoprotéines prend place. La colocalisation de ces différents facteurs, exprimés en fusion avec une étiquette GFP, avec des ARNm de sélénoprotéines détectés par hybridation d'ARN *in situ* pourrait être analysée. Le stress oxydant provoque l'accumulation de SBP2 dans le noyau et stimule l'assemblage des mRNP de sélénoprotéines (de Jesus et al., 2006; Papp et al., 2006; Touat-Hamici et al., 2014). Il pourrait également stimuler le recrutement de l'isoforme nucléaire de Tgs1 sur l'ARNm de sélénoprotéine via Nop58.

Les caractéristiques des ARNm de sélénoprotéines requises pour le recrutement de Tgs1 et l'hyperméthylation de la coiffe n'ont pas encore été étudiées. Puisque SBP2 semble jouer un rôle central dans ce mécanisme, il est vraisemblable que cela soit le cas par l'intermédiaire de son interaction avec l'élément SECIS. Pour tester cette hypothèse, des extraits de cellules exprimant soit la sélénoprotéine GPx1, soit un mutant dont le codon UGA^{Sec} a été remplacé par un codon UGU^{Cys}, soit un mutant dont l'élément SECIS a été délété, seront utilisés et la nature de la coiffe de ces ARNm sera analysée par immunoprécipitation à l'aide d'anticorps anti-m₃G. Ainsi, si l'élément SECIS ou le codon UGA^{Sec} sont essentiels au recrutement de Tgs1, les ARNm des deux mutants de GPx1 ne seront pas immunoprécipités.

Rôle des complexes SMN et du méthylosome dans la biogenèse des mRNP de sélénoprotéines

Dans le cas des snRNP, le complexe SMN fonctionne de manière coordonnée avec le complexe du méthylosome pour méthyler et assembler les protéines Sm en anneau sur les snARN, et c'est la présence combinée des protéines Sm et du complexe SMN qui permet le recrutement de Tgs1 et l'hyperméthylation de la coiffe (Mouaikel et al., 2002; Neuenkirchen et al., 2015; Plessel et al., 1994). La coiffe m₃G et les protéines Sm assemblées sur le snARN sont alors reconnues par la snurportine-1 qui permet l'import des snRNP dans le noyau, où elles exerceront leurs fonctions dans l'épissage des pré-ARNm (Fischer and Lührmann, 1990; Meister et al., 2002; Strasser et al., 2004). Ni les protéines Sm, ni la snurportine-1 n'ont été retrouvées associées aux mRNP de sélénoprotéines, indiquant un devenir très différent de ces particules par

rapport aux snRNP (Figure 73). Nous avons cependant identifié un important réseau d'interactions entre la protéine SBP2 et les différentes sous-unités des complexes SMN et du méthylosome et montré que l'ensemble des composants de ces deux complexes pourrait être impliqué dans la biogenèse des mRNP de sélénoprotéines. En effet, par des expériences d'immunoprécipitation, de co-purification et par des tests d'interaction double-hybride dans la levure, nous avons montré que SBP2 interagit avec 6 des 9 protéines du complexe SMN (SMN, Gemin2, Gemin3, Gemin4, Gemin7 et Gemin8) et avec les trois protéines du méthylosome (PRMT5, MEP50 et pICln). Les domaines impliqués dans ces interactions devront être analysés en détail. Pour cela, l'interaction de chaque domaine individuel de ces différentes protéines avec les domaines de SBP2 sera testée par la technique de double-hybride dans la levure ou de manière plus directe par la co-purification de protéines tronquées *in vitro*.

L'analyse fonctionnelle des protéines des complexes SMN et du méthylosome par inactivation des ARNm des protéines individuelles, par siARN, nous a permis de mieux comprendre le mécanisme d'assemblage des mRNP de sélénoprotéines (Figure 74). L'inactivation de la protéine SMN entraîne une inhibition de l'expression de la sélénoprotéine HA-GPx1, indiquant que SMN, probablement via son rôle dans le recrutement de Tgs1, est impliquée dans la synthèse des sélénoprotéines. Cependant, l'inhibition des sous-unités Gemin4 et Gemin5 du complexe SMN, n'a aucun effet sur l'expression de HA-GPx1. Gemin5 est une protéine de liaison à l'ARN située en périphérie du complexe SMN (Gubitz et al., 2002). La protéine reconnaît et délivre les snARN au complexe SMN, permettant ainsi l'assemblage des protéines Sm (Battle et al., 2006). Nous avons montré précédemment que les protéines Sm ne sont pas associées aux ARNm de sélénoprotéines (cf Résultats I.3) et n'avons pas vu d'interaction entre SBP2 et Gemin5. De manière intéressante, une autre fonction de Gemin5, indépendamment du complexe SMN, a été montrée : la protéine serait impliquée dans l'initiation de la traduction (Piñeiro et al., 2015). En effet, Gemin5 interagit avec le facteur eIF4E (Fierro-Monti et al., 2006) et présente une activité de liaison de la coiffe m⁷G (Bradrick and Gromeier, 2009). Il a été montré que la protéine inhibe l'initiation de la traduction coiffe et IRES-dépendante en interférant potentiellement avec le recrutement du complexe eIF4F vers les ARNm (Pacheco et al., 2009; Piñeiro et al., 2015). Nous avons établi précédemment que le facteur eIF4E

n'est pas impliqué dans la synthèse *in vivo* de la sélénoprotéine HA-GPx1 dont l'ARNm porte une coiffe hyperméthylée m₃G. Ainsi, si la fonction d'inhibition de la traduction de Gemin5 implique son interaction avec le facteur eIF4E ou la coiffe m⁷G des ARNm, il n'est pas étonnant que son inactivation n'affecte pas la synthèse de la sélénoprotéine HA-GPx1, d'autant plus qu'il n'interagit pas avec SBP2. La protéine Gemin2 lie les protéines Sm et assure la spécificité de leur assemblage au niveau des snARN (Grimm et al., 2013; Zhang et al., 2011). Puisque les protéines Sm ne sont pas assemblées sur les ARNm de sélénoprotéines, l'interaction de Gemin2 avec la protéine SBP2 pourrait suggérer un rôle alternatif de cette protéine dans l'assemblage des mRNP de sélénoprotéines. Gemin3 est une hélicase qui assure l'activité chaperonne du complexe SMN ainsi que l'hydrolyse de l'ATP durant la réaction d'assemblage (Campbell, 2000; Charroux et al., 1999; Yan et al., 2003). Le rôle fonctionnel de toutes les protéines Gemin détectées en interaction avec SBP2 devra être testé sur la synthèse des sélénoprotéines, comme nous l'avons fait pour SMN, Gemin4 et Gemin5. L'analyse fonctionnelle de Gemin3 permettra de déterminer si l'activité ATP-dépendante du complexe SMN est requise durant l'assemblage des mRNP de sélénoprotéines.

Des études récentes ont montré que le complexe SMN joue un rôle central dans la biogenèse de nombreuses RNP (voir Introduction chapitre II.1.4.2), dont des snRNP (Meister et al., 2001b; Pellizzoni, 2002; Pillai, 2003; Tisdale et al., 2013a), des snoRNP (données non-publiées, Thèse de M. Dodré, S. Massenet, C. Allmang), la SRP (Signal Recognition Particle) (Piazzon et al., 2013) mais également des mRNP (Fallini et al., 2012b). SMN jouerait un rôle dans la stabilité et la localisation de certains ARNm en modulant leur association avec des protéines de liaison à l'ARN (Akten et al., 2011; Fallini et al., 2011, 2014; Hubers et al., 2011; Tadesse et al., 2008). Ainsi, il est possible qu'en plus de son rôle dans le recrutement de Tgs1, le complexe SMN puisse avoir une autre fonction dans l'assemblage, la stabilité ou la localisation des mRNP de sélénoprotéines. Pour étudier la stabilité des ARNm de sélénoprotéines, nous pouvons bloquer leur transcription grâce à l'Actinomycine D, afin de suivre leur dégradation en fonction du temps et déterminer leur temps de demi-vie. Si le temps de demi-vie des ARNm de sélénoprotéines est diminué après inactivation de SMN, cela indiquera l'implication de SMN dans le maintien de la

stabilité des ARNm de sélénoprotéines. En comparant l'effet de l'inactivation de SMN et de Tgs1 sur le temps de demi-vie des ARNm de sélénoprotéines, on pourra déterminer si le rôle de SMN dans la stabilité de ces ARNm est lié au recrutement de Tgs1 et à l'hyperméthylation de la coiffe, ou s'il intervient également dans une autre étape de la biogenèse des mRNP de sélénoprotéines. Par ailleurs, SMN colocalise avec plusieurs ARNm dans des granules activement transportés lors de processus neuronaux, indiquant qu'elle est impliquée dans le trafic neuronal des mRNP. De manière intéressante, l'ARNm de la sélénoprotéine W, abondante dans le cerveau, colocalise avec SMN dans les neurites et les axones et semble traduit localement dans les synapses (Rage et al., 2013; Raman et al., 2013). SMN pourrait donc également jouer un rôle dans le transport intracellulaire des ARNm de sélénoprotéines et dans leur traduction localisée.

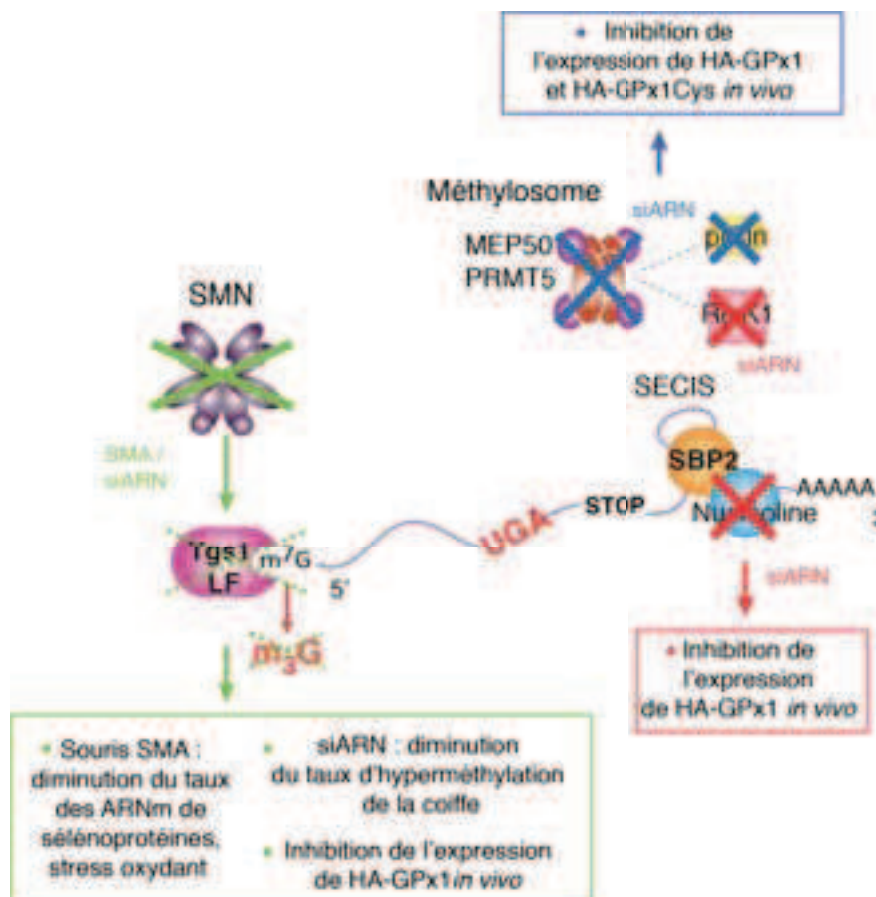


Figure 74. Analyse fonctionnelle du rôle des complexes SMN et du méthylosome dans la synthèse des sélénoprotéines.

Les croix représentent l'inactivation de l'expression des ARNm par siARN, ou l'invalidation du gene SMN dans les cellules de souris modèles de la pathologie SMA. Les croix en pointillés représentent les conséquences des inactivations. Leur impact sur l'expression des sélénoprotéines est indiqué dans les boîtes. La couleur verte est dédiée au complexe SMN, le bleu au méthylosome composé de PRMT5, MEP50, pICln et le rouge à RioK1 et à la nucléoline.

Deux formes du complexe du méthylosome ont été identifiées. L'arginine méthyltransférase PRMT5 associée à MEP50 constituent dans chaque cas le cœur du complexe. Une protéine adaptatrice additionnelle fournit au complexe une spécificité de substrat. Ainsi, associé à pICln, le méthylosome permet de recruter et de modifier les protéines Sm (Friesen et al., 2001). En complexe avec RioK1, le méthylosome a pour cible la nucléoline (Guderian et al., 2011). Nos tests d'interaction indiquent clairement que les protéines PRMT5/MEP50 sont associées à SBP2. pICln interagit avec ce complexe de manière plus labile et aucune interaction avec RioK1 n'a été observée. L'inactivation par siARN de PRMT5, MEP50 et pICln affecte fortement et de manière égale l'expression de HA-GPx1 et d'un mutant dont l'expression ne dépend pas du mécanisme de recodage (HA-GPx1Cys). Ce complexe ne semble donc pas être requis pour l'incorporation de la sélénocystéine mais participe sans doute à l'assemblage et/ou la méthylation d'une protéine cœur de la mRNP, différente des protéines Sm (Figure 74). Un des candidats potentiels pourrait être la protéine hnRNPA1 que nous avons co-purifiée avec l'ARNm HA-GPx1 *in vivo* et qui est connue comme cible de méthyltransférases (Jean-Philippe et al., 2013; Thandapani et al., 2013). La méthylation de hnRNPA1 par PRMT5/MEP50 sera testée *in vitro* afin de déterminer si cette protéine peut être un substrat du méthylosome. Son interaction avec les ARNm de sélénoprotéines devra être vérifiée par des expériences d'immunoprécipitation d'ARN ou *in vitro* par des expériences de retard sur gel. Son interaction avec la protéine SBP2 pourra également être testée par immunoprécipitation ou tests double-hybride dans la levure. Étonnamment, l'inactivation de RioK1 a un impact sur l'expression de HA-GPx1, similaire à celui observé pour l'inhibition de la nucléoline, une protéine de liaison au SECIS, impliquée dans la régulation positive de l'expression des sélénoprotéines (Miniard et al., 2010; Squires et al., 2007b; Wu et al., 2000) (Figure 73 et 74). Les protéines PRMT5/MEP50 pourraient donc, associées de façon transitoire à RioK1, permettre la méthylation et le recrutement de la nucléoline vers les mRNP de sélénoprotéines. Les interactions entre PRMT5, MEP50, RioK1, la nucléoline et SBP2 pourront être analysées comme décrit dans la partie Résultats II.1.2 : les différentes protéines seront co-exprimées dans des cellules d'insecte grâce au système d'expression baculovirus et les protéines en interaction avec GST-SBP2 seront identifiées à l'issue d'expériences de GST pull-down. Au vu de son interaction directe avec le méthylosome, et parce

qu'elle est au centre du mécanisme de recodage des sélénoprotéines, la protéine SBP2 pourrait être le substrat du méthylosome. Cependant, les tests de méthylation *in vitro* ont montré que SBP2 n'est pas méthylée par PRMT5/MEP50. La présence des protéines adaptatrices pICln ou RioK1 pourrait être requise pour le recrutement de SBP2 vers PRMT5/MEP50. Ainsi, la méthylation potentielle de SBP2 par PRMT5/MEP50 sera testée *in vitro* en présence de pICln ou de RioK1. Dans le cas des snRNP, le méthylosome joue un rôle indirect dans le recrutement de Tgs1 du fait de son rôle dans l'assemblage des protéines Sm. Son influence sur l'hyperméthylation de la coiffe des ARNm de sélénoprotéines sera également testée par inactivation de l'ARNm de PRMT5 par siARN suivi de l'immunoprécipitation des ARN à l'aide d'anticorps anti-m₃G.

Pour conclure, nous avons montré que le complexe SMN et la protéine Nop58 participent au recrutement de l'enzyme Tgs1 vers les ARNm et donc à l'hyperméthylation de leur coiffe. Ce réseau pourrait constituer un point de régulation de l'expression des sélénoprotéines, *via* le contrôle de la biogenèse de la coiffe m₃G. Cette modification pourrait avoir lieu dans le noyau ou le cytoplasme et être régulée par le niveau de sélénium ou le stress oxydant. Le méthylosome quant à lui pourrait être impliqué dans la méthylation et l'assemblage d'une ou plusieurs protéines constitutives de la RNP, telles que hnRNPA1 ou la nucléoline. L'assemblage correct est une étape essentielle à l'obtention de mRNP matures et donc à l'expression des sélénoprotéines.

Durant ma thèse, nous avons révélé l'existence d'un lien inattendu entre l'expression des sélénoprotéines et la pathologie de l'amyotrophie spinale (SMA). Cette pathologie est provoquée par des défauts d'expression de la protéine SMN et se caractérise par une dégénérescence des motoneurons de la moelle épinière menant progressivement à l'atrophie musculaire. Il s'agit d'une des causes principales de mortalité infantile d'origine génétique. Nous avons montré que le taux de la majorité des ARNm de sélénoprotéines est diminué dans les cellules de moelle épinière de souris modèles de la pathologie SMA. Nos résultats indiquent que SMN pourrait influencer sur le niveau des ARNm de sélénoprotéines, *via* son rôle dans le recrutement de l'enzyme Tgs1 et l'hyperméthylation de leur coiffe. Le complexe SMN lie également des protéines de

liaison aux ARNm qui sont impliquées dans multiples aspects de la régulation post-transcriptionnelle, incluant le transport et la stabilité des ARNm, ainsi que la traduction localisée dans les neurones (Akten et al., 2011; Fallini et al., 2011, 2014; Hubers et al., 2011; Tadesse et al., 2008). SMN jouerait, entre-autres, un rôle dans la diminution de la stabilité de transcrits tels que ceux contenant des motifs de liaison ARE (AU-Rich Element) en modulant leur association avec des protéines de liaison à l'ARN (pour revue, Fallini et al., 2012). Ainsi, la baisse du taux des ARNm de sélénoprotéines observés dans les cellules de souris SMA pourrait s'expliquer par un défaut du recrutement, dépendant de SMN, d'une protéine de liaison à l'ARN impliquée dans leur stabilité. De manière intéressante, il a été montré que la protéine SMN est associée à des complexes RNP contenant la nucléoline (Lefebvre et al., 2002). Ces complexes sont dissociés dans les fibroblastes de patients SMA indiquant que SMN joue un rôle dans l'assemblage ou la stabilité de ces RNP et que cela pourrait contribuer au phénotype de la pathologie SMA. Le complexe SMN et le méthylosome (associé à RioK1) pourraient donc, ensemble, méthyliser et assembler la nucléoline sur les mRNP de sélénoprotéines. Cette protéine jouerait alors un rôle dans la stabilité et le mécanisme de synthèse des sélénoprotéines (Miniard et al., 2010; Squires et al., 2007b; Wu et al., 2000). Une étude préalable avait montré que les motoneurons de patients SMA sont soumis à un stress oxydant important pouvant être impliqué dans leur dégénérescence (Hayashi et al., 2012). Il est possible que ce stress oxydant soit lié à des défauts de synthèse des sélénoprotéines et qu'il contribuerait aux symptômes de la pathologie SMA. Ce résultat révèle une fois de plus la complexité du mécanisme de synthèse des sélénoprotéines et le caractère essentiel de ces protéines pour la santé humaine.

Rôle biologique de la coiffe m₃G sur la synthèse des sélénoprotéines

Nous avons montré que les ARNm de sélénoprotéines portant une coiffe m₃G sont retrouvés associés aux polysomes dans le cytoplasme, indiquant qu'ils seraient traduits dans les cellules (Wurth et al., 2014). La coiffe m₃G pourrait jouer un rôle lors de l'initiation de la traduction des ARNm de sélénoprotéines ou du mécanisme de recodage (Figure 75). Elle pourrait également augmenter la stabilité de ces ARNm en

leur permettant d'échapper au mécanisme du NMD, ou être impliquée dans l'assemblage et la localisation des ARNm de sélénoprotéines.

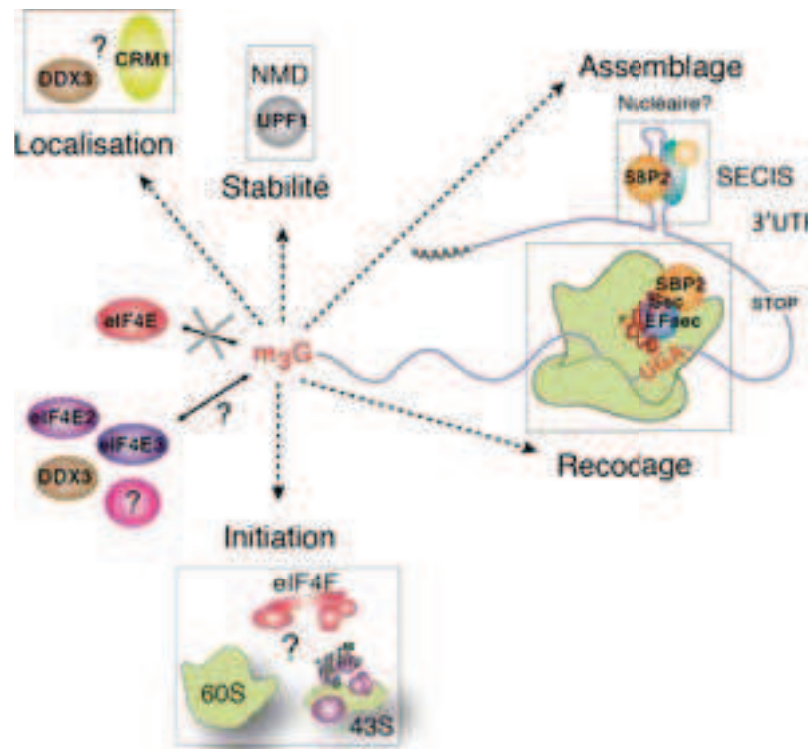


Figure 75. Rôles potentiels de la coiffe m_3G dans la synthèse des sélénoprotéines.

Les ARNm de sélénoprotéines portant une coiffe m_3G ne sont pas associés au facteur d'initiation eIF4E. Cette coiffe pourrait cependant être reconnue par d'autres facteurs tels que les homologues eIF4E2 ou eIF4E3, l'hélicase DDX3 ou un facteur encore inconnu. Les rôles potentiels de la coiffe m_3G dans la synthèse des sélénoprotéines sont indiqués d'une flèche en pointillés.

Rôle de la coiffe m_3G dans l'initiation de la traduction

La présence d'une coiffe m_3G en 5' d'ARNm voués à être traduits pose la question de leur mécanisme d'initiation de la traduction. En effet, dans le cas d'ARNm classiques, la coiffe m^7G est reconnue par le facteur eIF4E qui permet le recrutement de la machinerie d'initiation de la traduction. Or nous avons montré que le facteur eIF4E ne semblait pas essentiel à la traduction des ARNm de sélénoprotéines portant une coiffe m_3G *in vivo* (Wurth et al., 2014). Ainsi, d'autres facteurs pourraient être impliqués dans cette reconnaissance. Paradoxalement, la traduction des ARNm de SelM et GPx1 coiffés par du m_3G est moins efficace que lorsqu'ils sont coiffés par du m^7G dans les lysats de réticulocytes de lapin. De plus, des expériences d'assemblage de complexes traductionnels dans ces extraits sur ces mêmes ARN indiquent que la présence de la coiffe m_3G conduit à une plus faible prise en charge par les ribosomes. Des facteurs de reconnaissance de la coiffe m_3G , impliqués dans l'initiation de la traduction des ARNm de sélénoprotéines pourraient être absents dans les RRL. En effet, il a été

montré que dans ces extraits plusieurs protéines sont limitantes dont SBP2, essentielle au mécanisme de recodage des sélénoprotéines (Copeland et al., 2000; Low et al., 2000b; Panthu et al., 2015; Takeuchi et al., 2009). Par ailleurs, l'ajout de SBP2 stimule le recodage quelque soit la nature de la coiffe des ARNm de sélénoprotéines. Ainsi, il est tout à fait envisageable que la coiffe m₃G joue un rôle dans le mécanisme d'initiation de la traduction. Un des trois modèles du mécanisme de recodage des sélénoprotéines propose l'existence de ribosomes spécialisés, associés à la protéine SBP2 et dédiés à la synthèse des sélénoprotéines (Donovan et al., 2008; Kinzy et al., 2005). La présence d'une coiffe m₃G en 5' des ARNm de sélénoprotéines pourrait permettre le recrutement exclusif de ces ribosomes spécialisés au détriment des ribosomes classiques afin d'augmenter l'efficacité de synthèse des sélénoprotéines. Dans ce cas, des facteurs devraient être impliqués dans la reconnaissance de la coiffe m₃G et dans la modulation de son activité dans l'initiation de la traduction. Enfin, il est également envisageable que la traduction des ARNm de sélénoprotéines soit indépendante de la coiffe.

Reconnaissance de la coiffe m₃G par une protéine homologue de eIF4E ?

De nombreux homologues de la protéine eIF4E sont trouvés chez les eucaryotes, mais la fonction de la plupart reste encore inconnue (Joshi et al., 2004, 2005; Rhoads, 2009). La forme la plus abondante est exprimée de manière ubiquitaire dans tous les organismes et est représentée par eIF4E1 (appelé communément eIF4E) chez les mammifères. Elle est responsable de l'initiation de la traduction de la majorité des ARNm et interagit avec les protéines 4E-BP (eIF4E Binding Proteins), une famille de protéines qui agissent comme répresseur de la traduction coiffe-dépendante en empêchant l'interaction entre eIF4E1 et eIF4G, essentielle au recrutement du ribosome (Mader et al., 1995). Les autres formes de eIF4E sont vraisemblablement impliquées dans des fonctions spécialisées et exprimées à des étapes précises du développement et de la différenciation cellulaire (Cho et al., 2005).

Des études d'interaction de eIF4E1 avec des analogues de la coiffe montrent que eIF4E interagit au moins 8 fois (Carberry et al., 1990) et jusqu'à 760 fois (Niedzwiecka et al., 2002) plus fortement avec les coiffes m⁷G qu'avec les coiffes m₃G *in vitro* (Cai et al., 1999; Niedzwiecka et al., 2002). Ceci est en accord avec nos

résultats montrant que eIF4E n'est pas requis pour la traduction des ARNm de sélénoprotéines qui portent une coiffe m₃G (Wurth et al., 2014). Il existe cinq isoformes du facteur d'initiation eIF4E chez *C. elegans* (IFE-1, 2, 3, 4 et 5) dont trois sont capables de reconnaître la coiffe m₃G avec forte affinité (Keiper et al., 2000). La majorité des ARNm de cet organisme portent une coiffe triméthylée m₃G acquise par un mécanisme de trans-épissage (voir discussion Résultats I.4), et sont traduits (Liou and Blumenthal, 1990). Les facteurs IFE-1, 2 et 5 interagissent avec les coiffes m⁷G et m₃G, tandis que IFE-3 et 4 ne se lient qu'aux coiffes m⁷G (Keiper et al., 2000). Il a été montré que IFE-5 présente une cavité plus ouverte que celle des facteurs eIF4E conventionnels ce qui lui permet d'interagir avec la coiffe m₃G, plus volumineuse que la coiffe m⁷G (Miyoshi et al., 2002). Il apparaît que IFE-1, 2 et 5 résultent de duplications de IFE-3 et font donc partie de la classe des facteurs eIF4E1, qui ont donné naissance à une sous-classe spécialisée d'eIF4E pouvant reconnaître des structures de coiffe alternatives. Aucun homologue de cette sous-classe n'a été retrouvé chez les mammifères (Joshi et al., 2005).

Chez les mammifères on trouve trois formes majeures de eIF4E : eIF4E1, eIF4E2 (4EHP) et eIF4E3 (Joshi et al., 2004, 2005; Rhoads, 2009). eIF4E2 n'interagit pas avec eIF4G et est considéré comme un compétiteur de eIF4E1 (Rom et al., 1998). eIF4E2 possède une affinité 30-100 fois plus faible pour la coiffe m⁷G que eIF4E1 (Rosettani et al., 2007; Zuberek et al., 2007). Cette affinité peut être augmentée en présence de partenaires protéiques tel que le facteur bicoïde par exemple, à la manière de eIF4G pour eIF4E1 (Haghighat and Sonenberg, 1997). En complexe avec les protéines GIGYF2 et ZNF598, eIF4E2 réprime la traduction de certains ARNm spécifiques durant l'embryogenèse chez la souris (Morita et al., 2012). Cependant dans certaines conditions d'hypoxie, eIF4E2 serait impliquée dans la régulation positive de la traduction d'autres ARNm (Uniacke et al., 2012). L'action stimulatrice ou inhibitrice des homologues de eIF4E sur la traduction d'ARNm particuliers semble ainsi déterminée par les partenaires protéiques avec lesquels ils interagissent. Le 3^{ème} homologue, eIF4E3, forme un complexe avec eIF4G mais ne lie pas les protéines 4E-BP (Joshi et al., 2004). Bien que son affinité pour la coiffe m⁷G soit 10 à 40 fois plus faible que celle de eIF4E1, la surexpression de eIF4E3 inhibe l'interaction de eIF4E1 avec la coiffe (Osborne et al., 2013). L'organisation structurale de son site de liaison à

la coiffe est étonnamment distincte de celle de eIF4E1 et eIF4E2 (Joshi et al., 2004; Osborne et al., 2013). En effet, un des résidus aromatiques conservé, impliqué dans l'empilement de la coiffe m⁷G, est absent. eIF4E3 est exprimée de manière tissu-spécifique suggérant un rôle dans la traduction d'ARNm particuliers (Cha et al., 2011). Dans le cadre de mon sujet, il serait intéressant de déterminer si les facteurs eIF4E2 ou eIF4E3 sont capables de reconnaître la coiffe m₃G des ARNm de sélénoprotéines (Figure 75). Pour tester cela, les protéines eIF4E2 et eIF4E3 endogènes seraient immunoprécipitées, et les ARN associés seraient analysés par RT-PCR quantitative dans le but de rechercher les ARNm de sélénoprotéines. Dans l'hypothèse d'un résultat positif, l'interaction *in vitro* de eIF4E2 et eIF4E3 recombinantes avec un ARNm de sélénoprotéine radiomarqué et coiffé par du m₃G serait testée par retard sur gel.

Enfin, il est tout à fait envisageable que la liaison d'un partenaire protéique puisse induire un changement conformationnel conduisant à une augmentation d'affinité d'un homologue de eIF4E pour la coiffe m₃G. Ce partenaire pourrait être un des composant de la mRNP de sélénoprotéines par exemple. Dans un premier temps, des expériences *in vivo* seraient réalisées. L'inactivation de l'expression de eIF4E2 et eIF4E3 par siARN serait effectuée. Si ces facteurs sont impliqués dans la synthèse des sélénoprotéines, l'expression de HA-GPx1 devrait être inhibée. eIF4E3 ne lie pas les répresseurs de l'initiation canonique 4EBP. Dans ce cas, l'inactivation ou la surexpression de 4EBP ne devrait pas avoir d'impact sur la traduction des sélénoprotéines.

Reconnaissance de la coiffe m₃G par un facteur spécialisé ?

Il est également envisageable que la coiffe m₃G des ARNm de sélénoprotéines puisse être reconnue par un autre facteur de liaison à la coiffe. Nous avons montré que l'inactivation de l'hélicase DDX3 a un impact sur la synthèse de la sélénoprotéine HA-GPx1 *in vivo*, indiquant un rôle potentiel de cette protéine dans la synthèse des sélénoprotéines. DDX3 a également été identifiée par spectrométrie de masse dans les complexes de traduction assemblés sur l'ARNm de SelM à partir de lysats de réticulocytes de lapin que nous avons purifiés (données non montrées). Cela indique que DDX3 pourrait être associée aux ARNm de sélénoprotéines et jouerait peut-être

un rôle lors de leur traduction. DDX3 a récemment été caractérisée comme une protéine de liaison à la coiffe qui est requise pour l'initiation de la traduction de l'ARNg du virus VIH-1 (Soto-Rifo et al., 2013). Une autre étude a montré que la coiffe de cet ARNg peut également être hyperméthylée par Tgs1 (Yedavalli and Jeang, 2010). Ainsi, DDX3 constitue un candidat idéal qui reconnaîtrait éventuellement la coiffe m₃G des ARNm de sélénoprotéines et participerait à l'initiation de leur traduction (Figure 75). L'interaction potentielle de DDX3 avec SBP2 ou eEFSec pourrait être analysée par immunoprécipitation ou par des tests d'interaction double hybride dans la levure. La liaison de DDX3 pour l'ARN SECIS serait testée par des expériences de retard sur gel. L'impact de DDX3 sur l'initiation de la traduction des ARNm de sélénoprotéines pourrait être évalué *in vitro*, en immunodéplétant DDX3 des lysats de réticulocytes et en analysant la capacité de ces extraits à assembler des complexes de traduction sur l'ARNm de selM.

Il est également envisageable qu'une autre protéine encore non-identifiée puisse être impliquée dans cette reconnaissance. Durant ma thèse, j'ai mis au point une méthode permettant de purifier les complexes ribonucléiques assemblés *in vivo* sur les ARNm de la sélénoprotéine HA-GPx1 exprimés *de novo*, basée sur l'utilisation d'oligonucléotides complémentaires biotinylés. Plusieurs facteurs impliqués dans la régulation de l'expression des sélénoprotéines, ont été trouvés associés à cet ARNm par spectrométrie de masse. En particulier eIF4A3, dont l'expression est régulée par le sélénium et qui interagit avec certains ARN SECIS, dont celui de GPx1. eIF4A3 empêche le recrutement de SBP2 et inhibe le mécanisme de recodage des sélénoprotéines (Budiman et al., 2009a, 2011). Les complexes que nous avons majoritairement purifiés seraient donc des complexes de régulation négative de la traduction ou soumis au NMD. L'objectif futur sera d'améliorer nos conditions expérimentales, en surexprimant SBP2 par exemple ou en faisant varier divers paramètres tels la concentration de sélénium, afin d'augmenter les chances de purifier des complexes de traduction ou de recodage actifs. La caractérisation de ces complexes nous permettra peut-être de déterminer la nature des facteurs associés à la coiffe m₃G, voire d'identifier de nouveaux facteurs impliqués dans le mécanisme de traduction et de recodage des ARNm de sélénoprotéines.

Rôle de la coiffe m₃G dans le mécanisme de recodage

L'inactivation de Tgs1, l'enzyme responsable de l'hyperméthylation de la coiffe, inhibe la synthèse de la sélénoprotéine HA-GPx1 mais pas d'un mutant dans lequel la sélénocystéine est mutée en cystéine et dont la synthèse ne dépend pas d'un mécanisme de recodage (Wurth et al., 2014). Ainsi, le rôle de Tgs1, via l'hyperméthylation de la coiffe, pourrait être de stimuler l'étape de recodage et d'incorporation de la sélénocystéine. Plusieurs études indiquent que le mécanisme de recodage des sélénoprotéines est intrinsèquement inefficace (Berry et al., 1994; Low et al., 2000a; Suppmann et al., 1999; Weiss Sachdev and Sunde, 2001) et il a été proposé que des facteurs encore inconnus pourraient être requis afin d'optimiser la synthèse des sélénoprotéines (Mehta et al., 2004a). La coiffe m₃G pourrait constituer l'un de ces facteurs. De plus, *in vitro*, nous avons montré que la coiffe m⁷G restait accessible après la formation des complexes de traduction et pouvait être hyperméthylée par Tgs1. Il est possible que, dans les cellules, la traduction soit initiée sur des ARNm de sélénoprotéines portant une coiffe m⁷G et que cette dernière, serait hyperméthylée dans un second temps, permettant la stimulation du mécanisme de recodage. La coiffe m₃G, s'associerait à d'autres éléments de la mRNP nécessaires au recodage tels que SBP2 ou eEFSec afin de stimuler l'incorporation de sélénocystéine (Figure 75). Des analogues de la coiffe m⁷G comportant un nucléotide photo-activable capable de ponter les protéines en interaction avec la coiffe m⁷G ont récemment été développés (Nowakowska et al., 2014). Le développement d'un analogue photoactivable de la coiffe m₃G et son ajout à l'extrémité 5' d'un ARNm de sélénoprotéine pourrait permettre d'identifier les protéines qui y seraient directement liées.

Rôle de la coiffe m₃G dans l'assemblage, la stabilité et la localisation des mRNP de sélénoprotéines

Nous avons montré que le recrutement de Tgs1 vers les ARNm de sélénoprotéines pouvait faire intervenir soit le complexe SMN soit la protéine Nop58. Le recrutement de Tgs1 par le biais de Nop58 suggère que l'hyperméthylation de la coiffe des ARNm de sélénoprotéines puisse se produire dans le noyau, comme pour les snoARN. Nos

expériences de fractionnement subcellulaires indiquent que les ARNm de sélénoprotéines qui portent une coiffe m₃G sont majoritairement cytoplasmiques (Wurth et al., 2014) mais que 25-30% des ARNm sont retrouvés dans le noyau. De plus, contrairement aux snARN, les ARNm de sélénoprotéines n'interagissent pas avec la protéine snurportine-1 qui est nécessaire au réimport des snRNP dans le noyau. Ainsi, les ARNm de sélénoprotéines coiffés de m₃G détectés dans le noyau ne proviendraient pas d'un réimport nucléaire d'ARN triméthylés dans le cytoplasme, mais pourraient résulter d'une hyperméthylation nucléaire. Plusieurs études, révèlent la présence des facteurs de recodage dans le noyau et proposent un mode d'assemblage nucléaire des mRNP de sélénoprotéines (Allmang et al., 2009; de Jesus et al., 2006; Papp et al., 2006; Touat-Hamici et al., 2014). L'hyperméthylation de la coiffe m₃G pourrait constituer une étape de maturation requise pour l'assemblage de facteurs nucléaires essentiels à l'expression des sélénoprotéines. Cet assemblage précoce de facteurs dans le noyau, pourrait être requis pour leur traduction, une fois les ARNm exportés dans le cytoplasme.

L'hyperméthylation de la coiffe est une étape critique dans la biogenèse des sn et snoRNP, et est requise à la formation des RNP matures et fonctionnelles (Matera et al., 2007a; Mouaikel et al., 2002). Elle joue également un rôle important dans leur stabilité (Hamm et al., 1990; Reddy et al., 1992; Speckmann et al., 2000). La coiffe m₃G des ARNm de sélénoprotéines pourrait augmenter leur stabilité en empêchant leur dégradation par le mécanisme du NMD. Elle permettrait par exemple un mode d'export cytoplasmique particulier, ou recruterait des facteurs spécifiques qui inhiberaient l'association de protéines du NMD (Figure 75). Il a été montré que la coiffe de certains ARN du virus du VIH-1 est également hyperméthylée par Tgs1 dans le noyau et que cette modification permet à des ARN viraux non-épissés d'échapper au NMD via un mécanisme d'export cytoplasmique particulier (Yedavalli and Jeang, 2010). En effet, les ARNm de la cellule ne peuvent habituellement quitter le noyau avant d'être pleinement fonctionnels et compétents pour la traduction. Ce n'est qu'une fois assemblés, maturés et épissés qu'ils sont exportés dans le cytoplasme via les facteurs Tap-p15 et le complexe TREX (Izaurralde, 2002). Puisque le cycle de réplication du VIH-1 requiert la traduction de l'ARNg non-épissé pour synthétiser des protéines structurales et accessoires, le virus doit déjouer le système de

surveillance de la cellule afin d'exporter ses ARN vers le cytoplasme. Dans ce cas, l'hyperméthylation de la coiffe stimule la prise en charge de ces ARN non-épissés par le facteur d'export CRM1, permettant leur export dans le cytoplasme et augmentant leur expression (Cullen, 2003; Pollard and Malim, 1998). Durant l'export cytoplasmique, les mRNP subissent une étape de remodelage : certaines protéines nucléaires sont ainsi remplacées par des protéines cytoplasmiques. Ce remodelage affecte de manière positive l'efficacité du mécanisme de contrôle-qualité du NMD (Popp and Maquat, 2014). En effet, durant ou directement après l'export des mRNP dans le cytoplasme, le complexe de liaison à la coiffe CBC est remplacé par eIF4E via l'importine- β (Dias et al., 2009; Sato and Maquat, 2009) et les protéines de liaison à la queue poly(A) nucléaires (PABPN1) sont remplacées par des protéines cytoplasmiques (PABPC1) (Sato and Maquat, 2009). On pourrait imaginer que les ARNm de sélénoprotéines, sensibles au NMD (Squires et al., 2007b), puissent échapper à ce mécanisme de contrôle qualité grâce à la coiffe m₃G et peut-être CRM1, afin d'être traduits dans le cytoplasme (Figure 75). Il serait intéressant de tester si l'inhibition de l'export nucléaire via CRM1, par la leptomycine B, provoque une accumulation des ARNm qui portent une coiffe m₃G dans le noyau. Dans le cas des ARN du VIH-1, l'export cytoplasmique implique la protéine DDX3 (Mahboobi et al., 2015). Ceci constitue une autre similitude avec notre mécanisme puisque nous avons montré que l'inactivation par siARN de DDX3 a un fort impact sur l'expression de HA-GPx1 *in vivo*. Des expériences d'immunolocalisation des facteurs impliqués dans la biogenèse de la coiffe m₃G, viseront à déterminer avec précision le compartiment dans lequel a lieu l'hyperméthylation (voir paragraphe précédent), ce qui permettra de mieux comprendre le rôle de la coiffe m₃G dans la synthèse des sélénoprotéines.

Par ailleurs, la question de stabilité n'a pas directement été testée. Cependant, nous avons montré que des défauts d'expression de la protéine SMN mènent à une baisse du taux de certains ARNm de sélénoprotéines dans les cellules de la moelle épinière de souris modèles de la pathologie SMA. Les ARNm de sélénoprotéines les plus sensibles à l'absence de SMN correspondent à ceux qui portent une coiffe hyperméthylée m₃G. Ainsi, puisque SMN est impliqué dans le recrutement de l'enzyme Tgs1, il est possible que la diminution du taux d'ARNm observé soit liée à une baisse du taux d'hyperméthylation de leur coiffe. Ces résultats sont cependant

trop indirects pour conclure que la coiffe m₃G puisse être impliquée dans la stabilité des ARNm de sélénoprotéines. Ainsi, l'analyse de l'impact de la coiffe m₃G sur la stabilité des ARNm de sélénoprotéines constitue un des objectifs futurs du laboratoire. Pour tester cela, l'impact de l'inactivation de Tgs1 sur la demi-vie des ARNm de sélénoprotéines sera analysé de la même manière que proposé précédemment pour SMN.

Pour conclure, plusieurs fonctions peuvent être envisagées pour la coiffe m₃G, au niveau de l'étape d'initiation ou de recodage de la traduction mais également au niveau de la régulation post-transcriptionnelle de la mRNP. La mise en place d'un processus alternatif d'initiation pourrait être une cause mais également une conséquence de la présence de cette coiffe.

Un mécanisme d'initiation atypique de la traduction des ARNm de sélénoprotéines ?

L'hypothèse d'un mécanisme d'initiation de la traduction atypique découle directement du fait de la présence d'une coiffe m₃G sur les ARNm de sélénoprotéines. Nous avons montré que, pour certains ARNm, ce processus ne fait pas intervenir le facteur d'initiation canonique eIF4E indiquant que d'autres facteurs seraient requis pour l'initiation des ARNm de sélénoprotéines qui portent une coiffe m₃G. Les expériences d'assemblage des complexes de traduction révèlent la présence de ribosome 80S sur l'ARNm en présence d'un inhibiteur spécifique, qui permet de bloquer l'étape d'initiation. Cependant, étonnamment nos expériences de toeprint ne nous ont pas permis de localiser le ribosome bloqué à l'initiation, au niveau du codon AUG sur les ARNm des sélénoprotéines SelM et GPx1. Cette singularité pourrait être le signe d'un mécanisme d'initiation non-canonique qui impliquerait un mode de recrutement du ribosome différent ou à un autre endroit sur l'ARNm des sélénoprotéines.

Nous avons analysé le mécanisme d'initiation de la traduction de l'ARNm de la sélénoprotéine M grâce à l'utilisation d'inhibiteurs spécifiques de cette étape durant des expériences d'assemblage de complexes traductionnels dans les lysats de

réticulocytes de lapins. Une singularité dans le mécanisme d'initiation de cet ARNm a pu être observée. En effet, aucun complexe de pré-initiation 48S n'a pu être observé sur l'ARNm de SelM en présence de GMP-PNP, un inhibiteur spécifique des GTPases, qui bloque la formation du complexe 80S mais pas du complexe 48S. Ceci pourrait s'expliquer par le fait que le mécanisme d'initiation de cet ARNm requiert une activité GTPasique supplémentaire par rapport au mécanisme d'initiation canonique et qui serait différente des facteurs d'initiation eIF2 et eIF5B (Wilson et al., 2000b). Il a été proposé que GAPsec, une protéine impliquée dans l'activation des GTPases (GTPase-activating protein, GAP), par son interaction avec EFSec, jouerait un rôle dans le mécanisme de redéfinition du codon UGA^{Sec} des sélénoprotéines (Hirosawa-Takamori et al., 2009). Cette protéine pourrait également être impliquée dans l'activation de la GTPase requise au mécanisme d'initiation des ARNm de sélénoprotéines de SelM. Une des protéines que nous avons purifiée avec l'ARNm de HA-GPx1, la protéine U5-116kD (U5S1 ou EFTuD2) possède une activité GTPase et constituerait le candidat idéal pour assurer la fonction que nous recherchons. En effet, cette même protéine avait précédemment été retrouvée en interaction avec la protéine SBP2 lors d'un crible double-hybride réalisé dans notre laboratoire (données de C. Allmang). Bien qu'impliquée dans l'épissage des ARNm (Fabrizio et al., 1997), cette protéine est également un homologue du facteur d'élongation EF-2 (translocase du ribosome). Ainsi, il serait intéressant d'analyser le rôle fonctionnel de GAPsec et de U5-116kD en évaluant l'effet de leur inactivation sur la synthèse de la sélénoprotéine HA-GPx1. Leur interaction avec les différents facteurs de la machinerie de recodage, dont SBP2, pourrait être testée. Enfin, les extraits de réticulocytes de lapin pourraient être immunodéplétés de U5-116kD. Si l'activité GTPase de U5-116kD est requise pour le recrutement des ribosomes sur l'ARNm de SelM, aucun complexe ne devrait être obtenu à partir de ces extraits.

Puisque nous n'avons pas vu de complexe 48S associé à l'ARNm de SelM, il est envisageable que le mécanisme d'initiation de cet ARNm soit indépendant de cette étape, et passe directement par l'assemblage du complexe 80S selon un mécanisme GTP-dépendant. Chez les procaryotes, certains ARNm sont capables de recruter directement le ribosome 70S, en absence de facteurs d'initiation, et ne recrutent pas la petite sous-unité 30S (Balakin et al., 1992; Moll et al., 2002). Ces ARNm appelés

« leaderless » recrutent le ribosome au niveau du codon AUG mais n'ont pas de séquence Shine-Dalgarno habituellement requise à cette étape d'initiation, chez ces organismes. Cette classe d'ARNm est également présente chez les archae et les mitochondries de mammifères (Li and Wang, 2004; Malys and McCarthy, 2011). Une étude a montré que le ribosome eucaryote est capable de s'assembler sur ces ARNm de la même manière que le ribosome procaryote (Andreev et al., 2006). L'IRES IGR du virus CrPv (*Cricket Paralysis Virus*) est également capable de recruter directement le ribosome 80S en absence de facteurs d'initiation de la traduction (Pisarev et al., 2005; Wilson et al., 2000b). Ainsi, on pourrait imaginer que le ribosome soit recruté sur l'ARNm de SelM selon un mécanisme similaire, c'est à dire directement sous forme de particule 80S, ce qui expliquerait l'impossibilité d'observer la présence de complexe 48S. Pour tester cette hypothèse, l'ARNm de SelM radiomarké sera incubé en présence de ribosome 80S ou de complexe de pré-initiation 43S purifiés, et leur assemblage sur l'ARNm sera analysé par gradient de sucrose. S'il existe, ce mécanisme d'initiation nécessiterait la présence d'une GTPase, puisqu'aucun complexe n'est formé sur l'ARNm de SelM en présence de l'inhibiteur édéine. Le processus d'initiation atypique de l'ARNm de SelM pourrait faire intervenir une structure de type IRES, le recrutement direct de la particule 80S et l'intervention d'une protéine qui apporte une activité GTPase, essentielle à l'assemblage du complexe 80S, telle que U5-116kD. Un autre argument en faveur d'un mécanisme d'initiation de ce type pour les ARNm de sélénoprotéines vient de nos données concernant la protéine DDX3. En effet, nous avons identifié DDX3 au sein des complexes de traduction associés à l'ARNm de SelM et nos résultats indiquent que cette hélicase ATP dépendante serait importante pour l'expression des sélénoprotéines. DDX3 est nécessaire à la traduction IRES-dépendante du virus du VIH-1 (Liu et al., 2011). En effet, DDX3 est capable de se substituer au facteur d'initiation eIF4E dans la liaison à la coiffe m⁷G de l'ARNg du virus VIH-1 où il génère la formation d'un complexe trimérique DDX3/PABP/eIF4G (Soto-Rifo et al., 2013). Ce complexe de pré-initiation ne contient pas les protéines de liaisons à la coiffe CBP20/CBP80 ou eIF4E et est requis pour l'initiation de la traduction de l'ARNg (Soto-Rifo et al., 2012, 2013). Une autre étude a montré que ces mêmes ARN du VIH-1 auraient une coiffe m₃G nécessaire à leur expression (Yedavalli and Jeang, 2010). Ainsi, DDX3 pourrait être impliquée dans une initiation de la traduction de

type IRES de l'ARNm de SelM, indépendamment du facteur d'initiation eIF4E et qui pourrait faire intervenir la coiffe m³G. La capacité de DDX3 à interagir directement avec les ARNm de sélénoprotéines sera vérifiée par des expériences de retard sur gel. Dans l'hypothèse d'un résultat positif, la zone d'interaction précise de la protéine avec l'ARNm pourrait être déterminée par des expériences de protection à la DNase I.

NSEP1 (ou YB-1), une protéine de liaison au SECIS est impliquée dans la synthèse des sélénoprotéines (Fagegaltier et al., 2000c; Shen et al., 1998, 2006), mais également dans la traduction IRES-dépendante. En effet, NSEP1 a été proposée comme ITAF (IRES *trans* acting factors), qui sont des protéines chaperon impliquées dans le recrutement du ribosome vers les IRES (Cobbold et al., 2008; King et al., 2010). NSEP1 interagirait avec les IRES cellulaires de la famille myc et régulerait de manière positive leur traduction IRES-dépendante (Cobbold et al., 2008). L'existence d'IRES cellulaire a été proposé par de nombreuses études (Bushell et al., 2006; Cobbold et al., 2008; Henis-Korenblit et al., 2000; Holcik et al., 1999; Macejak and Sarnow, 1991; Qin and Sarnow, 2004) mais reste encore largement controversée (Jackson, 2013; Shatsky et al., 2010, 2014). Les ARNm cellulaires contenant une IRES sont caractérisés par (i) une longue 5'UTR riche en résidus GC, prédite hautement structurée et (ii) leur traduction qui persiste en condition de stress, durant la mitose ou quand l'activité de eIF4E est compromise (Jackson, 2013; Vagner et al., 2001). L'analyse du transcriptome à grande échelle (Kapeli and Yeo, 2012) a dévoilé l'existence d'une diversité remarquable du mode de recrutement des ribosomes eucaryotes par les ARNm cellulaires (Shatsky et al., 2014). Ces données permettent une classification des ARNm en plusieurs groupes basés sur leur capacité à recruter les différents composants de eIF4F, leurs homologues ou d'autres facteurs (Shatsky et al., 2014). Ceci remet en question le concept strict, généralement admis, de la dépendance de l'initiation vis à vis de eIF4E/m⁷G et propose une alternative au concept d'IRES cellulaire. Ainsi, la traduction de certains ARNm est peu dépendante de eIF4E/m⁷G mais ces ARNm ont une plus forte affinité pour eIF4G/eIF4A et pourraient recruter ces facteurs via des structures secondaires présentes dans leur 5'UTR (Badura et al., 2012; Wallace et al., 2010). D'autres ARNm utilisent des variants de eIF4E ou eIF4G (Coldwell et al., 2012), ou d'autres facteurs pour assembler le ribosome (Shatsky et al., 2014).

Il serait donc essentiel de déterminer l'implication des facteurs d'initiation sur la traduction des ARNm de sélénoprotéines. Pour cela, les facteurs eIF4G ou eIF4A seraient inactivés par siARN et l'impact de leur inhibition sur la traduction de HA-GPx1 pourrait être analysée comme nous l'avons fait pour eIF4E. Ainsi, si la traduction des ARNm de sélénoprotéines dépend fortement de eIF4G/eIF4A, une inhibition de l'expression de HA-GPx1 devrait être observée. La présence potentielle d'un IRES dans la région 5'UTR d'un ARNm de sélénoprotéine pourrait être testée à l'aide d'un rapporteur bicistronique. Pour cela, la région d'intérêt serait clonée entre les gènes de la luciférase rénilla et de celle de la luciole, et leur expression serait analysée.

Les IRES constituent des séquences d'environ 200-300 nucléotides dans la région 5' non-codante du génome des virus. Les extrémités 5'UTR des ARNm de sélénoprotéines sont généralement courtes (SelM : 66 Nts, GPx1 : 80 Nts) et ne peuvent donc pas contenir une IRES au sens stricte. On pourrait imaginer la présence d'une structure de type IRES qui serait présente dans la région 5'UTR ou ailleurs dans l'ARNm et qui permettrait de recruter des facteurs d'initiation ou le ribosome ; un mécanisme qui présenterait des similarités avec le mode d'initiation de l'ARNm de l'histone H4 (Martin et al., 2011), dépendant de la coiffe m⁷G, et dont le recrutement des ribosomes se fait par "tethering". Ce processus pourrait éventuellement dépendre de la coiffe m₃G.

Plusieurs études ont montré que la région 3'UTR des ARNm peut être impliquée dans la régulation de leur traduction. En effet, l'expression de certains ARNm, dont la région 3'UTR est la cible de protéines qui inhibent l'initiation de la traduction, est spécifiquement régulée au cours des premières phases du développement et de la différenciation cellulaire (pour revue, De Moor et al., 2005). Chez certains virus de plantes, des éléments structuraux spécifiques appelés 3'CITE (3' cap-independent translation enhancer) sont retrouvés dans la région 3' des génomes viraux et sont capables de recruter des facteurs d'initiation ou des particules ribosomiques (Gao et al., 2014; Nicholson and White, 2011; Simon and Miller, 2013). Ces éléments sont essentiels au mécanisme d'initiation de la traduction de ces ARNg. Durant ma thèse j'ai pu montrer que la région 3'UTR de l'ARNm de SelM est indispensable au

recrutement des ribosomes. De plus, une tige-boucle prédite en aval de l'élément SECIS, semble impliquée dans une interaction avec des protéines (Figure 76). Cet élément pourrait, éventuellement avec le SECIS, être impliqué dans le recrutement de facteurs importants pour le mécanisme d'initiation de la traduction de ces ARNm et le recrutement de ribosomes.

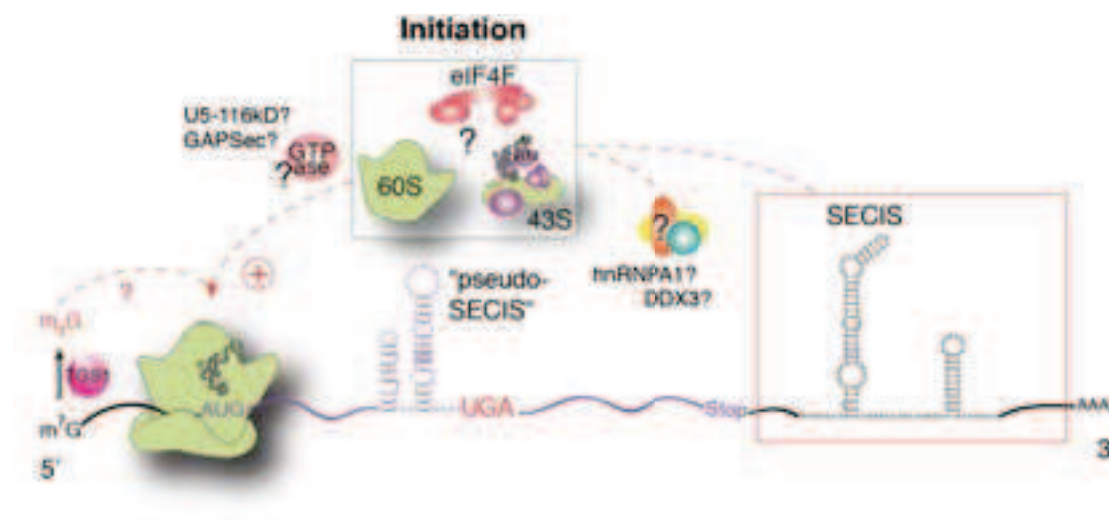


Figure 76. Le mécanisme d'initiation de la traduction des ARNm de sélénoprotéines est atypique. La formation du complexe de pré-initiation 48S sur l'ARNm de SelM requiert une GTPase supplémentaire (les protéines U5-116kD ou GAPSec constituent GTPases potentielles). La région 3'UTR est essentielle pour le recrutement des ribosomes sur les ARNm de sélénoprotéines. La coiffe m_3G et d'autres facteurs importants pour l'expression des sélénoprotéines tel que DDX3, pourraient jouer un rôle dans le mécanisme d'initiation de la traduction.

Dans un premier temps, l'implication du facteur d'initiation eIF4F dans le recrutement des ribosomes vers les ARNm de sélénoprotéines pourrait être testé. Pour cela, des expériences d'assemblage des ARNm de sélénoprotéines radiomarqués dans les lysats de réticulocytes de lapins seraient réalisées en présence d'inhibiteurs de eIF4F. Des inhibiteurs des trois composants du facteur d'initiation eIF4F ont été décrits dans la littérature et pourraient être utilisés. L'inhibiteur 4EGI inhibe eIF4E et induit sa dissociation de eIF4G, l'inhibiteur FL3 (fourni par Laurent Desaubry) est un inhibiteur de eIF4A et enfin, la protéase 2A du virus HRV (*Human Rhino Virus*) permet le clivage et l'inhibition spécifique de eIF4G. Si eIF4F participe au mécanisme de traduction des ARNm de sélénoprotéines, l'inhibition de ces sous-unités empêchera la formation des polysomes. Dans ce cas, la capacité de la région 3'UTR à les recruter devra être analysée. La liaison des facteurs d'initiation eIF4E,

eIF4G et eIF4A purifiés sur un ARN transcrit *in vitro* correspondant à la région 3'UTR isolée sera testée par des expériences de retard sur gel. La contribution potentielle de la tige-boucle prédite en aval de l'élément SECIS sur le recrutement de facteurs sera analysée plus en détail par des expériences de délétion. Si elle joue un rôle dans le recrutement de facteurs, un ARN biotinylé correspondant à cette tige-boucle pourra être utilisé comme appât et incubé en présence de lysat de réticulocytes de lapin ou d'extraits cellulaires et les protéines qui s'y associeront seront analysées par spectrométrie de masse. Les éléments 3'CITE sont connus pour interagir avec une structure en tige-boucle localisée à l'extrémité 5' de l'ARN viral via des interactions ARN-ARN en « kissing loop » (Nicholson et al., 2010, 2013; Treder et al., 2008). Ainsi, si un élément de type 3'CITE est impliqué dans l'initiation de la traduction de l'ARNm de SelM, l'impact de la délétion de la région 5'UTR sur le recrutement des ribosomes devra également être testé.

Au vu de l'ensemble de nos résultats, il apparaît indispensable d'obtenir des informations concernant la structure de l'ARNm, au niveau des régions 5' et 3'UTR mais également aux alentours du codon UGA^{Sec}. La détermination de la structure secondaire de l'ARNm de SelM sera réalisée par des méthodes d'analyse structurale en solution telles le sondage chimique et enzymatique (Jaeger et al., 2006) ou la méthode de SHAPE (Selective 2'-Hydroxyl Acylation analyzed by Primer Extension) (Merino et al., 2005). Il serait également intéressant d'analyser la structure de l'ARNm pour lequel le codon UGA^{Sec} est muté en codon UGU^{cys} afin de déterminer si le changement d'un nucléotide a un impact sur la structure de l'ARNm et pourrait expliquer les différences que nous observons dans les mécanismes d'assemblage de ces deux ARNm. Il serait nécessaire de réaliser ces études pour différents ARNm de sélénoprotéines afin de vérifier si nos observations sont généralisables à l'ensemble des ARNm de sélénoprotéines. En effet, si ce processus a un lien avec la coiffe m₃G, des différences au sein d'ARNm de sélénoprotéines distincts devraient être observées.

Mécanisme de recodage

La protéine SBP2 se lie en amont du codon UGA^{Sec} de l'ARNm de SelM

La protéine SBP2, essentielle au mécanisme de recodage du codon UGA^{Sec}, interagit avec l'élément SECIS dans la région 3'UTR des ARNm des sélénoprotéines eucaryotiques (Caban et al., 2007; Copeland et al., 2000, 2001; Kinzy et al., 2005; Lescure et al., 2002c). Chez les bactéries, en revanche, l'élément SECIS est présent dans la région codante directement en aval du codon UGA^{Sec}, ce qui permet d'apporter les facteurs d'incorporation de la sélénocystéine au niveau du site de recodage (Böck et al., 2006; Stock and Rother, 2009; Yoshizawa and Böck, 2009). Nos expériences indiquent que la protéine SBP2 serait également capable d'interagir en amont du codon UGA^{Sec} des ARNm de SelM et GPx1 eucaryotes. Des prédictions de structure secondaires réalisées à l'aide du logiciel SECISaln (Chapple et al., 2009a) suggèrent la présence d'une structure semblable à un SECIS, que nous avons appelé "pseudo-SECIS". La présence d'une protéine (appelée alors SBP, pour SECIS Binding Protein) capable de se lier à la fois au SECIS en 3'UTR et à une séquence en amont du codon UGA^{Sec} avait été détectée avant la caractérisation de SBP2 (Mizutani and Fujiwara, 2000). En effet, SBP peut être déplacée de l'ARN SECIS par des expériences de compétition à l'aide d'un fragment d'ARN correspondant à la séquence en amont du codon UGA^{Sec}. Cette protéine, du fait de son faible poids moléculaire n'a pas été assimilée à SBP2 et son étude n'a pas été poursuivie. Il est envisageable que la protéine SBP identifiée par Mizutani and Fujiwara, 2000 puisse être une forme tronquée de SBP2. L'analyse de la structure secondaire de l'ARNm de SelM, proposée précédemment, permettra de confirmer la présence de l'élément "pseudo-SECIS" en amont du codon UGA^{Sec}. Dans le but de valider l'interaction de SBP2 avec l'élément "pseudo-SECIS" de l'ARNm de SelM, nous avons réalisé des expériences de retard sur gel et avons observé que SBP2 ne lie que faiblement le « pseudo-SECIS ». L'ARN utilisé présente cependant plusieurs conformations, pouvant indiquer un problème de repliement lié au choix de la séquence. Ces données préliminaires devront donc être confirmées et l'interaction du "pseudo-SECIS" de GPx1 avec SBP2 devra être vérifiée de la même manière. Dans les lysats de réticulocytes de lapin, nous avons observé une pause de la reverse transcriptase au

niveau du “pseudo-SECIS” différente de celle liée à SBP2. Cette pause pourrait être induite par l’interaction d’une protéine présente dans l’extrait. De manière intéressante, cette pause est observée au même endroit de la structure du “pseudo-SECIS” pour les deux ARNm de SelM et de GPx1. Ceci suggère l’existence d’une protéine qui pourrait lier le “pseudo-SECIS” et assister ou renforcer l’interaction de SBP2 au niveau de cet élément. Le “pseudo-SECIS” pourrait participer au mécanisme du recodage en assurant la communication entre le ribosome au niveau du codon UGA^{Sec} et l’élément SECIS dans la région 3’UTR. Dans les autres ARNm de sélénoprotéines, aucune structure “pseudo-SECIS” n’a pu être prédite directement en amont du codon UGA^{Sec} avec le logiciel SECISaln. Un “pseudo-SECIS” a été retrouvé cependant plus de 90 nucléotides en amont du codon UGA^{Sec} de l’ARNm de SelO. Il est possible que la position du “pseudo-SECIS” relative au codon UGA^{Sec} puisse varier d’un ARNm de sélénoprotéine à l’autre, de la même manière que varie celle du SECIS par rapport au codon UGA^{Sec} (Grundner-Culemann et al., 1999; Latreche et al., 2009). Ainsi, il serait intéressant de rechercher de manière systématique la présence d’éléments “pseudo-SECIS” au sein des ARNm de sélénoprotéines.

L’ARN SECIS adopte une structure en tige-boucle formée de deux hélices séparées d’une boucle interne. Le quartet de paires de bases non Watson-Crick, au pied de l’hélice II, et le motif conservé AAR, localisé dans la boucle apicale, sont deux éléments du SECIS essentiels au mécanisme de recodage du codon UGA^{Sec} (Fagegaltier et al., 2000a; Walczak et al., 1996, 1998). Le “pseudo-SECIS” de SelM, hormis sa position dans la région codante, présente ces mêmes caractéristiques. Cependant, dans les “pseudo-SECIS” de GPx1 et de SelO, malgré la présence du quartet de paires de bases non Watson-Crick, les deux adénines du motif AAR sont remplacés par CC et CA, de plus la longueur des hélices et la taille des boucles internes diffèrent légèrement de la structure consensus du SECIS. Ces éléments pourraient également constituer des variants d’éléments de type SRE (Selenocysteine Redefinition Element) qui sont des structures en tige-boucle présentes dans la séquence codante de certains ARNm de sélénoprotéines impliquées dans le mécanisme de recodage (Bubenik et al., 2013; Howard et al., 2005a, 2007b). Cependant, les éléments SRE sont, jusqu’à présent, trouvés en aval du codon UGA^{Sec}

et leur interaction avec SBP2 n'a pas été analysée. Par ailleurs, il a été montré qu'un élément SECIS, présent dans la région codante d'un ARNm de sélénoprotéine viral permettrait l'incorporation de sélénocystéine chez les mammifères (Mix et al., 2007). Cet élément est localisé dans la région codante de l'ARNm, 350 nucléotides en aval du codon UGA^{Sec}. Le déplacement de cet élément en amont du codon UGA^{Sec} n'affecte pas le mécanisme de recodage (Mix et al., 2007). D'après cette étude, la localisation de l'élément SECIS dans la région 3'UTR ne serait pas une nécessité fonctionnelle mais plutôt une adaptation évolutive qui permettrait une plus grande efficacité de synthèse des sélénoprotéines. On peut se demander si le "pseudo-SECIS" aurait une fonction redondante à celle du SECIS et permettrait le recodage du codon UGA^{Sec}. Nos expériences de traduction *in vitro*, en accord avec la littérature, montrent que, lorsque l'élément SECIS est délété de la région 3'UTR des ARNm de SelM et de GPx1, aucun produit de traduction n'est détecté, indiquant que le recodage n'a pas lieu. Le rôle de l'élément "pseudo-SECIS" pourrait cependant être d'assurer la communication entre le ribosome au niveau du codon UGA^{Sec} et l'élément SECIS dans la région 3'UTR (Figure 77). L'importance fonctionnelle du "pseudo-SECIS" sera étudiée. Pour cela, des mutations seront introduites dans la séquence de cet élément, et leur impact sur la traduction *in vitro* de l'ARNm sera analysé.



Figure 77. Rôle potentiel de la structure "pseudo-SECIS" dans le mécanisme de recodage des sélénoprotéines.

L'élément "pseudo-SECIS" pourrait assurer la communication entre l'élément SECIS, le codon UGA^{Sec} et le ribosome par le biais d'une interaction avec SBP2 et/ou d'autres protéines lors du mécanisme de recodage.

Vers une analyse structurale des complexes de traduction des ARNm de sélénoprotéines

Durant ma thèse, nous avons mis au point une méthode qui a permis de purifier pour la première fois des complexes traductionnels associés à un ARNm de sélénoprotéine, dans des conditions compatibles à une analyse structurale par cryo-microscopie électronique. Ce projet collaboratif avec l'équipe de Yaser Hashem (IBMC, Strasbourg) nous a permis d'obtenir la structure de trois complexes d'élongation. L'objectif futur sera de rechercher et d'améliorer les conditions permettant de purifier un complexe de recodage. Pour cela, les extraits de réticulocytes de lapin seront complétés par différents facteurs essentiels et spécifiques du mécanisme de recodage des sélénoprotéines tels que SBP2, eEFSec et le Sec-ARNt^{[Ser]Sec}. Des extraits réalisés à partir de types cellulaires dans lesquels l'expression des sélénoprotéines est élevée, seront testés pour leur capacité à assembler des complexes de traduction sur les ARNm transcrits *in vitro* et pour leur efficacité de recodage. Notre objectif est d'obtenir la structure tridimensionnelle du complexe de recodage associé à l'ARNm de SelM. Ces données structurales nous permettront d'avoir une meilleure connaissance du mécanisme de recodage des sélénoprotéines et peut-être de comprendre comment se fait la communication entre l'élément SECIS, le codon UGA^{Sec} et le ribosome durant ce processus.

Conclusion

Durant ma thèse, j'ai contribué à la mise en évidence et à la description d'un mécanisme particulier de maturation de la coiffe en 5' des ARNm de sélénoprotéines. En effet, nous avons montré que certains ARNm de sélénoprotéines portent une coiffe hyperméthylée m₃^{2,7}G à la manière des sn et snoRNP. L'étude du mécanisme de biogenèse de cette coiffe a révélé d'importantes similarités avec le mode d'assemblage de ces RNP non-codants : l'enzyme Tgs1 est responsable de la modification de la coiffe et serait recrutée par le complexe SMN et la protéine Nop58. Comme dans le cas des snRNP, le complexe SMN et le méthylosome semblent également participer à l'assemblage des mRNP de sélénoprotéines.

Tgs1, est requise pour l'expression *in vivo* des sélénoprotéines, probablement via son action directe sur l'hyperméthylation de la coiffe des ARNm, indiquant que la coiffe m₃G constituerait un élément important dans la synthèse des sélénoprotéines. Cependant, le rôle fonctionnel de la coiffe m₃G n'est toujours pas compris et elle pourrait être impliquée dans la stabilité et l'assemblage, ou l'initiation de la traduction et le mécanisme de recodage des mRNP de sélénoprotéines. Nous avons également révélé que le facteur d'initiation eIF4E n'est pas requis pour la traduction des ARNm de sélénoprotéines portant une coiffe m₃G, suggérant l'existence d'un mode d'initiation de la traduction différent. Mes résultats de thèse indiquent qu'il pourrait s'agir d'un processus d'initiation original qui nécessiterait une GTPase supplémentaire par rapport au mécanisme canonique. De plus, la région 3'UTR des ARNm de sélénoprotéines semble indispensable au processus de recrutement des ribosomes. L'ensemble de ces résultats va dans le sens de l'existence d'un mécanisme atypique d'initiation de la traduction des ARNm de sélénoprotéines.

Par ailleurs, j'ai également identifié la présence d'un élément de structure semblable à un SECIS présent dans la phase codante de certains ARNm de sélénoprotéines et qui pourrait contribuer au mécanisme de synthèse des sélénoprotéines. Ainsi, le processus de synthèse des sélénoprotéines semble bien plus sophistiqué qu'initialement anticipé. Cette complexité pourrait permettre une régulation fine de l'expression des sélénoprotéines, adaptée aux besoins de la cellule.

Durant ma thèse, j'ai développé des outils qui ont permis d'évaluer l'impact fonctionnel de facteurs sur la synthèse *in vivo* de sélénoprotéines. D'autres outils permettront d'identifier de nouveaux facteurs impliqués dans la traduction et le recodage des ARNm de sélénoprotéines *in vivo*. J'ai également mis au point une méthode de purification des complexes de traduction associés à un ARNm de sélénoprotéine dans le but de résoudre la structure tridimensionnelle du complexe de recodage par cryo-microscopie électronique. L'objectif est d'élucider l'ensemble des étapes de la traduction et du recodage des ARNm de sélénoprotéines.

Matériel et méthodes

I Matériel

I.1 Lignées cellulaires eucaryotes

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

Flp-In™ T-REx™ 293 : lignée cellulaire humaine dérivée des cellules HEK293 (Life technologies). Cette lignée permet de créer rapidement une lignée cellulaire stable exprimant de manière homogène une protéine d'intérêt à partir d'un vecteur d'expression Flp-In™. Ces cellules contiennent un site de recombinaison FRT dans un locus transcriptionnellement actif du génome. Elles contiennent également deux séquences plasmidiques intégrées stablement dans leur génome (pFRT/lacZeo et pcDNA™6/TR). La co-transfection des cellules Flp-In™ T-REx™ 293 avec un vecteur Flp-In™ (pcDNA™5/FRT/TO) exprimant le gène d'intérêt et un vecteur codant pour la recombinase Flp, pOG44, résulte en une intégration stable du vecteur d'expression au même locus pour toutes les cellules.

HeLa : lignée cellulaire humaine issue de cellules cancéreuses du col de l'utérus (ATCC)

I.2 Plasmides

Tableau M1 : liste des plasmides

Nom	Insert	Etiquette	Type/Utilisation
pHSBP2	hSBP2	6His	Vecteur d'expression procaryote de type pET32b (promoteur T7)
pHSΔ343	hSBP2 344-854	6His	pET32b
pHSΔ514	hSBP2 514-854	6His	pET32b
pHSBP2 1-344	hSBP2 1-344	6His	pET32b

pTHSBP2	hSBP2	Trx-6His	pET32b
pTHSΔ514	hSBP2 514-854	Trx-6His	pET32b
pLenti/V5-SBP2	SBP2	V5	Vecteur d'expression eucaryote de type pLenti/V5-D-TOPO pour transduction et transfection
pEGFP-SMN	SMN	GFP	Plasmide de type pEGFP-C2, pour expression eucaryotique (promoteur CMV)
pEGFP-PRMT5	PRMT5	GFP	pEGFP-C2
pSelM	5' et 3'UTR SelM	-	Plasmide de type pCDNA3.1, pour expression eucaryote (promoteur CMV)
pSelMCys	5' et 3'UTR SelMCys	-	pCDNA3.1
pSelMΔSECIS	5' et 3'UTR SelMΔSECIS	-	pCDNA3.1
pGPx1	5' et 3'UTR GPx1	-	pCDNA3.1
pGPx1Cys	5' et 3'UTR GPx1Cys	-	pCDNA3.1
pGPx1ΔSECIS	5' et 3'UTR GPx1ΔSECIS	-	pCDNA3.1
pcDNA TM 5/FRT/TO	-	-	Plasmide utilisé pour la création des lignées stables Flp-In TM T-REx TM 293. Promoteur CMV, opérateur TetO ₂
pHA-GPx1	5' et 3'UTR GPx1	HA	pcDNA TM 5/FRT/TO
pHA-GPx1Cys	5' et 3'UTR GPx1Cys	HA	pcDNA TM 5/FRT/TO
pHA-GPx1CysΔSECIS	5' et 3'UTR GPx1CysΔSECIS	HA	pcDNA TM 5/FRT/TO
pHA-GPx1-intron	5' et 3'UTR GPx1-intron	HA	pcDNA TM 5/FRT/TO
pOG44	FLP recombinase	-	Plasmide utilisé pour la création des lignées stables Flp-In TM T-REx TM 293.
pGADT7-SBP2	SBP2	AD Gal4	pGADT7, plasmide utilisé pour les tests d'interaction double-hybride chez <i>S. cerevisiae</i> . Genère une protéine en fusion avec le domaine d'activation de Gal4 (AD Gal4)
pGBKT7-SBP2	SBP2	BD Gal4	pGBKT7, plasmide utilisé pour les tests d'interaction double-hybride chez <i>S. cerevisiae</i> . Genère une protéine en fusion avec le domaine de liaison à l'ADN de Gal4 (BD Gal4)
pGADT7-PRMT5	PRMT5	AD Gal4	pGADT7
pGBKT7-PRMT5	PRMT5	BD Gal4	pGBKT7
pGADT7-MEP50	MEP50	AD Gal4	pGADT7
pGBKT7-MEP50	MEP50	BD Gal4	pGBKT7
pGADT7-pICln	pICln	AD Gal4	pGADT7
pGBKT7-pICln	pICln	BD Gal4	pGBKT7

pACT2-Nufip	Nufip	AD Gal4	pACT2-AD, plasmide utilisé pour les tests d'interaction double-hybride chez <i>S. cerevisiae</i> . Génère une protéine en fusion avec le domaine d'activation de Gal4 (AD Gal4)
pAS2-Nufip	Nufip	BD Gal4	pAS2 BD, plasmide utilisé pour les tests d'interaction double-hybride chez <i>S. cerevisiae</i> . Génère une protéine en fusion avec le domaine de liaison à l'ADN de Gal4
pACT2-Alix	Alix	AD Gal4	pACT2-AD
pAS2-Alix	Alix	BD Gal4	pAS2 BD
pACT2-Gemin-2	Gemin-2	AD Gal4	pACT2-AD
pAS2-Gemin-2	Gemin-2	BD Gal4	pAS2 BD
pACT2-Gemin-3	Gemin-3	AD Gal4	pACT2-AD
pAS2-Gemin-3	Gemin-3	BD Gal4	pAS2 BD
pACT2-Gemin-4	Gemin-4	AD Gal4	pACT2-AD
pAS2-Gemin-4	Gemin-4	BD Gal4	pAS2 BD
pACT2-Gemin-5	Gemin-5	AD Gal4	pACT2-AD
pAS2-Gemin-5	Gemin-5	BD Gal4	pAS2 BD
pACT2-Gemin-6	Gemin-6	AD Gal4	pACT2-AD
pAS2-Gemin-6	Gemin-6	BD Gal4	pAS2 BD
pACT2-Gemin-7	Gemin-7	AD Gal4	pACT2-AD
pAS2-Gemin-7	Gemin-7	BD Gal4	pAS2 BD
pACT2-Gemin-8	Gemin-8	AD Gal4	pACT2-AD
pAS2-Gemin-8	Gemin-8	BD Gal4	pAS2 BD

I.3 Anticorps

Tableau M2 : liste des anticorps

Nom de l'anticorps	Poly/monoclonal, source	fournisseur
Anticorps primaires :		
Anti-SMN (2B1)	Monoclonal, souris	Santa Cruz
Anti-PRMT5	Polyclonal, chèvre	Santa Cruz
Anti-Mep50	Monoclonal, souris	Santa Cruz
Anti-Gemin2 (2E17)	Monoclonal, souris	Fournis par Severine Massenet
Anti-Sm (Y12)	Monoclonal, souris	Abcam

Anti-SBP2	Polyclonal, lapin	NeoMPS
Anti-GFP (B-2)	Monoclonal, souris	Santa Cruz
Anti-His	Monoclonal, souris	Santa Cruz
Anti-Tgs1	Monoclonal, souris	Fournis par Rémy Bordonné
Anti-rpS21	Polyclonal, chèvre	Santa Cruz
Anti-AspRS	Polyclonal, lapin	Fournis par Magalie Frugier
Anti-Nop58 (C-20)	Polyclonal, chèvre	Santa Cruz
Anti-pICln (Q-10)	Polyclonal, chèvre	Santa Cruz
Anti-HA (F-7)	Monoclonal, souris	Santa Cruz
Anti-m ₃ G cap (R1131)	Polyclonal, lapin	Synaptic Systems
Anticorps secondaires :		
Anti-IgG de souris	Monoclonal, chèvre	Bio-Rad
Anti-IgG de chèvre	Monoclonal, âne	Bethyl
Anti-IgG de lapin	Monoclonal, chèvre	Bio-Rad

I.4 Oligonucléotides

Oligonucléotides utilisés pour les expériences de RT-qPCR

Les oligonucléotides utilisés lors de la qPCR ciblent des jonctions exon-exon permettant d'amplifier uniquement les ADNc et non l'ADN génomique.

Tableau M3 : liste des Oligonucléotides (RT-qPCR)

Cible de l'oligonucléotide	Séquence de l'oligonucléotide (5'-3')
SelT	Fw : TCCAGATTTGTGTTTCCTGAGG Rev : CTGGGTACCGCTGGCTAATA
Sel15	Fw : CGGACAGTTCAACCTGCTTC Rev : AATACAGGGTCTGAACCACG
SelM	Fw : TCCCGATGAGCCTCCTGTTG Rev : ATGGAATGTCCTGCGTGACG
SelW	Fw : ACGTGGACACAGAAAGCAAG Rev : ACAGCAGCCACGAGAACATC
SelN	Fw : AGCTTCATCAGCACCTGGTC Rev : CGGAGGTGATGTCCAAGAAG
SelR	Fw : CAGGTTTTCCAGAATCACTTTG Rev : GGCCATGGAGACGAGTGT
SelK	Fw : ATCTGATTCCAGATATGATGAT Rev : TGATTGATTCTACCCATTCTTC
SelO	Fw : GAGGAGTTTGACGCCGAGTTC Rev : GCTCAGCAAGTAGAAGGTGTTTGTG
GPx3	Fw : GGAGTACATCCCCTTCAAGC Rev : CGAATGGTGCAAGCTCTTCC
GPx4	Fw : CGGGCTACAACGTCAAATTCG Rev : GGGGCAGGTCCTTCTCTATCA
GPx1	Fw : TGCAACCAGTTTGGGCATCA Rev : ACCGTTCACCTCGCACTTC

TrxR1	Fw : CCTTATCATCATTGGAGGTGG Rev : AAGAGGGGTGGGAGTGACAAA
SMN	Fw : CTGATGCTTTGGGAAGTAT Rev : GCCAGCATTTCTCCTTAA
Nop58	Fw : CCATGAACCGCTGGCAGTAG Rev : CCAATCCAAGACACATAGCTG
TGS1	Fw : GGCTATTACATCAGAGACAGTG Rev : GAATCAAGTTCACCTTCATCCA
snARN U2	Fw : TTCTCGGCCTTTTGGCTAAG Rev : CTCCCTGCTCCAAAAATCCA
snoARN U3	Fw : TTCTCTGAACGTGTAGAGCACC Rev : GATCATCAATGGCTGACGGCAG
SBP2	Fw : GGTGAGGACAAGTCTGAAGAGCC Rev : CTTTGAGTAGGTCGGTAACACAAGC
eIF4E	Fw : GGCCCACTCTGTAATAGTTC Rev : CCATCAGCAAGAGTACAGC
β -actin	Fw : CTTTCTACAATGAGCTGCGT Rev : CTCCTTAATGTCACGCACGAT
LDHA	Fw : TGGCAGCCTTTTCCTTAGAACA Rev : ACGGCTTTCTCCCTCTTGCTGA
HPRT	Fw : TGACACTGGCAAAACAATGCA Rev : GGTCCCTTTTACCAGCAAGCT
Nucléoline	Fw: GGACGATGAGGATGACGAAG Rev: CATCTTCATCCTCAGCCACG
Riok1	Fw: GTCAGGAGTTCAGAAGGTC Rev: GCTCTTCAGAGTCTGTGTCAG
pICln	Fw: CCACACTGGAGAGATTAGAAGG Rev: CCACCTCCATCCCATCTTCAT
Mep50	Fw: CTGGACTCAAGCCTTTCTG Rev: GGTAAGCAGGGAGTGATTGAGC
eEFSec	Fw: AGCTGAAGCACAAAGCATGG Rev: CACGAAGAGCTGGATGTTG
CBP20	Fw: CCATGCGGTACATAAATGGGAC Rev: CGATACTCATCCCGAACCTG
CBP80	Fw: CCACTCCTTCAGTGCTCTTG Rev: GATGGTTCCTCCAGACCTCA
NSEP1	Fw: TGATGGAGGGTGCTGACAAC Rev: CCTGCGGAATCGTGGTCTAT
DDX3X	Fw: GAAGTGCCGTCTTGTTAG Rev: CTTTGTCGGTAGTCTCTGGC
Gemin4	Fw: CTAGGACCCTTGAACATCTGTG Rev: CTTTGTTAATTCTGCCAGTGCC
Gemin5	FW: GGAGCTGCTCAAGTCAAACC Rev: CTCTTTCTAGGACGGTTCCC
dN9 (pour la RT)	NNNNNNNNN

Fw : Oligonucléotide sens (Forward)

Rev : Oligonucléotide anti-sens (Reverse)

Séquence des siARN

Une Library de siARN «à façon», On-TARGET Plus SMART pool Custom Library (Thermo scientific), dirigés contre 18 cibles différentes a été utilisée. Pour chaque gène, un mélange de 4 siARN ciblant des séquences différentes sur l'ARN cible, sont utilisés simultanément pour une meilleure inhibition et spécificité. Les siARN contrôles utilisés sont dirigés contre la luciférase firefly (FFL).

Tableau M4 : liste des siARN

Plaque personnalisée			
Cible	GENE ID	Accession du gène	Séquence (5' -> 3')
SECISBP2	79048	NM_024077	AAACAGUGAAGCAUCGAAA
			GCUCAAUGUGGCAUGGUUA
			GUUAUCAAAGGGUGAAAUA
			GCUUCAACCUCUCAAAUGA
eEFSEC	60678	NM_021937	GUUCCGAGGUGCACCGAUU
			GCCUGACUUUCAAGCGUUA
			GCUCCAGGCUAGAUGC GGA
			GAACUGGGCAUCAUCGACA
NSEP1	4904	NM_004559	CUGAGUAAAUGCCGGCUUA
			CGACGCAGACGCCCAGAAA
			GUAAGGAACGGAUAUGGUU
			GCGGAGGCAGCAAUGUUA
NCL	4691	NM_005381	GCAAAGAAGGUGGUCGUUU
			GAUAGUUACUGACCGGGAA
			CAAAUCUGCUCCUGAAUUA
			GAAAGAAGACGAAGUUUGA
CBP80	4686	NM_002486	CUGAGAAGCUGACAAUUUA
			CAACCUGCAUUUACAAGUA
			GAAAGGAGUUGUACGAAAA
			CAUCAGCGUAUAUUAGAUU
CBP20	22916	NM_007362	GAAUAUUACUCACGCGCAG
			CAACUGAAGAACAAUUCUA
			GAAAAGCUGUACGUUAUAU
			GCCAUGCGGUACAUAAAUG
EIF4E	1977	NM_001968	GACGAUGGCUAAUUACAUAU
			CAUAUCCAGUUGUCUAGUA
			GUGAUAAAGAUAGCAAUAUG
			GAGACGAAGUGACCUCGAU
PRMT5	10419	NM_006109	UGGCACAACUUCGGGACUU
			CAACAGAGAUCCUAUGAUU
			CGAAAUAGCUGACACACUA
			CCAAGUGACCGUAGUCUCA
pICiN	1207	NM_001293	GGACCGAAGUGACUGUCUA
			GUGAGCAGCCAGUAUAAUA

			GUUAGAUGGCUCUGGAUUA
			AACAGUUGCUGGACAGUUU
MEP50	79084	NM_024102	CUGAGUAGAUUGGAUUUAA
			GUACUGAGUUCAUACCGAG
			GCAAAGACAUCUGCAUCAA
			GAUGAGAAUGAGACACUUA
RIOK1	83732	NM_153005	GCGCCAACGUCAAUGAUUU
			GAACAUGGAUGCUUAUCUC
			GGAGGCGUGUAUAUCAUUG
			CCAAUAAUGCUAAGAAGUC
GEMIN5	25929	NM_015465	CAGGGAAGCUAUUGCGAUU
			CUGUAAAGCUGGAAUCGAU
			AGAAAUACGGCAACACGAA
			CAGAAGAAGGUGAGCGAAU
GEMIN4	50628	NM_015721	CGUUGACACUUCUGCCGAA
			UACCUGUGCGGGAGCGAAA
			GGUCAUAGUGCACCCGGAA
			AUGCAGAACUAGAGCGCUU
DDX20	11218	NM_007204	CAAGAAAUGAUGCAUAGUA
			GAAUUAAGCAACUCAUAGA
			GGUUGUAAAUCUGGAUGUA
			CCGCUUCUCAUUCAUUUA
SMN1	6606	NM_022874	GCUCACAUUCCUAAAUAUA
			GCUGAUGCUUUGGGAAGUA
			GAUGAUACAGCACUGAUAA
			UAAGCAUGCUCUAAAGAAU
TGS1	96764	NM_024831	GCGGGAGGAUUGUAAGAUU
			UAUCGAUCCUGUUAAGAUU
			GCACGACGCUUCCACAAGU
			UAGGAUAACUGCACAUAAG
NOP58	51602	NM_015934	CCUCAAAUCUAGACGGGAU
			CAUUAUGCGAUGUAGAGAA
			ACGCAAAUAGAACAGGUA
			AGAGAGAGAACGAGGAUUA
DDX3X	1654	NM_001193417	CAGAUUUAGUGGAGGGUUU
			GCAACACUGGGAUUAAUUU
			GAUGCUGGCUCGUGAUUUC
			GGUAUUAGCACCAACGAGA
Thermo Scientific™ Anti-Luc siARN 1			
siARN à façon Eurogentec :			
FFL (Ctrl luciférase)			CUUACGCUGAGUACUUCGA
Tgs1	96764	NM_024831	GUAUCACAGCUUUGCGAGG
Thermo Scientific™ Anti-Luc siARN 1			

Oligonucléotides utilisés pour les expériences de toeprint

Tableau M5 : liste des Oligonucléotides (toeprint)

Cible de l'oligonucléotide	Séquence de l'oligonucléotide (5'-3')
SelM-AUG	GTCCGGCCGGTAGGCAGTGG

SelM-UGA ^{Sec}	CATCACCAGGTTGTGATAG
SelM-SECIS	GAAGGAAGAAAGTGGGGTTGG
GPx1-AUG	CAGGCTCACAGGCTCCCCGCC
GPx1- UGA ^{Sec}	CGCCGCTGCAGCTCGTTCATC
GPx1-SECIS	CTTTATTAGTGGGGAAACTCGCC

Oligonucléotides utilisés pour réaliser les matrices PCR, utilisées pour transcription

Tableau M6 : liste des Oligonucléotides (transcription)

Cible de l'oligonucléotide	Séquence de l'oligonucléotide (5'-3')
SelM entier	T7Fw: TAATACGACTCACTATAGGGT GGCGCAGCGACTCGGAGGTT Rev: TTAAGGCAGGAGGGAATTTAGTGGA
GPx1 entier	T7Fw: TAATACGACTCACTATAG CAGTTAAAAGGAGGCGCC TGCTG Rev : TTCTGCTGACACCCGGCACTTTAT
SelM entier (pour Complexes CryoEM)	T7Fw: TAATACGACTCACTATAGGGT GGCGCAGCGACTCGG AGGTT Rev : GTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTAAGGCA GGAGGGAATTTAGTGGA
SelM UGA ^{Sec} (pour Gelshift)	T7Fw: TAATACGACTCACTATAGGG CGCTTGTGGCCCCAG CCACAG Rev : CCGCAGGTCTCTACCCGGGC
SelM SECIS (pour Gelshift)	T7Fw: TAATACGACTCACTATAGGGG GAGACAGAATGAAG CGCTC Rev: GGGGACGGGCATCGGCTCTC

La séquence du promoteur T7 est indiquée en vert. La séquence permettant la ligation de l'oligonucléotide A-biot à l'ARNm de SelM est indiquée en gris.

Oligonucléotides utilisés pour purifier les complexes de traduction associés aux ARNm de SelM

Tableau M7 : liste des Oligonucléotides (purification des complexes, ARNm de SelM)

Nom de l'oligonucléotide	Séquence de l'oligonucléotide (5'-3')
oligonucléotide A-biot	CTCCTCCTCCTCAAAAAA-Biotine
oligonucléotide Splint 1	GAGGAGGAGGAGGTGTTGTTGTTGTT
oligonucléotide Splint 2	GAGGAGGAGGAGNGTGTTGTTGTTGTT
oligonucléotide anti-Splint 1	AACAACAACAACCTCCTCCTCCTC
oligonucléotide anti-Splint 2	AACAACAACAACNCTCCTCCTCCTC

Oligonucléotides utilisés pour la purification des complexes associés *in vivo* aux ARNm HA-GPx1

Tableau M8 : Liste des Oligonucléotides (purification des complexes, ARNm de HA-GPx1 *in vivo*)

Nom de l'oligonucléotide	Séquence de l'oligonucléotide (5'-3')
oligonucléotide ARN HA-Biot	mCmAmTmAmGmGmGmAmTmAmGmCmCmCmGmCmAmTmAmG/TEG//3Bio/
oligonucléotide 5' HA	CAGGAACATCGTATGGG
oligonucléotide 3' HA	CACTGAGCAGCGTAATCTGG

m : 2'O-méthylation

3Bio : 3 groupements biotine

TEG : tétraéthylène glycol, bras permettant d'éloigner le groupement biotine de l'Oligonucléotide

II Méthodes

II.1 Surproduction et purification de protéines (hSBP2-His et dérivés) dans *E. coli*

Les bactéries BL21(DE3)RIL sont transformées par 100 ng des plasmides pHisSBP2, pTHSBP2, pHΔ343, pHΔ514, ou pHisSBP2 1-344. Les clones sont sélectionnés sur milieu 2YT solide contenant 100 $\mu\text{g}/\mu\text{L}$ d'ampicilline et 30 $\mu\text{g}/\mu\text{L}$ de chloramphénicol. Une culture de 200 ml (2YT double antibiotique) est incubée à 37°C. Lorsque la densité optique à 600nm ($\text{DO}_{600\text{nm}}$) atteint 0,6, l'expression de la protéine est induite par ajout de 2 ml d'IPTG 100 mM. Les bactéries sont incubées à 18°C pendant une nuit puis centrifugées 10 min à 3500 rpm et lavées dans 50 ml d'eau milliQ. Le culot est repris dans 5 ml de tampon de lyse (50 mM phosphate de sodium pH 8 ; 300 mM NaCl ; 10 mM imidazole) auxquels sont ajoutés 5 μL de cocktail d'inhibiteurs de protéases 100X (Sigma). Les bactéries sont lysées par sonication (4 fois 15 sec) et maintenues dans la glace. Après clarification, le lysat est déposé sur une colonne contenant 1 ml d'agarose Ni-NTA, préalablement lavée avec 10 ml de tampon de lyse. Le tout est placé sous agitation à 4°C pendant 1 h pour une purification en « batch ». Après sédimentation de la résine, l'effluent est collecté et la colonne est lavée 2 fois avec 4 ml de tampon de lavage (50 mM phosphate de sodium pH 8 ; 300 mM NaCl ; 20 mM imidazole). Les protéines étiquetées His sont éluées

avec 4 ml de tampon d'élution (50 mM phosphate de sodium pH 8 ; 300 mM NaCl ; 250 mM imidazole) et récupérées sous forme de fractions de 500 μ L.

II.2 Culture cellulaire : lignées HEK 293FT et Flp-In™ T-REx™ 293

Multiplication et maintien des cellules

Les cellules HEK 293FT sont cultivées dans du milieu de culture DMEM (Dulbecco's Modified Eagle's Medium GlutaMAX, GIBCO), additionné de 10% de sérum de veau fœtal (FCS), 1% de généticine (GIBCO, 50 mg/ml), 1% de pénicilline/ streptavidine (GIBCO, 10⁴ U/ml pénicilline, 10⁵ μ g/ml streptavidine) en présence ou non de 10 nM Na₂SeO₃ (sélénite de sodium). Des cellules stock maintenues dans l'azote liquide (3.10⁶ cellules) sont décongelées etensemencées stérilement dans 8 ml de milieu de culture 30% FCS (boîte de culture de 25 cm²). Après environ 24 h d'incubation dans l'étuve à 37°C, sous 5% CO₂ et hydratation contrôlée, les cellules arrivées à confluence sont transférées dans des boîtes de 75 cm². Pour cela, le milieu est d'abord éliminé et les cellules sont délicatement lavées puis dissociées dans 1 ml de trypsine-EDTA 0,25% (GIBCO) ; la réaction est arrêtée par ajout de 9 ml de milieu de culture. Les cellules sont diluées 10 fois et réparties dans des nouvelles boîtes de culture en vue de leur utilisation. Elles sont maintenues en culture jusqu'à un maximum de 20 passages successifs.

Transfection de cellules et inactivation de genes par ARN interférence

Les cellules HEK 293FT et Flp-In™ T-REx™ 29 ont été transfectées par divers plasmides ou inactivées par siARN à l'aide de lipofectamine 2000 (Invitrogen) comme décrit par le fournisseur. 48h après transfection, les cellules sont été lavées avec du PBS (137 mM NaCl ; 26,8 mM KCl ; 40,2 mM Na₂HPO₄ ; 17,6 mM KH₂PO₄ ; pH7,4) et les extraits peuvent être récoltés par les différentes méthodes décrites ci-dessous.

Préparation d'extraits cellulaires

En vue de leur récupération, les cellules sont lavées puis grattées dans 1ml de PBS froid. Après 10 min de centrifugation à 1000 g, le culot de cellules est repris dans un volume de tampon « polysome » (10 mM HEPES-NaOH pH 7 ; 100 mM KCl ; 5 mM MgCl₂ ; 0,5% NP40 ; 1 mM DTT ; 100 U/ml RNasin (Promega) ; 400 μ M VRC

(Vanadyl Ribonucleotide Complex, Sigma) ; 1X cocktail d'inhibiteurs de protéases (Sigma) et incubé 5 min dans la glace. Les extraits sont congelés dans l'azote liquide puis conservés à -80°C. Les extraits ainsi récupérés pourront directement être utilisés dans des expériences d'immunoprécipitation de complexes ARN/protéines.

Extraction d'ARN à partir d'extraits cellulaires totaux

Les extraits cellulaires sont additionnés de 10 volumes de TriReagent (MRC) et de 2 volumes de chloroforme, vortexés 3 min, incubés 10 min à température ambiante puis centrifugés 15 min à 4°C à 13000 rpm. La phase aqueuse est récupérée et précipitée par 1 volume d'isopropanol. Après centrifugation à 13000 rpm, le culot est lavé à l'éthanol 70% puis repris dans 25 µL d'eau milliQ.

Génération des lignées stables HA-GPx1 à partir des cellules Flp-In™ T-REx™ 293 et induction de l'expression des protéines

Le système Flp-In™ T-REx™ 293 (Invitrogen) a été utilisé en suivant les instructions du fournisseur pour créer des lignées stables exprimant de manière inductible les protéines HA-GPx1, HA-GPx1Cys, HA-GPx1CysΔSECIS et HA-GPx1-intron. Les séquences d'intérêt ont été clonées dans le plasmide pcDNA™ 5/FRT/TO et co-transfectées dans les cellules T-REx™ 293 avec le plasmide pOG44, qui code pour la recombinaise FLP. L'expression de la recombinaise permet d'intégrer, par recombinaison homologue, le gène d'intérêt dans le site FRT présent dans un locus transcriptionnellement actif du génome des cellules, en aval du codon initiateur du gène LacZ en fusion au gène de résistance de la zéocine. Ainsi, si l'intégration a lieu dans le bon locus, les cellules deviennent sensibles à la zéocine, résistantes à l'hygromycine et n'expriment plus la β-galactosidase. Les clones ayant intégré stablement ces différentes constructions ont été sélectionnés à l'aide d'un milieu de culture sélectif contenant 250 µg/mL d'hygromycine (InvivoGen) et 15 µg/mL de blasticidine (InvivoGen). Les colonies résistantes ont été mises en culture puis amplifiées et l'intégration a été testée par analyse de l'expression de la β-galactosidase (test bleu/blanc). Les gènes d'intérêts sont sous contrôle du promoteur hybride CMV/TetO₂, permettant leur expression de manière inductible. L'expression de la protéine d'intérêt est induite par ajout de 0,5 à 1 µg/mL de doxycycline dans le

milieu de culture et se fait en présence de 3 à 10 nM de sélénite de sodium, conditions optimales pour la synthèse des sélénoprotéines.

II.3 Transfection des lignées cellulaires par la librairie de siARN et analyse de l'expression de HA-GPx1 par western blot

Les lignées cellulaires décrites précédemment ont été utilisées pour analyser l'impact de l'inhibition de différents facteurs par siARN sur l'expression *de novo* des protéines HA-GPx1 et HA-GPx1Cys. 12 h avant transfection, 10^5 cellules issues des lignées HA-GPx1 et HA-GPx1Cys sont mises en cultures dans 200 μ L de DMEM/10% FCS dans des boîtes de 48 puits. Lorsque les cellules ont atteint environ 70% de confluence, l'inactivation de chacune cible est réalisée par transfection de 12 pmol d'un mélange de 4 siARN par puits (On-TARGET Plus SMART pool Custom Library (Thermo scientific), voir Tableau M4) à l'aide de lipofectamine 2000 (Invitrogen) comme décrit par le fournisseur. Pour chaque cible, 4 répliquas biologiques sont réalisés. L'induction de l'expression des protéines HA-GPx1 et HA-GPx1-Cys est réalisée 48h après transfection. Pour cela, le milieu est remplacé par 200 μ L de DMEM/10% FCS contenant 0,5 μ g/mL de Doxycycline et 3 nM de Na₂SeO₃. 7 h après l'induction, les cellules sont lavées avec 250 μ L de PBS et collectées dans 50 μ L de tampon « polysome », congelées dans l'azote liquide, puis conservées à -80°C. 6 μ L d'extrait de cellules clarifiés sont déposés sur gel SDS-PAGE 8% et analysés par western blot à l'aide d'anticorps anti-HA et anti-AspRS.

Analyse de la traduction de novo de protéines par chasse à la ³⁵S méthionine (pulse-chase)

Les cellules des lignées HA-GPx1, HA-GPx1Cys et HA-GPx1Cys Δ SECIS ont été préparées, transfectées et induites comme décrit dans le paragraphe précédent. Pour chaque cible, 2 répliquats biologiques sont réalisés. 3 h après induction, le milieu est éliminé et les cellules sont incubées pendant 1 h dans 200 μ L de milieu Pulse (DMEM Glutamax Gibco® dépourvu des acides aminés méthionine, leucine, cystéine et alanine (Life Technologies) complété par 10% de FCS, 0,1 mCi/mL de ³⁵S méthionine, 3,97 mM d'alanine, 3,97 mM de glutamine, 0,45 mM de cystéine, 0,5 μ g/mL de Doxycycline et 3 nM de Na₂SeO₃). Les cellules sont lavées avec 250 μ L de PBS, collectées dans 50 μ L de tampon « polysome », congelées dans l'azote liquide

puis conservées à -80°C. La totalité des extraits clarifiés, ajoutés de 100 μ L de tampon de lyse (150 mM NaCl, 1% Triton® X-100, 50 mM Tris-HCl pH 8) et les protéines HA-GPx1 sont purifiées à l'aide de 25 μ L de billes magnétiques μ MACs couplées aux anticorps anti-HA (μ MACS HA isolation Kit, Miltenyi Biotec) en suivant les instructions du fournisseur. L'élution est réalisée dans un volume de 30 μ L dont 12 μ L sont déposés sur gel SDS-PAGE 15%. Après migration le gel est séché, exposé 48h et révélé par autoradiographie.

Durant la purification des protéines, le surnageant est récupéré et les ARN en sont extraits et analysés par RT-qPCR pour déterminer l'efficacité de l'inhibition des cibles par les siARN.

II.4 Immunoprécipitation de particules ribonucléoprotéiques (RNP)

L'immunoprécipitation de RNP permet de récupérer les protéines et ARN qui sont associés à une protéine d'intérêt à l'aide d'un anticorps spécifique.

Préparation des billes

Pour chaque immunoprécipitation, 50 μ L de billes Protéine A-sépharose 50% (Sigma), préalablement gonflées et lavées 3 fois dans du PBS, sont saturées par 20 μ g de BSA et 20 μ g d'ARNt total de levure et incubées 30 min à 4°C en rotation. 4 μ g d'anticorps (ou non dans le cas du contrôle négatif) sont ajoutés sur les billes en présence de 500 μ L de PBS et incubés une nuit à 4°C en rotation. Les billes sont lavées 3X dans 500 μ L de PBS puis 3X dans 500 μ L de tampon d'immunoprécipitation NT2 (50 mM Tris-HCl pH 7,5 ; 150 mM NaCl ; 1 mM $MgCl_2$; 0,05% NP40 ; 1 mM DTT ; 400 μ M VRC ; 200 U/mL RNasin).

Clarification des extraits et immunoprécipitation

Afin de clarifier l'extrait cellulaire avant immunoprécipitation et d'éliminer les ARN et protéines interagissant de façon aspécifique avec les billes, 300 μ L d'extrait cellulaire sont mis en présence de 150 μ L de billes sans anticorps dans 3 mL de tampon NT2. Les billes ont été préalablement saturées et lavées comme décrit ci-dessus. Après 30 min d'incubation à 4°C, le surnageant clarifié est récupéré par centrifugation et divisé en 3 fractions égales. L'une sert de référence (Input), les deux autres sont déposées sur les billes avec et sans anticorps (IP et Ctrl). Après 2h

d'incubation à 4°C en rotation, les billes sont récupérées par centrifugation et lavées avec 5X 1 mL de NT2 avant d'être transférées, pour un dernier lavage, dans un nouveau tube Eppendorf afin d'éliminer les rétentions aspécifiques d'ARN sur les parois. 1/4 des billes en solution sont prélevées, centrifugées et stockées à -20°C en vue d'une analyse par Western blot. Le reste des billes sont reprises dans 40 µL de tampon protéinase K (50 mM Tris-HCl pH 7,5 ; 10 mM EDTA ; 1% SDS ; 750 µg/mL Protéinase K) et incubées 20 min à 37°C afin d'hydrolyser les protéines. Les ARN sont ensuite extraits.

II.5 Détection des ARN par RT-PCR et RT-PCR quantitative

Les ARN récupérés après immunoprécipitation ou directement extraits des cellules, sont traités à la DNase I, soumis à une transcription inverse et analysés par qPCR ou PCR.

Synthèse des ADNc

- Traitement DNase I

Afin d'éliminer l'ADN génomique résiduel qui interférerait avec la synthèse des cDNA. 2µg d'ARN sont incubés en présence de 2,5 U de DNase I RNase free (Fermentas) et de tampon DNase (40 mM Tris-HCl pH 7,9 ; 10 mM NaCl ; 6 mM MgCl₂ ; 1 mM CaCl₂) dans un volume final de 20µL pendant 15 min à 25°C. La réaction est arrêtée par ajout de 8 mM final d'EDTA final et incubation 10 min à 70°C.

- Réverse transcription

Cette étape a pour objectif d'obtenir les ADN simple-brins complémentaires (ADNc) de l'ensemble des ARN présents dans l'échantillon. Pour cela, 10 µL des échantillons traités à la DNase I sont dénaturés 5 min à 70°C en présence de 100 pmol d'oligonucléotides dégénérés dN9 (Sigma) et 10 nmol de dNTP, puis mis dans la glace. La reverse transcription s'effectue avec 5 U de reverse transcriptase AMV (MP Biomedicals) dans le tampon AMV (50 mM Tris-HCl pH 8,3 ; 6 mM MgCl₂ ; 60 mM NaCl ; 100 mM DTT) et en présence de 40 U de RNasin. L'hybridation est réalisée à 25°C pendant 15min et l'élongation 1h à 42°C. Les ADNc sont stockés à -20°C.

Quantification des ADNc par qPCR

La qPCR permet la quantification précise des ADNc présents dans l'échantillon en mesurant en temps réel l'amplification des ADN. Nous avons utilisé le kit Master Mix Maxima SYBR Green qPCR de Fermentas, contenant du SYBR Green et une Hot Start Taq ADN polymérase. L'intensité du signal émis, proportionnel à la quantité d'ADN, est mesuré en temps réel par le thermocycleur (BioRad CFX96/384 ou Stratagene MX3005P). 3 μ L des échantillons d'ADNc dilués au 1/20^{ème} sont additionnés de 5 μ L de réactif Master mix 2X et de 0,25 μ M final de chaque oligonucléotide sens et antisens dans un volume final de 10 μ L. Les échantillons sont dénaturés pendant 10 min à 95°C puis soumis à 55 cycles de qPCR : 15 sec à 95°C (dénaturation), 30 sec à 58°C (hybridation) et 30 sec à 72°C (élongation). Une courbe de fusion des produits de PCR finaux est réalisée dans une dernière étape de qPCR : 1 min à 95°C, 30 sec à 55°C et 30 sec à 95°C. Les données sont traitées à l'aide du logiciel CFX Manager ou REST 2009 (Relative Expression Software Tool).

Amplification des ADNc par réaction de polymérisation en chaîne (PCR)

Dans des conditions de PCR semi-quantitatives, la réaction de PCR est réalisée à l'aide par une Taq ADN polymérase thermostable (DyNAzyme EXT, Finnzymes). 1 μ L d'ADNc dilué est additionné de 49 μ L de tampon PCR contenant 50 mM Tris-HCl pH 9 ; 1.5 mM MgCl₂ ; 15 mM (NH₄)₂SO₄ ; 0.1 % Triton X100 ; 0,5 mM de dNTP ; 2 μ M de chaque oligonucléotide et 1 U de DyNAzyme EXT. Les échantillons sont dénaturés pendant 2 min à 94°C puis soumis à 35 cycles de PCR : 30 sec à 94°C (dénaturation), 30 sec à 58°C (hybridation) et 1 min à 72°C (élongation) ; puis 5 min à 72°C pour une dernière phase d'élongation. 10 μ L de produit PCR additionnés de 2 μ L de bleu type III sont analysés sur gel d'agarose 1%.

II.6 Expériences d'immunoprécipitation d'ARN à l'aide des anticorps anti-m₃G

Les ARN dont la coiffe est hyperméthylée ont été spécifiquement immunoprécipités en utilisant les anticorps anti-m₃G (R1131, Synaptic Systems). 10 μ L d'anticorps ont été couplés à 50 μ L de billes protéine A-sépharose saturées avec 20 μ g de BSA et 20 μ g de tRNA total de levure dans du tampon NT2 (50 mM Tris-HCl pH 7,5 ; 150 mM NaCl ; 1 mM MgCl₂ ; 0,05% NP40 ; 1 mM DTT ; 400 μ M VRC ; 200 U/mL RNasin)

pendant 18h à 4°C. Après 6 lavages dans du PBS, les anticorps immobilisés ont été incubés avec 200 μ g d'ARN totaux qui ont préalablement été extraits de cellules HEK293FT (Tri-reagent, Euromedex) et clarifiés (par incubation avec des billes sans anticorps), dans un volume total de 1mL de tampon NT2, pendant 2h à 4°C. Les billes ont été lavées 6 fois dans du tampon NT2 et les ARN liés aux billes ont été récoltés par extraction phénol/chloroforme et précipitation à l'éthanol. Après traitement à la DNase I (Roche), les ARN ont été reverse-transcrits en utilisant la RT-AMV (Q-Biogen) et les ADNc ont été amplifiés et analysés par qPCR.

II.7 Expériences de co-immunoprécipitation

Des cellules HEK293FT ont été cultivées dans du milieu additionné de 10% de FCS et de 1% d'antibiotiques streptomycine/pénicilline. Les cellules ont été co-transfectées à l'aide de lipofectamine 2000 (Invitrogen) par les plasmides codant pour les protéines d'intérêts. 24 heures après transfection, les cellules ont été lavées dans du PBS puis récoltées et lysées par congélation-décongélation dans du tampon Tris-HCl 100 mM pH 8 ou additionné d'un mélange d'inhibiteurs de protéases (Sigma). Après centrifugation, le surnageant a été récupéré. 100 à 300 μ L d'extraits de cellules co-transfectées ont été incubés en présence de 25 μ L de billes magnétiques μ MACS (Miltenyi Biotec) ou de 50 μ L de billes protéine A-sépharose, couplées aux anticorps correspondants 2 h à 4°C dans du tampon NT2. Les expériences de co-immunoprécipitation ont été effectuées d'après le protocole du kit (μ MACS Epitope Tag Protein Isolation kits). Les billes protéine A-sépharose ont été lavées 6 fois dans du tampon NT2. Les protéines co-immunoprécipitées ont été déposées sur gel SDS-PAGE 8% puis analysées par Western blot, comme décrit ci-dessus.

II.8 Analyse de protéines sur gel de polyacrylamide SDS-PAGE

Les protéines sont fractionnées par électrophorèse sur un gel de polyacrylamide 8% en condition dénaturante (SDS-PAGE). Gel de séparation : 7,8% acrylamide ; 0,2% bisacrylamide ; 375 mM Tris-HCl pH 8,9 ; 0,1% SDS. Gel de concentration : 4,8% acrylamide ; 0,13% bisacrylamide ; 125 mM Tris-HCl pH 6,7 ; 0,1% SDS. Les échantillons sont dénaturés pendant 3 min à environ 100 °C dans le tampon de charge (62,5 mM Tris-HCl pH 8 ; 2% SDS ; 5% β -mercaptoéthanol ; 10% glycérol ; 0,005%

bleu de bromophénol). La migration est effectuée pendant 60 min à 120 V dans le tampon de migration Tris-glycine pH 8 (25 mM Tris ; 200 mM glycine ; 0,1% SDS). Les protéines sont visualisées par coloration du gel au bleu de Coomassie (50% éthanol ; 10% acide acétique ; 0,03% bleu de Coomassie) pendant 15 min, suivie de trois étapes de décoloration (25% d'éthanol ; 10% d'acide acétique) d'environ 30 min.

II.9 Immuno-détection de protéines par Western blot

Après fractionnement sur un gel SDS-PAGE, les protéines sont transférées sur une membrane Immobilion-P (Millipore). Le transfert s'effectue à 200 mA, sur une membrane préalablement rendue hydrophile dans du méthanol, dans le tampon de transfert pH 8 (25 mM Tris ; 200 mM glycine ; 20% éthanol) à 4°C pendant 2h. La membrane est alors immergée pendant 2 sec dans du méthanol puis rincée par du tampon de transfert.

Pour le Western blot, la membrane sur laquelle sont immobilisées les protéines est saturée pendant 30 min sous agitation en présence du tampon de blocage (1X PBS ; 3% lait en poudre ; 0,1% Tween-20). L'hybridation de l'anticorps primaire, dilué selon les instructions du fournisseur dans le tampon de blocage est réalisée 18h à 4°C. Les anticorps primaires non hybridés sont éliminés par 3 lavages successifs dans du tampon de lavage (PBS ; 0,1% Tween 20). La membrane est ensuite incubée 1 h en présence de l'anticorps secondaire couplé à la HRP, dilué au 1/10000 de tampon de blocage, puis à nouveau lavée 3 fois. L'interaction anticorps-antigène est enfin révélée par chimioluminescence à l'aide du Kit ECL Plus Western Blotting Detection System (GE Healthcare). Le signal lumineux est détecté à l'aide d'une station d'imagerie ChemiDoc (Biorad).

II.10 Test d'interaction SBP2-méthylosome à partir du système d'expression baculovirus

Différents bacmides permettant l'expression de GST-hSBP2, hPRMT5, hMEP50 et HA-hpICln ont été utilisés pour infecter des cellules d'insecte Sf9 (*Spodoptera frugiperda*). Les cellules ont été cultivées dans du milieu TNM-FH (Sigma) supplémenté de 10% de FCS et 50 µg/mL de gentamycine à 27 °C pendant 72 h. Ce travail a été réalisé par Nathalie Troffer-Charlier du service baculovirus de l'IGBMC

(Strasbourg), en collaboration avec l'équipe de Jean Cavarelli. 5 mL de culture sont centrifugés 1 min à 1200 rpm et le culot de cellules est lavé avec 5 mL de glycérol 10% puis stocké à -20°C. Le culot est repris dans 2 mL de tampon NT2, lysé par sonication et l'extrait est clarifié par 10 min de centrifugation à 9500 rpm. Pour les expériences de GST pull-down, 50 μ L de billes GSH-agarose 50% ont été incubés pendant 3 h à 4°C avec 300 μ L d'extrait cellulaire dans un volume final de 1 mL de tampon NT2. Après 5 lavages dans le tampon NT2, les billes sont mises en suspension dans 25 μ L de tampon de charge. 15 μ L sont déposés sur gel SDS-PAGE 8% suivis et analysés par Western blot.

Pour les expériences de co-immunoprécipitation à partir de HA-hpICln, 50 μ L de billes magnétiques μ MACs couplées aux anticorps anti-HA (μ MACS HA isolation Kit, Miltenyi Biotec) ont été utilisées pour 300 μ L d'extrait cellulaire suivant les instructions du fournisseur.

II.11 Test de méthylation *in vitro* de SBP2 par les protéines PRMT5/Mep50 du méthylosome

La capacité de SBP2 à être méthylée par le complexe PRMT5/Mep50 a été testée en présence de S-adénosyl-méthionine ^{14} C (SAM) et comparée à un substrat connu de PRMT5, le dimère d'histones H3/H4 (contrôle positif). L'enzyme PRMT5 copurifiée avec Mep50 et le dimère d'histones H3/H4 nous ont généreusement été fournis par Nathalie Troffer-Charlier et Vincent Cura (Equipe de Jean Cavarelli, IGBMC). 30 à 54 pmol de protéine SBP2 (entière ou tronquée) ou 40 pmol d'histone H3/H4 sont mis en présence de 5 pmol de PRMT5/Mep50 dans du tampon de méthylation (50 mM Tris-HCl pH 8 ; 200 mM NaCl ; 0,25 mM EDTA ; 0,25 mM DTT) et de 0,08 μ g de BSA dans un volume final de 8 μ L. La réaction est initiée par ajout de 1 μ L de SAM ^{14} C (10 μ M à 50 mCi/mmol), suivie d'une incubation de 30 min à 37°C. La réaction est arrêtée par ajout de 3 μ L de tampon de charge 4X. Des contrôles négatifs de méthylation sont réalisés en absence de PRMT5/Mep50. Les échantillons sont ensuite dénaturés 2 min à 100°C puis déposés sur un gel SDS-PAGE à gradient 4-12% de polyacrylamide (NuPAGE, Invitrogen). Le gel est coloré au bleu de Coomassie puis séché. La présence d'un signal radioactif de méthylation est détectée par autoradiographie.

II.12 Expériences de double hybride

Les cellules de la souche de levure *S. cerevisiae* AH109 (Clontech Laboratories, Inc.) servant aux tests d'interaction protéine-protéine par double-hybride ont été cultivées et co-transformées par des techniques standard. Un plasmide codant pour une protéine liée au domaine de liaison à l'ADN de Gal4 (pGBKT7 ou pAS) a été co-transformé avec un plasmide codant pour une protéine liée au domaine d'activation de la transcription de Gal4 (pGAD ou pACTII). Le domaine de liaison à l'ADN de Gal4 reconnaît le promoteur Gal1-10 qui est fusionné au gène de sélection HIS3. Les transformants ont été cultivés sur un milieu dépourvu en acides aminés leucine et tryptophane (-L-W). Les transformants qui poussent sur -L-W ont été transférés sur un milieu dépourvu de leucine, tryptophane, et histidine (-L-W-H). Une croissance indique l'existence d'une interaction entre les deux protéines testées, et donc la synthèse de l'histidine. Pour les expériences de double hybride bridge, un plasmide ADE2 codant pour une troisième protéine, Nufip (p422Nufip), a été transformé, la sélection des interactants se faisant sur (-L-W-H-ADE).

II.13 Transcription in vitro des ARN SelM et GPx1 par la T7 ARN polymérase

Transcription

Les matrices de transcription ont été réalisées par amplification PCR à partir des plasmides pSelM ou pGPx1, sauvages ou mutants. Les oligonucléotides utilisés en 5' sont allongés de la séquence du promoteur T7 afin de permettre la transcription du produit amplifié.

Les réactions de transcription ont été réalisées dans 100 μ L final contenant 100 μ g/mL de matrice ADN (produit de PCR), 40 mM de Tris-HCl pH 8,1, 22 mM de MgCl₂, 0,01% de Triton X-100, 1 mM de Spermidine et 5 mM de DTT, 5 mM de chaque rNTP, 0,4 U/ μ L de RNasine et 0,5 mg/mL de T7 ARN polymérase. Après 1 h d'incubation à 37°C, 0,01 mg/mL de pyrophosphatase (Roche) est ajoutée et l'incubation est poursuivie pour 30 minutes afin de digérer le pyrophosphate libre et déplacer l'équilibre vers la synthèse d'ARN. La DNase I est ensuite ajoutée (20 U/mL final) afin de digérer les brins d'ADN matrices pendant 1h. La réaction est arrêtée par extraction des ARN au phénol puis au phénol-chloroforme et précipitation à l'éthanol.

Purification des transcrits sur gel de polyacrylamide dénaturant

Les culots d'ARN repris dans 40 μ L de tampon Formamide (95% (v/v) formamide, 0,025% (p/v) bleu de bromophénol, 0,025% (p/v) bleu de xylène cyanol), dénaturés 5 min à 95°C sont séparés par électrophorèse sur gel de polyacrylamide 4% (3,8 % acrylamide, 0,2% bisacrylamide), Urée 8M, tampon TBE (89 mM Tris-Base, 89 mM acide borique, 2,5 mM EDTA). Les ARN ont été visualisés sous lumière UV et la bande correspondant à l'ARN d'intérêt, une fois coupée, est mise à éluer dans des cuves d'électro-élution « Biotrap » (Schleicher & Schuell), 2 fois 1 h à 120 V dans du tampon TBE. Les ARN élués sont précipités à l'éthanol.

II.14 Addition de la coiffe des ARN***Addition de la coiffe m⁷G***

Les ARN ont été coiffés de manière post-transcriptionnelle à l'aide du Kit ScriptCap™ m⁷G Capping System (C-SCCE0625, Cellscript) comme décrit par le fournisseur. Ce Kit contient l'enzyme VCE (Vaccinia Virus Capping Enzyme) qui a la particularité de catalyser successivement les 3 réactions nécessaires à la formation de la coiffe m⁷G, aboutissant à la structure (m⁷G(5')ppp(5')N1-pN2...).

Le coiffage "froid", nécessaire pour les expériences de traduction *in vitro*, de toeprint ou pour la purification de complexes traductionnels (étude de Cryo-microscopie électronique), est réalisé sur 60 μ g d'ARN en présence d'1 mM de GTP. Les ARN sont purifiés sur une colonne G25 et précipités à l'éthanol.

Le coiffage "chaud", nécessaire pour les expériences d'assemblage de complexes traductionnels analysés par fractionnement sur gradient de sucrose, est réalisé sur 10 μ g d'ARN en présence de 25 μ Ci d' [α^{32} P]-GTP. Les ARN sont ensuite purifiés sur gel de polyacrylamide 4% dénaturant, élués de manière passive 18h à 4°C dans du tampon E (10mM Tris-HCl pH 7.5, 0.3M NaCl, 0.5mM EDTA), puis précipités.

Addition de la coiffe m₃G

L'hyperméthylation est réalisée en même temps que le coiffage m⁷G, par ajout de 1,8 μ M final de GST-Tgs1₆₁₈₋₈₅₃ et de 5 mM de DTT dans le milieu réactionnel. A la fin de la réaction, 0,1 mM de SAM supplémentaires sont ajoutés et la réaction est poursuivie pendant 30 minutes. L'enzyme Tgs1 hyperméthyle la coiffe m⁷G

préalablement additionnée par la VCE, aboutissant à la structure ($m_3^{2,2,7}G(5')ppp(5')N1-pN2\dots$).

Caractérisation des coiffes par Chromatographie sur couche mince (CCM)

Dans le but de vérifier l'efficacité de la réaction d'hyperméthylation de la coiffe, les ARN sont digérés par 5 U/ μ L de nucléase S1 (Fermentas) pendant 20 min à 30°C dans du tampon contenant 40 mM d'acétate de sodium pH 4,5, 300 mM de NaCl et 2 mM de $ZnSO_4$, libérant ainsi des nucléosides. La réaction est arrêtée par l'ajout de 50 mM d'EDTA pH 8 et chauffage 10 minutes à 70°C. 2 μ L d'ARN digérés sont alors déposés sur une plaque de cellulose imprégnée de polyéthylèneimine (Polygram CEL 300 PEI, Macherey-Nagel). La chromatographie est réalisée dans une cuve contenant 100mL de sulfate d'ammonium 0,1M. Lorsque l'éluant atteint le haut de la plaque par capillarité, celle-ci est séchée puis révélée par autoradiographie.

II.15 Traduction *in vitro* des ARNm

250 nM d'ARNm de HA-SelM ou HA-GPx1 transcrits *in vitro* et coiffés par du m_3G ou du m^7G sont incubés en présence de 10 μ L de lysats de réticulocytes de lapins (RRL) non-traités à la nucléase, en présence de 0,22 mM d'un mélange d'acide aminés sans méthionine, 0,125 μ Ci/ μ L de méthionine ^{35}S , 1 U/ μ L de RNasin, 80 mM de KCl et 0,5 mM de $MgCl_2$ dans un volume final de 20 μ L, 1 h à 30°C.

Les protéines HA-SelM traduites sont ensuite purifiées à l'aide de billes magnétiques μ MACs couplées à des anticorps anti-HA (μ MACS HA isolation Kit, Miltenyi Biotec). Les produits de purification sont déposés sur un gel SDS-PAGE 12%, séchés puis analysés par autoradiographie après 3 jours d'exposition.

II.16 Cinétiques de traduction

40 μ L de milieu traductionnel (voir paragraphe précédent) contenant 85 nM d'ARNm HA-SelM ou HA-GPx1 non-coiffés, coiffés par du m_3G ou du m^7G et supplémentés de 160 nM de protéine SBP2 344-854 sont incubés à 30°C pendant 30 min. Toutes les 5 min, 5 μ L de milieu réactionnel sont prélevés, ajoutés de 45 μ L d'une solution 1 M NaOH, 2% H_2O_2 et incubés 10 min à 37°C afin d'arrêter la traduction. Les protéines sont précipitées sur du papier 3MM par une incubation de 10 min dans une solution

de TCA (acide trichloroacétique) 5%. Un lavage de 10 min dans du TCA 5% chauffé préalablement à 95°C permet d'éliminer les ARNt aminoacylés. Après une étape supplémentaire de lavage au TCA 5% froid, suivie de deux lavages à l'éthanol 100%, le papier 3MM est séché et la radioactivité est comptée à l'aide d'un compteur à scintillation liquide.

II.17 Assemblage de complexes traductionnels sur un ARNm à partir d'extraits RRL

Les ARNm de SelM ou de GPx1 transcrits *in vitro* sont coiffés par du m₃G ou du m⁷G avec de l'[α³²P]-GTP. En parallèle, 30 μL d'extraits de RRL non-traités à la nucléase ont été activés en présence de 0,22 mM d'un mélange d'acide aminés, 1 U/μL de RNasin, 80 mM de KCl et de 0,5 mM de MgCl₂ dans un volume final de 70 μL, 5 min à 30°C puis incubés 10 min dans la glace. 250000 cpm d'ARN marqué au niveau de la coiffe sont ensuite incubés 5 min à 30°C en présence du mélange de traduction contenant les RRL. Les complexes sont stabilisés par 15mM final d'acétate de magnésium et incubés dans la glace pendant 30 min. La totalité des complexes est ensuite déposée sur un gradient de saccharose 7-47% préparé dans le tampon G (25 mM Tris-HCl pH 7,4, 50 mM KCl, 5 mM MgCl₂, 1 mM DTT) et centrifugé à 37000rpm (Rotor SW41, Beckmann) pendant 2h30. Les gradients sont ensuite collectés en 44 fractions et la radioactivité présente dans chaque fraction est quantifiée à l'aide d'un compteur Cerenkov.

II.18 Assemblage de complexes traductionnels sur un ARNm à partir d'extraits de cellules HeLa

Des extraits de cellules HeLa traductionnellement actifs ont été préparés de la même manière que (Rakotondrafara and Hentze, 2011). 250000 cpm d'ARNm SelM ou de GPx1 coiffés par du m₃G ou du m⁷G avec de l'[α³²P]-GTP sont incubés pendant 7 min à 37°C en présence de 30 μL d'extraits dans un volume final de 75 μL contenant 1,6 mM d'HEPES pH 7,6, 2 mM de creatine phosphate, 0,01 μg/μL de creatine kinase, 0,01 mM spermidine, 40 mM d'acétate de potassium, 2 mM d'acétate de magnésium, 10 μM d'un mélange d'acide aminés et 1,6 U/μL de RNasin. Les complexes sont

stabilisés par 15 mM final d'acétate de magnésium et incubés dans la glace pendant 30 min. Ils sont analysés comme décrit dans le paragraphe précédent.

II.19 Purification de complexes traductionnels en vue d'analyse par Cryo-microscopie électronique

Purification des complexes

Le protocole suivi a été établi dans notre laboratoire par Prongidi-Fix et al. (2013). Le transcrit d'ARNm de SelM a été étendu, en 3' d'une queue 5'(CAA)₉CAC3' afin d'y liguer un oligonucléotide biotinyté (voir Tableau M7 pour les oligonucléotides spécifiques servant à générer la matrice de transcription). Après transcription, 60 µg d'ARNm de SelM sont coiffés par du m⁷G et ajoutés de 6,25 µM d'oligonucléotide A-biot, 6,25 µM d'oligonucléotide Splint 1, 6,25 µM d'oligonucléotide Splint 2 dans un volume final de 80 µL de tampon T4 (50 mM Tris-HCl pH 7,5, 10 mM MgCl₂, 1 mM ATP et 10 mM DTT) (Figure M1). L'ensemble est incubé 1 min à 95°C puis 10 min à température ambiante afin de permettre l'hybridation d'oligonucléotides. 0,5 U/µL de T4 DNA ligase ainsi que 1 mM d'ATP frais sont ajoutés pour un volume final de 100 µL. La réaction est incubée 18 h à 16°C. Les molécules hybrides ARN/ADN issues de la ligation sont ensuite fixées aux billes magnétiques streptavidine MagSi-STA 600 (MagnaMedics). Pour cela, la totalité des hybrides sont additionnés de 8,3 µM d'oligonucléotide anti-Splint1 et de 8,3 µM d'oligonucléotide anti-Splint2 dans un volume final de 120 µL, incubés 2 min à 95°C puis 10 min à température ambiante. 150 µL de billes MagSi-STA, préalablement lavées avec 2 fois 150 µL de tampon BB (100 mM phosphate de potassium pH 7,2, 150 mM NaCl), sont ajoutées aux hybrides et le mélange est incubé 30 min à température ambiante. Les ARN non liés aux billes sont ensuite éliminés par une succession de 5 lavages avec 400 µL d'H₂O (Figure M1). En parallèle, 85,7 µL d'extraits de RRL non-traités à la nucléase ont été activés en présence de 0,22 mM d'un mélange d'acide aminés, 1 U/µL de RNasin, 80 mM de KCl et de 0,5 mM de MgCl₂ dans un volume final de 200 µL, 5 min à 30°C puis incubés 10 min dans la glace. Les hybrides ARN/ADN immobilisés sur les billes sont ensuite incubés 5 min à 30°C en présence du mélange de traduction contenant les RRL. Les complexes sont incubés en présence de 17 mM d'acétate de magnésium 3 min à dans la glace puis 200 µL (1 volume) de tampon A (20 mM Tris-HCl pH 7,5, 100 mM acétate de potassium, 2,5 mM acétate de

magnésium, 2 mM DTT, 1 mM ATP, 0,1 mM GMP-PNP, 0,25 mM spermidine) sont ajoutés et l'incubation est poursuivie dans la glace pendant 5 min.

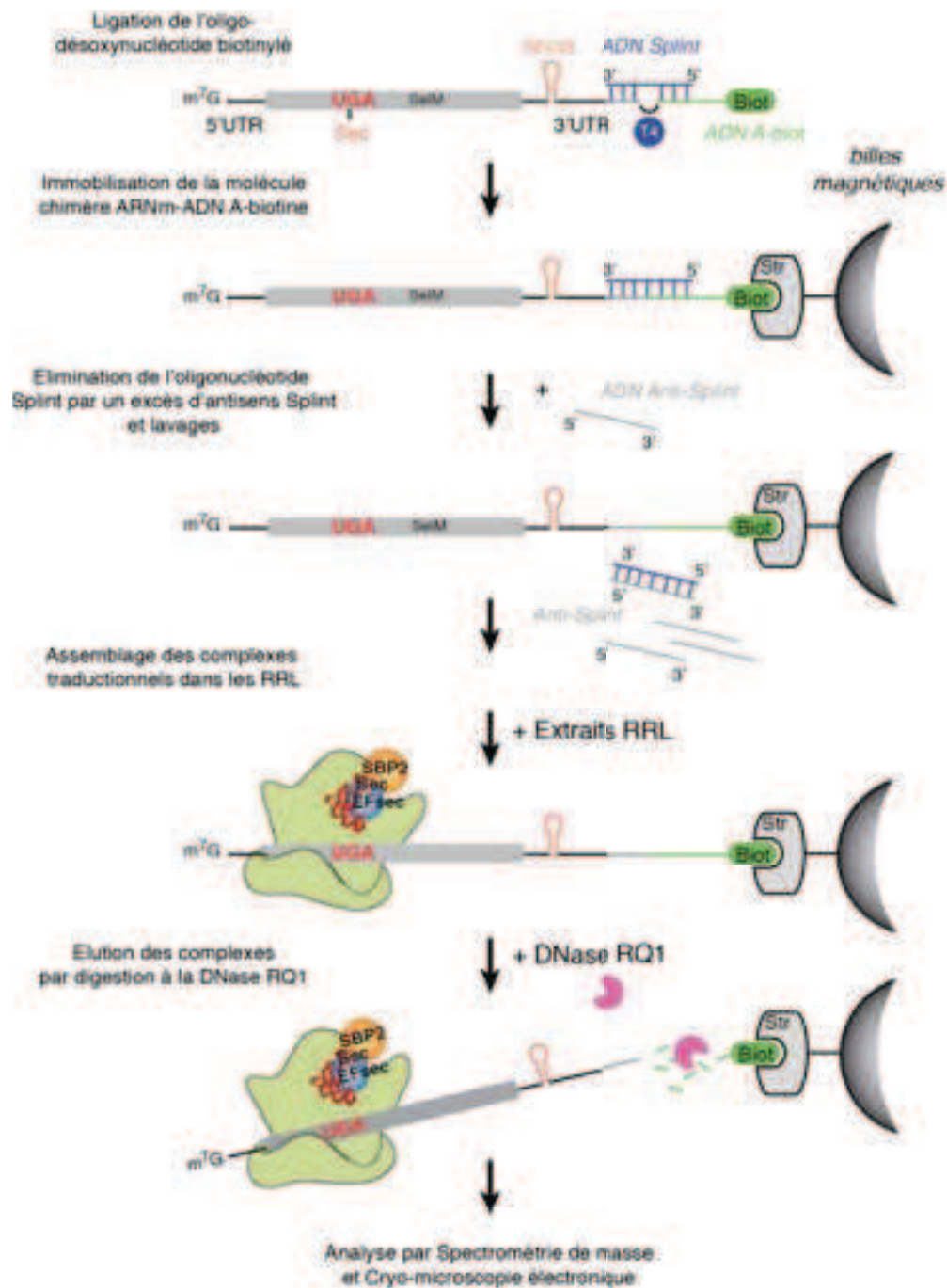


Figure M1. Schéma expérimental de la purification des complexes associés à l'ARNm de SelM *in vitro* en vue d'une analyse structurale par cryo-microscopie électronique.

Les billes sont ensuite lavées avec 2X 400 μ L de tampon 250 (50 mM Tris-HCl pH 7,5, 250 mM acétate de potassium, 25 mM acétate de magnésium, 250 mM de sucrose, 0,1% Igepal, 5 mM β -mercaptoéthanol), puis 3X 400 μ L de tampon 100 (20

mM Tris-HCl pH 7,5, 100 mM acétate de potassium, 10 mM acétate de magnésium, 1 mM de chlorure de calcium et 1 mM de DTT). Les complexes sont ensuite élués des billes, par digestion de la partie ADN de la molécule hybride ARNm de SelM/ ADN. Pour cela, 0,08 U/ μ L de DNase RQ1 (Promega) et 0,3 U/ μ L de RNasin sont ajoutés sur les billes présentes dans un volume de 120 μ L de tampon 100. L'ensemble est incubé 30 min à 37°C. Les billes sont éliminées et le surnageant, contenant les complexes élués, est ultra-centrifugé 1 h à 108 Krpm à 4°C (rotor S140AT, Sorvall-Hitachi). Le culot final est repris dans 6 μ L de tampon B (20 mM HEPES pH 7,5, 100 mM acétate de potassium, 10 mM acétate de magnésium, 0,2 mM EDTA et 1 mM DTT) à 4°C. L'absorbance à 260 nm est mesurée, 1 unité de DO correspond à 0,02 μ M de complexe 80S. Les complexes sont analysés par spectrométrie de masse par la plateforme de protéomique de l'IBMC.

II.20 Expériences de toeprint

Séquences de l'ARN à la RT-AMV

4 μ L d'ARN [2 μ M], sont dénaturés en présence de 400000 cpm d'oligonucléotides marqués en 5' au [γ^{32} P]-ATP dans un volume de 18 μ L final par 2 min d'incubation à 95°C suivis de 2 min dans la glace. L'hybridation est ensuite réalisée pendant 15 min à 20°C dans 20 μ L final de tampon AMV. Le mélange est ensuite réparti en 4X 5 μ L auxquels sont ajoutés 0,33 μ M d'un des 4 ddNTP (ddATP ou ddTTP ou ddCTP ou ddGTP), 2,5 μ M du dNTP correspondant, 10 μ M des 3 autres dNTP et 0,67 U/ μ L de RT AMV dans un volume final de 15 μ L de tampon AMV. L'élongation est réalisée 30 min à 37°C et les séquences sont ensuite précipitées à l'éthanol. Le culot est repris dans 20 μ L de tampon Formamide et 2 μ L seront déposés sur le gel.

Toeprint

7 μ L d'extraits de RRL non-traités à la nucléase ont été activés 5 min à 30°C en présence de 0,22 mM d'un mélange d'acide aminés, 1 U/ μ L de RNasin, 80 mM de KCl et 0,5 mM de MgCl₂ dans un volume final de 15 μ L, puis incubés 10 min dans la glace. 0,4 μ M final d'ARN coiffés par du m⁷G sont incubés 5 min à 30°C en présence du mélange de traduction contenant les RRL et, si indiqué, de 1 mg/mL de cycloheximide et/ou 5 mg/mL d'hygromycine, ou de 1 μ M de SBP2 344-854. 15 μ L (1 volume) de Tampon A (20 mM Tris-HCl pH 7,5, 100 mM acétate de potassium,

2,5 mM acétate de magnésium, 2 mM DTT, 1 mM ATP, 0,25 mM spermidine) froid sont ajoutés et les complexes sont centrifugés 1 h à 100000 rpm à 4°C (rotor S100AT3, Sorvall-Hitachi). Le surnageant, et le culot repris dans un volume de 30 μ L de tampon A froid, sont ensuite soumis à l'étape de réverse-transcription. Pour cela, ils sont incubés 3 min à 30°C avec 1 μ L (200000 cpm) d'oligonucléotides marqués en 5' au [γ^{32} P]-ATP par la T4 polynucleotide kinase (Fementas). 10 μ L de mix d'extension (9,26 mM dNTP, 29 μ M MgCl₂, 5,72 U/ μ L RNasin, 1,38 U/ μ L RT AMV (Q-Biogen)) sont ajoutés aux complexes et le mélange est incubé 1 h à 30°C. Les ARN sont ensuite extraits au phénol, puis phénol-chloroforme et précipités à l'éthanol. Les culots sont lavés 2 fois à l'éthanol 70% avant d'être séchés et repris dans 10 μ L de tampon Formamide. Ils seront déposés avec les séquences sur un gel de polyacrylamide 8% (7,6% acrylamide, 0,4% bisacrylamide), Urée 8 M, tampon TBE (89 mM Tris-Base, 89 mM acide borique, 2,5 mM EDTA). Après migration, le gel est séché et exposé 18 h avant d'être révélé par autoradiographie.

II.21 Expériences de retard sur gel

Les expériences de retard sur gel ont été réalisées comme décrit par (Takeuchi et al., 2009). Les ARN ont été transcrits *in vitro* et marqués de manière interne à l'aide d' [α^{32} P]-ATP. 30000 cpm d'ARN marqué au ^{32} P a été incubé 30 min à 30°C avec différentes concentrations de protéines hSBP2 515-854 (0 à 9 μ M) dans 7,5 μ L de tampon PBS pH 7,4, 2mM DTT. Les complexes ARN-SBP2 formés sont ensuite séparés sur un gel de polyacrylamide 6% non-dénaturant (5,7% acrylamide, 0,3% bisacrylamide), dans du tampon d'électrophorèse 0,5X TBE, 5% glycérol. Le gel est séché, exposé pendant 30 min et révélé par autoradiographie.

II.22 Purification des complexes associés à l'ARNm de HA-GPx1 *in vivo*

Durant ma thèse, j'ai développé une méthode permettant la purification de l'ARN HA-GPx1 à partir d'extraits cellulaires basée sur l'utilisation d'oligonucléotides biotinylés. Cette méthode et sa mise au point sont détaillées ci-dessous (voir aussi Figure M2).

Mise au point *in vitro*

Les conditions de rétention et d'élution de l'ARNm HA-GPx1 ont été mises au point sur le transcrit *in vitro* (Figure M2).

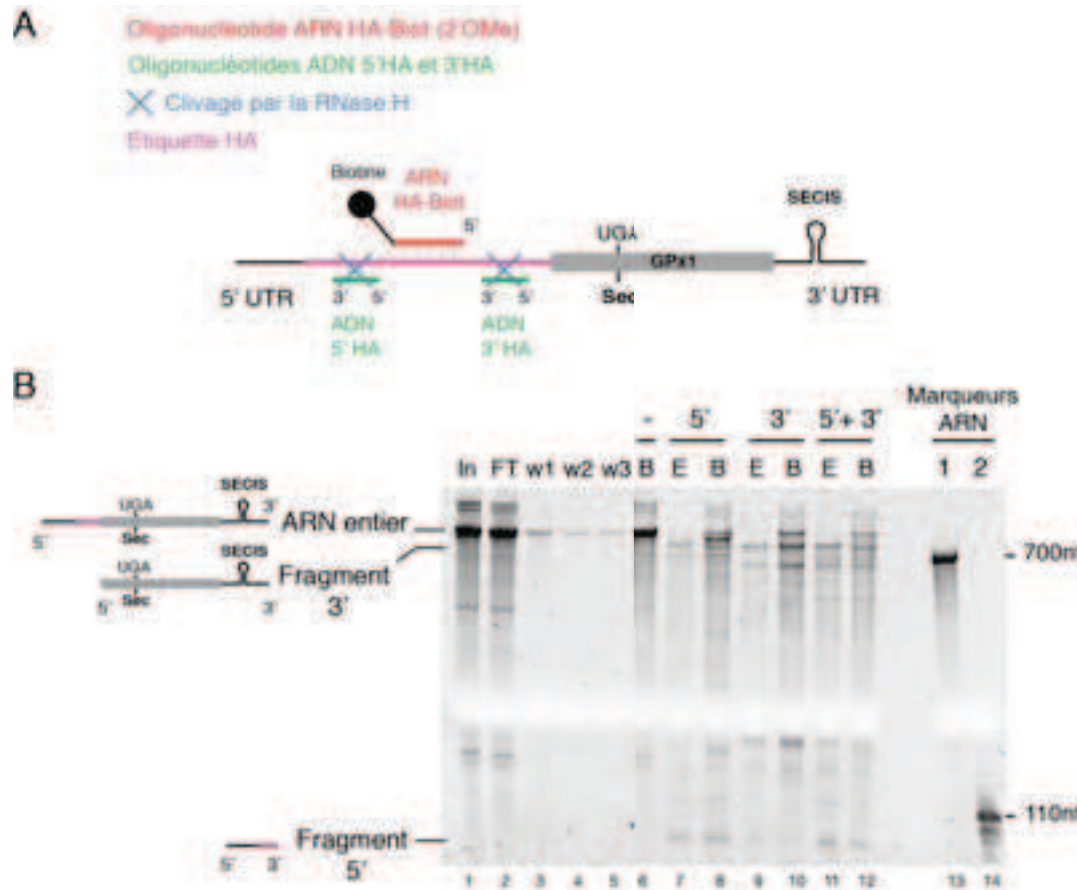


Figure M2. Mise au point *in vitro* de la purification de l'ARNm HA-GPx1.

(A) Schéma expérimental de la purification de l'ARNm HA-GPx1. (B) Suivi de purification de l'ARNm HA-GPx1 par électrophorèse sur gel d'acrylamide 4%, urée 8M, coloration au BET. In, input ; FT, surnageant (flow-through), w, wash ; B, billes ; E, éluat. L'expérience a été réalisée en absence d'élution (-), ou avec une étape d'élution en présence de l'oligonucléotide ADN 5' HA (5'), de l'oligonucléotide ADN 3' HA (3') ou avec des deux (5'+3').

Un oligonucléotide ARN 2'O-méthylé et biotinylé, complémentaire à une région de 20 nucléotides de l'étiquette HA (Oligonucléotide ARN HA-Biot, voir tableau M8) fixé aux billes MagSi-STA (Fournisseur) a été utilisé pour purifier l'ARN. Les oligonucléotides 2'O-méthylé sont plus résistants aux dégradations par les ribonucléases et forment des hybrides plus stables avec les ARN que les oligonucléotides ADN. Afin d'éluer spécifiquement les complexes des billes deux désoxy-oligonucléotides (5' HA et/ou 3' HA, voir tableau M8) sont hybridés de part et d'autre de l'oligonucléotide ARN HA-Biot. La digestion des duplexes ARN/ADN résultants par la RNase H (Figure M1-A) permet de libérer spécifiquement l'ARN

HA-GPx1. Les ARN présents dans les différentes fractions (inputs, lavages, billes et éluions) ont été extraits et précipités avant d'être analysés sur un gel de polyacrylamide 4%, 8 M urée et visualisés au BET (Figure M2-B). A la fin des étapes de purification, l'ARN HA-GPx1 est bien retenu par les billes via l'oligonucléotide ARN HA-Biot (Figure M2-B, piste 6). L'éluion par la RNase H conduit à la libération des fragments d'ARN en 5' (Figure M2-B, piste 7) et en 3' (Figure M2-B, piste 9) du site d'hybridation de l'oligonucléotide biotinylé. C'est l'utilisation conjointe des deux désoxy-oligonucléotides ADN et leur clivage par la RNase H qui permet d'obtenir un meilleur rendement d'éluion (Figure M2-B, piste 11).

Purification in vivo

Après validation de la méthode de clivage et d'éluion de l'ARN nous avons mis au point la purification de l'ARN HA-GPx1 *in vivo*. Notre objectif est de purifier les complexes associés à l'ARNm de HA-GPx1, exprimé *de novo* dans la lignée cellulaire correspondante selon la méthode exposée dans le paragraphe précédent, et d'analyser leur composition protéique par spectrométrie de masse. Dans ce but 10⁶ cellules issues des lignées HA-GPx1 inductibles sont mises en cultures dans 12 mL de DMEM/10% FCS dans des boîtes de 10 cm. Lorsque les cellules ont atteint 90% de confluence, l'expression de l'ARNm de HA-GPx1 est induite (ou non) pendant 24 h en remplaçant le milieu par 12 mL de DMEM/10% FCS contenant 0,5 µg/mL de Doxycycline et 10 nM de Na₂SeO₃. Les cellules sont ensuite lavées et collectées dans 2 mL de PBS contenant 150 µg/mL de cycloheximide. Les cellules sont centrifugées 10 min à 1000g, le culot est repris dans 1 volume de tampon « polysome » contenant 150 µg/mL de cycloheximide, congelé dans l'azote liquide, puis conservé à -80°C. Après élimination des débris cellulaires par centrifugation, 180 µL d'extrait cellulaire sont additionnés de 450 µL de tampon LS (20 mM HEPES pH 7,5, 100 mM KCl, 4 mM MgCl₂, 0,01% NP40, 1 mM DTT). Afin d'éliminer les ARN et protéines interagissant de façon aspécifique avec les billes, les extraits cellulaires sont clarifiés par incubation en présence de 25 µL de billes MagSi-STA (préalablement lavées avec du tampon LS) pendant 1 h à 4°C. En parallèle, 25 µL de billes MagSi-STA destinées à la purification de l'ARNm HA-GPx1 sont lavées 2 fois avec 100 µL de tampon LS et incubées 30 min à 4°C en présence de 50 µL de tampon LS contenant 10 µM d'oligonucléotide ARN HA-Biot (ou non, pour le contrôle sans oligonucléotide). Les

billes sont ensuite lavées 2 fois avec 200 μ L de tampon LS puis incubées 2 h à 4°C avec 600 μ L d'extrait cellulaire clarifié. Les billes sont alors lavées avec 600 μ L de tampon HS (20 mM HEPES pH 7,5, 250 mM KCl, 4 mM MgCl₂, 0,01% NP40, 1 mM DTT) puis 3 fois 600 μ L de tampon LS. Les complexes associés à l'ARNm de HA-GPx1 sont élués par incubation des billes pendant 45 min à 37°C, en présence de 10 μ M de chacun des 2 oligonucléotides 5' HA et 3' HA (voir tableau M8) et de 4 U de RNase H dans un volume de 25 μ L final de tampon H (20 mM Tris-HCl pH 7,8, 40 mM KCl, 8 mM MgCl₂, 1 mM DTT). 1/3 des fractions sont analysées par RT-qPCR après digestion à la protéinase K et extraction des ARN. 1/3 sont déposés sur un gel SDS-PAGE 12%, coloré au nitrate d'argent. Le reste est analysé par spectrométrie de masse afin d'identifier la composition des complexes associés à l'ARNm HA-GPx1.

Bibliographie

- Aachmann, F.L., Fomenko, D.E., Soragni, A., Gladyshev, V.N., and Dikiy, A. (2007). Solution structure of selenoprotein W and NMR analysis of its interaction with 14-3-3 proteins. *J. Biol. Chem.* 282, 37036–37044.
- Aachmann, F.L., Sal, L.S., Kim, H.Y., Marino, S.M., Gladyshev, V.N., and Dikiy, A. (2010). Insights into function, catalytic mechanism, and fold evolution of selenoprotein methionine sulfoxide reductase B1 through structural analysis. *J. Biol. Chem.* 285, 33315–33323.
- Acker, M.G., and Lorsch, J.R. (2008). Mechanism of ribosomal subunit joining during eukaryotic translation initiation. *Biochem. Soc. Trans.* 36, 653–657.
- Agamy, O., Ben Zeev, B., Lev, D., Marcus, B., Fine, D., Su, D., Narkis, G., Ofir, R., Hoffmann, C., Leshinsky-Silver, E., et al. (2010). Mutations disrupting selenocysteine formation cause progressive cerebello-cerebral atrophy. *Am. J. Hum. Genet.* 87, 538–544.
- Akten, B., Kye, M.J., Hao, L.T., Wertz, M.H., Singh, S., Nie, D., Huang, J., Merianda, T.T., Twiss, J.L., Beattie, C.E., et al. (2011). Interaction of survival of motor neuron (SMN) and HuD proteins with mRNA cpg15 rescues motor neuron axonal deficits. *Proc. Natl. Acad. Sci.* 108, 10337–10342.
- Algire, M.A., and Lorsch, J.R. (2006). Where to begin? The mechanism of translation initiation codon selection in eukaryotes. *Curr. Opin. Chem. Biol.* 10, 480–486.
- Allamand, V., Richard, P., Lescure, A., Ledeuil, C., Desjardin, D., Petit, N., Gartioux, C., Ferreira, A., Krol, A., Pellegrini, N., et al. (2006). A single homozygous point mutation in a 3'untranslated region motif of selenoprotein N mRNA causes SEPNI-related myopathy. *EMBO Rep* 7, 450–454.
- Allmang, C., and Krol, A. (2006a). Selenoprotein synthesis: UGA does not end the story. *Biochimie* 88, 1561–1571.
- Allmang, C., and Krol, A. (2006b). SECIS RNAs and K-turn binding proteins. A survey of evolutionary conserved RNA and protein motifs. *Selenium, Its Mol. Biol. Role Hum. Heal.* 2nd Ed. DL Hatf. Kluwer Acad. Publ. 5, 51–61.
- Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *Embo J* 18, 5399–5410.
- Allmang, C., Carbon, P., and Krol, A. (2002a). The SBP2 and 15.5 kD/Snu13p proteins share the same RNA binding domain: identification of SBP2 amino acids important to SECIS RNA binding. *RNA* 8, 1308–1318.
- Allmang, C., Carbon, P., and Krol, A. (2002b). The SBP2 and 15.5 kD/Snu13p proteins share the same RNA binding domain: identification of SBP2 amino acids important to SECIS RNA binding. *RNA* 8, 1308–1318.
- Allmang, C., Wurth, L., and Krol, A. (2009). The selenium to selenoprotein pathway in eukaryotes: More molecular partners than anticipated. *Biochim. Biophys. Acta - Gen. Subj.* 1790, 1415–1423.
- Altmann, M., Edery, I., Sonenberg, N., and Trachsel, H. (1985). Purification and characterization of protein synthesis initiation factor eIF-4E from the yeast *Saccharomyces cerevisiae*. *Biochemistry* 24, 6085–6089.
- Amaral, A.F.S., Cantor, K.P., Silverman, D.T., and Malats, N. (2010). Selenium and bladder cancer risk: a meta-analysis. *Cancer Epidemiol. Biomarkers Prev.* 19, 2407–2415.
- Ancelin, K., Lange, U.C., Hajkova, P., Schneider, R., Bannister, A.J., Kouzarides, T., and Surani, M.A. (2006). Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. *Nat. Cell Biol.* 8, 623–630.
- Andreev, D.E., Terenin, I.M., Dunaevsky, Y.E., Dmitriev, S.E., and Shatsky, I.N. (2006). A leaderless mRNA can bind to mammalian 80S ribosomes and direct polypeptide synthesis in the absence of translation initiation factors. *Mol. Cell. Biol.* 26, 3164–3169.
- Anger, A.M., Armache, J.-P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N., and Beckmann, R. (2013). Structures of the human and *Drosophila* 80S ribosome. *Nature* 497, 80–85.
- Antal, M., Boros, E., Solymosy, F., and Kiss, T. (2002). Analysis of the structure of human telomerase RNA in vivo. *Nucleic Acids Res.* 30, 912–920.
- Antonyamy, S., Bonday, Z., Campbell, R.M., Doyle, B., Druzina, Z., Gheyi, T., Han, B., Jungheim, L.N., Qian, Y., Rauch, C., et al. (2012). Crystal structure of the human PRMT5:MEP50 complex. *Proc. Natl. Acad. Sci.* 109, 17960–17965.
- Araiso, Y., Palioura, S., Ishitani, R., Sherrer, R.L., O'Donoghue, P., Yuan, J., Oshikane, H., Domae, N., Defranco, J., Soll, D., et al. (2008). Structural insights into RNA-dependent eukaryal and archaeal selenocysteine formation. *Nucleic Acids Res* 36, 1187–1199.
- Ariumi, Y. (2014). Multiple functions of DDX3 RNA helicase in gene regulation, tumorigenesis, and viral infection. *Front. Genet.* 5, 423.
- Armache, J.-P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., et al. (2010). Localization of eukaryote-specific ribosomal proteins in a 5.5-Å cryo-EM map of the 80S eukaryotic ribosome. *Proc. Natl. Acad. Sci. U. S. A.* 107, 19754–19759.
- Arnér, E.S.J. (2009). Focus on mammalian thioredoxin reductases - Important selenoproteins with versatile functions. *Biochim. Biophys. Acta - Gen. Subj.* 1790, 495–526.
- Arnér, E.S.J., and Holmgren, A. (2000). Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem.* 267, 6102–6109.

- Arsova-Sarafinovska, Z., Matevska, N., Eken, A., Petrovski, D., Banev, S., Dzikova, S., Georgiev, V., Sikole, A., Erdem, O., Sayal, A., et al. (2009). Glutathione peroxidase 1 (GPX1) genetic polymorphism, erythrocyte GPX activity, and prostate cancer risk. *Int. Urol. Nephrol.* *41*, 63–70.
- Asano, K., Phan, L., Valásek, L., Schoenfeld, L.W., Shalev, A., Clayton, J., Nielsen, K., Donahue, T.F., and Hinnebusch, A.G. (2001). A multifactor complex of eIF1, eIF2, eIF3, eIF5, and tRNAⁱ Met promotes initiation complex assembly and couples GTP hydrolysis to AUG recognition. In *Cold Spring Harbor Symposia on Quantitative Biology*, pp. 403–415.
- Askjaer, P., Bachi, A., Wilm, M., Bischoff, F.R., Weeks, D.L., Ogniewski, V., Ohno, M., Niehrs, C., Kjems, J., Mattaj, J.W., et al. (1999). RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. *Mol. Cell. Biol.* *19*, 6276–6285.
- Au, H.H., and Jan, E. (2014). Novel viral translation strategies. *Wiley Interdiscip. Rev. RNA*.
- Badura, M., Braunstein, S., Zavadil, J., and Schneider, R.J. (2012). DNA damage and eIF4G1 in breast cancer cells reprogram translation for survival and DNA repair mRNAs. *Proc. Natl. Acad. Sci.*
- Balakin, A.G., Skripkin, E.A., Shatsky, I.N., and Bogdanov, A.A. (1992). Unusual ribosome binding properties of mRNA encoding bacteriophage lambda repressor. *Nucleic Acids Res.* *20*, 563–571.
- Banerjee, A.K. (1980). 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol. Rev.* *44*, 175–205.
- Barbarossa, A., Antoine, E., Neel, H., Gostan, T., Soret, J., and Bordonné, R. (2014). Characterization and in vivo functional analysis of the *Schizosaccharomyces pombe* ICLN gene. *Mol. Cell. Biol.* *34*, 595–605.
- Bardia, A., Tleyjeh, I.M., Cerhan, J.R., Sood, A.K., Limburg, P.J., Erwin, P.J., and Montori, V.M. (2008). Efficacy of antioxidant supplementation in reducing primary cancer incidence and mortality: systematic review and meta-analysis. *Mayo Clin. Proc.* *83*, 23–34.
- Baron, C., and Böck, A. (1991). The length of the aminoacyl-acceptor stem of the selenocysteine-specific tRNA^{sec} of *Escherichia coli* is the determinant for binding to elongation factors SELB or Tu. *J. Biol. Chem.* *266*, 20375–20379.
- Battle, D.J., Lau, C.K., Wan, L., Deng, H., Lotti, F., and Dreyfuss, G. (2006). The Gemin5 Protein of the SMN Complex Identifies snRNAs. *Mol. Cell* *23*, 273–279.
- Battle, D.J., Kasim, M., Wang, J., and Dreyfuss, G. (2007). SMN-independent subunits of the SMN complex: Identification of a small nuclear ribonucleoprotein assembly intermediate. *J. Biol. Chem.* *282*, 27953–27959.
- Bebee, T.W., Dominguez, C.E., and Chandler, D.S. (2012). Mouse models of SMA: Tools for disease characterization and therapeutic development. *Hum. Genet.* *131*, 1277–1293.
- Beckett, G.J., and Arthur, J.R. (2005). Selenium and endocrine systems. *J. Endocrinol.* *184*, 455–465.
- Behne, D., Hilmert, H., Scheid, S., Gessner, H., and Elger, W. (1988). Evidence for specific selenium target tissues and new biologically important selenoproteins. *Biochim. Biophys. Acta* *966*, 12–21.
- Bellinger, F.P., Raman, A. V., Reeves, M.A., and Berry, M.J. (2009). Regulation and function of selenoproteins in human disease. *Biochem J* *422*, 11–22.
- 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.
- Ben-Shem, A., Jenner, L., Yusupova, G., and Yusupov, M. (2010). Crystal structure of the eukaryotic ribosome. *Science* *330*, 1203–1209.
- Bentley, D.L. (2005). Rules of engagement: Co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* *17*, 251–256.
- Berlow, R.B., Dyson, H.J., and Wright, P.E. (2015). Functional advantages of dynamic protein disorder. *FEBS Lett.*
- Bermano, G., Pagmantidis, V., Holloway, N., Kadri, S., Mowat, N.A.G., Shiel, R.S., Arthur, J.R., Mathers, J.C., Daly, A.K., Broom, J., et al. (2007). Evidence that a polymorphism within the 3'UTR of glutathione peroxidase 4 is functional and is associated with susceptibility to colorectal cancer. *Genes Nutr.* *2*, 225–232.
- Berry, M.J., Banu, L., Chen, Y.Y., Mandel, S.J., Kieffer, J.D., Harney, J.W., and Larsen, P.R. (1991). Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3' untranslated region. *Nature* *353*, 273–276.
- Berry, M.J., Banu, L., Harney, J.W., and Larsen, P.R. (1993). Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J.* *12*, 3315–3322.
- Berry, M.J., Harney, J.W., Ohama, T., and Hatfield, D.L. (1994). Selenocysteine insertion or termination: factors affecting UGA codon fate and complementary anticodon:codon mutations. *Nucleic Acids Res.* *22*, 3753–3759.
- Bifano, A.L., Atassi, T., Ferrara, T., and Driscoll, D.M. (2013). Identification of nucleotides and amino acids that mediate the interaction between ribosomal protein L30 and the SECIS element. *BMC Mol. Biol.* *14*, 12.
- Biterova, E.I., Turanov, A.A., Gladyshev, V.N., and Barycki, J.J. (2005). Crystal structures of oxidized and reduced mitochondrial thioredoxin reductase provide molecular details of the reaction mechanism. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 15018–15023.
- Blanchet, S., Rowe, M., Von der Haar, T., Fabret, C., Demais, S., Howard, M.J., and Namy, O. (2015). New insights into stop codon recognition by eRF1. *Nucleic Acids Res.* *43*, 3298–3308.
- Blumenthal, T. (2012). Trans-splicing and operons in *C. elegans*. *Wormb. Online Rev. C Elegans Biol.* 1–11.
- Blumenthal, T., and Gleason, K.S. (2003). *Caenorhabditis elegans* operons: form and function. *Nat. Rev. Genet.* *4*, 112–120.

- Böck, A. (2001). Selenium metabolism in bacteria. D. L. Hatf. (Ed), Selenium, Its Mol. Biol. Role Hum. Heal. *Kluwer Aca*, 7–22.
- Böck, A., Rother, M., Leibundgut, M., and Ban, N. (2006). Selenium metabolism in prokaryotes. In Selenium, pp. 9–28.
- Boisvert, F.-M., van Koningsbruggen, S., Navascués, J., and Lamond, A.I. (2007). The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* 8, 574–585.
- Boon, K.-L., Pearson, M.D., and Koš, M. (2015). Self-association of Trimethylguanosine Synthase Tgs1 is required for efficient snRNA/snoRNA trimethylation and pre-rRNA processing. *Sci. Rep.* 5, 11282.
- Borg, R.M., Bordonne, R., Vassallo, N., and Cauchi, R.J. (2015). Genetic Interactions between the Members of the SMN-Gemins Complex in *Drosophila*. *PLoS One* 10, e0130974.
- Bosl, M.R., Takaku, K., Oshima, M., Nishimura, S., and Taketo, M.M. (1997). Early embryonic lethality caused by targeted disruption of the mouse selenocysteine tRNA gene (Trsp). *Proc Natl Acad Sci U S A* 94, 5531–5534.
- Boulisfane, N., Choleza, M., Rage, F., Neel, H., Soret, J., and Bordonné, R. (2011). Impaired minor tri-snRNP assembly generates differential splicing defects of U12-type introns in lymphoblasts derived from a type I SMA patient. *Hum. Mol. Genet.* 20, 641–648.
- Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jady, B.E., Rothe, B., Pescia, C., Robert, M.C., Kiss, T., et al. (2008). The Hsp90 chaperone controls the biogenesis of L7Ae RNPs through conserved machinery. *J Cell Biol* 180, 579–595.
- Bradrick, S.S., and Gromeier, M. (2009). Identification of gemin5 as a novel 7-methylguanosine cap-binding protein. *PLoS One* 4.
- Brahms, H., Meheus, L., DE BRABANDERE, V., FISCHER, U.T.Z., and LÜHRMANN, R. (2001a). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* 7, 1531–1542.
- Brahms, H., Meheus, L., de Brabandere, V., Fischer, U., and Lührmann, R. (2001b). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* 7, 1531–1542.
- Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Haendler, B., and Jacob, M. (1982). U2 RNA shares a structural domain with U1, U4 and U5 RNAs. *EMBO J.* 1, 1259–1265.
- Branscombe, T.L., Frankel, A., Lee, J.H., Cook, J.R., Yang, Z.H., Pestka, S., and Clarke, S. (2001). PRMT5 (Janus Kinase-binding Protein 1) Catalyzes the Formation of Symmetric Dimethylarginine Residues in Proteins. *J. Biol. Chem.* 276, 32971–32976.
- Braun, K. a, and Young, E.T. (2014). Coupling mRNA synthesis and decay. *Mol. Cell. Biol.* 34, 4078–4087.
- De Breyne, S., Yu, Y., Unbehauen, A., Pestova, T. V, and Hellen, C.U.T. (2009). Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proc. Natl. Acad. Sci. U. S. A.* 106, 9197–9202.
- Brigelius-Flohé, R., and Maiorino, M. (2013). Glutathione peroxidases. *Biochim. Biophys. Acta - Gen. Subj.* 1830, 3289–3303.
- Brinkman, M., Reulen, R.C., Kellen, E., Buntinx, F., and Zeegers, M.P. (2006). Are men with low selenium levels at increased risk of prostate cancer? *Eur. J. Cancer* 42, 2463–2471.
- Bubenik, J.L., and Driscoll, D.M. (2007). Altered RNA binding activity underlies abnormal thyroid hormone metabolism linked to a mutation in selenocysteine insertion sequence-binding protein 2. *J Biol Chem* 282, 34653–34662.
- Bubenik, J.L., Ladd, A.N., Gerber, C.A., Budiman, M.E., and Driscoll, D.M. (2009). Known turnover and translation regulatory RNA-binding proteins interact with the 3' UTR of SECIS-binding protein 2. *RNA Biol* 6, 73–83.
- Bubenik, J.L., Miniard, A.C., and Driscoll, D.M. (2013). Alternative Transcripts and 3'UTR Elements Govern the Incorporation of Selenocysteine into Selenoprotein S. *PLoS One* 8.
- Budiman, M.E., Bubenik, J.L., Miniard, A.C., Middleton, L.M., Gerber, C.A., Cash, A., and Driscoll, D.M. (2009a). Eukaryotic initiation factor 4a3 is a selenium-regulated RNA-binding protein that selectively inhibits selenocysteine incorporation. *Mol Cell* 35, 479–489.
- Budiman, M.E., Bubenik, J.L., Miniard, A.C., Middleton, L.M., Gerber, C.A., Cash, A., and Driscoll, D.M. (2009b). Eukaryotic Initiation Factor 4a3 Is a Selenium-Regulated RNA-Binding Protein that Selectively Inhibits Selenocysteine Incorporation. *Mol. Cell* 35, 479–489.
- Budiman, M.E., Bubenik, J.L., and Driscoll, D.M. (2011). Identification of a signature motif for the eIF4a3-SECIS interaction. *Nucleic Acids Res.* 39, 7730–7739.
- Budkevich, T., Giesebrecht, J., Altman, R.B., Munro, J.B., Mielke, T., Nierhaus, K.H., Blanchard, S.C., and Spahn, C.M.T. (2011). Structure and dynamics of the mammalian ribosomal pretranslocation complex. *Mol. Cell* 44, 214–224.
- Burghes, A.H.M., and Beattie, C.E. (2009a). Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? *Nat. Rev. Neurosci.* 10, 597–609.
- Burghes, A.H.M., and Beattie, C.E. (2009b). Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? *Nat. Rev. Neurosci.* 10, 597–609.
- Burk, R.F., and Hill, K.E. (2005). Selenoprotein P: an extracellular protein with unique physical characteristics and a role in selenium homeostasis. *Annu. Rev. Nutr.* 25, 215–235.
- Burk, R.F., and Hill, K.E. (2009). Selenoprotein P-Expression, functions, and roles in mammals. *Biochim. Biophys. Acta - Gen. Subj.* 1790, 1441–1447.

- Burk, R.F., and Hill, K.E. (2015). Regulation of Selenium Metabolism and Transport. *Annu. Rev. Nutr.*
- Burnett, B.G., Muñoz, E., Tandon, A., Kwon, D.Y., Sumner, C.J., and Fischbeck, K.H. (2009). Regulation of SMN protein stability. *Mol. Cell. Biol.* 29, 1107–1115.
- Bushell, M., Stoneley, M., Kong, Y.W., Hamilton, T.L., Spriggs, K.A., Dobbyn, H.C., Qin, X., Sarnow, P., and Willis, A.E. (2006). Polypyrimidine Tract Binding Protein Regulates IRES-Mediated Gene Expression during Apoptosis. *Mol. Cell* 23, 401–412.
- Caban, K., and Copeland, P.R. (2012). Selenocysteine Insertion Sequence (SECIS)-binding protein 2 alters conformational dynamics of residues involved in tRNA accommodation in 80 S ribosomes. *J. Biol. Chem.* 287, 10664–10673.
- Caban, K., Kinzy, S.A., and Copeland, P.R. (2007). The L7Ae RNA binding motif is a multifunctional domain required for the ribosome-dependent Sec incorporation activity of Sec insertion sequence binding protein 2. *Mol Cell Biol* 27, 6350–6360.
- Cabañas, M.J., Vázquez, D., and Modolell, J. (1978). Dual interference of hygromycin B with ribosomal translocation and with aminoacyl-tRNA recognition. *Eur. J. Biochem.* 87, 21–27.
- Cai, A., Jankowska-Anyszka, M., Centers, A., Chlebicka, L., Stepinski, J., Stolarski, R., Darzynkiewicz, E., and Rhoads, R.E. (1999). Quantitative assessment of mRNA cap analogues as inhibitors of in vitro translation. *Biochemistry* 38, 8538–8547.
- Campbell, L. (2000). Direct interaction of Smn with dp103, a putative RNA helicase: a role for Smn in transcription regulation? *Hum. Mol. Genet.* 9, 1093–1100.
- Campbell, L., Potter, A., Ignatius, J., Dubowitz, V., and Davies, K. (1997). Genomic variation and gene conversion in spinal muscular atrophy: implications for disease process and clinical phenotype. *Am. J. Hum. Genet.* 61, 40–50.
- Campion, Y., Neel, H., Gostan, T., Soret, J., and Bordonné, R. (2010). Specific splicing defects in *S. pombe* carrying a degtron allele of the Survival of Motor Neuron gene. *EMBO J.* 29, 1817–1829.
- Carberry, S.E., Darzynkiewicz, E., Stepinski, J., Tahara, S.M., Rhoads, R.E., and Goss, D.J. (1990). A spectroscopic study of the binding of N-7-substituted cap analogues to human protein synthesis initiation factor 4E. *Biochemistry* 29, 3337–3341.
- Carissimi, C., Saieva, L., Gabanella, F., and Pellizzoni, L. (2006). Gemin8 is required for the architecture and function of the survival motor neuron complex. *J. Biol. Chem.* 281, 37009–37016.
- Carlson, B.A., Xu, X.-M., Kryukov, G. V., Rao, M., Berry, M.J., Gladyshev, V.N., and Hatfield, D.L. (2004). Identification and characterization of phosphoserine-tRNA[Ser]Sec kinase. *PNAS* 101, 12848–12853.
- Carlson, B.A., Xu, X.M., Gladyshev, V.N., and Hatfield, D.L. (2005). Selective rescue of selenoprotein expression in mice lacking a highly specialized methyl group in selenocysteine tRNA. *J. Biol. Chem.* 280, 5542–5548.
- Carlson, B.A., Moustafa, M.E., Sengupta, A., Schweizer, U., Shrimali, R., Rao, M., Zhong, N., Wang, S., Feigenbaum, L., Byeong, J.L., et al. (2007). Selective restoration of the selenoprotein population in a mouse hepatocyte selenoproteinless background with different mutant selenocysteine tRNAs lacking Um34. *J. Biol. Chem.* 282, 32591–32602.
- Cassidy, P.B., Edes, K., Nelson, C.C., Parsawar, K., Fitzpatrick, F.A., and Moos, P.J. (2006). Thioredoxin reductase is required for the inactivation of tumor suppressor p53 and for apoptosis induced by endogenous electrophiles. *Carcinogenesis* 27, 2538–2549.
- Castellano, S., Gladyshev, V.N., Guigo, R., and Berry, M.J. (2008). SelenoDB 1.0: a database of selenoprotein genes, proteins and SECIS elements. *Nucleic Acids Res* 36, D332–D338.
- Castets, P., Lescure, A., Guicheney, P., and Allamand, V. (2012). Selenoprotein N in skeletal muscle: From diseases to function. *J. Mol. Med.* 90, 1095–1107.
- Cha, J.-D., Kim, H.J., and Cha, I.-H. (2011). Genetic alterations in oral squamous cell carcinoma progression detected by combining array-based comparative genomic hybridization and multiplex ligation-dependent probe amplification. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* 111, 594–607.
- Chambers, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W., and Harrison, P.R. (1986). The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the “termination” codon, TGA. *EMBO J.* 5, 1221–1227.
- Chamond, N., Deforges, J., Ulryck, N., and Sargueil, B. (2014). 40S recruitment in the absence of eIF4G/4A by EMCV IRES refines the model for translation initiation on the archetype of Type II IRESs. *Nucleic Acids Res.* 42, 10373–10384.
- Chan, C.C., Dostie, J., Diem, M.D., Feng, W., Mann, M., Rappsilber, J., and Dreyfuss, G. (2004). eIF4A3 is a novel component of the exon junction complex. *RNA* 10, 200–209.
- Chan, S., Choi, E.-A., and Shi, Y. (2011). Pre-mRNA 3'-end processing complex assembly and function. *Wiley Interdiscip. Rev. RNA* 2, 321–335.
- Chandramouli, P., Topf, M., Ménétret, J.F., Eswar, N., Cannone, J.J., Gutell, R.R., Sali, A., and Akey, C.W. (2008). Structure of the Mammalian 80S Ribosome at 8.7 Å Resolution. *Structure* 16, 535–548.
- Chang, Y.-F., Imam, J.S., and Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. *Annu. Rev. Biochem.* 76, 51–74.
- Chao, J.A., and Williamson, J.R. (2004). Joint X-ray and NMR refinement of the yeast L30e-mRNA complex. *Struct.* 12, 1165–1176.
- Chapple, C.E., Guigo, R., and Krol, A. (2009a). SECISaln, a web-based tool for the creation of structure-based alignments of eukaryotic SECIS elements. *Bioinformatics* 25, 674–675.

- Chapple, C.E., Guigó, R., and Krol, A. (2009b). SECISalr, a web-based tool for the creation of structure-based alignments of eukaryotic SECIS elements. *Bioinformatics* 25, 674–675.
- Chari, A., and Fischer, U. (2010). Cellular strategies for the assembly of molecular machines. *Trends Biochem. Sci.* 35, 676–683.
- Chari, A., Golas, M.M., Klingenhäger, M., Neuenkirchen, N., Sander, B., Englbrecht, C., Sickmann, A., Stark, H., and Fischer, U. (2008). An Assembly Chaperone Collaborates with the SMN Complex to Generate Spliceosomal SnRNPs. *Cell* 135, 497–509.
- Charroux, B., Pellizzoni, L., Perkinson, R.A., Shevchenko, A., Mann, M., and Dreyfuss, G. (1999). Gemin3: A novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems. *J. Cell Biol.* 147, 1181–1193.
- Chavatte, L., Brown, B.A., and Driscoll, D.M. (2005). Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes. *Nat Struct Mol Biol.*
- Chen, W., and Moore, M.J. (2014). The spliceosome: Disorder and dynamics defined. *Curr. Opin. Struct. Biol.* 24, 141–149.
- Chen, P., Wang, R.R., Ma, X.J., Liu, Q., and Ni, J.Z. (2013a). Different forms of selenoprotein M differentially affect A?? aggregation and ROS generation. *Int. J. Mol. Sci.* 14, 4385–4399.
- Chen, Y.-C., Prabhu, K.S., and Mastro, A.M. (2013b). Is selenium a potential treatment for cancer metastasis? *Nutrients* 5, 1149–1168.
- Cheng, Q., Sandalova, T., Lindqvist, Y., and Arnér, E.S.J. (2009). Crystal structure and catalysis of the selenoprotein thioredoxin reductase 1. *J. Biol. Chem.* 284, 3998–4008.
- Cheng, W.H., Ho, Y.S., Ross, D.A., Han, Y., Combs, G.F., and Lei, X.G. (1997). Overexpression of cellular glutathione peroxidase does not affect expression of plasma glutathione peroxidase or phospholipid hydroperoxide glutathione peroxidase in mice offered diets adequate or deficient in selenium. *J. Nutr.* 127, 675–680.
- Chicher, J., Simonetti, A., Kuhn, L., Schaeffer, L., Hammann, P., Eriani, G., and Martin, F. (2015). Purification of mRNA-programmed translation initiation complexes suitable for mass spectrometry analysis. *Proteomics* 15, 2417–2425.
- Chittum, H.S., Hill, K.E., Carlson, B.A., Lee, B.J., Burk, R.F., and Hatfield, D.L. (1997). Replenishment of selenium deficient rats with selenium results in redistribution of the selenocysteine tRNA population in a tissue specific manner. *Biochim. Biophys. Acta - Mol. Cell Res.* 1359, 25–34.
- Chiu, S.Y., Lejeune, F., Ranganathan, A.C., and Maquat, L.E. (2004). The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex. *Genes Dev.* 18, 745–754.
- Cho, I.-M., Lai, L.B., Susanti, D., Mukhopadhyay, B., and Gopalan, V. (2010). Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14573–14578.
- Cho, P.F., Poulin, F., Cho-Park, Y.A., Cho-Park, I.B., Chicoine, J.D., Lasko, P., and Sonenberg, N. (2005). A new paradigm for translational control: Inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. *Cell* 121, 411–423.
- Cléry, A., Bourguignon-Igel, V., Allmang, C., Krol, A., and Branlant, C. (2007). An improved definition of the RNA-binding specificity of SECIS-binding protein 2, an essential component of the selenocysteine incorporation machinery. *Nucleic Acids Res.* 35, 1868–1884.
- Cobbold, L.C., Spriggs, K.A., Haines, S.J., Dobbryn, H.C., Hayes, C., de Moor, C.H., Lilley, K.S., Bushell, M., and Willis, A.E. (2008). Identification of Internal Ribosome Entry Segment (IRES)-trans-Acting Factors for the Myc Family of IRESs. *Mol. Cell. Biol.* 28, 40–49.
- Coldwell, M.J., Sack, U., Cowan, J.L., Barrett, R.M.A., Vlasak, M., Sivakumaran, K., and Morley, S. (2012). Multiple isoforms of the translation initiation factor eIF4GII are generated via use of alternative promoters, splice sites and a non-canonical initiation codon.
- Colgan, D.F., and Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* 11, 2755–2766.
- Conrad, M. (2009). Transgenic mouse models for the vital selenoenzymes cytosolic thioredoxin reductase, mitochondrial thioredoxin reductase and glutathione peroxidase 4. *Biochim. Biophys. Acta - Gen. Subj.* 1790, 1575–1585.
- Conti, E., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: Molecular insights and mechanistic variations across species. *Curr. Opin. Cell Biol.* 17, 316–325.
- Cooper, M.L., Adami, H.-O., Grönberg, H., Wiklund, F., Green, F.R., and Rayman, M.P. (2008). Interaction between single nucleotide polymorphisms in selenoprotein P and mitochondrial superoxide dismutase determines prostate cancer risk. *Cancer Res.* 68, 10171–10177.
- Copeland, P.R. (2003). Regulation of gene expression by stop codon recoding: Selenocysteine. *Gene* 312, 17–25.
- Copeland, P.R., and Driscoll, D.M. (1999). Purification, redox sensitivity, and RNA binding properties of SECIS-binding protein 2, a protein involved in selenoprotein biosynthesis. *J Biol Chem* 274, 25447–25454.
- Copeland, P.R., and Driscoll, D.M. (2001). RNA binding proteins and selenocysteine. *Biofactors* 14, 11–16.
- Copeland, P.R., Fletcher, J.E., Carlson, B.A., Hatfield, D.L., and Driscoll, D.M. (2000). A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *Embo J* 19, 306–314.
- Copeland, P.R., Stepanik, V.A., and Driscoll, D.M. (2001). Insight into mammalian selenocysteine insertion: domain structure and ribosome binding properties of Sec insertion sequence binding protein 2. *Mol Cell Biol* 21, 1491–1498.

- Di Cosmo, C., McLellan, N., Liao, X.H., Khanna, K.K., Weiss, R.E., Papp, L., and Refetoff, S. (2009). Clinical and molecular characterization of a novel selenocysteine insertion sequence-binding protein 2 (SBP2) gene mutation (R128X). *J Clin Endocrinol Metab.*
- Costantino, D.A., Pfingsten, J.S., Rambo, R.P., and Kieft, J.S. (2008). tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nat. Struct. Mol. Biol.* 15, 57–64.
- Cougot, N., Van Dijk, E., Babajko, S., and Séraphin, B. (2004). “Cap-tabolism.” *Trends Biochem. Sci.* 29, 436–444.
- Cowling, V.H. (2010). Regulation of mRNA cap methylation. *Biochem. J.* 425, 295–302.
- Cox, D.G., Tamimi, R.M., and Hunter, D.J. (2006). Gene x Gene interaction between MnSOD and GPX-1 and breast cancer risk: a nested case-control study. *BMC Cancer* 6, 217.
- Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M., and Borden, K.L.B. (2005). eIF4E promotes nuclear export of cyclin D1 mRNAs via an element in the 3'UTR. *J. Cell Biol.* 169, 245–256.
- Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M., and Borden, K.L.B. (2006). eIF4E is a central node of an RNA regulon that governs cellular proliferation. *J. Cell Biol.* 175, 415–426.
- Culjkovic, B., Topisirovic, I., and Borden, K.L.B. (2007). Controlling gene expression through RNA regulons: The role of the eukaryotic translation initiation factor eIF4E. *Cell Cycle* 6, 65–69.
- Cullen, B.R. (2003). Nuclear mRNA export: Insights from virology. *Trends Biochem. Sci.* 28, 419–424.
- Curcio, C., Baqui, M.M.A., Salvatore, D., Rihn, B.H., Mohr, S., Harney, J.W., Larsen, P.R., and Bianco, A.C. (2001). The Human Type 2 Iodothyronine Deiodinase Is a Selenoprotein Highly Expressed in a Mesothelioma Cell Line. *J. Biol. Chem.* 276, 30183–30187.
- Dabeva, M.D., and Warner, J.R. (1987). The yeast ribosomal protein L32 and its gene. *J. Biol. Chem.* 262, 16055–16059.
- Dacwag, C.S., Ohkawa, Y., Pal, S., Sif, S., and Imbalzano, A.N. (2007). The protein arginine methyltransferase Prmt5 is required for myogenesis because it facilitates ATP-dependent chromatin remodeling. *Mol. Cell. Biol.* 27, 384–394.
- Dacwag, C.S., Bedford, M.T., Sif, S., and Imbalzano, A.N. (2009). Distinct protein arginine methyltransferases promote ATP-dependent chromatin remodeling function at different stages of skeletal muscle differentiation. *Mol. Cell. Biol.* 29, 1909–1921.
- Dahlberg, J.E., Yang, H., Neuman De Vegvar, H., and Lund, E. (1990). Formation of the 3' end of U1 snRNA. *Mol. Biol. Rep.* 14, 161–162.
- Danthinne, X., Seurinck, J., Meulewaeter, F., Van Montagu, M., and Cornelissen, M. (1993). The 3' untranslated region of satellite tobacco necrosis virus RNA stimulates translation in vitro. *Mol. Cell. Biol.* 13, 3340–3349.
- Darzacq, X., Jády, B.E., Verheggen, C., Kiss, A.M., Bertrand, E., and Kiss, T. (2002). Cajal body-specific small nuclear RNAs: A novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J.* 21, 2746–2756.
- Darzynkiewicz, E., Stepinski, J., Ekiel, I., Jin, Y., Haber, D., Sijuwade, T., and Tahara, S.M. (1988). Beta-globin mRNAs capped with m7G, m2.7(2)G or m2.2.7(3)G differ in intrinsic translation efficiency. *Nucleic Acids Res.* 16, 8953–8962.
- 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.
- Decatur, W.A., and Fournier, M.J. (2003). RNA-guided nucleotide modification of ribosomal and other. RNAs, *J. Biol. Chem.* 278, 695–698.
- Deniz, N., Lenarcic, E.M., Landry, D.M., and Thompson, S.R. (2009). Translation initiation factors are not required for Dicistroviridae IRES function in vivo. *RNA* 15, 932–946.
- Derry, M.C., Yanagiya, A., Martineau, Y., and Sonenberg, N. (2006). Regulation of poly(A)-binding protein through PABP-interacting proteins. In *Cold Spring Harbor Symposia on Quantitative Biology*, pp. 537–543.
- Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb. Perspect. Biol.* 4, 1–16.
- Diamond, A., Dudock, B., and Hatfield, D. (1981). Structure and properties of a bovine liver UGA suppressor serine tRNA with a tryptophan anticodon. *Cell* 25, 497–506.
- Diamond, A.M., Choi, I.S., Crain, P.F., Hashizume, T., Pomerantz, S.C., Cruz, R., Steer, C.J., Hill, K.E., Burk, R.F., McCloskey, J.A., et al. (1993). Dietary selenium affects methylation of the wobble nucleoside in the anticodon of selenocysteine tRNA[Ser]Sec. *J. Biol. Chem.* 268, 14215–14223.
- Dias, S.M.G., Wilson, K.F., Rojas, K.S., Ambrosio, A.L.B., and Cerione, R.A. (2009). The molecular basis for the regulation of the cap-binding complex by the importins. *Nat. Struct. Mol. Biol.* 16, 930–937.
- Dikiy, A., Novoselov, S. V., Fomenko, D.E., Sengupta, A., Carlson, B.A., Cerny, R.L., Ginalski, K., Grishin, N. V., Hatfield, D.L., and Gladyshev, V.N. (2007). SelT, SelW, SelH, and Rdx12: genomics and molecular insights into the functions of selenoproteins of a novel thioredoxin-like family. *Biochemistry* 46, 6871–6882.
- Ding, F., and Grabowski, P.J. (1999). Identification of a protein component of a mammalian tRNA(Sec) complex implicated in the decoding of UGA as selenocysteine. *RNA* 5, 1561–1569.
- Dinos, G., Wilson, D.N., Teraoka, Y., Szaflarski, W., Fucini, P., Kalpaxis, D., and Nierhaus, K.H. (2004). Dissecting the Ribosomal Inhibition Mechanisms of Edeine and Pactamycin: The Universally Conserved Residues G693 and C795 Regulate P-Site RNA Binding. *Mol. Cell* 13, 113–124.
- Donovan, J., and Copeland, P.R. (2009). Evolutionary history of selenocysteine incorporation from the perspective of SECIS binding proteins. *BMC Evol. Biol.* 9:229, doi:10.1186/1471-2148-9-229.

- Donovan, J., and Copeland, P.R. (2010). The efficiency of selenocysteine incorporation is regulated by translation initiation factors. *J. Mol. Biol.* **400**, 659–664.
- Donovan, J., and Copeland, P.R. (2012). Selenocysteine insertion sequence binding protein 2L is implicated as a novel post-transcriptional regulator of selenoprotein expression. *PLoS One* **7**.
- Donovan, J., Caban, K., Ranaweera, R., Gonzalez-Flores, J.N., and Copeland, P.R. (2008). A novel protein domain induces high affinity selenocysteine insertion sequence binding and elongation factor recruitment. *J Biol Chem* **283**, 35129–35139.
- Van Doren, K., and Hirsh, D. (1990). mRNAs that mature through trans-splicing in *Caenorhabditis elegans* have a trimethylguanosine cap at their 5' termini. *Mol. Cell. Biol.* **10**, 1769–1772.
- Dreher, I., Schmutzler, C., Jakob, F., and Köhrle, J. (1997). Expression of selenoproteins in various rat and human tissues and cell lines. *J. Trace Elem. Med. Biol.* **11**, 83–91.
- Dumitrescu, A.M., Liao, X.H., Abdullah, M.S., Lado-Abeal, J., Majed, F.A., Moeller, L.C., Boran, G., Schomburg, L., Weiss, R.E., and Refetoff, S. (2005). Mutations in SECISBP2 result in abnormal thyroid hormone metabolism. *Nat Genet* **37**, 1247–1252.
- Durand, S., and Lykke-Andersen, J. (2013). Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nat. Struct. Mol. Biol.* **20**, 702–709.
- Eliseeva, I. a, Lyabin, D.N., and Ovchinnikov, L.P. (2013). Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochem. Biokhimiia* **78**, 1377–1391.
- Emmink, B.L., Laoukili, J., Kipp, A.P., Koster, J., Govaert, K.M., Fatrai, S., Verheem, A., Steller, E.J. a, Brigelius-Flohé, R., Jimenez, C.R., et al. (2014). GPx2 Suppression of H2O2 Stress Links the Formation of Differentiated Tumor Mass to Metastatic Capacity in Colorectal Cancer. *Cancer Res.* **74**, 6717–6730.
- Epp, O., Ladenstein, R., and Wendel, A. (1983). The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution. *Eur. J. Biochem.* **133**, 51–69.
- Erzberger, J.P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C.H.S., Cimermančič, P., Boehringer, D., Sali, A., Aebersold, R., et al. (2014). Molecular architecture of the 40S·eIF1·eIF3 translation initiation complex. *Cell* **158**, 1123–1135.
- Etchison, D., Milburn, S.C., Edery, 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.
- Fabrizio, P., Laggenbauer, B., Lauber, J., Lane, W.S., and Luhrmann, R. (1997). An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. *Embo J* **16**, 4092–4106.
- Pagegaltier, D., Lescure, A., Walczak, R., Carbon, P., and Krol, A. (2000a). Structural analysis of new local features in SECIS RNA hairpins. *Nucleic Acids Res* **28**, 2679–2689.
- Pagegaltier, D., Hubert, N., Yamada, K., Mizutani, T., Carbon, P., and Krol, A. (2000b). Characterization of mSelB, a novel mammalian elongation factor for selenoprotein translation. *Embo J* **19**, 4796–4805.
- Pagegaltier, D., Hubert, N., Carbon, P., and Krol, A. (2000c). The selenocysteine insertion sequence binding protein SBP is different from the Y-box protein dbpB. *Biochimie* **82**, 117–122.
- Pagegaltier, D., Hubert, N., Carbon, P., and Krol, A. (2000d). The selenocysteine insertion sequence binding protein SBP is different from the Y-box protein dbpB. *Biochimie* **82**, 117–122.
- Fallini, C., Zhang, H., Su, Y., Silani, V., Singer, R.H., Rossoll, W., and Bassell, G.J. (2011). The survival of motor neuron (SMN) protein interacts with the mRNA-binding protein HuD and regulates localization of poly(A) mRNA in primary motor neuron axons. *J. Neurosci.* **31**, 3914–3925.
- Fallini, C., Bassell, G.J., and Rossoll, W. (2012a). Spinal muscular atrophy: The role of SMN in axonal mRNA regulation. *Brain Res.* **1462**, 81–92.
- Fallini, C., Bassell, G.J., and Rossoll, W. (2012b). Spinal muscular atrophy: The role of SMN in axonal mRNA regulation. *Brain Res.* **1462**, 81–92.
- Fallini, C., Rouanet, J.P., Donlin-Asp, P.G., Guo, P., Zhang, H., Singer, R.H., Rossoll, W., and Bassell, G.J. (2014). Dynamics of survival of motor neuron (SMN) protein interaction with the mRNA-binding protein IMP1 facilitates its trafficking into motor neuron axons. *Dev. Neurobiol.* **74**, 319–332.
- Feldkötter, M., Schwarzer, V., Wirth, R., Wienker, T.F., and Wirth, B. (2002). Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am. J. Hum. Genet.* **70**, 358–368.
- Feng, W., Gubit, A.K., Wan, L., Battle, D.J., Dostie, J., Golembe, T.J., and Dreyfuss, G. (2005). Gemin5 modulate the expression and activity of the SMN complex. *Hum. Mol. Genet.* **14**, 1605–1611.
- Feoktistova, K., Tuvshintogs, E., Do, A., and Fraser, C.S. (2013). Human eIF4E promotes mRNA restructuring by stimulating eIF4A helicase activity. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13339–13344.
- Ferguson, A.D., Labunskyy, V.M., Fomenko, D.E., Araç, D., Chelliah, Y., Amezcua, C.A., Rizo, J., Gladyshev, V.N., and Deisenhofer, J. (2006). NMR structures of the selenoproteins Sep15 and SelM reveal redox activity of a new thioredoxin-like family. *J. Biol. Chem.* **281**, 3536–3543.
- Fernández, I.S., Bai, X.C., Murshudov, G., Scheres, S.H.W., and Ramakrishnan, V. (2014). Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* **157**, 823–831.
- Fierro-Monti, I., Mohammed, S., Matthiesen, R., Santoro, R., Burns, J.S., Williams, D.J., Proud, C.G., Kassem, M., Jensen, O.N., and Roepstorff, P. (2006). Quantitative proteomics identifies Gemin5, a scaffolding protein

- involved in ribonucleoprotein assembly, as a novel partner for eukaryotic initiation factor 4E. *J. Proteome Res.* 5, 1367–1378.
- Fischer, U., and Lührmann, R. (1990). An essential signaling role for the m3G cap in the transport of U1 snRNP to the nucleus. *Science* 249, 786–790.
- Fischer, U., Sumpter, V., Sekine, M., Satoh, T., and Lührmann, R. (1993). Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. *EMBO J.* 12, 573–583.
- Fischer, U., Liu, Q., and Dreyfuss, G. (1997). The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* 90, 1023–1029.
- Fischer, U., Englbrecht, C., and Chari, A. (2011). Biogenesis of spliceosomal small nuclear ribonucleoproteins. *Wiley Interdiscip. Rev. RNA* 2, 718–731.
- Fixsen, S.M., and Howard, M.T. (2010). Processive selenocysteine incorporation during synthesis of eukaryotic selenoproteins. *J. Mol. Biol.* 399, 385–396.
- Flanagan IV, J.F., Namy, O., Brierley, I., and Gilbert, R.J.C. (2010). Direct Observation of Distinct A/P Hybrid-State tRNAs in Translocating Ribosomes. *Structure* 18, 257–264.
- Fletcher, J.E., Copeland, P.R., and Driscoll, D.M. (2000). Polysome distribution of phospholipid hydroperoxide glutathione peroxidase mRNA: evidence for a block in elongation at the UGA/selenocysteine codon. *Rna* 6, 1573–1584.
- Fletcher, J.E., Copeland, P.R., Driscoll, D.M., and Krol, A. (2001). The selenocysteine incorporation machinery: interactions between the SECIS RNA and the SECIS-binding protein SBP2. *Rna* 7, 1442–1453.
- Flohe, L., Günzler, W.A., and Schock, H.H. (1973). Glutathione peroxidase: a selenoenzyme. *FEBS Lett.* 32, 132–134.
- Flohe, L., Andreessen, J.R., Brigelius-Flohe, R., Maiorino, M., and Ursini, F. (2000). Selenium, the element of the moon, in life on earth. *IUBMB Life* 49, 411–420.
- Flohé, L. (2009). The labour pains of biochemical selenology: The history of selenoprotein biosynthesis. *Biochim. Biophys. Acta - Gen. Subj.* 1790, 1389–1403.
- Fornerod, M., Ohno, M., Yoshida, M., and Mattaj, I.W. (1997a). CRM1 Is an Export Receptor for Leucine-Rich Nuclear Export Signals. *Cell* 90, 1051–1060.
- Fornerod, M., Ohno, M., Yoshida, M., and Mattaj, I.W. (1997b). CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* 90, 1051–1060.
- Frank, J., and Agrawal, R.K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* 406, 318–322.
- Frederiks, W.M., Bosch, K.S., Hoeben, K.A., van Marle, J., and Langbein, S. (2010). Renal cell carcinoma and oxidative stress: The lack of peroxisomes. *Acta Histochem.* 112, 364–371.
- French, R.L., Gupta, N., Copeland, P.R., and Simonović, M. (2014). Structural asymmetry of the terminal catalytic complex in selenocysteine synthesis. *J. Biol. Chem.* 289, 28783–28794.
- Friend, L.R., Landsberg, M.J., Nouwens, A.S., Wei, Y., Rothnagel, J.A., and Smith, R. (2013). Arginine Methylation of hnRNP A2 Does Not Directly Govern Its Subcellular Localization. *PLoS One* 8.
- Friesen, W.J. (2000). Specific Sequences of the Sm and Sm-like (Lsm) Proteins Mediate Their Interaction with the Spinal Muscular Atrophy Disease Gene Product (SMN). *J. Biol. Chem.* 275, 26370–26375.
- Friesen, W.J., and Dreyfuss, G. (2000). Specific sequences of the Sm and Sm-like (Lsm) proteins mediate their interaction with the spinal muscular atrophy disease gene product (SMN). *J. Biol. Chem.* 275, 26370–26375.
- Friesen, W.J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G.S., Van Duyne, G., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001). The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol. Cell. Biol.* 21, 8289–8300.
- Fu, Y., Cheng, W.H., Porres, J.M., Ross, D.A., and Lei, X.G. (1999). Knockout of cellular glutathione peroxidase gene renders mice susceptible to diquat-induced oxidative stress. *Free Radic. Biol. Med.* 27, 605–611.
- Gamarnik, A. V., and Andino, R. (1996). Replication of poliovirus in *Xenopus* oocytes requires two human factors. *EMBO J.* 15, 5988–5998.
- Ganichkin, O.M., Xu, X.M., Carlson, B.A., Mix, H., Hatfield, D.L., Gladyshev, V.N., and Wahl, M.C. (2008). Structure and catalytic mechanism of eukaryotic selenocysteine synthase. *J Biol Chem* 283, 5849–5865.
- Gao, F., Kasprzak, W.K., Szarko, C., Shapiro, B.A., and Simon, A.E. (2014). The 3' Untranslated Region of Pea Enation Mosaic Virus Contains Two T-Shaped, Ribosome-Binding, Cap-Independent Translation Enhancers. *J. Virol.* 88, 11696–11712.
- Gasdaska, J.R., Harney, J.W., Gasdaska, P.Y., Powis, G., and Berry, M.J. (1999). Regulation of human thioredoxin reductase expression and activity by 3'- untranslated region selenocysteine insertion sequence and mRNA instability elements. *J. Biol. Chem.* 274, 25379–25385.
- Gautier, T., Bergès, T., Tollervey, D., and Hurt, E. (1997). Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. *Mol. Cell. Biol.* 17, 7088–7098.
- Gehring, N.H., Kunz, J.B., Neu-Yilik, G., Breit, S., Viegas, M.H., Hentze, M.W., and Kulozik, A.E. (2005). Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. *Mol. Cell* 20, 65–75.
- 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.
- Girard, C., Verheggen, C., Neel, H., Cammas, A., Vagner, S., Soret, J., Bertrand, E., and Bordonne, R. (2008). Characterization of a short isoform of human Tgs1 hypermethylase associating with small nucleolar

- ribonucleoprotein core proteins and produced by limited proteolytic processing. *J Biol Chem* 283, 2060–2069.
- Gobler, C.J., Lobanov, A. V, Tang, Y.-Z., Turanov, A. a, Zhang, Y., Doblin, M., Taylor, G.T., Sañudo-Wilhelmy, S. a, Grigoriev, I. V, and Gladyshev, V.N. (2013). The central role of selenium in the biochemistry and ecology of the harmful pelagophyte, *Aureococcus anophagefferens*. *ISME J.* 7, 1333–1343.
- Gonatos-Pournatzis, T., and Cowling, V.H. (2014). Cap-binding complex (CBC). *Biochem. J.* 457, 231–242.
- Gonsalvez, G.B., Tian, L., Ospina, J.K., Boisvert, F.M., Lamond, A.I., and Matera, A.G. (2007). Two distinct arginine methyltransferases are required for biogenesis of Sm-class ribonucleoproteins. *J. Cell Biol.* 178, 733–740.
- Gonzalez-Flores, J.N., Gupta, N., DeMong, L.W., and Copeland, P.R. (2012). The selenocysteine-specific elongation factor contains a novel and multi-functional domain. *J. Biol. Chem.* 287, 38936–38945.
- Gonzalez-Moreno, O., Boque, N., Redrado, M., Milagro, F., Campion, J., Endermann, T., Takahashi, K., Saito, Y., Catena, R., Schomburg, L., et al. (2011). Selenoprotein-P is down-regulated in prostate cancer, which results in lack of protection against oxidative damage. *Prostate* 71, 824–834.
- 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.
- Gribun, A., Cheung, K.L., Huen, J., Ortega, J., and Houry, W.A. (2008). Yeast Rvb1 and Rvb2 are ATP-dependent DNA helicases that form a heterohexameric complex. *J Mol Biol* 376, 1320–1333.
- Grimm, C., Chari, A., Pelz, J.-P., Kuper, J., Kisker, C., Diederichs, K., Stark, H., Schindelin, H., and Fischer, U. (2013). Structural Basis of Assembly Chaperone-Mediated snRNP Formation. *Mol. Cell* 49, 692–703.
- Grimmler, M., Otter, S., Peter, C., Müller, F., Chari, A., and Fischer, U. (2005). Unrip, a factor implicated in cap-independent translation, associates with the cytosolic SMN complex and influences its intracellular localization. *Hum. Mol. Genet.* 14, 3099–3111.
- Gross, J.D., Moerke, N.J., Von Der Haar, T., Lugovskoy, A.A., Sachs, A.B., McCarthy, J.E.G., and Wagner, G. (2003). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115, 739–750.
- Grundner-Culemann, E., Martin 3rd, G.W., Harney, J.W., and Berry, M.J. (1999). Two distinct SECIS structures capable of directing selenocysteine incorporation in eukaryotes. *Rna* 5, 625–635.
- Gu, M., and Lima, C.D. (2005). Processing the message: Structural insights into capping and decapping mRNA. *Curr. Opin. Struct. Biol.* 15, 99–106.
- Gubitz, A.K., Mourelatos, Z., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). Gemin5, a novel WD repeat protein component of the SMN complex that binds Sm proteins. *J. Biol. Chem.* 277, 5631–5636.
- Gubitz, A.K., Feng, W., and Dreyfuss, G. (2004). The SMN complex. *Exp. Cell Res.* 296, 51–56.
- Guderian, G., Peter, C., Wiesner, J., Sickmann, A., Schulze-Osthoff, K., Fischer, U., and Grimmler, M. (2011). RioK1, a new interactor of protein arginine methyltransferase 5 (PRMT5), competes with pICln for binding and modulates PRMT5 complex composition and substrate specificity. *J. Biol. Chem.* 286, 1976–1986.
- Guerriero, G., Trocchia, S., Abdel-Gawad, F.K., and Ciarcia, G. (2014). Roles of reactive oxygen species in the spermatogenesis regulation. *Front. Endocrinol. (Lausanne)*. 5.
- Guimaraes, M.J., Peterson, D., Vicari, A., Cocks, B.G., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Ferrick, D.A., Kastelein, R.A., Bazan, J.F., et al. (1996). Identification of a novel selD homolog from eukaryotes, bacteria, and archaea: is there an autoregulatory mechanism in selenocysteine metabolism? *Proc Natl Acad Sci U S A* 93, 15086–15091.
- Guo, L., Allen, E.M., and Miller, W.A. (2001). Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. *Mol. Cell* 7, 1103–1109.
- Gupta, M., and Copeland, P.R. (2007). Functional analysis of the interplay between translation termination, selenocysteine codon context, and selenocysteine insertion sequence-binding protein 2. *J Biol Chem* 282, 36797–36807.
- Gupta, K., Martin, R., Sharp, R., Sarachan, K.L., Ninan, N.S., and Van Duyne, G.D. (2015). Oligomeric Properties of Survival Motor Neuron-Gemin2 Complexes. *J. Biol. Chem.* jbc.M115.667279.
- Gupta, N., Demong, L.W., Banda, S., and Copeland, P.R. (2013). Reconstitution of selenocysteine incorporation reveals intrinsic regulation by SECIS elements. *J. Mol. Biol.* 425, 2415–2422.
- Von der Haar, T., Gross, J.D., Wagner, G., and McCarthy, J.E.G. (2004). The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat. Struct. Mol. Biol.* 11, 503–511.
- Haghighat, A., and Sonenberg, N. (1997). eIF4g dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. *J. Biol. Chem.* 272, 21677–21680.
- Haghighat, A., Svitkin, Y., Novoa, I., Kuechler, E., Skern, T., and Sonenberg, N. (1996). The eIF4G-eIF4E complex is the target for direct cleavage by the rhinovirus 2A proteinase. *J. Virol.* 70, 8444–8450.
- Haimovich, G., Choder, M., Singer, R.H., and Treck, T. (2013). The fate of the messenger is pre-determined: A new model for regulation of gene expression. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1829, 643–653.
- Halic, M., Becker, T., Frank, J., Spahn, C.M., and Beckmann, R. (2005). Localization and dynamic behavior of ribosomal protein L30e. *Nat Struct Mol Biol* 12, 467–468.
- Hamm, J., Darzynkiewicz, E., Tahara, S.M., and Mattaj, I.W. (1990). The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. *Cell* 62, 569–577.

- Hashem, Y., Des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., Hellen, C.U.T., Pestova, T. V., and Frank, J. (2013). XStructure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* 153.
- Hastings, K.E.M. (2005). SL trans-splicing: Easy come or easy go? *Trends Genet.* 21, 240–247.
- Hatfield, D.L., and Gladyshev, V.N. (2002). How selenium has altered our understanding of the genetic code. *Mol. Cell. Biol.* 22, 3565–3576.
- Hatfield, D.L., Yoo, M.H., Carlson, B.A., and Gladyshev, V.N. (2009). Selenoproteins that function in cancer prevention and promotion. *Biochim Biophys Acta.*
- Hatfield, D.L., Tsuji, P.A., Carlson, B.A., and Gladyshev, V.N. (2014). Selenium and selenocysteine: Roles in cancer, health, and development. *Trends Biochem. Sci.* 39, 112–120.
- Hatin, I., Fabret, C., Rousset, J.P., and Namy, O. (2009). Molecular dissection of translation termination mechanism identifies two new critical regions in eRF1. *Nucleic Acids Res.* 37, 1789–1798.
- Hausmann, S., and Shuman, S. (2005). Specificity and mechanism of RNA cap guanine-N2 methyltransferase (Tgs1). *J. Biol. Chem.*
- Hayashi, M., Miyata, R., and Tanuma, N. (2012). Oxidative stress in developmental brain disorders. *Adv. Exp. Med. Biol.* 724, 278–290.
- He, W., and Parker, R. (2000). Functions of Lsm proteins in mRNA degradation and splicing. *Curr. Opin. Cell Biol.* 12, 346–350.
- Henis-Korenblit, S., Strumpf, N.L., Goldstaub, D., and Kimchi, A. (2000). A novel form of DAP5 protein accumulates in apoptotic cells as a result of caspase cleavage and internal ribosome entry site-mediated translation. *Mol. Cell. Biol.* 20, 496–506.
- Hentze, M.W. (1997). eIF4G: a multipurpose ribosome adapter? *Science* 275, 500–501.
- Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahms, H., and Lührmann, R. (1995). snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J.* 14, 2076–2088.
- Hernandez, N. (2001). Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. *J. Biol. Chem.* 276, 26733–26736.
- Hershey, J.W.B. (2010). Regulation of protein synthesis and the role of eIF3 in cancer. *Brazilian J. Med. Biol. Res.* 43, 920–930.
- Hesselberth, J.R. (2013). Lives that introns lead after splicing. *Wiley Interdiscip. Rev. RNA* 4, 677–691.
- Hill, K.E., Lloyd, R.S., Yang, J.G., Read, R., and Burk, R.F. (1991). The cDNA for rat selenoprotein P contains 10 TGA codons in the open reading frame. *J. Biol. Chem.* 266, 10050–10053.
- Hill, K.E., Lyons, P.R., and Burk, R.F. (1992). Differential regulation of rat liver selenoprotein mRNAs in selenium deficiency. *Biochem. Biophys. Res. Commun.* 185, 260–263.
- Hill, K.E., Lloyd, R.S., and Burk, R.F. (1993). Conserved features of selenoprotein P cDNA. *Biochem. Soc. Trans.* 21, 832–835.
- Hill, K.E., Zhou, J., McMahan, W.J., Motley, A.K., Atkins, J.F., Gesteland, R.F., and Burk, R.F. (2003). Deletion of selenoprotein P alters distribution of selenium in the mouse. *J. Biol. Chem.* 278, 13640–13646.
- Hinnebusch, A.G. (2014). The scanning mechanism of eukaryotic translation initiation. *Annu. Rev. Biochem.* 83, 779–812.
- Le Hir, H., and Séraphin, B. (2008). EJC at the Heart of Translational Control. *Cell* 133, 213–216.
- Hiremath, L.S., Hiremath, S.T., Rychlik, W., Joshi, S., Domier, L.L., and Rhoads, R.E. (1989). In vitro synthesis, phosphorylation, and localization on 48 S initiation complexes of human protein synthesis initiation factor 4E. *J. Biol. Chem.* 264, 1132–1138.
- Hirosawa-Takamori, M., Ossipov, D., Novoselov, S. V., Turanov, A.A., Zhang, Y., Gladyshev, V.N., Krol, A., Vorbruggen, G., and Jackle, H. (2009). A novel stem loop control element-dependent UGA read-through system without translational selenocysteine incorporation in Drosophila. *Faseb J* 23, 107–113.
- Hirose, Y., and Manley, J.L. (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev.* 14, 1415–1429.
- Ho, C.K., Sriskanda, V., McCracken, S., Bentley, D., Schwer, B., and Shuman, S. (1998). The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* 273, 9577–9585.
- Ho, M.-C., Wilczek, C., Bonanno, J.B., Xing, L., Seznec, J., Matsui, T., Carter, L.G., Onikubo, T., Kumar, P.R., Chan, M.K., et al. (2013). Structure of the Arginine Methyltransferase PRMT5-MEP50 Reveals a Mechanism for Substrate Specificity. *PLoS One* 8, e57008.
- Holcik, M., Lefebvre, C., Yeh, C., Chow, T., and Korneluk, R.G. (1999). A new internal-ribosome-entry-site motif potentiates XIAP-mediated cytoprotection. *Nat. Cell Biol.* 1, 190–192.
- Van Hoof, A., Lennertz, P., and Parker, R. (2000). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol. Cell. Biol.* 20, 441–452.
- Houseley, J., and Tollervey, D. (2009). The Many Pathways of RNA Degradation. *Cell* 136, 763–776.
- Howard, M.T., Aggarwal, G., Anderson, C.B., Khatri, S., Flanigan, K.M., and Atkins, J.F. (2005a). Recoding elements located adjacent to a subset of eukaryal selenocysteine-specifying UGA codons. *Embo J* 24, 1596–1607.
- Howard, M.T., Aggarwal, G., Anderson, C.B., Khatri, S., Flanigan, K.M., and Atkins, J.F. (2005b). Recoding elements located adjacent to a subset of eukaryal selenocysteine-specifying UGA codons. *EMBO J.* 24, 1596–1607.

- Howard, M.T., Moyle, M.W., Aggarwal, G., Carlson, B.A., and Anderson, C.B. (2007a). A recoding element that stimulates decoding of UGA codons by Sec tRNA[Ser]Sec. *RNA* 13, 912–920.
- Howard, M.T., Moyle, M.W., Aggarwal, G., Carlson, B.A., and Anderson, C.B. (2007b). A recoding element that stimulates decoding of UGA codons by Sec tRNA[Ser]Sec. *Rna* 13, 912–920.
- Howard, M.T., Carlson, B.A., Anderson, C.B., and Hatfield, D.L. (2013). Translational redefinition of UGA codons is regulated by selenium availability. *J. Biol. Chem.* 288, 19401–19413.
- Hu, J., Lutz, C.S., Wilusz, J., and Tian, B. (2005). Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA* 11, 1485–1493.
- Huang, L., and Lilley, D.M.J. (2013). The molecular recognition of kink-turn structure by the L7Ae class of proteins. *RNA* 19, 1703–1710.
- Huang, Q., and Pederson, T. (1999). A human U2 RNA mutant stalled in 3' end processing is impaired in nuclear import. *Nucleic Acids Res.* 27, 1025–1031.
- Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M., and Lührmann, R. (1998). Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *EMBO J.* 17, 4114–4126.
- Hubers, L., Valderrama-Carvajal, H., Laframboise, J., Timbers, J., Sanchez, G., and Côté, J. (2011). HuD interacts with survival motor neuron protein and can rescue spinal muscular atrophy-like neuronal defects. *Hum. Mol. Genet.* 20, 553–579.
- Hubert, N., Sturchler, C., Westhof, E., Carbon, P., and Krol, A. (1998). The 9/4 secondary structure of eukaryotic selenocysteine tRNA: more pieces of evidence. *Rna* 4, 1029–1033.
- 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.
- Hurst, R., Hooper, L., Norat, T., Lau, R., Aune, D., Greenwood, D.C., Vieira, R., Collings, R., Harvey, L.J., Sterne, J.A.C., et al. (2012). Selenium and prostate cancer: systematic review and meta-analysis. *Am. J. Clin. Nutr.* 96, 111–122.
- Hüttenhofer, A., and Böck, A. (1998). Selenocysteine inserting RNA elements modulate GTP hydrolysis of elongation factor SelB. *Biochemistry* 37, 885–890.
- Hwang, J., Sato, H., Tang, Y., Matsuda, D., and Maquat, L.E. (2010). UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps. *Mol. Cell* 39, 396–409.
- Ichimura, Y., Habuchi, T., Tsuchiya, N., Wang, L., Oyama, C., Sato, K., Nishiyama, H., Ogawa, O., and Kato, T. (2004). Increased risk of bladder cancer associated with a glutathione peroxidase 1 codon 198 variant. *J. Urol.* 172, 728–732.
- Imai, H., Hakkaku, N., Iwamoto, R., Suzuki, J., Suzuki, T., Tajima, Y., Konishi, K., Minami, S., Ichinose, S., Ishizaka, K., et al. (2009). Depletion of selenoprotein GPx4 in spermatocytes causes male infertility in mice. *J. Biol. Chem.* 284, 32522–32532.
- Ishii, T.M., Kotlova, N., Tapsoba, F., and Steinberg, S. V. (2013). The long D-stem of the selenocysteine tRNA provides resilience at the expense of maximal function. *J. Biol. Chem.* 288, 13337–13344.
- Itoh, Y., Chiba, S., Sekine, S.I., and Yokoyama, S. (2009a). Crystal structure of human selenocysteine tRNA. *Nucleic Acids Res.*
- Itoh, Y., Sekine, S. ichi, Matsumoto, E., Akasaka, R., Takemoto, C., Shirouzu, M., and Yokoyama, S. (2009b). Structure of Selenophosphate Synthetase Essential for Selenium Incorporation into Proteins and RNAs. *J. Mol. Biol.* 385, 1456–1469.
- Itoh, Y., Sekine, S.I., Suetsugu, S., and Yokoyama, S. (2013). Tertiary structure of bacterial selenocysteine tRNA. *Nucleic Acids Res.* 41, 6729–6738.
- Izaurralde, E. (2002). Nuclear export of messenger RNA. *Results Probl. Cell Differ.* 35, 133–150.
- Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, L.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* 78, 657–668.
- Jablonka, S., Rossoll, W., Schrank, B., and Sendtner, M. (2000). The role of SMN in spinal muscular atrophy. *J. Neurol.* 247 Suppl, I37–I42.
- Jablonska, E., Gromadzinska, J., Sobala, W., Reszka, E., and Wasowicz, W. (2008). Lung cancer risk associated with selenium status is modified in smoking individuals by Sep15 polymorphism. *Eur. J. Nutr.* 47, 47–54.
- Jackson, R.J. (2013). The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harb. Perspect. Biol.* 5.
- Jackson, R.J., and Kaminski, A. (1995). Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. *RNA* 1, 985–1000.
- Jackson, R.J., 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., Hellen, C.U.T., and Pestova, T. V (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* 11, 113–127.
- Jaeger, S., Barends, S., Giegé, R., Eriani, G., and Martin, F. (2005). Expression of metazoan replication-dependent histone genes. In *Biochimie*, pp. 827–834.
- Jaeger, S., Martin, F., Rudinger-Thirion, J., Giegé, R., and Eriani, G. (2006). Binding of human SLBP on the 3'-UTR of histone precursor H4-12 mRNA induces structural rearrangements that enable U7 snRNA anchoring. *Nucleic Acids Res.* 34, 4987–4995.
- Jan, E., Kinzy, T.G., and Sarnow, P. (2003). Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15410–15415.

- Jang, S.K., Kräusslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., and Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *J. Virol.* **62**, 2636–2643.
- Jean-Philippe, J., Paz, S., and Caputi, M. (2013). hnRNP A1: The Swiss Army Knife of gene expression. *Int. J. Mol. Sci.* **14**, 18999–19024.
- Jefferies, S., Kote-Jarai, Z., Goldgar, D., Houlston, R., Frazer-Williams, M.J., A'Hern, R., Henk, J., Gore, M., Rhys-Evans, P., Archer, D., et al. (2005). Association between polymorphisms of the GPX1 gene and second primary tumours after index squamous cell cancer of the head and neck. *Oral Oncol.* **41**, 455–461.
- Jenner, L., Melnikov, S., de Loubresse, N.G., Ben-Shem, A., Iskakova, M., Urzhumtsev, A., Meskauskas, A., Dinman, J., Yusupova, G., and Yusupov, M. (2012). Crystal structure of the 80S yeast ribosome. *Curr. Opin. Struct. Biol.* **22**, 759–767.
- De Jesus, L.A., Hoffmann, P.R., Michaud, T., Forry, E.P., Small-Howard, A., Stillwell, R.J., Morozova, N., Harney, J.W., and Berry, M.J. (2006). Nuclear Assembly of UGA Decoding Complexes on Selenoprotein mRNAs: a Mechanism for Eluding Nonsense-Mediated Decay? *Mol Cell Biol* **26**, 1795–1805.
- 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.
- Johansson, L., Gafvelin, G., and Arnér, E.S.J. (2005). Selenocysteine in proteins-properties and biotechnological use. *Biochim. Biophys. Acta* **1726**, 1–13.
- Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3**, 415–428.
- Joshi, B., Cameron, A., and Jagus, R. (2004). Characterization of mammalian eIF4E-family members. *Eur. J. Biochem.* **271**, 2189–2203.
- Joshi, B., Lee, K., Maeder, D.L., and Jagus, R. (2005). Phylogenetic analysis of eIF4E-family members. *BMC Evol. Biol.* **5**, 48.
- Kaiser, J.T., Gromadski, K., Rother, M., Engelhardt, H., Rodnina, M. V, and Wahl, M.C. (2005). Structural and functional investigation of a putative archaeal selenocysteine synthase. *Biochemistry* **44**, 13315–13327.
- Kakihara, Y., and Houry, W.A. (2012). The R2TP complex: Discovery and functions. *Biochim. Biophys. Acta - Mol. Cell Res.* **1823**, 101–107.
- Kambach, C., Walke, S., and Nagai, K. (1999). Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. *Curr. Opin. Struct. Biol.* **9**, 222–230.
- 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.
- Kapeli, K., and Yeo, G.W. (2012). Genome-Wide Approaches to Dissect the Roles of RNA Binding Proteins in Translational Control: Implications for Neurological Diseases. *Front. Neurosci.* **6**.
- Karkhanis, V., Hu, Y.J., Baiocchi, R.A., Imbalzano, A.N., and Sif, S. (2011). Versatility of PRMT5-induced methylation in growth control and development. *Trends Biochem. Sci.* **36**, 633–641.
- Katahira, J. (2015). Nuclear Export of Messenger RNA. *Genes (Basel)*. **6**, 163–184.
- Keiper, B.D., Lamphear, B.J., Deshpande, A.M., Jankowska-Anyszka, M., Aamodt, E.J., Blumenthal, T., and Rhoads, R.E. (2000). Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans*. *J. Biol. Chem.* **275**, 10590–10596.
- Khusial, P., Plaag, R., and Zieve, G.W. (2005). LSm proteins form heptameric rings that bind to RNA via repeating motifs. *Trends Biochem. Sci.* **30**, 522–528.
- Kieft, J.S., Zhou, K., Jubin, R., and Doudna, J.A. (2001). Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* **7**, 194–206.
- Kim, H.Y., Fomenko, D.E., Yoon, Y.E., and Gladyshev, V.N. (2006). Catalytic advantages provided by selenocysteine in methionine-S-sulfoxide reductases. *Biochemistry* **45**, 13697–13704.
- Kim, M.-J., Lee, B.C., Hwang, K.Y., Gladyshev, V.N., and Kim, H.-Y. (2015). Selenium utilization in thioredoxin and catalytic advantage provided by selenocysteine. *Biochem. Biophys. Res. Commun.* **461**, 648–652.
- King, H.A., Cobbold, L.C., and Willis, A.E. (2010). The role of IRES trans-acting factors in regulating translation initiation. *Biochem. Soc. Trans.* **38**, 1581–1586.
- Kinzy, S.A., Caban, K., and Copeland, P.R. (2005). Characterization of the SECIS binding protein 2 complex required for the co-translational insertion of selenocysteine in mammals. *Nucleic Acids Res* **33**, 5172–5180.
- Kiss, T. (2006). SnoRNP Biogenesis Meets Pre-mRNA Splicing. *Mol. Cell* **23**, 775–776.
- Kiss, A.M., Jády, B.E., Bertrand, E., and Kiss, T. (2004). Human box H/ACA pseudouridylation guide RNA machinery. *Mol. Cell. Biol.* **24**, 5797–5807.
- Kiss, T., Fayet, E., Jády, B.E., Richard, P., and Weber, M. (2006). Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs. In *Cold Spring Harbor Symposia on Quantitative Biology*, pp. 407–417.
- Kiss, T., Fayet-Lebaron, E., and Jády, B.E. (2010). Box H/ACA Small Ribonucleoproteins. *Mol. Cell* **37**, 597–606.
- Kiss-László, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA: A novel function for small nucleolar RNAs. *Cell* **85**, 1077–1088.
- Kiss-László, Z., Henry, Y., and Kiss, T. (1998). Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J.* **17**, 797–807.
- Klein, D.J., Schmeing, T.M., Moore, P.B., and Steitz, T.A. (2001). The kink-turn: a new RNA secondary structure motif. *Embo J* **20**, 4214–4221.

- Knight, J.A., Onay, U.V., Wells, S., Li, H., Shi, E.J.Q., Andrulis, I.L., and Ozelik, H. (2004). Genetic Variants of GPX1 and SOD2 and Breast Cancer Risk at the Ontario Site of the Breast Cancer Family Registry. *Cancer Epidemiol. Biomarkers Prev.* 13, 146–149.
- Köhler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.* 8, 761–773.
- Köhrle, J. (2013). Selenium and the thyroid. *Curr. Opin. Endocrinol. Diabetes. Obes.* 20, 441–448.
- Kondo, N., Nakamura, H., Masutani, H., and Yodoi, J. Redox regulation of human thioredoxin network. *Antioxid. Redox Signal.* 8, 1881–1890.
- Korotkov, K. V., Novoselov, S. V., Hatfield, D.L., and Gladyshev, V.N. (2002a). Mammalian selenoprotein in which selenocysteine (Sec) incorporation is supported by a new form of Sec insertion sequence element. *Mol. Cell Biol.* 22, 1402–1411.
- Korotkov, K. V., Novoselov, S. V., Hatfield, D.L., and Gladyshev, V.N. (2002b). Mammalian selenoprotein in which selenocysteine (Sec) incorporation is supported by a new form of Sec insertion sequence element. *Mol. Cell Biol.* 22, 1402–1411.
- Kossinova, O., Malygin, A., Krol, A., and Karpova, G. (2013). A novel insight into the mechanism of mammalian selenoprotein synthesis. *RNA*.
- Kossinova, O., Malygin, A., Krol, A., and Karpova, G. (2014). The SBP2 protein central to selenoprotein synthesis contacts the human ribosome at expansion segment 7L of the 28S rRNA. *RNA*.
- Kote-Jarai, Z., Durocher, F., Edwards, S.M., Hamoudi, R., Jackson, R.A., Ardern-Jones, A., Murkin, A., Dearnaley, D.P., Kirby, R., Houlston, R., et al. (2002). Association between the GCG polymorphism of the selenium dependent GPX1 gene and the risk of young onset prostate cancer. *Prostate Cancer Prostatic Dis.* 5, 189–192.
- Kotini, S.B., Peske, F., and Rodnina, M. V (2015). Partitioning between recoding and termination at a stop codon-selenocysteine insertion sequence. *Nucleic Acids Res.*
- 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.
- Kraus, R.J., Prohaska, J.R., and Ganther, H.E. (1980). Oxidized forms of ovine erythrocyte glutathione peroxidase. Cyanide inhibition of a 4-glutathione:4-selenoenzyme. *Biochim. Biophys. Acta* 615, 19–26.
- Kroiss, M., Schultz, J., Wiesner, J., Chari, A., Sickmann, A., and Fischer, U. (2008). Evolution of an RNP assembly system: a minimal SMN complex facilitates formation of UsnRNPs in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 10045–10050.
- Krol, A. (2002). Evolutionarily different RNA motifs and RNA-protein complexes to achieve selenoprotein synthesis. *Biochimie* 84, 765–774.
- Kryukov, G. V., Kryukov, V.M., and Gladyshev, V.N. (1999). New mammalian selenocysteine-containing proteins identified with an algorithm that searches for selenocysteine insertion sequence elements. *J Biol Chem* 274, 33888–33897.
- Kryukov, G. V., Kumar, R.A., Koc, A., Sun, Z., and Gladyshev, V.N. (2002). Selenoprotein R is a zinc-containing stereo-specific methionine sulfoxide reductase. *Proc. Natl. Acad. Sci. U. S. A.* 99, 4245–4250.
- Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R., and Gladyshev, V.N. (2003a). Characterization of mammalian selenoproteomes. *Science* (80-.). 300, 1439–1443.
- Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigó, R., and Gladyshev, V.N. (2003b). Characterization of mammalian selenoproteomes. *Science* 300, 1439–1443.
- Kufel, J., Allmang, C., Verdone, L., Beggs, J., and Tollervey, D. (2003). A complex pathway for 3' processing of the yeast U3 snoRNA. *Nucleic Acids Res* 31, 6788–6797.
- Kuhn, J.F., Tran, E.J., and Maxwell, E.S. (2002). Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kd/Snu13p snoRNP core protein. *Nucleic Acids Res.* 30, 931–941.
- Kühn-Hölsken, E., Lenz, C., Dickmanns, A., Hsiao, H.H., Richter, F.M., Kastner, B., Ficner, R., and Urlaub, H. (2010). Mapping the binding site of snurportin 1 on native u1 snRNP by cross-linking and mass spectrometry. *Nucleic Acids Res.* 38, 5581–5593.
- Kunkel, G.R., Maser, R.L., Calvet, J.P., and Pederson, T. (1986). U6 small nuclear RNA is transcribed by RNA polymerase III. *Proc. Natl. Acad. Sci. U. S. A.* 83, 8575–8579.
- Labunskyy, V.M., Hatfield, D.L., and Gladyshev, V.N. (2007). The Sep15 protein family: roles in disulfide bond formation and quality control in the endoplasmic reticulum. *IUBMB Life* 59, 1–5.
- Labunskyy, V.M., Yoo, M.H., Hatfield, D.L., and Gladyshev, V.N. (2009). Sep15, a thioredoxin-like selenoprotein, is involved in the unfolded protein response and differentially regulated by adaptive and acute ER stresses. *Biochemistry* 48, 8458–8465.
- Labunskyy, V.M., Hatfield, D.L., and Gladyshev, V.N. (2014). Selenoproteins: Molecular Pathways and Physiological Roles. *Physiol. Rev.* 94, 739–777.
- Ladenstein, R., Epp, O., Bartels, K., Jones, A., Huber, R., and Wendel, A. (1979). Structure analysis and molecular model of the selenoenzyme glutathione peroxidase at 2.8 Å resolution. *J. Mol. Biol.* 134, 199–218.
- Lafontaine, D.L., and Tollervey, D. (1999). Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. *RNA* 5, 455–467.
- Lafontaine, D.L., and Tollervey, D. (2000). Synthesis and assembly of the box C+D small nucleolar RNPs. *Mol. Cell Biol.* 20, 2650–2659.

- Lai, M.-C., Lee, Y.-H.W., and Tarn, W.-Y. (2008). The DEAD-box RNA helicase DDX3 associates with export messenger ribonucleoproteins as well as tip-associated protein and participates in translational control. *Mol. Biol. Cell* 19, 3847–3858.
- Lall, S., Friedman, C.C., Jankowska-Anyszka, M., Stepinski, J., Darzynkiewicz, E., and Davis, R.E. (2004). Contribution of trans-splicing, 5'-leader length, cap-poly(A) synergism, and initiation factors to nematode translation in an *Ascaris suum* embryo cell-free system. *J. Biol. Chem.* 279, 45573–45585.
- Lamond, A.I., and Sleeman, J.E. (2003). Nuclear substructure and dynamics. *Curr. Biol.* 13, R825–R828.
- Lamphear, B.J., Kirchweiger, 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.
- Latreche, L., Jean-Jean, O., Driscoll, D.M., and Chavatte, L. (2009). Novel structural determinants in human SECIS elements modulate the translational recoding of UGA as selenocysteine. *Nucleic Acids Res.*
- Latrèche, L., Duhieu, S., Touat-Hamici, Z., Jean-Jean, O., and Chavatte, L. (2012). The differential expression of glutathione peroxidase 1 and 4 depends on the nature of the SECIS element. *RNA Biol.* 9, 681–690.
- Lau, C., Bachorik, J.L., and Dreyfuss, G. (2009). Gemin5-snRNA interaction reveals an RNA binding function for WD repeat domains. *Nat. Struct. & Mol. Biol.* 16, 486–491.
- Lee, K.H., and Jeong, D. (2012). Bimodal actions of selenium essential for antioxidant and toxic pro-oxidant activities: The selenium paradox (Review). *Mol. Med. Rep.* 5, 299–304.
- Lee, B., Péterfi, Z., Hoffmann, F.W., Moore, R.E., Kaya, A., Avanesov, A., Tarrago, L., Zhou, Y., Weerapana, E., Fomenko, D.E., et al. (2013). MsrB1 and MICALs Regulate Actin Assembly and Macrophage Function via Reversible Stereoselective Methionine Oxidation. *Mol. Cell* 51, 397–404.
- Lee, B.J., Worland, P.J., Davis, J.N., Stadtman, T.C., and Hatfield, D.L. (1989). Identification of a selenocysteyl-tRNA(Ser) in mammalian cells that recognizes the nonsense codon, UGA. *J. Biol. Chem.* 264, 9724–9727.
- Lee, C.H., Lee, K.Y., Choe, K.H., Hong, Y.C., Noh, S. Il, Eom, S.Y., Ko, Y.J., Zhang, Y.W., Yim, D.H., Kangu, J.W., et al. (2006). Effects of Oxidative DNA damage and genetic polymorphism of the Glutathione Peroxidase 1 (GPX1) and 8-Oxoguanine Glycosylase 1 (hOGG1) on lung cancer. *J. Prev. Med. Public Heal.* 39, 130–134.
- Lee, E.-H., Myung, S.-K., Jeon, Y.-J., Kim, Y., Chang, Y.J., Ju, W., Seo, H.G., and Huh, B.Y. (2011). Effects of Selenium Supplements on Cancer Prevention: Meta-analysis of Randomized Controlled Trials. *Nutr. Cancer* 63, 1185–1195.
- 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.
- Lee, S.R., Bar-Noy, S., Kwon, J., Levine, R.L., Stadtman, T.C., and Rhee, S.G. (2000). Mammalian thioredoxin reductase: oxidation of the C-terminal cysteine/selenocysteine active site forms a thioselenide, and replacement of selenium with sulfur markedly reduces catalytic activity. *Proc. Natl. Acad. Sci. U. S. A.* 97, 2521–2526.
- Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., and Zeviani, M. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80, 155–165.
- Lefebvre, S., Burlet, P., Viollet, L., Bertrand, S., Huber, C., Belser, C., and Munnich, A. (2002). A novel association of the SMN protein with two major non-ribosomal nucleolar proteins and its implication in spinal muscular atrophy. *Hum. Mol. Genet.* 11, 1017–1027.
- Lei, X.G., Evenson, J.K., Thompson, K.M., and Sunde, R.A. (1995). Glutathione peroxidase and phospholipid hydroperoxide glutathione peroxidase are differentially regulated in rats by dietary selenium. *J. Nutr.* 125, 1438–1446.
- Leibundgut, M., Frick, C., Thanbichler, M., Bock, A., and Ban, N. (2005). Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors. *Embo J* 24, 11–22.
- Lemm, I., Girard, C., Kuhn, A.N., Watkins, N.J., Schneider, M., Bordonné, R., and Lüthmann, R. (2006). Ongoing U snRNP biogenesis is required for the integrity of Cajal bodies. *Mol. Biol. Cell* 17, 3221–3231.
- Lescoute, A., Leontis, N.B., Massire, C., and Westhof, E. (2005). Recurrent structural RNA motifs, Isostericity Matrices and sequence alignments. *Nucleic Acids Res* 33, 2395–2409.
- Lescure, A., Gautheret, D., Carbon, P., and Krol, A. (1999). Novel selenoproteins identified in silico and in vivo by using a conserved RNA structural motif. *J Biol Chem* 274, 38147–38154.
- Lescure, A., Gautheret, D., and Krol, A. (2002a). Novel selenoproteins identified from genomic sequence data. *Methods Enzym.* 347, 57–70.
- Lescure, A., Allmang, C., Yamada, K., Carbon, P., and Krol, A. (2002b). cDNA cloning, expression pattern and RNA binding analysis of human selenocysteine insertion sequence (SECIS) binding protein 2. *Gene* 291, 279–285.
- Lescure, A., Fagegaltier, D., Carbon, P., and Krol, A. (2002c). Protein factors mediating selenoprotein synthesis. *Curr Protein Pept Sci* 3, 143–151.
- Lescure, A., Allmang, C., Yamada, K., Carbon, P., and Krol, A. (2002d). cDNA cloning, expression pattern and RNA binding analysis of human selenocysteine insertion sequence (SECIS) binding protein 2. *Gene* 291, 279–285.
- Lescure, A., Deniziak, M., Rederstorff, M., and Krol, A. (2008). Molecular basis for the role of selenium in muscle development and function. *Chem Biodivers* 5, 408–413.

- Lewis, J.D., and Izaurralde, E. (1997). The role of the cap structure in RNA processing and nuclear export. *Eur. J. Biochem.* 247, 461–469.
- Li, L., and Wang, C.C. (2004). Capped mRNA with a Single Nucleotide Leader Is Optimally Translated in a Primitive Eukaryote, *Giardia lamblia*. *J. Biol. Chem.* 279, 14656–14664.
- Li, D.K., Tisdale, S., Lotti, F., and Pellizzoni, L. (2014a). SMN control of RNP assembly: From post-transcriptional gene regulation to motor neuron disease. *Semin. Cell Dev. Biol.* 32, 22–29.
- Li, D.K., Tisdale, S., Lotti, F., and Pellizzoni, L. (2014b). SMN control of RNP assembly: From post-transcriptional gene regulation to motor neuron disease. *Semin. Cell Dev. Biol.* 32, 22–29.
- Li, Q., Imataka, H., Morino, S., Rogers, G.W., Richter-Cook, N.J., Merrick, W.C., and Sonenberg, N. (1999). Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol. Cell. Biol.* 19, 7336–7346.
- Lightfoot, T.J., Skibola, C.F., Smith, A.G., Forrest, M.S., Adamson, P.J., Morgan, G.J., Bracci, P.M., Roman, E., Smith, M.T., and Holly, E.A. (2006). Polymorphisms in the oxidative stress genes, superoxide dismutase, glutathione peroxidase and catalase and risk of non-Hodgkin's lymphoma. *Haematologica* 91, 1222–1227.
- Liou, R.F., and Blumenthal, T. (1990). trans-spliced *Caenorhabditis elegans* mRNAs retain trimethylguanosine caps. *Mol. Cell. Biol.* 10, 1764–1768.
- Liu, Q., and Dreyfuss, G. (1996). A novel nuclear structure containing the survival of motor neurons protein. *EMBO J.* 15, 3555–3565.
- Liu, J., Henao-Mejia, J., Liu, H., Zhao, Y., and He, J.J. (2011). Translational regulation of HIV-1 replication by HIV-1 rev cellular cofactors Sam68, eIF5A, hRIP, and DDX3. *J. Neuroimmune Pharmacol.* 6, 308–321.
- Lobanov, A. V., Hatfield, D.L., and Gladyshev, V.N. (2008). Selenoproteinless animals: selenophosphate synthetase SPS1 functions in a pathway unrelated to selenocysteine biosynthesis. *Protein Sci* 17, 176–182.
- Lobanov, A. V., Hatfield, D.L., and Gladyshev, V.N. (2009). Eukaryotic selenoproteins and selenoproteomes. *Biochim Biophys Acta*.
- Locker, N., Easton, L.E., and Lukavsky, P.J. (2006). Affinity purification of eukaryotic 48S initiation complexes. *RNA* 12, 683–690.
- 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.
- Lodygin, D., Epanchintsev, A., Menssen, A., Diebold, J., and Hermeking, H. (2005). Functional epigenomics identifies genes frequently silenced in prostate cancer. *Cancer Res.* 65, 4218–4227.
- Lorsch, J.R., and Herschlag, D. (1999). Kinetic dissection of fundamental processes of eukaryotic translation initiation in vitro. *EMBO J.* 18, 6705–6717.
- De Loubresse, N.G., Prokhorova, I., Holtkamp, W., Rodnina, M. V., Yusupova, G., and Yusupov, M. (2014). Structural basis for the inhibition of the eukaryotic ribosome. *Nature*.
- Low, S.C., Harney, J.W., and Berry, M.J. (1995). Cloning and Functional Characterization of Human Selenophosphate Synthetase, an Essential Component of Selenoprotein Synthesis. *J. Biol. Chem.* 270, 21659–21664.
- Low, S.C., Grundner-Culemann, E., Harney, J.W., and Berry, M.J. (2000a). SECIS-SBP2 interactions dictate selenocysteine incorporation efficiency and selenoprotein hierarchy. *Embo J* 19, 6882–6890.
- Low, S.C., Grundner-Culemann, E., Harney, J.W., and Berry, M.J. (2000b). SECIS-SBP2 interactions dictate selenocysteine incorporation efficiency and selenoprotein hierarchy. *EMBO J.* 19, 6882–6890.
- Lozano, G., and Martínez-Salas, E. (2015). Structural insights into viral IRES-dependent translation mechanisms. *Curr. Opin. Virol.* 12, 113–120.
- Lu, J., and Holmgren, A. (2009). Selenoproteins. *J Biol Chem* 284, 723–727.
- Lubos, E., Loscalzo, J., and Handy, D.E. (2011). Glutathione peroxidase-1 in health and disease: from molecular mechanisms to therapeutic opportunities. *Antioxid. Redox Signal.* 15, 1957–1997.
- Lui, L., and Lowe, T. (2013). Small nucleolar RNAs and RNA-guided post-transcriptional modification. *Essays Biochem.* 54, 53–77.
- Macejak, D.G., and Sarnow, P. (1991). Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature* 353, 90–94.
- Macías, S., Bragulat, M., Tardiff, D.F., and Vilardell, J. (2008). L30 Binds the Nascent RPL30 Transcript to Repress U2 snRNP Recruitment. *Mol. Cell* 30, 732–742.
- Mader, S., Lee, H., Pause, A., and Sonenberg, N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol. Cell. Biol.* 15, 4990–4997.
- Mahboobi, S.H., Javanpour, A.A., and Mofrad, M.R.K. (2015). The interaction of RNA helicase DDX3 with HIV-1 Rev-CRM1-RanGTP complex during the HIV replication cycle. *PLoS One* 10, e0112969.
- Mailman, M.D., Heinz, J.W., Papp, A.C., Snyder, P.J., Sedra, M.S., Wirth, B., Burghes, A.H.M., and Prior, T.W. (2002). Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. *Genet. Med.* 4, 20–26.
- Maiti, B., Arbogast, S., Allamand, V., Moyle, M.W., Anderson, C.B., Richard, P., Guicheney, P., Ferreira, A., Flanigan, K.M., and Howard, M.T. (2008). A mutation in the SEPNI1 selenocysteine redefinition element (SRE) reduces selenocysteine incorporation and leads to SEPNI1 related myopathy. *Hum. Mutat. doi: 10.10*
- 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.

- Malys, N., and McCarthy, J.E.G. (2011). Translation initiation: Variations in the mechanism can be anticipated. *Cell. Mol. Life Sci.* 68, 991–1003.
- Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci.* 65, 1099–1122.
- Maquat, L.E., Hwang, J., Sato, H., and Tang, Y. (2010a). CBP80-promoted mRNP rearrangements during the pioneer round of translation, nonsense-mediated mRNA decay, and thereafter. *Cold Spring Harb. Symp. Quant. Biol.* 75, 127–134.
- Maquat, L.E., Tarn, W.Y., and Isken, O. (2010b). The pioneer round of translation: Features and functions. *Cell* 142, 368–374.
- Mariotti, M., Lobanov, A. V., Guigo, R., and Gladyshev, V.N. (2013). SECISearch3 and Sebastian: New tools for prediction of SECIS elements and selenoproteins. *Nucleic Acids Res.* 41.
- Marmier-Gourrier, N., Clery, A., Senty-Segault, V., Charpentier, B., Schlotter, F., Leclerc, F., Fournier, R., and Brantant, C. (2003). A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA. *Rna* 9, 821–838.
- Maroney, P.A., Denker, J.A., Darzynkiewicz, E., Laneve, R., and Nilsen, T.W. (1995). Most mRNAs in the nematode *Ascaris lumbricoides* are trans-spliced: a role for spliced leader addition in translational efficiency. *RNA* 1, 714–723.
- Martin, F., Schaller, A., Eglite, S., Schümperli, D., and Müller, B. (1997). The gene for histone RNA hairpin binding protein is located on human chromosome 4 and encodes a novel type of RNA binding protein. *EMBO J.* 16, 769–778.
- Martin, F., Barends, S., Jaeger, S., Schaeffer, L., Prongidi-Fix, L., and Eriani, G. (2011). Cap-Assisted Internal Initiation of Translation of Histone H4. *Mol. Cell* 41, 197–209.
- Martin, G.W., Harney, J.W., and Berry, M.J. (1996). Selenocysteine incorporation in eukaryotes: insights into mechanism and efficiency from sequence, structure, and spacing proximity studies of the type 1 deiodinase SECIS element. *RNA* 2, 171–182.
- Martin, G.W., Harney, J.W., and Berry, M.J. (1998). Functionality of mutations at conserved nucleotides in eukaryotic SECIS elements is determined by the identity of a single nonconserved nucleotide. *RNA* 4, 65–73.
- Martin, R., Gupta, K., Ninan, N.S., Perry, K., and Van Duyne, G.D. (2012). The survival motor neuron protein forms soluble glycine zipper oligomers. *Structure* 20, 1929–1939.
- Massenet, S., Pellizzoni, L., Paushkin, S., Mattaj, I.W., and Dreyfuss, G. (2002). The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol. Cell. Biol.* 22, 6533–6541.
- Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* 15, 108–121.
- Matera, A.G., Terns, R.M., and Terns, M.P. (2007a). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* 8, 209–220.
- Matera, A.G., Terns, R.M., and Terns, M.P. (2007b). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.* 8, 209–220.
- Matsumoto, J., Dewar, K., Wasserscheid, J., Wiley, G.B., Macmil, S.L., Roe, B.A., Zeller, R.W., Satou, Y., and Hastings, K.E.M. (2010). High-throughput sequence analysis of *Ciona intestinalis* SL trans-spliced mRNAs: Alternative expression modes and gene function correlates. *Genome Res.* 20, 636–645.
- Mattaj, I.W. (1986). Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell* 46, 905–911.
- Mazza, C., Ohno, M., Segref, A., Mattaj, I.W., and Cusack, S. (2001). Crystal structure of the human nuclear cap binding complex. *Mol. Cell* 8, 383–396.
- McCormack, J.C., Yuan, X., Yingling, Y.G., Kasprzak, W., Zamora, R.E., Shapiro, B.A., and Simon, A.E. (2008). Structural domains within the 3' untranslated region of Turnip crinkle virus. *J. Virol.* 82, 8706–8720.
- McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D.L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* 11, 3306–3318.
- Meerovitch, K., Svitkin, Y. V., Lee, H.S., Lejbkovicz, 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.
- Mehta, A., Rebsch, C.M., Kinzy, S.A., Fletcher, J.E., and Copeland, P.R. (2004a). Efficiency of mammalian selenocysteine incorporation. *J. Biol. Chem.* 279, 37852–37859.
- Mehta, A., Rebsch, C.M., Kinzy, S.A., Fletcher, J.E., and Copeland, P.R. (2004b). Efficiency of mammalian selenocysteine incorporation. *J Biol Chem* 279, 37852–37859.
- Meister, G., and Fischer, U. (2002). Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *EMBO J.* 21, 5853–5863.
- Meister, G., Eggert, C., Bühler, D., Brahms, H., Kambach, C., and Fischer, U. (2001a). Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Curr. Biol.* 11, 1990–1994.
- Meister, G., Bühler, D., Pillai, R., Lottspeich, F., and Fischer, U. (2001b). A multiprotein complex mediates the ATP-dependent assembly of spliceosomal U snRNPs. *Nat. Cell Biol.* 3, 945–949.
- Meister, G., Eggert, C., and Fischer, U. (2002). SMN-mediated assembly of RNPs: A complex story. *Trends Cell Biol.* 12, 472–478.
- Méplan, C., Crosley, L.K., Nicol, F., Beckett, G.J., Howie, A.F., Hill, K.E., Horgan, G., Mathers, J.C., Arthur, J.R., and Hesketh, J.E. (2007). Genetic polymorphisms in the human selenoprotein P gene determine the

- response of selenoprotein markers to selenium supplementation in a gender-specific manner (the SELGEN study). *FASEB J.* 21, 3063–3074.
- Méplán, C., Crosley, L.K., Nicol, F., Horgan, G.W., Mathers, J.C., Arthur, J.R., and Hesketh, J.E. (2008). Functional effects of a common single-nucleotide polymorphism (GPX4c718t) in the glutathione peroxidase 4 gene: interaction with sex. *Am. J. Clin. Nutr.* 87, 1019–1027.
- Méplán, C., Hughes, D.J., Pardini, B., Naccarati, A., Soucek, P., Vodickova, L., Hlavatá, I., Vrána, D., Vodicka, P., and Hesketh, J.E. (2010). Genetic variants in selenoprotein genes increase risk of colorectal cancer. *Carcinogenesis* 31, 1074–1079.
- Merino, E.J., Wilkinson, K.A., Coughlan, J.L., and Weeks, K.M. (2005). RNA structure analysis at single nucleotide resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). *J. Am. Chem. Soc.* 127, 4223–4231.
- Miniard, A.C., Middleton, L.M., Budiman, M.E., Gerber, C.A., and Driscoll, D.M. (2010). Nucleolin binds to a subset of selenoprotein mRNAs and regulates their expression. *Nucleic Acids Res.* 38, 4807–4820.
- Misra, S., Boylan, M., Selvam, A., Spallholz, J., and Björnstedt, M. (2015). Redox-Active Selenium Compounds—From Toxicity and Cell Death to Cancer Treatment. *Nutrients* 7, 3536–3556.
- Mitchell, P., and Tollervey, D. (2001). mRNA turnover. *Curr. Opin. Cell Biol.* 13, 320–325.
- Mix, H., Lobanov, A. V., and Gladyshev, V.N. (2007). SECIS elements in the coding regions of selenoprotein transcripts are functional in higher eukaryotes. *Nucleic Acids Res.* 35, 414–423.
- Miyoshi, H., Dwyer, D.S., Keiper, B.D., Jankowska-Anyszka, M., Darzynkiewicz, E., and Rhoads, R.E. (2002). Discrimination between mono- and trimethylated cap structures by two isoforms of *Caenorhabditis elegans* eIF4E. *EMBO J.* 21, 4680–4690.
- Mizutani, T., and Fujiwara, T. (2000). SBP, SECIS binding protein, binds to the RNA fragment upstream of the Sec UGA codon in glutathione peroxidase mRNA. *Mol Biol Rep* 27, 99–105.
- Moazed, D., and Noller, H.F. (1989). Intermediate states in the movement of transfer RNA in the ribosome. *Nature* 342, 142–148.
- Moghadaszadeh, B., Petit, N., Jaillard, C., Brockington, M., Roy, S.Q., Merlini, L., Romero, N., Estournet, B., Desguerre, I., Chaigne, D., et al. (2001). Mutations in SEPNI cause congenital muscular dystrophy with spinal rigidity and restrictive respiratory syndrome. *Nat Genet* 29, 17–18.
- Moll, I., Grill, S., Gualerzi, C.O., and Bläsi, U. (2002). Leaderless mRNAs in bacteria: Surprises in ribosomal recruitment and translational control. *Mol. Microbiol.* 43, 239–246.
- Monani, U.R., Coovert, D.D., and Burghes, A.H. (2000). Animal models of spinal muscular atrophy. *Hum. Mol. Genet.* 9, 2451–2457.
- Monecke, T., Dickmanns, A., and Ficner, R. (2009). Structural basis for m7G-cap hypermethylation of small nuclear, small nucleolar and telomerase RNA by the dimethyltransferase TGS1. *Nucleic Acids Res.* 37, 3865–3877.
- De Moor, C.H., Meijer, H., and Lissenden, S. (2005). Mechanisms of translational control by the 3'UTR in development and differentiation. *Semin. Cell Dev. Biol.* 16, 49–58.
- Moore, T., Zhang, Y., Fenley, M.O., and Li, H. (2004). Molecular basis of box C/D RNA-protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Struct.* 12, 807–818.
- Moran, S.J., Flanagan, J.F., Namy, O., Stuart, D.I., Brierley, I., and Gilbert, R.J.C. (2008). The Mechanics of Translocation: A Molecular “Spring-and-Ratchet” System. *Structure* 16, 664–672.
- Moriarty, P.M., Reddy, C.C., and Maquat, L.E. (1998). Selenium deficiency reduces the abundance of mRNA for Se-dependent glutathione peroxidase 1 by a UGA-dependent mechanism likely to be nonsense codon-mediated decay of cytoplasmic mRNA. *Mol. Cell Biol.* 18, 2932–2939.
- Morita, M., Ler, L.W., Fabian, M.R., Siddiqui, N., Mullin, M., Henderson, V.C., Alain, T., Fonseca, B.D., Karashchuk, G., Bennett, C.F., et al. (2012). A Novel 4EHP-GIGYF2 Translational Repressor Complex Is Essential for Mammalian Development. *Mol. Cell Biol.* 32, 3585–3593.
- Morley, S.J., and Willett, M. (2009). Kinky Binding and SECsy Insertions. *Mol. Cell* 35, 396–398.
- Morris, G.E. (2008). The Cajal body. *Biochim. Biophys. Acta - Mol. Cell Res.* 1783, 2108–2115.
- Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., and Bordonné, R. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell* 9, 891–901.
- Mouaikel, J., Bujnicki, J.M., Tazi, J., and Bordonné, R. (2003a). Sequence-structure-function relationships of Tgs1, the yeast snRNA/snoRNA cap hypermethylase. *Nucleic Acids Res.* 31, 4899–4909.
- Mouaikel, J., Narayanan, U., Verheggen, C., Matera, A.G., Bertrand, E., Tazi, J., and Bordonné, R. (2003b). Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Rep.* 4, 616–622.
- Muhs, M., Hilal, T., Mielke, T., Skabkin, M.A., Sanbonmatsu, K.Y., Pestova, T. V., and Spahn, C.M.T. (2015a). Cryo-EM of Ribosomal 80S Complexes with Termination Factors Reveals the Translocated Cricket Paralysis Virus IRES. *Mol. Cell* 57, 422–432.
- Muhs, M., Hilal, T., Mielke, T., Skabkin, M.A., Sanbonmatsu, K.Y., Pestova, T. V., and Spahn, C.M.T. (2015b). Cryo-EM of ribosomal 80S complexes with termination factors reveals the translocated cricket paralysis virus IRES. *Mol. Cell* 57, 422–432.
- Mukherjee, A., and Martin, S.G. (2008). The thioredoxin system: A key target in tumour and endothelial cells. In *British Journal of Radiology*.
- Mustacich, D., and Powis, G. (2000). Thioredoxin reductase. *Biochem. J.* 346 Pt 1, 1–8.
- Muttenthaler, M., and Alewood, P.F. (2008). Selenopeptide chemistry. *J Pept Sci* 14, 1223–1239.

- Nagy, E., and Maquat, L.E. (1998). A rule for termination-codon position within intron-containing genes: When nonsense affects RNA abundance. *Trends Biochem. Sci.* 23, 198–199.
- Nano, N., and Houry, W.A. (2013). Chaperone-like activity of the AAA+ proteins Rvb1 and Rvb2 in the assembly of various complexes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 368, 20110399.
- Narayanan, U., Ospina, J.K., Frey, M.R., Hebert, M.D., and Matera, A.G. (2002). SMN, the spinal muscular atrophy protein, forms a pre-import snRNP complex with snurportin1 and importin beta. *Hum. Mol. Genet.* 11, 1785–1795.
- Narita, T., Yung, T.M.C., Yamamoto, J., Tsuboi, Y., Tanabe, H., Tanaka, K., Yamaguchi, Y., and Handa, H. (2007). NELF Interacts with CBC and Participates in 3' End Processing of Replication-Dependent Histone mRNAs. *Mol. Cell* 26, 349–365.
- Neuenkirchen, N., Chari, A., and Fischer, U. (2008). Deciphering the assembly pathway of Sm-class U snRNPs. *FEBS Lett.* 582, 1997–2003.
- Neuenkirchen, N., Englbrecht, C., Ohmer, J., Ziegenhals, T., Chari, A., and Fischer, U. (2015). Reconstitution of the human U snRNP assembly machinery reveals stepwise Sm protein organization. *EMBO J.*
- Newburn, L.R., and White, K.A. (2015). Cis-acting RNA elements in positive-strand RNA plant virus genomes. *Virology* 479–480, 434–443.
- Nicholson, B.L., and White, K.A. (2011). 3' Cap-independent translation enhancers of positive-strand RNA plant viruses. *Curr. Opin. Virol.* 1, 373–380.
- Nicholson, B.L., Wu, B., Chevtchenko, I., and White, K.A. (2010). Tombusvirus recruitment of host translational machinery via the 3' UTR. *RNA* 16, 1402–1419.
- Nicholson, B.L., Zaslaver, O., Mayberry, L.K., Browning, K.S., and White, K.A. (2013). Tombusvirus Y-shaped translational enhancer forms a complex with eIF4F and can be functionally replaced by heterologous translational enhancers. *J. Virol.* 87, 1872–1883.
- Nicole, S., Diaz, C.C., Frugier, T., and Melki, J. (2002). Spinal muscular atrophy: recent advances and future prospects. *Muscle Nerve* 26, 4–13.
- Niedzwiecka, A., Marcotrigiano, J., Stepinski, J., Jankowska-Anyszka, M., Wyslouch-Cieszyńska, A., Dadlez, M., Gingras, A.C., Mak, P., Darzynkiewicz, E., Sonenberg, N., et al. (2002). Biophysical studies of eIF4E cap-binding protein: Recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J. Mol. Biol.* 319, 615–635.
- Niepmann, M. (2013). Hepatitis C virus RNA translation. *Curr. Top. Microbiol. Immunol.* 369, 143–166.
- Noller, H.F., Yusupov, M.M., Yusupova, G.Z., Baucom, A., and Cate, J.H.D. (2002). Translocation of tRNA during protein synthesis. *FEBS Lett.* 514, 11–16.
- Novoselov, S. V., Lobanov, A. V., Hua, D., Kasaikina, M. V., Hatfield, D.L., and Gladyshev, V.N. (2007). A highly efficient form of the selenocysteine insertion sequence element in protozoan parasites and its use in mammalian cells. *Proc Natl Acad Sci U S A* 104, 7857–7862.
- Nowakowska, M., Kowalska, J., Martin, F., D'Orchymont, A., Zuberek, J., Lukaszewicz, M., Darzynkiewicz, E., and Jemielity, J. (2014). Cap analogs containing 6-thioguanosine – reagents for the synthesis of mRNAs selectively photo-crosslinkable with cap-binding biomolecules. *Org. Biomol. Chem.* 12, 4841–4847.
- Ohno, M., Segref, A., Bachi, A., Wilm, M., and Mattaj, I.W. (2000). PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* 101, 187–198.
- Olieric, V., Wolff, P., Takeuchi, A., Bec, G., Birck, C., Vitorino, M., Kieffer, B., Beniaminov, A., Cavigliolo, G., Theil, E., et al. (2009). SECIS-binding protein 2, a key player in selenoprotein synthesis, is an intrinsically disordered protein. *Biochimie* 91, 1003–1009.
- Osborne, M.J., Volpon, L., Kornblatt, J. a, Culjkovic-Kraljacic, B., Baguet, A., and Borden, K.L.B. (2013). eIF4E3 acts as a tumor suppressor by utilizing an atypical mode of methyl-7-guanosine cap recognition. *Proc. Natl. Acad. Sci. U. S. A.* 110, 3877–3882.
- Otter, S., Grimmmler, M., Neuenkirchen, N., Chari, A., Sickmann, A., and Fischer, U. (2007). A comprehensive interaction map of the human survival of motor neuron (SMN) complex. *J. Biol. Chem.* 282, 5825–5833.
- Otto, G.A., and Puglisi, J.D. (2004). The pathway of HCV IRES-mediated translation initiation. *Cell* 119, 369–380.
- Ou, Y., Jiang, B., Wang, X., Ma, W., and Guo, J. (2012). Selenium and colorectal adenomas risk: a meta-analysis. *Nutr. Cancer* 64, 1153–1159.
- Pabis, M., Neufeld, N., Shav-Tal, Y., and Neugebauer, K.M. (2010). Binding properties and dynamic localization of an alternative isoform of the cap-binding complex subunit CBP20. *J.* 412–421.
- Pacheco, A., De Quinto, S.L., Ramajo, J., Fernández, N., and Martínez-Salas, E. (2009). A novel role for Gemin5 in mRNA translation. *Nucleic Acids Res.* 37, 582–590.
- Pal, S., Vishwanath, S.N., Erdjument-Bromage, H., Tempst, P., and Sif, S. (2004). Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes. *Mol. Cell. Biol.* 24, 9630–9645.
- Palacios, I., Hetzer, M., Adam, S.A., and Mattaj, I.W. (1997). Nuclear import of U snRNPs requires importin beta. *EMBO J.* 16, 6783–6792.
- Palacios, I.M., Gatfield, D., St Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature* 427, 753–757.

- Palioura, S., Sherrer, R.L., Steitz, T.A., Soll, D., and Simonovic, M. (2009). The human SepSecS-tRNA^{Sec} complex reveals the mechanism of selenocysteine formation. *Science* (80-.). 325, 321–325.
- Panthu, B., Decimo, D., Balvay, L., and Ohlmann, T. (2015). In vitro translation in a hybrid cell free lysate with exogenous cellular ribosomes. *Biochem. J.*
- Papp, L. V., Lu, J., Striebel, F., Kennedy, D., Holmgren, A., and Khanna, K.K. (2006). The redox state of SECIS binding protein 2 controls its localization and selenocysteine incorporation function. *Mol Cell Biol* 26, 4895–4910.
- Papp, L. V., Lu, J., Holmgren, A., and Khanna, K.K. (2007). From selenium to selenoproteins: synthesis, identity, and their role in human health. *Antioxid Redox Signal* 9, 775–806.
- Papp, L. V., Wang, J., Kennedy, D., Boucher, D., Zhang, Y., Gladyshev, V.N., Singh, R.N., and Khanna, K.K. (2008). Functional characterization of alternatively spliced human SECISBP2 transcript variants. *Nucleic Acids Res* 36, 7192–7206.
- Parsyan, A., Shahbazian, D., Martineau, Y., Petroulakis, E., Alain, T., Larsson, O., Mathonnet, G., Tettweiler, G., Hellen, C.U., Pestova, T. V., et al. (2009). The helicase protein DHX29 promotes translation initiation, cell proliferation, and tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* 106, 22217–22222.
- 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.
- Patel, A.A., and Steitz, J.A. (2003). Splicing double: insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* 4, 960–970.
- Patel, S.B., and Bellini, M. (2008). The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res.* 36, 6482–6493.
- Pause, A., Belsham, G.J., Gingras, A.C., Donzé, O., Lin, T.A., Lawrence, J.C., and Sonenberg, N. (1994). Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* 371, 762–767.
- Paushkin, S., Gubitz, A.K., Massenet, S., and Dreyfuss, G. (2002). The SMN complex, an assemblyosome of ribonucleoproteins. *Curr. Opin. Cell Biol.* 14, 305–312.
- Pearn, J. (1978). A study of environmental factors in relation to fatal infantile spinal muscular atrophy (SMA type I). An analysis of birth order and parental age effects, social class, seasonal influence, and clustering in time and place. *J. Neurol. Sci.* 38, 39–45.
- 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.
- Pellizzoni, L. (2002). Essential Role for the SMN Complex in the Specificity of snRNP Assembly. *Science* (80-.). 298, 1775–1779.
- Pellizzoni, L. (2007). Chaperoning ribonucleoprotein biogenesis in health and disease. *EMBO Rep.* 8, 340–345.
- Pellizzoni, L., Charroux, B., and Dreyfuss, G. (1999). SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. *Proc. Natl. Acad. Sci. U. S. A.* 96, 11167–11172.
- Pérez-Ortín, J.E., Alepuz, P., Chávez, S., and Choder, M. (2013). Eukaryotic mRNA decay: Methodologies, pathways, and links to other stages of gene expression. *J. Mol. Biol.* 425, 3750–3775.
- Persson-Moschos, M.E., Stavenow, L., Akesson, B., and Lindgärde, F. (2000). Selenoprotein P in plasma in relation to cancer morbidity in middle-aged Swedish men. *Nutr. Cancer* 36, 19–26.
- Pesiridis, G.S., Diamond, E., and Van Duyne, G.D. (2009). Role of pICln in Methylation of Sm Proteins by PRMT5. *J. Biol. Chem.* 284, 21347–21359.
- Peske, F., Savelsbergh, A., Katunin, V.I., Rodnina, M. V., and Wintermeyer, W. (2004). Conformational changes of the small ribosomal subunit during elongation factor G-dependent tRNA-mRNA translocation. *J. Mol. Biol.* 343, 1183–1194.
- 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., and Hellen, C.U.T. (2003). Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev.* 17, 181–186.
- 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., 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., Lomakin, I.B., Lee, J.H., Choi, S.K., Dever, T.E., and Hellen, C.U. (2000). The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* 403, 332–335.
- Pestova, T. V., Shatsky, I.N., Fletcher, S.P., Jackson, R.J., and Hellen, C.U.T. (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.
- Peters, U., Chatterjee, N., Hayes, R.B., Schoen, R.E., Wang, Y., Chanock, S.J., and Foster, C.B. (2008). Variation in the selenoenzyme genes and risk of advanced distal colorectal adenoma. *Cancer Epidemiol. Biomarkers Prev.* 17, 1144–1154.
- Pettitt, J., Harrison, N., Stansfield, I., Connolly, B., and Müller, B. (2010). The evolution of spliced leader trans-splicing in nematodes. *Biochem. Soc. Trans.* 38, 1125–1130.

- Piazzon, N., Rage, F., Schlotter, F., Moine, H., Branlant, C., and Massenet, S. (2008). In vitro and in Cellulo evidences for association of the survival of motor neuron complex with the fragile X mental retardation protein. *J. Biol. Chem.* 283, 5598–5610.
- Piazzon, N., Schlotter, F., Lefebvre, S., Dodré, M., Méreau, A., Soret, J., Besse, A., Barkats, M., Bordonné, R., Branlant, C., et al. (2013). Implication of the SMN complex in the biogenesis and steady state level of the Signal Recognition Particle. *Nucleic Acids Res.* 41, 1255–1272.
- Pillai, R.S. (2003). Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev.* 17, 2321–2333.
- Pillai, R.S., Grimmer, M., Meister, G., Will, C.L., Lührmann, R., Fischer, U., and Schümperli, D. (2003). Unique Sm core structure of U7 snRNPs: Assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev.* 17, 2321–2333.
- Pillutla, R.C., Shimamoto, A., Furuichi, Y., and Shatkin, A.J. (1998). Human mRNA capping enzyme (RNGTT) and cap methyltransferase (RNMT) map to 6q16 and 18p11.22-p11.23, respectively. *Genomics* 54, 351–353.
- Piñeiro, D., Fernandez-Chamorro, J., Francisco-Velilla, R., and Martinez-Salas, E. (2015). Gemin5: A Multitasking RNA-Binding Protein Involved in Translation Control. *Biomolecules* 5, 528–544.
- Piontkivska, H., Rooney, A.P., and Nei, M. (2002). Purifying selection and birth-and-death evolution in the histone H4 gene family. *Mol. Biol. Evol.* 19, 689–697.
- Pisarev, A. V., Shirokikh, N.E., and Hellen, C.U.T. (2005). Translation initiation by factor-independent binding of eukaryotic ribosomes to internal ribosomal entry sites. *Comptes Rendus - Biol.* 328, 589–605.
- Plessel, G., Fischer, U., and Lührmann, R. (1994). m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) in vitro: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain. *Mol. Cell. Biol.* 14, 4160–4172.
- Pollard, V.W., and Malim, M.H. (1998). The HIV-1 Rev protein. *Annu. Rev. Microbiol.* 52, 491–532.
- Popp, M.W.-L., and Maquat, L.E. (2013). Organizing principles of mammalian nonsense-mediated mRNA decay. *Annu. Rev. Genet.* 47, 139–165.
- Popp, M.W.-L., and Maquat, L.E. (2014). The dharma of nonsense-mediated mRNA decay in mammalian cells. *Mol. Cells* 37, 1–8.
- Pradet-Balade, B., Girard, C., Boulon, S., Paul, C., Azzag, K., Bordonné, R., Bertrand, E., and Verheggen, C. (2011). CRM1 controls the composition of nucleoplasmic pre-snoRNA complexes to licence them for nucleolar transport. *EMBO J.* 30, 2205–2218.
- Prongidi-Fix, L., Schaeffer, L., Simonetti, A., Barends, S., Ménétret, J.-F., Klaholz, B.P., Eriani, G., and Martin, F. (2013). Rapid purification of ribosomal particles assembled on histone H4 mRNA: a new method based on mRNA-DNA chimaeras. *Biochem. J.* 449, 719–728.
- Pu, W.T., Krapivinsky, G.B., Krapivinsky, L., and Clapham, D.E. (1999). pICln inhibits snRNP biogenesis by binding core spliceosomal proteins. *Mol. Cell. Biol.* 19, 4113–4120.
- Qichang Shen, L.F.P.E.N. (2006). Nuclease sensitive element binding protein 1 associates with the selenocysteine insertion sequence and functions in mammalian selenoprotein translation. *J. Cell. Physiol.* 207, 775–783.
- Qin, X., and Sarnow, P. (2004). Preferential Translation of Internal Ribosome Entry Site-containing mRNAs during the Mitotic Cycle in Mammalian Cells. *J. Biol. Chem.* 279, 13721–13728.
- Raaschou-Nielsen, O., Sørensen, M., Hansen, R.D., Frederiksen, K., Tjønneland, A., Overvad, K., and Vogel, U. (2007). GPX1 Pro198Leu polymorphism, interactions with smoking and alcohol consumption, and risk for lung cancer. *Cancer Lett.* 247, 293–300.
- Rage, F., Boulsifane, N., Rihan, K., Neel, H., Gostan, T., Bertrand, E., Bordonné, R., and Soret, J. (2013). Genome-wide identification of mRNAs associated with the protein SMN whose depletion decreases their axonal localization. *RNA* 19, 1755–1766.
- Raker, V.A., Plessel, G., and Lührmann, R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *EMBO J.* 15, 2256–2269.
- Rakotondrafara, A.M., and Hentze, M.W. (2011). An efficient factor-depleted mammalian in vitro translation system. *Nat. Protoc.* 6, 563–571.
- Raman, A. V., Pitts, M.W., Seyedali, A., Hashimoto, A.C., Bellinger, F.P., and Berry, M.J. (2013). Selenoprotein W expression and regulation in mouse brain and neurons. *Brain Behav.* 3, 562–574.
- Ratnasinghe, D., Tangrea, J.A., Andersen, M.R., Barrett, M.J., Virtamo, J., Taylor, P.R., and Albanes, D. (2000). Glutathione peroxidase codon 198 polymorphism variant increases lung cancer risk. *Cancer Res.* 60, 6381–6383.
- Ravn-Haren, G., Olsen, A., Tjønneland, A., Dragsted, L.O., Nexø, B.A., Wallin, H., Overvad, K., Raaschou-Nielsen, O., and Vogel, U. (2006). Associations between GPX1 Pro198Leu polymorphism, erythrocyte GPX activity, alcohol consumption and breast cancer risk in a prospective cohort study. *Carcinogenesis* 27, 820–825.
- Rayman, M.P. (2000). The importance of selenium to human health. *Lancet* 356, 233–241.
- Rayman, M.P. (2012). Selenium and human health. *Lancet* 379, 1256–1268.
- Rayman, M.P., and Rayman, M.P. (2002). The argument for increasing selenium intake. *Proc. Nutr. Soc.* 61, 203–215.

- Read, R., Bellew, T., Yang, J.G., Hill, K.E., Palmer, I.S., and Burk, R.F. (1990). Selenium and amino acid composition of selenoprotein P, the major selenoprotein in rat serum. *J. Biol. Chem.* 265, 17899–17905.
- Reddy, R., Henning, D., Das, G., Harless, M., and Wright, D. (1987). The capped U6 small nuclear RNA is transcribed by RNA polymerase III. *J. Biol. Chem.* 262, 75–81.
- Reddy, R., Singh, R., and Shimba, S. (1992). Methylated cap structures in eukaryotic RNAs: Structure, synthesis and functions. *Pharmacol. Ther.* 54, 249–267.
- Rederstorff, M., Krol, A., and Lescure, A. (2006). Understanding the importance of selenium and selenoproteins in muscle function. *Cell Mol Life Sci* 63, 52–59.
- Reeves, M.A., Bellinger, F.P., and Berry, M.J. (2010). The neuroprotective functions of selenoprotein M and its role in cytosolic calcium regulation. *Antioxid. Redox Signal.* 12, 809–818.
- Reibarkh, M., Yamamoto, Y., Singh, C.R., Del Rio, F., Fahmy, A., Lee, B., Luna, R.E., Ii, M., Wagner, G., and Asano, K. (2008). Eukaryotic initiation factor (eIF) 1 carries two distinct eIF5-binding faces important for multifactor assembly and AUG selection. *J. Biol. Chem.* 283, 1094–1103.
- Reichow, S.L., Hamma, T., Ferré-D'Amaré, A.R., and Varani, G. (2007). The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res.* 35, 1452–1464.
- Ren, J., Wang, Y., Liang, Y., Zhang, Y., Bao, S., and Xu, Z. (2010). Methylation of ribosomal protein S10 by protein-arginine methyltransferase 5 regulates ribosome biogenesis. *J. Biol. Chem.* 285, 12695–12705.
- Reynolds, J.E., Kaminski, A., Kettinen, H.J., Grace, K., Clarke, B.E., Carroll, A.R., Rowlands, D.J., and Jackson, R.J. (1995). Unique features of internal initiation of hepatitis C virus RNA translation. *EMBO J.* 14, 6010–6020.
- Rhoads, R.E. (2009). EIF4E: New family members, new binding partners, new roles. *J. Biol. Chem.* 284, 16711–16715.
- Ritchie, M., and Yeap, B.B. (2015). Thyroid hormone: Influences on mood and cognition in adults. *Maturitas*.
- Rodnina, M. V., and Wintermeyer, W. (2009). Recent mechanistic insights into eukaryotic ribosomes. *Curr. Opin. Cell Biol.* 21, 435–443.
- Rodor, J., Letelier, I., Holuigue, L., and Echeverria, M. (2010). Nucleolar RNPs: from genes to functional snoRNAs in plants. *Biochem. Soc. Trans.* 38, 672–676.
- Rollenhagen, C., and Panté, N. (2006). Nuclear import of spliceosomal snRNPs. *Can. J. Physiol. Pharmacol.* 84, 367–376.
- Rom, E., Kim, H.C., Gingras, A.-C., Marcotrigiano, J., Favre, D., Olsen, H., Burley, S.K., and Sonenberg, N. (1998). Cloning and Characterization of 4EHP, a Novel Mammalian eIF4E-related Cap-binding Protein. *J. Biol. Chem.* 273, 13104–13109.
- Romagné, F., Santesmasses, D., White, L., Sarangi, G.K., Mariotti, M., Hübler, R., Weihmann, A., Parra, G., Gladyshev, V.N., Guigó, R., et al. (2014). SelenoDB 2.0: Annotation of selenoprotein genes in animals and their genetic diversity in humans. *Nucleic Acids Res.* 42.
- Roman, M., Jitaru, P., and Barbante, C. (2014). Selenium biochemistry and its role for human health. *Metallomics* 6, 25–54.
- Rosettani, P., Knapp, S., Vismara, M.G., Rusconi, L., and Cameron, A.D. (2007). Structures of the Human eIF4E Homologous Protein, h4EHP, in its m7GTP-bound and Unliganded Forms. *J. Mol. Biol.* 368, 691–705.
- Rotruck, J.T., Pope, A.L., Ganther, H.E., Swanson, A.B., Hafeman, D.G., and Hoekstra, W.G. (1973). Selenium: biochemical role as a component of glutathione peroxidase. *Science* (80-.). 179, 588–590.
- Rovet, J.F. (2014). The Role of Thyroid Hormones for Brain Development and Cognitive Function. pp. 26–43.
- 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.
- Rozhdestvensky, T.S., Tang, T.H., Tchirkova, I. V., Brosius, J., Bachellerie, J.P., and Hüttenhofer, A. (2003). Binding of L7Ae protein to the K-turn of archaeal snoRNAs: A shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res.* 31, 869–877.
- Rufener, S.C., and Mühlemann, O. (2013). eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nat. Struct. Mol. Biol.* 20, 710–717.
- Rundlöf, A.-K., and Arnér, E.S.J. (2004). Regulation of the mammalian selenoprotein thioredoxin reductase 1 in relation to cellular phenotype, growth, and signaling events. *Antioxid. Redox Signal.* 6, 41–52.
- Saito, Y., and Takahashi, K. (2002). Characterization of selenoprotein P as a selenium supply protein. *Eur. J. Biochem.* 269, 5746–5751.
- Samarsky, D.A., Fournier, M.J., Singer, R.H., and Bertrand, E. (1998). The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *EMBO J.* 17, 3747–3757.
- Sandalova, T., Zhong, L., Lindqvist, Y., Holmgren, A., and Schneider, G. (2001). Three-dimensional structure of a mammalian thioredoxin reductase: implications for mechanism and evolution of a selenocysteine-dependent enzyme. *Proc. Natl. Acad. Sci. U. S. A.* 98, 9533–9538.
- Sato, H., and Maquat, L.E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin ?? *Genes Dev.* 23, 2537–2550.
- Savaskan, N.E., Ufer, C., Kühn, H., and Borchert, A. (2007). Molecular biology of glutathione peroxidase 4: From genomic structure to developmental expression and neural function. *Biol. Chem.* 388, 1007–1017.
- Schaffert, N., Hossbach, M., Heintzmann, R., Achsel, T., and Lührmann, R. (2004). RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *EMBO J.* 23, 3000–3009.
- Scheres, S.H.W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* 180, 519–530.

- Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B., and Liu, J.O. (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat. Chem. Biol.* 6, 209–217.
- Schoenmakers, E., Agostini, M., Mitchell, C., Schoenmakers, N., Papp, L., Rajanayagam, O., Padidela, R., Ceron-Gutierrez, L., Doffinger, R., Prevosto, C., et al. (2010). Mutations in the selenocysteine insertion sequence-binding protein 2 gene lead to a multisystem selenoprotein deficiency disorder in humans. *J. Clin. Invest.* 120, 4220–4235.
- Schomburg, L., and Schweizer, U. (2009). Hierarchical regulation of selenoprotein expression and sex-specific effects of selenium. *Biochim. Biophys. Acta - Gen. Subj.* 1790, 1453–1462.
- Schultz, A., Nottrott, S., Watkins, N.J., and Lührmann, R. (2006). Protein-protein and protein-RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs. *Mol. Cell. Biol.* 26, 5146–5154.
- Seng, C.O., Magee, C., Young, P.J., Lorson, C.L., and Allen, J.P. (2015). The SMN structure reveals its crucial role in snRNP assembly. *Hum. Mol. Genet.* 24, 2138–2146.
- Sengottuvelan, M., Deeptha, K., and Nalini, N. (2009). Resveratrol ameliorates DNA damage, prooxidant and antioxidant imbalance in 1,2-dimethylhydrazine induced rat colon carcinogenesis. *Chem. Biol. Interact.* 181, 193–201.
- Seth-Vollenweider, T., Joshi, S., Dhawan, P., Sif, S., and Christakos, S. (2014). Novel Mechanism of Negative Regulation of 1,25-Dihydroxyvitamin D₃-induced 25-Hydroxyvitamin D₃ 24-Hydroxylase (Cyp24a1) Transcription. *J. Biol. Chem.* 289, 33958–33970.
- Seyedali, A., and Berry, M.J. (2014). Nonsense-mediated decay factors are involved in the regulation of selenoprotein mRNA levels during selenium deficiency. *RNA*.
- Seyedali, A., Berry, M.J., and Seyedali, A.L.I. (2014). Nonsense-mediated decay factors are involved in the regulation of selenoprotein mRNA levels during selenium deficiency Nonsense-mediated decay factors are involved in the regulation of selenoprotein mRNA levels during selenium deficiency. 0–9.
- Sharathchandra, A., Katoch, A., and Das, S. (2014). IRES mediated translational regulation of p53 isoforms. *Wiley Interdiscip. Rev. RNA* 5, 131–139.
- Shatkin, A.J. (1976). Capping of eucaryotic mRNAs. *Cell* 9, 645–653.
- Shatkin, A.J., and Manley, J.L. (2000). The ends of the affair: capping and polyadenylation. *Nat. Struct. Biol.* 7, 838–842.
- Shatsky, I.N., Dmitriev, S.E., Terenin, I.M., and Andreev, D.E. (2010). Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs. *Mol. Cells* 30, 285–293.
- Shatsky, I.N., Dmitriev, S.E., Andreev, D.E., and Terenin, I.M. (2014). Transcriptome-wide studies uncover the diversity of modes of mRNA recruitment to eukaryotic ribosomes. *Crit. Rev. Biochem. Mol. Biol.* 49, 164–177.
- Shen, Q., Leonard, J.L., and Newburger, P.E. (1995). Structure and function of the selenium translation element in the 3'-untranslated region of human cellular glutathione peroxidase mRNA. *RNA* 1, 519–525.
- Shen, Q., Wu, R., Leonard, J.L., and Newburger, P.E. (1998). Identification and molecular cloning of a human selenocysteine insertion sequence-binding protein. A bifunctional role for DNA-binding protein B. *J Biol Chem* 273, 5443–5446.
- Shen, Q., Fan, L., Newburger, P.E., Qichang Shen, L.F.P.E.N., Shen, Q., Fan, L., and Newburger, P.E. (2006). Nuclease sensitive element binding protein 1 associates with the selenocysteine insertion sequence and functions in mammalian selenoprotein translation. *J. Cell. Physiol.* 207, 775–783.
- Sheppard, K., Yuan, J., Hohn, M.J., Jester, B., Devine, K.M., and Söll, D. (2008). From one amino acid to another: tRNA-dependent amino acid biosynthesis. *Nucleic Acids Res.* 36, 1813–1825.
- Sherrer, R.L., O'Donoghue, P., and Soll, D. (2008). Characterization and evolutionary history of an archaeal kinase involved in selenocysteinyl-tRNA formation. *Nucleic Acids Res* 36, 1247–1259.
- Shetty, S.P., and Copeland, P.R. (2015). Selenocysteine incorporation: A trump card in the game of mRNA decay. *Biochimie*.
- Shetty, S.P., Shah, R., and Copeland, P.R. (2014). Regulation of selenocysteine incorporation into the selenium transport protein, Selenoprotein P. *J. Biol. Chem.*
- Shi, Y. (2012). Alternative polyadenylation: New insights from global analyses. *RNA*.
- Shi, Y., and Manley, J.L. (2015). The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes Dev.* 29, 889–897.
- Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., Frank, J., and Manley, J.L. (2009). Molecular Architecture of the Human Pre-mRNA 3' Processing Complex. *Mol. Cell* 33, 365–376.
- Shibata, T., Arisawa, T., Tahara, T., Ohkubo, M., Yoshioka, D., Maruyama, N., Fujita, H., Kamiya, Y., Nakamura, M., Nagasaka, M., et al. (2009). Selenoprotein S (SEPS1) gene -105G>A promoter polymorphism influences the susceptibility to gastric cancer in the Japanese population. *BMC Gastroenterol.* 9, 2.
- Shih, J., Wang, W., Tsai, T., Kuo, C., Li, H., and Wu Lee, Y. (2011). Critical roles of RNA helicase DDX3 and its interactions with eIF4E/PABP1 in stress granule assembly and stress response. *Biochem. J.* 441, 119–129.
- Shih, J.-W., Tsai, T.-Y., Chao, C.-H., and Wu Lee, Y.-H. (2008). Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. *Oncogene* 27, 700–714.
- Shisler, J.L., Senkevich, T.G., Berry, M.J., and Moss, B. (1998). Ultraviolet-induced cell death blocked by a selenoprotein from a human dermatotropic poxvirus. *Science* 279, 102–105.

- Shpargel, K.B., and Matera, A.G. (2005). Gemin proteins are required for efficient assembly of Sm-class ribonucleoproteins. *Proc. Natl. Acad. Sci. U. S. A.* 102, 17372–17377.
- Shuman, S. (2002). What messenger RNA capping tells us about eukaryotic evolution. *Nat. Rev. Mol. Cell Biol.* 3, 619–625.
- Simon, A.E. (2015). 3'UTRs of carmoviruses. *Virus Res.* 206, 27–36.
- Simon, A.E., and Miller, W.A. (2013). Translation Enhancers of Plant Viruses. *Annu. Rev. Microbiol.*
- Singh, R., and Reddy, R. (1989). Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proc. Natl. Acad. Sci. U. S. A.* 86, 8280–8283.
- Skabkin, M., Skabkina, O., Hellen, C.T., and Pestova, T. (2013). Reinitiation and other unconventional posttermination events during eukaryotic translation. *Mol. Cell* 51, 249–264.
- Sleeman, J.E., and Lamond, A.I. (1999). Nuclear organization of pre-mRNA splicing factors. *Curr. Opin. Cell Biol.* 11, 372–377.
- Small-Howard, A.L., and Berry, M.J. (2005). Unique features of selenocysteine incorporation function within the context of general eukaryotic translational processes. *Biochem. Soc. Trans.* 33, 1493–1497.
- Small-Howard, A., Morozova, N., Stoytcheva, Z., Forry, E.P., Mansell, J.B., Harney, J.W., Carlson, B.A., Xu, X.M., Hatfield, D.L., and Berry, M.J. (2006). Supramolecular complexes mediate selenocysteine incorporation in vivo. *Mol Cell Biol* 26, 2337–2346.
- Sonenberg, N. (2008). eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. *Biochem. Cell Biol.* 86, 178–183.
- 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.
- Soo-Kyung, O.H., 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.
- Soto-Rifo, R., Rubilar, P.S., Limousin, T., de Brejne, S., Décimo, D., and Ohlmann, T. (2012). DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. *EMBO J.* 31, 3745–3756.
- Soto-Rifo, R., Rubilar, P.S., and Ohlmann, T. (2013). The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA. *Nucleic Acids Res.* 41, 6286–6299.
- Speckmann, W.A., Terns, R.M., and Terns, M.P. (2000). The box C/D motif directs snoRNA 5'-cap hypermethylation. *Nucleic Acids Res.* 28, 4467–4473.
- Spector, D.L. (1993). Nuclear organization of pre-mRNA processing. *Curr. Opin. Cell Biol.* 5, 442–447.
- Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* 3, 1–12.
- Squires, J.E., and Berry, M.J. (2008). Eukaryotic selenoprotein synthesis: mechanistic insight incorporating new factors and new functions for old factors. *IUBMB Life* 60, 232–235.
- Squires, J.E., Stoytchev, I., Forry, E.P., and Berry, M.J. (2007a). SBP2 binding affinity is a major determinant in differential selenoprotein mRNA translation and sensitivity to nonsense-mediated decay. *Mol. Cell. Biol.* 27, 7848–7855.
- Squires, J.E., Stoytchev, I., Forry, E.P., and Berry, M.J. (2007b). SBP2 binding affinity is a major determinant in differential selenoprotein mRNA translation and sensitivity to nonsense-mediated decay. *Mol Cell Biol* 27, 7848–7855.
- Squires, J.E., Davy, P., Berry, M.J., and Allsopp, R. (2009). Attenuated expression of SECIS binding protein 2 causes loss of telomeric reserve without affecting telomerase. *Exp Gerontol.*
- Stade, K., Ford, C.S., Guthrie, C., and Weis, K. (1997). Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* 90, 1041–1050.
- Staněk, D., Rader, S.D., Klingauf, M., and Neugebauer, K.M. (2003). Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *J. Cell Biol.* 160, 505–516.
- Steinbrenner, H., and Sies, H. (2009). Protection against reactive oxygen species by selenoproteins. *Biochim. Biophys. Acta - Gen. Subj.* 1790, 1478–1485.
- Steinbrenner, H., Al-Quraishy, S., Dkhil, M.A., Wunderlich, F., and Sies, H. (2015). Dietary selenium in adjuvant therapy of viral and bacterial infections. *Adv. Nutr.* 6, 73–82.
- Stock, T., and Rother, M. (2009). Selenoproteins in Archaea and Gram-positive bacteria. *Biochim. Biophys. Acta - Gen. Subj.* 1790, 1520–1532.
- Stoffaneller, R., and Morse, N. (2015). A Review of Dietary Selenium Intake and Selenium Status in Europe and the Middle East. *Nutrients* 7, 1494–1537.
- Stopa, N., Krebs, J.E., and Shechter, D. (2015). The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. *Cell. Mol. Life Sci.*
- Stoytcheva, Z., Tujebajeva, R.M., Harney, J.W., and Berry, M.J. (2006a). Efficient incorporation of multiple selenocysteines involves an inefficient decoding step serving as a potential translational checkpoint and ribosome bottleneck. *Mol Cell Biol* 26, 9177–9184.
- Stoytcheva, Z., Tujebajeva, R.M., Harney, J.W., and Berry, M.J. (2006b). Efficient incorporation of multiple selenocysteines involves an inefficient decoding step serving as a potential translational checkpoint and ribosome bottleneck. *Mol. Cell. Biol.* 26, 9177–9184.

- Strasser, A., Dickmanns, A., Schmidt, U., Penka, E., Urlaub, H., Sekine, M., Lührmann, R., and Ficner, R. (2004). Purification, crystallization and preliminary crystallographic data of the m3G cap-binding domain of human snRNP import factor snurportin 1. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60, 1628–1631.
- Stupina, V.A., Meskauskas, A., McCormack, J.C., Yingling, Y.G., Shapiro, B.A., Dinman, J.D., and Simon, A.E. (2008). The 3' proximal translational enhancer of Turnip crinkle virus binds to 60S ribosomal subunits. *RNA* 14, 2379–2393.
- Stupina, V.A., Yuan, X., Meskauskas, A., Dinman, J.D., and Simon, A.E. (2011). Ribosome binding to a 5' translational enhancer is altered in the presence of the 3' untranslated region in cap-independent translation of turnip crinkle virus. *J. Virol.* 85, 4638–4653.
- Sturchler, C., Westhof, E., Carbon, P., and Krol, A. (1993). Unique secondary and tertiary structural features of the eucaryotic selenocysteine tRNA(Sec). *Nucleic Acids Res* 21, 1073–1079.
- Sturchler, C., Lescure, A., Keith, G., Carbon, P., Krol, A., Ibba, M., and Soll, D. (1994). Base modification pattern at the wobble position of *Xenopus* selenocysteine tRNA(Sec). *Nucleic Acids Res* 22, 1354–1358. co – translationally inserted during protein.
- Sun, Q.A., Zappacosta, F., Factor, V.M., Wirth, P.J., Hatfield, D.L., and Gladyshev, V.N. (2001a). Heterogeneity within animal thioredoxin reductases. Evidence for alternative first exon splicing. *J. Biol. Chem.* 276, 3106–3114.
- Sun, X., Li, X., Moriarty, P.M., Henics, T., LaDuca, J.P., and Maquat, L.E. (2001b). Nonsense-mediated decay of mRNA for the selenoprotein phospholipid hydroperoxide glutathione peroxidase is detectable in cultured cells but masked or inhibited in rat tissues. *Mol. Biol. Cell* 12, 1009–1017.
- Sun, X., Li, X., Moriarty, P.M., Henics, T., LaDuca, J.P., and Maquat, L.E. (2001c). Nonsense-mediated decay of mRNA for the selenoprotein phospholipid hydroperoxide glutathione peroxidase is detectable in cultured cells but masked or inhibited in rat tissues. *Mol Biol Cell* 12, 1009–1017.
- Sunde, R.A., and Raines, A.M. (2011). Selenium Regulation of the Selenoprotein and Nonselenoprotein Transcriptomes in Rodents. *Adv. Nutr. An Int. Rev. J.* 2, 138–150.
- Suppmann, S., Persson, B.C., and Böck, A. (1999). Dynamics and efficiency in vivo of UGA-directed selenocysteine insertion at the ribosome. *EMBO J.* 18, 2284–2293.
- Sutton, A., Nahon, P., Pessayre, D., Rufat, P., Poiré, A., Zioli, M., Vidaud, D., Barget, N., Ganne-Carrié, N., Charnaux, N., et al. (2006). Genetic polymorphisms in antioxidant enzymes modulate hepatic iron accumulation and hepatocellular carcinoma development in patients with alcohol-induced cirrhosis. *Cancer Res.* 66, 2844–2852.
- Svitkin, Y. V., Evdokimova, V.M., Brasey, A., Pestova, T. V., Fantus, D., Yanagiya, A., Imataka, H., Skabkin, M.A., Ovchinnikov, L.P., Merrick, W.C., et al. (2009). General RNA-binding proteins have a function in poly(A)-binding protein-dependent translation. *EMBO J.* 28, 58–68.
- Sweeney, T.R., Abaeva, I.S., Pestova, T. V., and Hellen, C.U.T. (2014). The mechanism of translation initiation on Type 1 picornavirus IRESs. *EMBO J.* 33, 76–92.
- Tadesse, H., Deschênes-Furry, J., Boisvenue, S., and Côté, J. (2008). KH-type splicing regulatory protein interacts with survival motor neuron protein and is misregulated in spinal muscular atrophy. *Hum. Mol. Genet.* 17, 506–524.
- Tajirishi, M.M., Tuteja, R., and Tuteja, N. (2011). Nucleolin: The most abundant multifunctional phosphoprotein of nucleolus. *Commun. Integr. Biol.* 4, 267–275.
- Takeuchi, A., Schmitt, D., Chapple, C., Babaylova, E., Karpova, G., Guigo, R., Krol, A., and Allmang, C. (2009). A short motif in *Drosophila* SECIS Binding Protein 2 provides differential binding affinity to SECIS RNA hairpins. *Nucleic Acids Res* 37, 2126–2141.
- Taskov, K., Chapple, C., Kryukov, G. V., Castellano, S., Lobanov, A. V., Korotkov, K. V., Guigó, R., and Gladyshev, V.N. (2005). Nematode selenoproteome: The use of the selenocysteine insertion system to decode one codon in an animal genome? *Nucleic Acids Res.* 33, 2227–2238.
- Taylor, P., Bubenik, J.L., Miniard, A.C., Driscoll, D.M., Bubenik, J.L., Miniard, A.C., and Driscoll, D.M. (2015). binding activities of SECIS-binding protein 2 Characterization of the UGA-recoding and SECIS- binding activities of SECIS-binding protein 2. 37–41.
- Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG Motif. *Mol. Cell* 50, 613–623.
- Tharun, S. (2008). Roles of Eukaryotic Lsm Proteins in the Regulation of mRNA Function. *K.W.J.B.T.-I.R. of C. and M. Biology*, ed. (Academic Press), pp. 149–189.
- Tisdale, S., Lotti, F., Saieva, L., Van Meerbeke, J.P., Crawford, T.O., Sumner, C.J., Mentis, G.Z., and Pellizzoni, L. (2013a). SMN Is Essential for the Biogenesis of U7 Small Nuclear Ribonucleoprotein and 3'-End Formation of Histone mRNAs. *Cell Rep.* 5, 1187–1195.
- Tisdale, S., Lotti, F., Saieva, L., VanMeerbeke, J., Crawford, T., Sumner, C., Mentis, G., and Pellizzoni, L. (2013b). SMN is essential for the biogenesis of U7 Small nuclear ribonucleoprotein and 3'-end formation of Histone mRNAs. *Cell Rep.* 5, 1187–1195.
- Tollervey, D., Lehtonen, H., Carmo-Fonseca, M., and Hurt, E.C. (1991). The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *EMBO J.* 10, 573–583.
- Tollervey, D., Lehtonen, H., Jansen, R., Kern, H., and Hurt, E.C. (1993). Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* 72, 443–457.

- Topisirovic, I., Svitkin, Y. V., Sonenberg, N., and Shatkin, A.J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev. RNA* 2, 277–298.
- Touat-Hamici, Z., Legrain, Y., Bulteau, A.L., and Chavatte, L. (2014). Selective up-regulation of human selenoproteins in response to oxidative stress. *J. Biol. Chem.* 289, 14750–14761.
- 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.
- 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.
- Treder, K., Kneller, E.L.P., Allen, E.M., Wang, Z., Browning, K.S., and Miller, W.A. (2008). The 3' cap-independent translation element of Barley yellow dwarf virus binds eIF4F via the eIF4G subunit to initiate translation. *RNA* 14, 134–147.
- Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M., and Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat. Struct. Mol. Biol.* 18, 1414–1420.
- 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.
- Tujebajeva, R.M., Copeland, P.R., Xu, X.M., Carlson, B.A., Harney, J.W., Driscoll, D.M., Hatfield, D.L., and Berry, M.J. (2000). Decoding apparatus for eukaryotic selenocysteine insertion. *EMBO Rep* 1, 158–163.
- Turunen, J.J., Niemelä, E.H., Verma, B., and Frilander, M.J. (2013). The significant other: Splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* 4, 61–76.
- Uniacke, J., Holterman, C.E., Lachance, G., Franovic, A., Jacob, M.D., Fabian, M.R., Payette, J., Holcik, M., Pause, A., and Lee, S. (2012). An oxygen-regulated switch in the protein synthesis machinery. *Nature*.
- Ursini, F., Heim, S., Kiess, M., Maiorino, M., Roveri, A., Wissing, J., and Flohe, L. (1999). Dual function of the selenoprotein PHGPx during sperm maturation. *Science* (80-.). 285, 1393–1396.
- Uversky, V.N. (2015). The multifaceted roles of intrinsic disorder in protein complexes. *FEBS Lett.*
- Vagner, S., Galy, B., and Pyronnet, S. (2001). Irresistible IRES. Attracting the translation machinery to internal ribosome entry sites. *EMBO Rep.* 2, 893–898.
- Verheggen, C., and Bertrand, E. (2012). CRM1 plays a nuclear role in transporting snoRNPs to nucleoli in higher eukaryotes. *Nucleus* 3, 132–137.
- Verheggen, C., Mouaikel, J., Thiry, M., Blanchard, J.M., Tollervey, D., Bordonné, R., Lafontaine, D.L.J., and Bertrand, E. (2001). Box C/D small nucleolar RNA trafficking involves small nucleolar RNP proteins, nucleolar factors and a novel nuclear domain. *EMBO J.* 20, 5480–5490.
- Verheggen, C., Lafontaine, D.L.J., Samarsky, D., Mouaikel, J., Blanchard, J.M., Bordonné, R., and Bertrand, E. (2002). Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *EMBO J.* 21, 2736–2745.
- Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R., and Ficner, R. (2000). Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol Cell* 6, 1331–1342.
- Villa, N., Do, A., Hershey, J.W.B., and Fraser, C.S. (2013). Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. *J. Biol. Chem.* 288, 32932–32940.
- Vitte, J., Fassier, C., Tiziano, F.D., Dalard, C., Soave, S., Roblot, N., Brahe, C., Saugier-Verber, P., Bonnefont, J.P., and Melki, J. (2007). Refined characterization of the expression and stability of the SMN gene products. *Am. J. Pathol.* 171, 1269–1280.
- Walczak, R., Westhof, E., Carbon, P., and Krol, A. (1996). A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs. *Rna* 2, 367–379.
- Walczak, R., Carbon, P., and Krol, A. (1998). An essential non-Watson-Crick base pair motif in 3'UTR to mediate selenoprotein translation. *Rna* 4, 74–84.
- Wallace, A., Filbin, M.E., Veo, B., McFarland, C., Stepinski, J., Jankowska-Anyszka, M., Darzynkiewicz, E., and Davis, R.E. (2010). The nematode eukaryotic translation initiation factor 4E/G complex works with a trans-spliced leader stem-loop to enable efficient translation of trimethylguanosine-capped RNAs. *Mol. Cell. Biol.* 30, 1958–1970.
- Wang, Z., Treder, K., and Miller, W.A. (2009). Structure of a viral cap-independent translation element that functions via high affinity binding to the eIF4E subunit of eIF4F. *J. Biol. Chem.* 284, 14189–14202.
- Wang, Z., Parisien, M., Scheets, K., and Miller, W.A. (2011). The cap-binding translation initiation factor, eIF4E, binds a pseudoknot in a viral cap-independent translation element. *Structure* 19, 868–880.
- Watkins, N.J., and Bohnsack, M.T. (2012). The box C/D and H/ACA snoRNPs: Key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA* 3, 397–414.
- Watkins, N.J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Luhrmann, R. (2000). A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP. *Cell* 103, 457–466.
- Watkins, N.J., Dickmanns, A., and Luhrmann, R. (2002). Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol. Cell. Biol.* 22, 8342–8352.

- Weiss, S.L., and Sunde, R.A. (1998). Cis-acting elements are required for selenium regulation of glutathione peroxidase-1 mRNA levels. *RNA* 4, 816–827.
- Weiss Sachdev, S., and Sunde, R.A. (2001). Selenium regulation of transcript abundance and translational efficiency of glutathione peroxidase-1 and -4 in rat liver. *Biochem. J.* 357, 851–858.
- 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.
- Wen, W., Weiss, S.L., and Sunde, R.A. (1998). UGA codon position affects the efficiency of selenocysteine incorporation into glutathione peroxidase-1. *J. Biol. Chem.* 273, 28533–28541.
- Westin, G., Lund, E., Murphy, J.T., Pettersson, U., and Dahlberg, J.E. (1984). Human U2 and U1 RNA genes use similar transcription signals. *EMBO J.* 3, 3295–3301.
- Wilden, B., Savelsbergh, A., Rodnina, M. V., and Wintermeyer, W. (2006). Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13670–13675.
- Wilson, J.E., Powell, M.J., Hoover, S.E., and Sarnow, P. (2000a). Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Mol. Cell. Biol.* 20, 4990–4999.
- Wilson, J.E., Pestova, T. V., Hellen, C.U., and Sarnow, P. (2000b). Initiation of protein synthesis from the A site of the ribosome. *Cell* 102, 511–520.
- Wirth, B., Brichta, L., and Hahnen, E. (2006). Spinal Muscular Atrophy: From Gene to Therapy. *Semin. Pediatr. Neurol.* 13, 121–131.
- Wirth, E.K., Conrad, M., Winterer, J., Wozny, C., Carlson, B.A., Roth, S., Schmitz, D., Bornkamm, G.W., Coppola, V., Tessarollo, L., et al. (2010). Neuronal selenoprotein expression is required for interneuron development and prevents seizures and neurodegeneration. *FASEB J.* 24, 844–852.
- Worch, R., Niedzwiecka, A., Stepinski, J., Mazza, C., Jankowska-Anyszka, M., Darzynkiewicz, E., Cusack, S., and Stolarski, R. (2005). Specificity of recognition of mRNA 5' cap by human nuclear cap-binding complex. *RNA* 11, 1355–1363.
- Wu, X.Q., and Gross, H.J. (1994). The length and the secondary structure of the D-stem of human selenocysteine tRNA are the major identity determinants for serine phosphorylation. *Embo J* 13, 241–248.
- Wu, R., Shen, Q., and Newburger, P.E. (2000). Recognition and binding of the human selenocysteine insertion sequence by nucleolin. *J Cell Biochem* 77, 507–516.
- Wurth, L., Gribbling-Burrer, A.S., Verheggen, C., Leichter, M., Takeuchi, A., Baudrey, S., Martin, F., Krol, A., Bertrand, E., and Allmann, C. (2014). Hypermethylated-capped selenoprotein mRNAs in mammals. *Nucleic Acids Res.* 42, 8663–8677.
- Xu, X., Hoang, S., Mayo, M.W., and Bekiranov, S. (2010a). Application of machine learning methods to histone methylation ChIP-Seq data reveals H4R3me2 globally represses gene expression. *BMC Bioinformatics* 11, 396.
- Xu, X.-M., Mix, H., Carlson, B.A., Grabowski, P.J., Gladyshev, V.N., Berry, M.J., and Hatfield, D.L. (2005). Evidence for Direct Roles of Two Additional Factors, SECp43 and Soluble Liver Antigen, in the Selenoprotein Synthesis Machinery. *J. Biol. Chem.* 280, 41568–41575.
- Xu, X.-M., Carlson, B.A., Mix, H., Zhang, Y., Saira, K., Glass, R.S., Berry, M.J., Gladyshev, V.N., and Hatfield, D.L. (2007a). Biosynthesis of Selenocysteine on Its tRNA in Eukaryotes. *PLoS Biol.* 5, e4.
- Xu, X.-M., Turanov, A.A., Carlson, B.A., Yoo, M.-H., Everley, R.A., Nandakumar, R., Sorokina, I., Gygi, S.P., Gladyshev, V.N., and Hatfield, D.L. (2010b). Targeted insertion of cysteine by decoding UGA codons with mammalian selenocysteine machinery. *Proc. Natl. Acad. Sci. U. S. A.* 107, 21430–21434.
- Xu, X.M., Carlson, B.A., Irons, R., Mix, H., Zhong, N., Gladyshev, V.N., and DL, H. (2007b). Selenophosphate synthetase 2 is essential for selenoprotein biosynthesis. *Biochem J.* 404, 115–120.
- Yamada-Okabe, T., Doi, R., Shimmi, O., Arisawa, M., and Yamada-Okabe, H. (1998). Isolation and characterization of a human cDNA for mRNA 5'-capping enzyme. *Nucleic Acids Res.* 26, 1700–1706.
- Yan, X., Mouillet, J.-F., Ou, Q., and Sadovsky, Y. (2003). A Novel Domain within the DEAD-Box Protein DP103 Is Essential for Transcriptional Repression and Helicase Activity. *Mol. Cell. Biol.* 23, 414–423.
- Yang, J.G., Morrison-Plummer, J., and Burk, R.F. (1987). Purification and quantitation of a rat plasma selenoprotein distinct from glutathione peroxidase using monoclonal antibodies. *J. Biol. Chem.* 262, 13372–13375.
- Yang, P., Bamlet, W.R., Ebbert, J.O., Taylor, W.R., and de Andrade, M. (2004). Glutathione pathway genes and lung cancer risk in young and old populations. *Carcinogenesis* 25, 1935–1944.
- Yant, L.J., Ran, Q., Rao, L., Van Remmen, H., Shibata, T., Belter, J.G., Motta, L., Richardson, A., and Prolla, T.A. (2003). The selenoprotein GPX4 is essential for mouse development and protects from radiation and oxidative damage insults. *Free Radic. Biol. Med.* 34, 496–502.
- Yedavalli, V.S.R.K., and Jeang, K.-T. (2010). Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14787–14792.
- Yong, J., Wan, L., and Dreyfuss, G. (2004). Why do cells need an assembly machine for RNA-protein complexes? *Trends Cell Biol.* 14, 226–232.
- Yong, J., Kasim, M., Bachorik, J.L., Wan, L., and Dreyfuss, G. (2010). Gemin5 Delivers snRNA Precursors to the SMN Complex for snRNP Biogenesis. *Mol. Cell* 38, 551–562.

- Yoo, M.H., Xu, X.M., Carlson, B.A., Gladyshev, V.N., and Hatfield, D.L. (2006). Thioredoxin reductase 1 deficiency reverses tumor phenotype and tumorigenicity of lung carcinoma cells. *J. Biol. Chem.* **281**, 13005–13008.
- Yoshizawa, S., and Böck, A. (2009). The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta - Gen. Subj.* **1790**, 1404–1414.
- Yuan, J., Palioura, S., Salazar, J.C., Su, D., O'Donoghue, P., Hohn, M.J., Cardoso, A.M., Whitman, W.B., and Söll, D. (2006). RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 18923–18927.
- Yue, Z., Maldonado, E., Pillutla, R., Cho, H., Reinberg, D., and Shatkin, A.J. (1997). Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12898–12903.
- Zavacki, A.M., Mansell, J.B., Chung, M., Klimovitsky, B., Harney, J.W., and Berry, M.J. (2003). Coupled tRNA(Sec)-dependent assembly of the selenocysteine decoding apparatus. *Mol Cell* **11**, 773–781.
- Zhang, R., So, B.R., Li, P., Yong, J., Glisovic, T., Wan, L., and Dreyfuss, G. (2011). Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell* **146**, 384–395.
- Zhao, H., Liang, D., Grossman, H.B., and Wu, X. (2005a). Glutathione peroxidase 1 gene polymorphism and risk of recurrence in patients with superficial bladder cancer. *Urology* **66**, 769–774.
- Zhao, J., Hyman, L., and Moore, C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**, 405–445.
- Zhao, R., Davey, M., Hsu, Y.-C., Kaplanek, P., Tong, A., Parsons, A.B., Krogan, N., Cagney, G., Mai, D., Greenblatt, J., et al. (2005b). Navigating the Chaperone Network: An Integrative Map of Physical and Genetic Interactions Mediated by the Hsp90 Chaperone. *Cell* **120**, 715–727.
- Zhao, R., Kakiyama, Y., Gribun, A., Huen, J., Yang, G., Khanna, M., Costanzo, M., Brost, R.L., Boone, C., Hughes, T.R., et al. (2008). Molecular chaperone Hsp90 stabilizes Pih1/Nop17 to maintain R2TP complex activity that regulates snoRNA accumulation. *J Cell Biol* **180**, 563–578.
- Zhou, Z., Sun, X., Zou, Z., Sun, L., Zhang, T., Guo, S., Wen, Y., Liu, L., Wang, Y., Qin, J., et al. (2010). PRMT5 regulates Golgi apparatus structure through methylation of the golgin GM130. *Cell Res.* **20**, 1023–1033.
- Zhuo, P., and Diamond, A.M. (2009a). Molecular mechanisms by which selenoproteins affect cancer risk and progression. *Biochim. Biophys. Acta - Gen. Subj.* **1790**, 1546–1554.
- Zhuo, P., and Diamond, A.M. (2009b). Molecular mechanisms by which selenoproteins affect cancer risk and progression. *Biochim Biophys Acta*.
- Zhuo, H., Smith, A.H., and Steinmaus, C. (2004). Selenium and lung cancer: A quantitative analysis of heterogeneity in the current epidemiological literature. *Cancer Epidemiol. Biomarkers Prev.* **13**, 771–778.
- Zuberek, J., Kubacka, D., Jablonowska, A., Jemielity, J., Stepinski, J., Sonenberg, N., and Darzynkiewicz, E. (2007). Weak binding affinity of human 4EHP for mRNA cap analogs. *RNA* **13**, 691–697.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415.

Etude du mécanisme d'hyperméthylation de la coiffe des ARNm de sélénoprotéines et impact sur leur traduction

Les sélénoprotéines sont les effecteurs majeurs de la défense antioxydante des cellules. La synthèse des sélénoprotéines fait appel à un mécanisme de recodage traductionnel complexe. En effet, l'acide aminé sélénocystéine est incorporé de manière co-traductionnelle dans les sélénoprotéines en réponse à un codon UGA^{Sec}, habituellement signe de terminaison de la traduction. Chez les mammifères, ce processus est conditionné par le recrutement de facteurs spécialisés dans la région 3' non codante des ARNm des sélénoprotéines, au niveau d'un élément de structure secondaire appelé SECIS (Selenocysteine Insertion Sequence). Notre équipe a montré que le mécanisme d'assemblage des mRNP de sélénoprotéines, prérequis à leur traduction, obéit aux mêmes règles que celui de petits ARN non-codants. Mon travail de thèse a contribué à mettre en évidence l'existence d'un mécanisme particulier de maturation de la coiffe en 5' des ARNm de sélénoprotéines. Nous avons montré que certains ARNm de sélénoprotéines possèdent une coiffe hyperméthylée m₃^{2,2,7}G à la manière des sn et snoARN et ne sont pas reconnus efficacement par le facteur canonique d'initiation de la traduction eIF4E. Nous avons déterminé le mécanisme de biogenèse de cette coiffe, et montré qu'il fait appel à l'enzyme Tgs1 (Triméthyl-guanosine synthase), recrutée via la protéine SMN, à la manière des snARN, ou via Nop58 comme dans le cas des snoARN. Nous avons également montré que les ARNm de sélénoprotéines coiffés m₃^{2,2,7}G sont traduits *in vivo*. L'impact de la coiffe m₃^{2,2,7}G, mais également du codon UGA^{Sec} et de l'élément SECIS, sur le recrutement des ribosomes sur les ARNm de sélénoprotéines a été analysé *in vitro*. L'ensemble de mes résultats de thèse va dans le sens de l'existence d'un mécanisme d'initiation atypique de la traduction des ARNm de sélénoprotéines. Nous avons montré que ce mécanisme requiert la présence d'une GTPase supplémentaire par rapport au processus d'initiation canonique et que la région 3'UTR est essentielle au recrutement des ribosomes sur l'ARNm de SelM. Enfin, nous avons identifié un site de liaison de SBP2 dans la phase codante de l'ARNm à proximité du codon UGA^{Sec}. J'ai également développé des méthodes de purification des ARNm de sélénoprotéines qui permettront d'identifier l'ensemble des facteurs nécessaires à leur traduction et de résoudre la structure tridimensionnelle d'un complexe de recodage par cryo-microscopie électronique. L'objectif est d'élucider l'ensemble des étapes de la traduction et du recodage des ARNm de sélénoprotéines.

Etude du mécanisme d'hyperméthylation de la coiffe des ARNm de sélénoprotéines et impact sur leur traduction

Résumé

La synthèse des sélénoprotéines fait appel à un mécanisme de recodage traductionnel d'un codon UGA^{Sec}. Chez les mammifères, ce processus est conditionné par le recrutement de facteurs spécialisés dans la région 3'UTR des ARNm de sélénoprotéines, au niveau d'une tige-boucle appelée SECIS. Lors de ma thèse, nous avons montré que certains ARNm de sélénoprotéines possèdent une coiffe hyperméthylée m₃^{2,2,7}G à leur extrémité 5', à la manière d'ARN non-codants, et ne sont pas reconnus efficacement par le facteur canonique d'initiation de la traduction eIF4E. Nous avons déterminé le mécanisme de biogenèse de cette coiffe qui fait appel à la Triméthyl-guanosine synthase, et avons montré que les ARNm de sélénoprotéines coiffés m₃^{2,2,7}G sont traduits *in vivo*. Par ailleurs, nos résultats indiquent que l'initiation de la traduction des ARNm de sélénoprotéines suit un mécanisme atypique qui ferait intervenir des éléments structuraux de l'ARNm, la région 3'UTR et une GTPase encore inconnue.

Mots clés : Traduction des sélénoprotéines, recodage, assemblage de RNP, coiffe m₃^{2,2,7}G

Résumé en anglais

Selenoprotein synthesis requires co-translational recoding of in-frame UGA codons. In mammals, this process is governed by the recruitment of dedicated factors on a hairpin structure, called SECIS, in the 3'UTR of selenoprotein mRNAs. During my PhD, we showed that several selenoprotein mRNAs bear a hypermethylated m₃^{2,2,7}G cap and undergo a similar 5' end maturation pathway than non-coding RNAs. This cap biogenesis mechanism involves the enzyme Trimethyl-guanosine synthase, m₃^{2,2,7}G capped selenoprotein mRNAs are not efficiently recognized by the canonical translation initiation factor eIF4E but are translated *in vivo*. Furthermore, our results suggest the existence of an atypical mechanism of translation initiation for selenoprotein mRNAs. This process involves structural RNA determinants, the 3'UTR region and a GTPase that remains to be identified.

Keywords: Selenoprotein translation, recoding, RNP assembly, m₃^{2,2,7}G cap