



HAL
open science

Recherche des mécanismes impliqués dans les dérégulations de l'épissage alternatif à l'origine de la progéria et étude du rôle de l'étape d'épissage dans les changements globaux d'expression des gènes en réaction au choc thermique

Valentin Vautrot

► **To cite this version:**

Valentin Vautrot. Recherche des mécanismes impliqués dans les dérégulations de l'épissage alternatif à l'origine de la progéria et étude du rôle de l'étape d'épissage dans les changements globaux d'expression des gènes en réaction au choc thermique. Médecine humaine et pathologie. Université de Lorraine, 2013. Français. NNT : 2013LORR0321 . tel-01331636

HAL Id: tel-01331636

<https://theses.hal.science/tel-01331636>

Submitted on 14 Jun 2016

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

Mention : « Sciences de la Vie et de la Santé »

par **Valentin VAUTROT**

**Recherche des mécanismes impliqués dans les
dérégulations de l'épissage alternatif à l'origine de la
progéria et étude du rôle de l'étape d'épissage dans les
changements globaux d'expression des gènes en réaction
au choc thermique**

Le 12 décembre 2013

Membres du jury :

Rapporteurs :	Mme C. CARON	Chargée de recherche INSERM, Grenoble
	M. M. DUTERTRE	Chargé de recherche INSERM, Paris
Examineurs :	M. J. TAZI	Professeur, Université de Montpellier 2
	M. B. CHARPENTIER	Professeur, Université de Lorraine, Nancy
	Mme C. BRANLANT	Directeur de recherche CNRS, Nancy directeur de thèse
	Mme I. BEHM-ANSMANT	Chargée de recherche CNRS, Nancy co-directeur de thèse
Membres invités:	M. L. VALLAR	Docteur, CRP-santé, Luxembourg
	M. G. BIAMONTI	Docteur, Istituto di Genetica Molecolare, Pavia, Italie

**Laboratoire d'Ingénierie Moléculaire et Physiopathologie Articulaire (IMoPA)
UMR 7365 CNRS-Université de Lorraine
Bâtiment Biopôle, Faculté de médecine de Nancy, Vandoeuvre-les-Nancy, France**

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude envers les Docteurs Cécile Caron et Martin Dutertre pour l'intérêt qu'ils ont porté à ce travail et pour avoir accepté d'être rapporteurs de ce manuscrit. Je remercie aussi vivement les Docteurs Jamal Tazi, Laurent Vallar et Giuseppe Biamonti qui ont accepté de juger mon travail et pour le travail de collaboration que nous avons réalisé ensemble, ainsi que le Docteur Bruno Charpentier, et le Dr James Stévenin pour sa participation aux discussions.

Je tiens à remercier Isabelle Behm-Ansmant. Je pense sincèrement que j'aurais difficilement pu avoir une meilleure encadrante de thèse, car elle fait partie des gens qui savent vous tirer vers le haut à la fois grâce à ses connaissances scientifiques étendues, son propre attrait pour la science et les discussions intéressantes qui en découlent, mais aussi par ses qualités humaines. C'est une personne entre autres attentionnée, compréhensive et solide, avec qui on s'amuse aussi et que j'admire beaucoup, notamment car elle met un point d'honneur à prendre le temps d'aider les personnes qui en ont besoin, parfois même au détriment d'elle-même. Ces raisons à elles seules seraient suffisantes pour surmonter toutes les difficultés que l'on rencontre en faisant une thèse. Ce petit paragraphe ne saurait rendre justice à la gratitude que j'éprouve, mais de tout cœur merci.

Je suis aussi tout particulièrement reconnaissant envers Mme Branlant, à qui je dois beaucoup et auprès de qui j'ai énormément appris, notamment grâce à ses qualités de chef, sa rigueur et sa précision. Elle fait également partie des personnes qui vous aident à vous améliorer sans arrêt, dans le travail comme dans le reste. Je suis reconnaissant pour tout le temps qu'elle a accordé à mon encadrement malgré toutes ses responsabilités, et le support qu'elle m'a apporté tout au long de mon séjour au laboratoire MAEM-AREMS-IMoPA.

Je remercie également Christelle Aigueperse et Florence Oillo-Blanloeil pour toute l'aide très précieuse qu'elles m'ont apporté dans les expériences, pour leur gentillesse, et leurs personnalités qui ont aidé à faire de mon environnement de travail quotidien un endroit agréable. Pour cette même raison et aussi pour leur aide lors des discussions et les réunions, je remercie également les anciens et actuels membres du Box2/Labo4, parmi lesquels Lilia, Audrey, Virginie, Yoann, Georges, Nathalie, et en particulier Jean-Vincent, un précieux ami. Je n'oublie pas également Véronique Senty-Ségault et Nicolas Rolland ainsi que les autres membres de l'ex Box1, Christine, JBF, Valérie, Benjamin, qui ont su m'aider et me supporter lors de mon arrivée au laboratoire et avec qui j'ai passé de bons moments.

Je remercie nos collaborateurs, notamment A. Metz, pour son aide dans les expériences de RNA-FISH et J. Tazi et I. Lopez pour leur travail, leurs discussions et leur support précieux dans la progéria et leur accueil au laboratoire IGMM de Montpellier. Je remercie également J. Stevenin pour ses travaux et son aide précieuse lors de l'étude de la progéria. Je tiens à remercier les membres du CRP-Santé du Luxembourg, en particulier L. Vallar, N. Nicot, T. Kaoma et A. Muller pour toute leur aide dans les

expériences de puces haut débit, pour leur accueil et leur gentillesse. Je remercie aussi G. Biamonti pour le clone pHur98 et les conseils et discussions quant au développement du projet ARN sat III, ainsi que H. Kempf pour son aide lors d'une discussion sur le RNA-FISH.

Je remercie également les membres du laboratoire IMoPA que j'ai côtoyé pendant ma thèse et en qui j'ai pu parfois trouver de très bon amis, comme Anne-Sophie et les nombreuses autres personnes que j'ai envie de citer comme Alexandra, Anne, Guillaume, Thibaut, Antoine, Maxime, Thomas, Samia, Decebal, Raphaël, Clémence et plein d'autres, de la partie enzymologie comme ARN et aussi du 2e étage.

Enfin, je remercie les membres de ma famille pour leur soutien constant et leurs encouragements pendant toute ces années, et tout particulièrement ma femme, pour sa grande patience, qui m'a aussi offert un soutien et des encouragements inestimables, que j'aime très très fort et qui m'a permis d'aller jusqu'au bout. Là encore ces quelques mots ne suffisent pas à exprimer ce que je ressens, mais de toute façon elle le sait elle est juste à côté de moi à l'heure où j'écris ceci.

Ce sont des remerciements peut-être un peu formels (du coup je mets un smiley pour rattraper ça -> :D voilà) et peut-être aussi incomplets. Il est important pour moi malgré tout que toutes les personnes qui m'ont aidé de près ou de loin lors de cette thèse, et que j'aurais oublié pour cause de fatigue et d'étourderie aussi (une de mes nombreuses qualités), sachent que je les remercie chaleureusement.

Que la force soit avec vous.

ABREVIATIONS ET ACRONYMES

ABREVIATIONS ET ACRONYMES

°C : degré Celsius

aa : acide(s) amine(s)

ADN : Acide DésoxyriboNucléique

ADNc : ADN complémentaire

ARN : Acide RiboNucléique

ARNm : ARN messenger

ARNr : ARN ribosomique

ARNt : ARN de transfert

BSA : "Bovine Serum Albumin" (Albumine de sérum bovin)

CPD : "Cyclobutamine Pyrimidine Dimers" (dimères cyclobutaniques de pyrimidine)

cSG : "cytoplasmic Stress Granules" (granules de stress cytoplasmiques)

C-ter : C-terminal(e)

Da : Dalton

DAVID : "Database for Annotation, Visualization and Integrated Discovery"

DSB : "Double Strand Break" (cassures double brin)

EDMD : "Emery-Dreyfuss Muscular Dystrophy" (dystrophie musculaire d'Emery-Dreyfuss)

EGF : "Epidermal Growth Factor"

EJC : "Exon Junction Complex" (complexe des jonctions exon)

EMSA : "Electrophoretic Mobility Shift Assay" (expériences de shift)

ESE : "Exonic Splicing Enhancer"

ESS : "Exonic Splicing Silencer"

EST : "Expressed Sequence Tag"

FISH : "Fluorescence *In Situ* Hybridization" (hybridation de fluorescence *in situ*)

FITC : Fluorescéine IsoThioCyanate

FRAP : "Fluorescence Recovery After Photobleaching"

GO : "Gene Ontology")

GPX : Glutathion Peroxidase(s)

GSH : Glutathion

HAT : Histone AcétylTransférase

HDAC : Histone Déacétylase

HGPS : "Hutchinson-Gilford Progeria Syndrome"

HJAY : "Human Junction Array"

hnRNP : "heteronuclear RiboNucleic Particle"

hsiPG : "heat shock induced Perichromatin Granules"

HSP : "Heat Shock Protein"

HSF : "Heat Shock Factor"
IF : "Intermediate Filament"
IRES : "Internal Ribosome Entry Site"
ISE : "Intronic Splicing Enhancer(s)"
ISS : "Intronic Splicing Silencer(s)"
ITAF : "IRES Trans-Acting Factors"
kb : kilobase
kDa : kilo Dalton
KEGG : "Kyoto Encyclopedia of Genes and Genomes"
LNA : "Locked Nucleic Acid" ("acide nucléique verrouillé")
MBP : "Maltose Binding Protein"
MCS : "Multi Cloning Site" (site de clonage)
MEC : Matrice Extracellulaire
miR, miRNA : microRNA
MMS : méthylméthanesulfonate
NADPH : Nicotinamide Adenine Dinucleotide Phosphatase
NHEJ : "Non Homologous End Joining"
NLS : "Nuclear Localisation Signal" (signal de localisation nucléaire)
NPC : "Nuclear Pore Complex" (pores nucléaires)
NRS : "Nuclear Retention Signal", signal de rétention nucléaire
nSB : "nuclear Stress Bodies"
nSG : "nuclear Stress Granules"
nt : nucléotide(s)
N-ter : N-terminal(e)
OMIM : "Online Mendelian Inheritance in Man" (base de données)
PCR : "Polymerase Chain Reaction"
PDGF : "Platelet Derived Growth Factor"
PG : "Perichromatin Granules"
pol : polymérase
pré-ARNm : Acide Ribo Nucléique pré-messenger
PRX : Peroxiredoxine(s)
RE : Reticulum Endoplasmique
RISC : RNA Induced Silencing Complex
RMN : Résonance Magnétique Nucléaire
RNP : "Ribo Nucleic Particle"
RNS : "Reactive Nitrogen Species"
ROS : "Reactive Oxygen Species"

RT : Rétrotranscription

RT-PCR : Rétrotranscription suivie d'une PCR

sat : satellite

scaRNA : scavenger RNA

SELEX : "systematic evolution of ligands through exponential enrichment"

siRNA : "small interfering RNA", petits ARN interférents

SMA : "Spinal Muscular Atrophy"

SMN : "Survival Motor Neuron" (protéine ou complexe SMN)

SNB : "Sam68 Nuclear Bodies"

SNP : "Single Nucleotide Polymorphism"

snRNA : "small nuclear RiboNucleic Acid"

snRNP : "small nuclear RiboNucleic Particle"

snoRNA : small nucleolar RNA

SR : "Serin-Arginin Rich"

T_m : "melting Température" (température de fusion)

TNF- α : "Tumor Necrosis Factor- α "

TPR : "Tetratricopeptide Repeats" (domaine protéique)

TRX : Thioredoxine(s)

UPR : "Unfolded Proteins Response"

UsnRNA : "Uridine-rich Small Nuclear Ribonucleic Acids"

UsnRNP : "Uridine-rich Small Nuclear Ribonucleic Particle"

WB : "Western Blot" ou "Western Blotting"

WT : "Wild Type" (type sauvage)

SOMMAIRE

AVANT PROPOS
PARTIE 1 - PROGERIA ET EPISSAGE (Introduction) 1
I. EPISSAGE ET EPISSAGE ALTERNATIF DES PRE-ARNm CHEZ L'HOMME..... 1

I.1. LA REACTION D'ÉPISSAGE ET LES SEQUENCES DE DEFINITION DES INTRONS	1
I.2. FORMATION DES COMPLEXES SPLICEOSOMAUX D'ÉPISSAGE	3
<i>I.2.1. Les particules UsnRNP</i>	<i>3</i>
<i>I.2.2. La formation des complexes spliceosomaux.....</i>	<i>4</i>
I.2.2.a. Etapes d'assemblage du spliceosome.....	4
I.2.2.b. Complexité et dynamisme du spliceosome.....	5
<i>I.2.3. Rôle des protéines SR et SR-like dans la formation des complexes spliceosomaux.....</i>	<i>8</i>
I.2.3.a. Structure des protéines SR.....	9
I.2.3.b. Autres protéines à domaines SR.....	11
I.2.3.c. Interaction des protéines SR avec l'ARN	12
I.2.3.d. Rôle des protéines SR dans l'épissage	12
I.2.3.e. Localisation des protéines SR.....	14
<i>I.2.4. Autres fonctions des protéines SR</i>	<i>15</i>
I.3. LIENS EXISTANTS ENTRE L'ÉPISSAGE ET L'ÉTAT DE LA CHROMATINE, LA TRANSCRIPTION ET L'EXPORT DES ARNm	16
<i>I.3.1. Couplage fonctionnel entre transcription et épissage des ARNm</i>	<i>16</i>
<i>I.3.2. Couplage cinétique entre transcription et épissage des ARNm</i>	<i>17</i>
<i>I.3.3. Liens entre l'organisation de la chromatine et l'épissage</i>	<i>18</i>
I.3.3.a. Influence de l'état de la chromatine et des modifications épigénétiques sur la transcription.....	18
I.3.3.b. L'organisation de la chromatine intervient dans la définition des exons.....	19
I.3.3.c. La chromatine intervient dans le recrutement des facteurs d'épissage.....	19
<i>I.3.4. Couplage de l'épissage avec la stabilité et le transport des ARNm.....</i>	<i>20</i>
I.4. L'ÉPISSAGE ALTERNATIF.....	22
<i>I.4.1. Les différents types d'événements d'épissage alternatif.....</i>	<i>22</i>
<i>I.4.2. Séquences cis-régulatrices de l'épissage.....</i>	<i>24</i>
<i>I.4.3. Les protéines hnRNP</i>	<i>24</i>
I.4.3.a. Domaines d'interaction des protéines hnRNP avec leurs ARN cibles.....	25
I.4.3.b. Localisation des protéines hnRNP dans la cellule	27
<i>I.4.4. Rôle des protéines hnRNP dans l'épissage alternatif.....</i>	<i>28</i>
<i>I.4.5. Régulation de l'épissage alternatif par les facteurs activateurs et inhibiteurs de l'épissage.....</i>	<i>30</i>
I.4.5.a. Antagonisme entre les protéines SR et hnRNP dans l'épissage alternatif.....	30
I.4.5.b. Interdépendance complexe entre les différents facteurs de l'épissage pour la régulation de l'épissage alternatif	32
<i>I.4.6. Influence de la structure secondaire des ARN.....</i>	<i>33</i>
I.4.6.a. Implication de la structure secondaire dans l'inhibition de l'épissage	33
I.4.6.b. Implication de la structure secondaire dans l'activation de l'épissage	35
I.5. EPISSAGE ALTERNATIF ET PATHOLOGIES	37

1.5.1. Maladies génétiques liées à des mutations ponctuelles affectant directement l'épissage.....	37
1.5.2. Maladies génétiques affectant des facteurs d'épissage ou des protéines impliquées dans la biogenèse des spliceosomes.....	39
II. LA PROGERIA, UNE MALADIE GENETIQUE RARE AFFECTANT L'ÉPISSAGE DU PRE-ARNM DE LA LAMINE A	41
II.1. LE SYNDROME DE HUTCHINSON-GILFORD OU PROGERIA.....	41
II.2. PRESENTATION DE LA LAMINE NUCLEAIRE OU LAMINA	42
II.2.1. Composition, structure et organisation	42
II.2.1.a. Description de la lamine nucléaire.....	42
II.2.1.b. Les différents types de lamines	42
II.2.1.c. Structure, propriétés de polymérisation et de désassemblage des lamines.....	44
II.2.1.d. Autres composants de la lamine nucléaire	47
II.2.2. Laminopathies et syndromes progéroïdes.....	49
II.3. LAMINE A ET PROGERIA	50
II.3.1. Organisation du gène LMNA et des différentes isoformes protéiques codées par ce gène.....	50
II.3.2. La maturation de la lamine A et son intégration dans la lamina	51
II.3.3. Mutations du gène LMNA à l'origine de la progéria	52
II.4. LES FONCTIONS DE LA LAMINE A ET LES PERTURBATIONS ENTRAINEES PAR L'EXPRESSION DE PROGERINE.....	55
II.4.1. Atteintes à l'organisation du noyau et à sa dynamique.....	55
II.4.2. Perturbation du métabolisme de l'ADN et du cycle cellulaire	58
II.4.2.a. Perturbation de la réplication de l'ADN.....	58
II.4.2.b. Instabilité génomique	58
II.4.2.c. Ralentissement du cycle cellulaire, et entrée en sénescence des cellules	60
II.4.3. Influence de la progérine sur l'expression des gènes.....	61
II.4.4. Liens avec le phénotype observé à l'échelle de l'organisme.....	62
II.4.4. Potentialités thérapeutiques face à la progéria	63
II.4.5. La progéria en tant que modèle d'étude du vieillissement cellulaire.....	65
II.5. LIENS EXISTANT ENTRE LA PROGERIA, LE VIEILLISSEMENT, ET LES MODELES CELLULAIRES DU STRESS.....	65
II.5.1. Progéria et stress oxydant.....	65
II.5.2. Progéria et choc thermique.....	66
PARTIE 1 - PROGERIA ET EPISSAGE (Objectifs).....	68
PARTIE 1 - PROGERIA ET EPISSAGE (Résultats)	70
III. ETUDE EXPERIMENTALE DES MECANISMES A L'ORIGINE DES REGULATIONS ET DEREGULATIONS DE L'EPISSAGE DU TRANSCRIT LMNA EN LIEN AVEC LA PROGERIA	70
III.1. ETUDE DES MECANISMES INDUISANT L'ACTIVATION DU SITE 5' SS PROGERINE EN PRESENCE DES MUTATIONS 1824C>U.....	70
III.1.1. Constat initial.....	70
III.1.2. Mise au point d'un système d'étude in vitro de l'épissage de l'intron 11 du transcrit LMNA.....	71

III.1.3. Etude de la structure secondaire en solution d'une grande région du transcrit LMNA contenant les sites 5'SS lamine A et progérine.....	74
III.1.3.a. Stratégie d'analyse de la structure secondaire.....	74
III.1.3.b. Structure secondaire adoptée par l'ARN et influence des mutations sur cette structure.....	76
III.1.4. Etude in vitro de l'effet des protéines SR sur l'activité des sites 5'SS lamine A et 5'SS progérine.....	80
III.1.4.a. Etude de l'association des protéines contenues dans un extrait nucléaire de cellules HeLa avec le site 5'SS progérine	80
III.1.4.c. Test de l'effet des protéines SRSF1 et SRSF6 sur l'utilisation des sites 5' progérine et lamine A in vitro.....	82
PUBLICATION N°1	
A CONSERVED SPLICING MECHANISM OF THE LMNA GENE CONTROLS PREMATURE AGING.	71
III.2. ETUDE DE L'INFLUENCE DES MUTATIONS 1868C>G ET 1822G>A SUR LES TAUX RELATIFS D'UTILISATION DES SITES 5'SS DU PRE-ARNm LMNA : PRESENTATION DE LA PUBLICATION N°2.	85
III.2.1. Analyse de l'effet sur la structure secondaire du pré-ARNm LMNA des mutations 1822G>A et 1868C>G	85
III.2.2. Identification des protéines SR impliquées dans la régulation de l'utilisation des sites 5' progérine et LA Δ 35 en absence et en présence des mutations 1822G>A et 1868C>G	86
MANUSCRIT PUBLICATION N°2	
IDENTIFICATION OF SRSF5 AS A REGULATOR OF LMNA EXON 11 SPLICING IN THE CONTEXT OF PROGERIA INDUCING MUTATIONS.....	85
III.3. CONCLUSIONS ET PERSPECTIVES	89
PARTIE 2 - STRESS THERMIQUE ET EPISSAGE (Introduction)	92
I. STRESS THERMIQUE ET EPISSAGE	92
I.1. DEFINITION ET ORIGINES D'UN STRESS THERMIQUE	92
I.2. DOMMAGES CAUSES PAR LE STRESS THERMIQUE	93
I.3. LES PRINCIPAUX MECANISMES DE REPONSE AU STRESS THERMIQUE	95
I.3.1. Protéines dont l'expression est induite par le stress thermique	95
I.3.2. Les protéines du choc thermique HSP	96
I.3.3. Principaux gènes et isoformes des protéines HSP et leurs fonctions associées	98
I.3.4. Régulation de l'expression des protéines de réponse au choc thermique par les facteurs HSF lors d'un stress	99
I.3.5. Importance des protéines de transduction du signal dans la réponse au stress thermique	101
I.3.6. Arrêt du cycle cellulaire et apoptose induits par le choc thermique	104
I.4. VALORISATION DE L'ENERGIE ANABOLIQUE APRES UN STRESS THERMIQUE.....	105
I.4.1. Granules de stress cytoplasmiques et blocage transitoire de la traduction de certains ARNm	105
I.4.1.a. Formation des granules de stress cytoplasmiques.....	105
I.4.1.b. Traduction indépendante de la coiffe et stress thermique	108

1.4.2. Blocage de l'épissage de certains ARNm.....	108
1.4.3. Les régulations de l'épissage et de l'épissage alternatif suite à divers stress.....	111
1.5. GRANULES DE STRESS NUCLEAIRES ET CONTROLE DE L'EPISSAGE ALTERNATIF.....	112
1.5.1. ADN et ARN satellites péricentromériques.....	112
1.5.2. Description des ARN et des ADN satellites III.....	113
1.5.3. Transcription des ARN satellite III par HSF1 ou TonEBP/NFAT5.....	116
1.5.4. Organisation et fonction des nSB.....	118
1.5.5. Protéines relocalisées dans les nSB.....	122
1.5.5.a. Protéines hnRNP et interactions avec la matrice nucléaire.....	122
1.5.5.b. Réseau d'interactions protéiques dans les nSB.....	123
1.6. LIENS ENTRE LES MECANISMES DE DEFENSE AU STRESS THERMIQUE ET VIEILLISSEMENT.....	126
PARTIE 2 - STRESS THERMIQUE ET EPISSAGE (Objectifs).....	127
PARTIE 2 - STRESS THERMIQUE ET EPISSAGE (Résultats).....	128
II. ETUDE DES ARN SATELLITES III ET RECHERCHE DE LEURS PARTENAIRES PROTEIQUES.....	128
II.1. ANALYSE DE LA STRUCTURE SECONDAIRE DE L'ARN SATELLITE III.....	128
II.1.1. Stratégie d'analyse et mise au point.....	128
II.1.2. Analyse de la structure secondaire de l'ARN sat III pHur98.....	129
II.2. RECHERCHE DE PARTENAIRES PROTEIQUES DES ARN SAT III.....	131
II.2.1. Utilisation de programmes informatiques pour prédire la présence de sites de fixations des protéines SR aux ARN sat III.....	131
II.2.2. Purification et identification des partenaires d'un ARN sat III in vitro.....	132
II.3. CONFIRMATION PAR WESTERN BLOT DE LA PRESENCE DES PROTEINES IDENTIFIEES DANS LES FRACTIONS DE PURIFICATION.....	137
II.4. ETUDE DE LA LOCALISATION DES PROTEINES CANDIDATES DANS DES CELLULES HE LA SOUMISES A UN CHOC THERMIQUE.....	139
II.4.1. Mise au point des conditions de choc thermique.....	139
II.5.2. Test de la présence des protéines purifiées avec l'ARN sat III in vitro dans les granules de stress nucléaires.....	142
II.5.2.a. Mise au point de la détection des granules de stress nucléaires par RNA-FISH.....	142
II.5.2.b. Test de la présence des partenaires protéiques potentiels de l'ARN sat III dans les granules de stress nucléaires par immunofluorescence.....	146
II.6. TEST PRELIMINAIRES DE L'INTERACTION DES PROTEINES SRSF1, HNRNP Q, HNRNP H ET HNRNP U AVEC LES ARN SAT III IN CELLULO.....	153
II.7 BILAN ET PERSPECTIVES.....	154
II.7.1 Structuration des ARN sat III.....	156
II.7.2 Interaction des ARN sat III avec des protéines.....	157
III. ETUDES DES CONSEQUENCES DU CHOC THERMIQUE SUR L'EPISSAGE ALTERNATIF.....	161
III.1. DESCRIPTION DES PUCES A ADN UTILISEES SPECIFIQUES DES JONCTIONS EXONIQUES.....	161

III.2. NATURE DES ECHANTILLONS ANALYSES SUR LES PUCES A JONCTIONS EXON-EXON	163
III.3. DONNEES CONCERNANT LES VARIATIONS D'EXPRESSION DES GENES	164
<i>III.3.1. Identification préliminaire des gènes induits et réprimés</i>	164
<i>III.3.2. Analyse des voies de régulation et des groupes fonctionnels auxquels appartiennent les gènes dont l'expression est modifiée lors d'un choc thermique</i>	166
III.3.2.a. Analyses des voies de régulations affectées par emploi des outils d'analyse des bases de données KEGG et REACTOME.....	166
III.3.2.b. Regroupement fonctionnel des termes ontologiques grâce au logiciel DAVID	167
III.3.2.d. Analyse par le logiciel Ingenuity.....	169
III.4. DONNEES DES PUCES CONCERNANT L'EPISSAGE ALTERNATIF	171
<i>III.4.1. Analyse globale des transcrits soumis à l'épissage alternatif</i>	171
<i>III.4.2. Confirmation par RT-PCR des événements d'épissage alternatif de type exon cassette détectés après un choc thermique</i>	173
III.4.2.a. Confirmation de la présence des d'événements d'épissage alternatif décrits comme étant associés au stress thermique	174
III.4.2.b. Validation d'événements d'épissage encore non décrits comme associés à un choc thermique et concernant les ARNm de protéines impliquées dans le transport nucléocytoplasmique.....	178
III.4.2.c. Validation de la variation de l'épissage du pré-ARNm Nop58.....	181
III.4.2.d. Validation de variations d'épissage touchant les pré-ARNm de protéines impliquées dans l'inhibition de la traduction et la dégradation des ARNm	183
III.4.2.d.1. Présentation des protéines impliquées dans la formation des P-bodies et des granules de stress cytoplasmiques.....	183
III.4.2.d.2. Validation des variations d'épissage des pré-ARNm des protéines du complexe de déadénylation par RT-PCR.....	186
III.4.2.d.3. Validation des variations d'épissage des pré-ARNm des protéines des P-bodies 4E-T et GW182.....	187
<i>III.4.3. Etude de l'influence de la présence de l'exon 13 dans les ARNm GW182 sur les propriétés de la protéine GW182</i>	190
III.4.3.a. Influence de l'inclusion de l'exon 13 sur la localisation de GW182	190
III.4.3.b. Influence de la présence de la séquence codée par l'exon 13 sur les fonctions de GW182	192
III.5. ETUDE DES MICROARN REGULES APRES UN CHOC THERMIQUE PAR ANALYSE HAUT DEBIT.....	195
III.6. CONCLUSIONS ET PERSPECTIVES	199
<i>III.6.1. Analyse globale de l'augmentation ou la diminution de l'expression des gènes après un choc thermique</i>	199
<i>III.6.2. Implications de la synthèse de nouvelles isoformes des protéines des P-bodies et des cSG dans la réponse au stress thermique</i>	200
<i>III.6.3. Analyse globale des nouveaux événements d'épissages apparaissant après un choc thermique</i>	202
<i>III.6.4. Analyse globale de la variation de l'expression des microARN après un choc thermique</i>	203

ANNEXE 1 : RESULTAT DE L'ANALYSE GRACE A CATEGORIZER DES TERMES GO ASSOCIES AUX GENES DONT L'EXPRESSION VARIE APRES LE CHOC THERMIQUE

ANNEXE 2 : RESULTAT DE L'ANALYSE EN AMONT DES DONNEES DE PUCES JONCTIONS EXON-EXON REALISEE AVEC LE LOGICIEL INGENUITY (P-VALUE<0,01)

ANNEXE 3 : REGROUPEMENT FONCTIONNEL GRACE AU LOGICIEL DAVID DES TERMES GO (BIOLOGICAL PROCESS) ASSOCIES AUX VARIATIONS D'EPISSAGE DETECTEES PAR PUCES A JONCTIONS EXON-EXON

ANNEXE 4 : REGROUPEMENT FONCTIONNEL GRACE AU LOGICIEL DAVID DES TERMES ASSOCIES AUX VARIATIONS D'EPISSAGE DETECTEES PAR PUCES A JONCTIONS EXON-EXON

MATERIEL ET METHODES	204
I. MATERIEL	204
I.1. SOUCHES BACTERIENNES, CELLULES HUMAINES ET PLASMIDES	204
<i>I.1.1. Souches bactériennes</i>	<i>204</i>
<i>I.1.2. Cellules humaines</i>	<i>204</i>
<i>I.1.3. Plasmides et constructions génétiques</i>	<i>204</i>
I.2. OLIGODESOXYRIBONUCLEOTIDES	206
I.3. MILIEUX DE CULTURE ET TAMPONS	215
I.4. ANTICORPS	217
I.5. MICROPUCES JONCTIONS EXON-EXON ET MICROARN	218
II. METHODES D'ANALYSE <i>IN VITRO</i>	225
II.1. INTRODUCTION DES MUTATIONS PAR MUTAGENESE DIRIGEE	225
II.2. PRODUCTION DE MATRICES POUR LA TRANSCRIPTION	226
II.3. PRODUCTION D'ARN PAR TRANSCRIPTION <i>IN VITRO</i>	226
II.4. ETUDE DE LA STRUCTURE SECONDAIRE DES ARN PAR EMPLOI DE SONDAS ENZYMATIQUES	227
<i>II.4.1. Marquage radioactif en 5' d'oligodésoxyribonucléotides ou d'ARN</i>	<i>227</i>
<i>II.4.2. Hydrolyse des ARN par les RNases T1, T2 et V1</i>	<i>228</i>
<i>II.4.3. Identification des positions modifiées par rétro-transcription des ARN non radiomarqués</i>	<i>228</i>
II.5. ETUDE DE L'INTERACTION DES PROTEINES AVEC LES ARN	230
<i>II.5.1. Purification par chromatographie d'affinité MBP-MS2</i>	<i>230</i>
II.5.1.a. Production et purification de la protéine de fusion MBP-MS2	230
II.5.1.b. Chromatographie d'affinité MBP-MS2	230
<i>II.5.2. Estimation des quantités de protéines par « western blot »</i>	<i>233</i>
<i>II.5.3. Expériences de retard sur gel</i>	<i>233</i>
II.6. EPISSAGE <i>IN VITRO</i> DE L'ARN LMNA ET DE SES VARIANTES	234
<i>II.6.1. Préparation des matrices de transcription par hydrolyse de plasmide</i>	<i>234</i>
<i>II.6.2. Production in vitro de transcrits uniformément radiomarqués</i>	<i>234</i>
<i>II.6.3. Epissage in vitro des transcrits uniformément radiomarqués</i>	<i>235</i>
III. METHODES D'ANALYSE DE CELLULES HUMAINES EN CULTURE	237
III.1. CULTURE ET TRANSFECTION DES CELLULES	237

III.1.1. Culture à 37°C et déroulement du choc thermique	237
III.1.2. Transfection des cellules.....	237
III.2. EXTRACTION DES ARN TOTAUX.....	237
III.3. ANALYSES TRANSCRIPTOMIQUES PAR PUCES A JONCTIONS EXON-EXON ET PUCES MICROARN	238
III.3.1. Marquage des ARN totaux et hybridation sur les puces	238
III.3.1.a. Hybridation sur les puces à jonctions exon-exon HJAY	238
III.3.1.b. Hybridation sur les puces microARN GeneChip® miRNA 3.0 Array	239
III.3.2. Protocole de traitement du signal et d'analyse des données des puces	239
III.3.2.a. Analyse des données des puces à jonctions exon-exon HJAY.....	239
III.3.2.b. Analyse des puces microARN GeneChip® miRNA 3.0 Array	240
III.4. RT-PCR ET RT-QPCR	240
III.4.1. Digestion des extraits d'ARN à la DNase RQ1	240
III.4.2. Réaction de rétrotranscription	240
III.4.3. Détection des produits rétrotranscrits par PCR point final.....	241
III.4.4. Détection des produits rétrotranscrits par qPCR	242
III.4.5. Détection des microARN par RT-qPCR TaqMan.....	243
III.5. MICROSCOPIE DE FLUORESCENCE	244
III.5.1. Visualisation des ARN satellite III après un stress thermique par RNA-FISH.....	244
III.5.1.a. Préparation des sondes biotinylées.....	244
III.5.1.b. Fixation des cellules et incubation avec les sondes	245
III.5.2. Visualisation des protéines par immunofluorescence	246
III.5.3. Montage des lamelles.....	246
III.6. EXPERIENCES DE TETHERING GW182	247
BIBLIOGRAPHIE	248

AVANT-PROPOS

Dans les cellules eucaryotes, les gènes nucléaires codant des protéines sont transcrits par l'ARN polymérase II sous forme d'ARN pré-messagers (pré-ARNm) non fonctionnels. La grande majorité de ces pré-ARNm comporte en effet des séquences excédentaires ou introns, entre les segments (exons) correspondants à l'ARN messager (ARNm) mature. Des séquences introniques sont également présentes dans des gènes codant des ARN stables (ARN de transfert, ARN ribosomiques), dans des gènes d'organites cellulaires (chloroplaste, mitochondrie), ainsi que dans certains gènes bactériens et d'archaea, mais le mécanisme d'élimination de ces introns est différent de celui utilisé pour l'élimination des introns des gènes nucléaires de protéines. Les introns de ces gènes sont éliminés au sein des pré-ARNm par un processus nucléaire en 2 étapes appelé épissage. Les pré-ARNm épissés, coiffés en 5' et polyadénylés en 3' constituent les ARNm qui sont ensuite transportés dans le cytoplasme où ils sont traduits en protéines.

Chez l'homme, la plupart des gènes contiennent au moins un intron et l'épissage constitue une étape clef de l'expression des gènes. En effet, le génome humain contient plus de 200 000 introns. Contrairement aux exons, dont la taille moyenne est d'environ 120 nt, la taille moyenne des introns est de 1800 nt, et varie de moins de 100 à plus de 700 000 nt (Deutsch et Long, 1999 ; Rowen et al., 2002)(pour revue, Hoskins and Moore, 2012). La reconnaissance des introns et des exons est donc au centre du mécanisme d'épissage, et leur nombre et leur diversité de taille reflètent toute la complexité et la précision de la machinerie cellulaire mise en place pour réaliser l'épissage. Ce processus est d'autant plus complexe qu'un mécanisme de régulation, nommé épissage alternatif, permet l'inclusion ou l'exclusion de certains exons dits alternatifs, ainsi que l'utilisation de sites d'épissage alternatifs. Aujourd'hui, le développement des approches à haut débit a révélé que chez l'homme environ 90 à 95% des transcrits de l'ARN pol II générés sont soumis à épissage alternatif et que ce phénomène qui permet de produire différents ARNm matures à partir d'un pré-ARNm donné participe largement à la diversification des protéines produites à partir de la trentaine de milliers de gènes connus chez l'homme (Nilsen and Graveley, 2010; Pan et al., 2008; Wang et al., 2008a). Il apparaît également que de nombreuses maladies génétiques humaines sont liées à des défauts d'épissage alternatif. En effet, entre 20 et 30 % des mutations à l'origine de maladies génétiques sont connues pour affecter l'épissage alternatif (pour revue, Corcos and Solier, 2005; Singh and Cooper, 2012).

Parmi ces maladies, on compte la progéria, ou syndrome de Hutchinson-Gilford, qui est une maladie génétique très rare dont les symptômes peuvent s'apparenter à un vieillissement accéléré. Cette maladie est causée par des mutations ponctuelles dans le gène *LMNA*, qui code la protéine lamine A, appartenant à une structure protéique de la face interne de l'enveloppe nucléaire appelée *lamina*. Ces mutations provoquent des défauts dans l'épissage alternatif du pré-ARNm de cette protéine, qui se traduisent par des défauts de maturation de la protéine LMNA ayant de graves conséquences sur le métabolisme et le phénotype de la cellule. Dans la première partie de ma thèse, je me suis intéressé aux mécanismes à l'origine des dérégulations de l'épissage causés par ces mutations du gène *LMNA*. Je

présenterai donc dans un premier temps l'état des connaissances sur les mécanismes généraux de l'épissage des ARN pré-messagers (pré-ARNm) chez les eucaryotes. Puis je présenterai les lames et décrirai la maladie progéria, et en particulier les mutations de la lame A qui en sont à l'origine. Dans un souci de simplicité, cette introduction sera immédiatement suivie des résultats que nous avons obtenus concernant les mécanismes de l'épissage mis en jeu dans la progéria.

Au cours de la deuxième partie de ma thèse, nous nous sommes intéressés à un tout autre phénomène lié aux régulations de l'épissage. En effet, des granules de stress nucléaires (les nSB) sont formés dans les cellules de patients atteints de progéria. Ces granules sont similaires à ceux retrouvés dans les cellules d'individus sains lorsqu'elles sont soumises à un choc thermique, mais aussi à différents autres types de stress tel que l'exposition aux métaux lourds. Ils sont formés autour d'ARN transcrits à partir de régions péri-centromériques des chromosomes, appelés ARN satellite III (ARN sat III), qui restent associés à leur site de transcription et semblent provoquer la séquestration de plusieurs facteurs impliqués dans l'épissage et l'épissage alternatif au sein des nSB. Cela changerait le taux de ces protéines régulatrices de l'épissage et serait susceptible de modifier le profil global d'épissage alternatif des transcrits dans la cellule. Afin de vérifier cet effet, nous avons utilisé le modèle du stress thermique couplé à une étude haut débit effectuée à l'aide de micropuces de type jonction exon. Dans cette seconde partie, je présenterai donc les résultats de notre étude de l'ARN satellite III et de ses partenaires ainsi que les résultats obtenus avec les micropuces exon jonction. Ces résultats seront précédés d'une partie introductive décrivant plus en détails les mécanismes de réponse au stress thermique et tout particulièrement ce qui est connu sur la formation des nSB dans la littérature.

PARTIE 1

PROGERIA ET EPISSAGE

(Introduction)

I. Epissage et épissage alternatif des pré-ARNm chez l'homme

I.1. La réaction d'épissage et les séquences de définition des introns

Les introns des pré-ARNm sont éliminés par 2 étapes successives de trans-estérifications (figure I-1). Au cours de la première étape, l'hydroxyle en 2' d'un résidu adénosine conservé situé dans l'intron (boîte de branchement) réalise une attaque nucléophile sur le phosphate situé à l'extrémité 5' de l'intron (site 5' donneur). Deux intermédiaires réactionnels en résultent, d'une part, l'exon 5', ayant une extrémité 3'OH libre, d'autre part, un intermédiaire réactionnel contenant l'exon 3' et l'intron structuré sous forme de lasso (figure I-1). Dans la seconde étape de trans-estérification, l'attaque nucléophile de l'extrémité 3' OH de l'exon 5', sur le phosphate situé à l'extrémité 3' de l'intron (site 3' accepteur d'épissage) conduit à la ligation des deux exons et à l'élimination de l'intron. Par conséquent, les produits finaux de la réaction d'épissage sont l'intron sous forme de lasso et l'ARNm mature (figure I-1) (Moore and Sharp, 1993)(pour revue, Chen and Cheng, 2012; Hoskins and Moore, 2012).

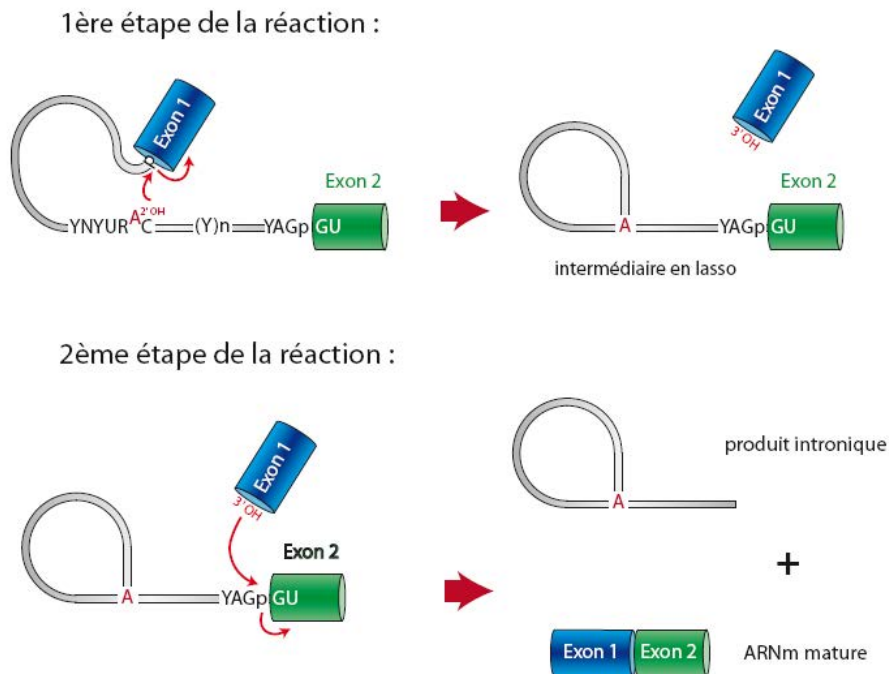


Figure I-1 : La réaction d'épissage. Représentation schématique des deux étapes de trans-estérification de la réaction. Dans la première étape, l'attaque du 2' OH du résidu A de la boîte de branchement sur le phosphate du résidu à l'extrémité 5' de l'intron provoque la formation d'un intermédiaire en lasso. Dans la seconde étape, l'attaque du 3' OH de l'exon 1 sur le phosphate à l'extrémité 3' de l'intron permet la libération de l'intron et de l'ARNm épissé. Y : Pyrimidine. R : Purine. N : un des 4 résidus A, C, G ou U.

Les introns contiennent un ensemble de séquences fortement conservées, au niveau des sites donneurs et accepteurs d'épissage, permettant leur reconnaissance par la machinerie d'épissage. Les séquences consensus de ces sites, obtenus par comparaison de séquences chez les vertébrés, sont données dans la figure I-2. La séquence consensus au niveau du site 5' donneur d'épissage est AG/GURAGU, où "/" représente la jonction exon 1/intron et R une purine. Les deux nucléotides **GU** à l'extrémité 5' de l'intron sont les plus conservés et sont très importants pour l'efficacité d'épissage. On peut noter qu'il existe également une autre catégorie d'introns, dits mineurs, peu représentés et dont la séquence consensus du site donneur est A/(G ou A)UAUCCU.

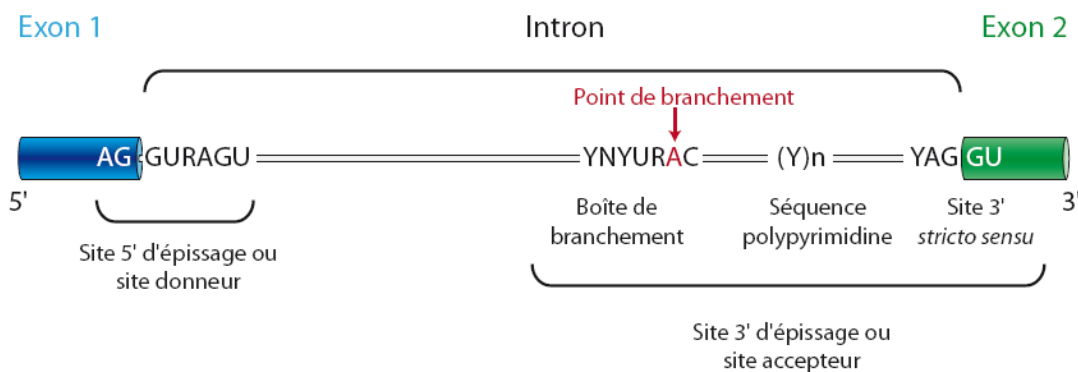


Figure I-2 : Séquences nécessaires pour l'épissage. Représentation schématique de l'organisation d'un intron situé entre deux exons. Les séquences conservées jouant un rôle important dans la mise en place de la machinerie d'épissage sont indiquées (Moore and Sharp, 1993)(pour revue, Chen and Cheng, 2012; Hoskins and Moore, 2012). Y : Pyrimidine. R : Purine. N : un des 4 résidus A, C, G ou U.

Quant aux sites accepteurs d'épissage, ils sont constitués par plusieurs éléments essentiels. La boîte de branchement est une séquence plus dégénérée de consensus YNYURAC (avec R = purine, Y = pyrimidine et N = A, C, G ou U) située environ 20 à 40 nt en amont du site 3' accepteur d'épissage proprement dit. Comme décrit ci-dessus, c'est le résidu **A** de cette boîte de branchement qui est le "point de branchement" à l'origine de la première attaque nucléophile de la réaction d'épissage, avec exceptionnellement l'utilisation possible d'un résidu autre qu'un A. La séquence consensus du site 3' d'épissage *stricto sensu* est YAG/GU, où Y est une pyrimidine, et "/" est la jonction intron/exon2. Dans le cas des introns mineurs, cette séquence est plus dégénérée (YA(C ou G)/N), par contre le consensus de la boîte de branchement, CCUUAACU (où le point de branchement est le deuxième A) est lui plus conservé que pour les introns majeurs. Enfin, entre la boîte de branchement et le site 3' d'épissage se trouve une séquence polypyrimidique contenant de 10 à 20 résidus C ou U, les enchaînements de résidus U conférant une plus forte efficacité du site.

Le taux de dégénérescence de toutes ces différentes séquences servant à définir les introns détermine l'efficacité intrinsèque de leur reconnaissance par la machinerie d'épissage et donc l'efficacité d'épissage aux sites correspondants.

I.2. Formation des complexes spliceosomaux d'épissage

I.2.1. Les particules UsnRNP

L'élimination des introns est catalysée par un macrocomplexe très dynamique, le spliceosome. Son assemblage fait intervenir des particules ribonucléoprotéiques appelées UsnRNP ou snRNP (pour "Uridine-rich Small Nuclear Ribonucleic Particle". Les snRNP U1, U2, U5 et U4·U6, les snRNP U4 et U6 étant associées dans une seule snRNP. Les complexes spliceosomaux comportent en plus de très nombreuses protéines. Comme nous le verrons plus en détails, le spliceosome est une structure complexe et dynamique dont la composition varie au cours des étapes d'épissage, ces changements de composition et de structure permettent la catalyse des 2 étapes successives de trans-esterification de la réaction d'épissage. Les 4 UsnRNP sont composées d'un ou deux petits ARN appelés UsnRNA (pour "Uridine-rich small nuclear ribonucleic acids"), associés à un ensemble de 7 protéines appartenant à la famille des protéines Sm, ainsi qu'à plusieurs protéines spécifiques de chaque UsnRNP. Les UsnRNA sont des ARN non codants de 100 à 300 nt, riches en résidus U. Leur séquence et leur structure bidimensionnelle sont très conservées chez les eucaryotes (Branlant et al., 1980, 1982, 1983; Krol et al., 1981). Il est à noter que dans les spliceosomes mineurs, intervenant dans l'épissage des introns mineurs, un autre ensemble de UsnRNA est utilisé, les snRNA U11, U12, U4atac et U6 atac. Seule la particule snRNP U5 est commune aux spliceosomes mineurs et majeurs.

Après la transcription des UsnRNA par l'ARN polymérase II, les UsnRNP sont assemblées selon un processus séquentiel (pour revue, Will and Luhrmann, 2001; Workman et al., 2012). Dans un premier temps, les protéines Sm sont associées sur le motif Sm des snRNA U1, U2, U4 et U5 dans le cytoplasme grâce à un complexe nommé SMN (Survival Motor Neuron). Cette association est nécessaire au réimport dans le noyau, où la maturation des particules est achevée, ainsi qu'à la stabilité et la fonctionnalité des UsnRNP. La maturation finale des UsnRNA, soit la 2'-O-méthylation et la pseudo-urydilation de certains résidus, est réalisée au sein de corpuscules nucléaires appelés corps de Cajal (Cajal Bodies). La plupart des protéines spécifiques sont associées aux UsnRNA dans ces corpuscules. Le processus de production de la snRNP U6 est entièrement différent, il a lieu dans le noyau. En effet, le snRNA U6 est transcrit par l'ARN polymérase III et s'associe avec 7 protéines Lsm (ou Sm-like), ayant des propriétés similaires à celles des protéines Sm classiques (Mayes et al., 1999; Salgado-Garrido et al., 1999). De plus le snRNA U6 est modifié post-transcriptionnellement dans le nucléole et non dans les corps de Cajal (Ganot et al., 1999; Lange and Gerbi, 2000; Tycowski et al., 1998). Ce snRNA U6 va néanmoins dans les corps de Cajal pour s'associer au snRNA U4 déjà associé aux protéines Sm (Achsel et al., 1999). A la fin de leur maturation, les particules snRNP U1, U2, U4/U6 et U5 sont relocalisées dans des sous-structures du noyau appelées "speckles", formées par des agglomérats de structures ribonucléoprotéiques visibles au microscope électronique, appelées granules interchromatiniens ou IG (pour "Interchromatin Granules") (pour revue, Lamond and Spector, 2003; Spector and Lamond, 2011). Ces "speckles" sont le plus

souvent localisés à proximité des sites de transcription active, et serviraient de lieu de stockage des protéines impliquées dans la transcription et l'épissage des pré-ARNm, et seraient aussi le lieu où se produirait l'épissage co-transcriptionnel (Girard et al., 2012).

1.2.2. La formation des complexes spliceosomaux

1.2.2.a. Etapes d'assemblage du spliceosome

Les complexes spliceosomaux s'assemblent de manière séquentielle et hiérarchisée au niveau de l'intron à éliminer, et leur structure et leur composition sont en constante évolution au cours de la réaction d'épissage. Par ailleurs, différentes étapes de l'assemblage du spliceosome nécessitent l'hydrolyse d'ATP lorsque des changements de conformation importants s'opèrent. Les complexes spliceosomaux contiennent également des protéines n'appartenant pas aux UsnRNP, qui s'associent de façon stable ou transitoire. Avant l'assemblage du spliceosome, les pré-ARNm sont d'ailleurs associés à de nombreuses protéines, dont des facteurs régulant l'épissage, et forment un complexe appelé complexe H. L'assemblage du spliceosome chez l'homme est en effet précédé par une première étape de reconnaissance des exons grâce à des facteurs d'épissage annexes, les protéines SR, dont les fonctions sont détaillées un peu plus loin.

L'assemblage du spliceosome sur le pré-ARNm est initié par la reconnaissance du site 5' donneur par la snRNP U1 grâce à la complémentarité de bases entre l'extrémité 5' du snRNA U1 et ce site (figure I-3). Cette reconnaissance est facilitée par des facteurs protéiques appartenant à la snRNP U1 tel que la protéine U1-C (Heinrichs et al., 1990; Muto et al., 2004), mais également par d'autres facteurs, tels que le facteur U2AF, composé de 2 sous unités U2AF35 et U2AF65, et qui reconnaissent respectivement la séquence polypyrimidine et le site accepteur d'épissage (figure I-3), ainsi que la protéine SF1/mBBP ("Splicing Factor 1/ mammalian Branchpoint Binding Protein"), reconnaissant la séquence de branchement. Le complexe ainsi formé est appelé complexe E (complexe d'engagement). Suite à la formation de ce complexe, la snRNP U2 reconnaît la boîte de branchement et remplace la protéine SF1, grâce à des interactions entre le snRNA U2 et le pré-ARNm assistées par des protéines appartenant aux familles SF3a et SF3b. Cela conduit au positionnement favorable du résidu A de la boîte de branchement pour réaliser l'attaque du site accepteur d'épissage. Ce complexe est appelé pré-spliceosome ou complexe A (figure I-3). Sa formation nécessite, entre autres, l'intervention d'ARN hélicases et l'hydrolyse d'ATP, qui permettent des changements de conformation. Ces changements facilitent le rapprochement spatial des snRNP U1 et U2 et de ce fait des exons situés à chaque extrémité de l'intron à épisser. Enfin, la tri-particule snRNP U4/U6·U5 est incorporée au pré-spliceosome et interagit avec le pré-ARNm pour former le complexe B. A ce stade, des réarrangements structuraux ATP-dépendants majeurs se produisent afin de former un complexe catalytiquement actif (pour revue, Chen and Cheng, 2012; Hoskins and Moore, 2012). Ces réarrangements consistent principalement en la

dissociation du duplex formé par les snRNA U4/U6 au profit de l'établissement d'une interaction entre les snRNA U2 et U6. Suite à cela la snRNP U1 se dissocie du pré-ARNm, ce qui permet la mise en place d'une interaction entre le snRNA U6 et le site 5'. La snRNP U5 favorise alors le rapprochement des deux exons par appariement de bases entre le snRNA U5 et les deux exons (figure I-3). C'est dans ce complexe contenant les snRNP U2, U5 et U6, parfois appelé complexe B*, qu'a lieu la première réaction de trans-estérification. Suite à cette étape, de nouvelles transitions structurales ont lieu et conduisent à la formation du complexe C. C'est dans cette conformation qu'a lieu la deuxième réaction de trans-estérification permettant la libération de l'intron (Jurica et al., 2004). Le rapprochement des deux exons par la snRNP U5 favorise l'attaque nucléophile de l'extrémité 3' OH de l'exon 5' sur le phosphate situé à l'extrémité 3' de l'intron (Collins and Guthrie, 2000; Siatecka et al., 1999).

1.2.2.b. Complexité et dynamisme du spliceosome

La fixation des UsnRNP sur le pré-ARNm et leur activité fait donc intervenir des réseaux complexes d'interactions ARN-ARN, ARN-protéine et protéine-protéine impliquant des protéines des UsnRNP et d'autres protéines. Comme le montre la figure I-4, cela s'accompagne d'une étonnante dynamique d'associations et de dissociations de protéines si l'on compare à d'autres machineries. Les spliceosomes sont en effet des structures extrêmement dynamiques dont la composition évolue sans cesse tout au long des 2 étapes de la réaction d'épissage : au total, les différents complexes spliceosomaux humains mettent en jeu plus de 300 protéines différentes (Jurica and Moore, 2003)(pour revue, Chen and Cheng, 2012; Hoskins and Moore, 2012; Wahl et al., 2009). Ce fonctionnement dynamique est nécessaire à la reconnaissance des sites d'épissage et à la catalyse, afin d'agir avec précision malgré la grande diversité des séquences permettant la définition des sites d'épissage. Les protéines des spliceosomes humains appartenant au noyau fonctionnel minimum nécessaire pour la reconnaissance des introns constitutifs et la catalyse des deux étapes d'épissage sont complétées par des protéines régulatrices. Ces protéines sont nécessaires du fait de la dégénérescence importante des sites d'épissage et à la régulation de l'épissage alternatif (Fabrizio et al., 2009). Ces facteurs de régulation sont d'une importance capitale car ils participent à la reconnaissance des sites 5' et 3' d'épissage, la mise en place des complexes spliceosomaux et pour certains à l'architecture des complexes spliceosomaux, ainsi qu'à leurs transitions structurales.

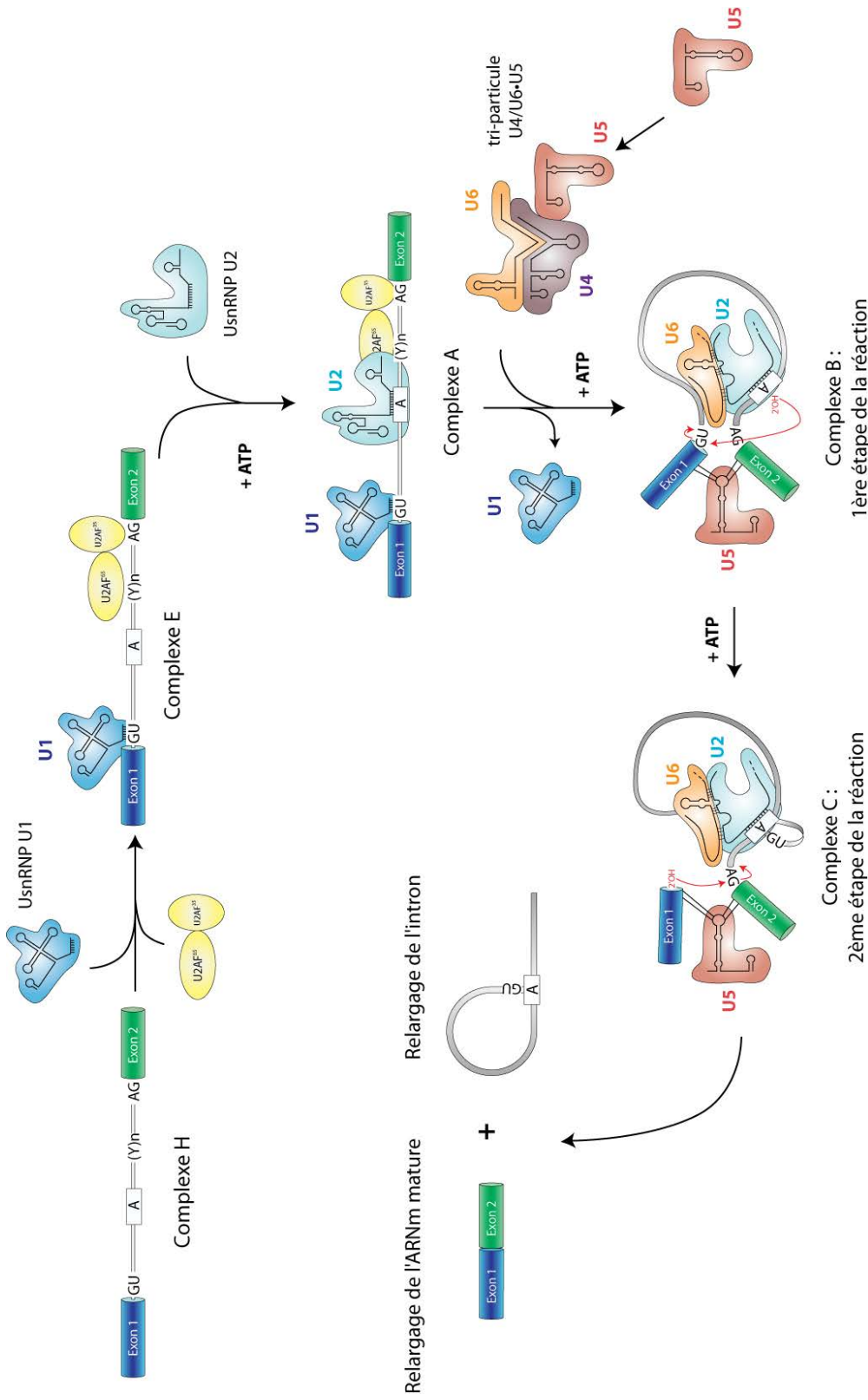


Figure I-3 : L'assemblage séquentiel du spliceosome. L'assemblage du spliceosome commence par l'association de la snRNP U1 avec le site 5' d'épissage et du facteur U2AF avec la séquence polypyrimidine située au site 3'. Ce facteur facilite ensuite la fixation de la snRNP U2, qui a lieu en présence d'ATP, comme toutes les étapes suivantes. Après association de la tripartite U4/U6•U5, la snRNP U1 est dissociée, ainsi que le snRNA U4 (formation du complexe B, qui réalise la première étape de la réaction d'épissage). Les snRNA U2 et U6 interagissent entre eux et avec l'intron (au niveau de la boîte de branchement et du site 5' d'épissage). Après la première réaction d'épissage des réarrangements structuraux ont lieu conduisant à la formation du complexe C dans lequel le snRNA U5 aligne les extrémités des deux exons pour faciliter leur ligature au cours de la seconde étape de la réaction d'épissage. L'ARNm formé et l'intron sous forme de lasso sont libérés. Y : Pyrimidine.

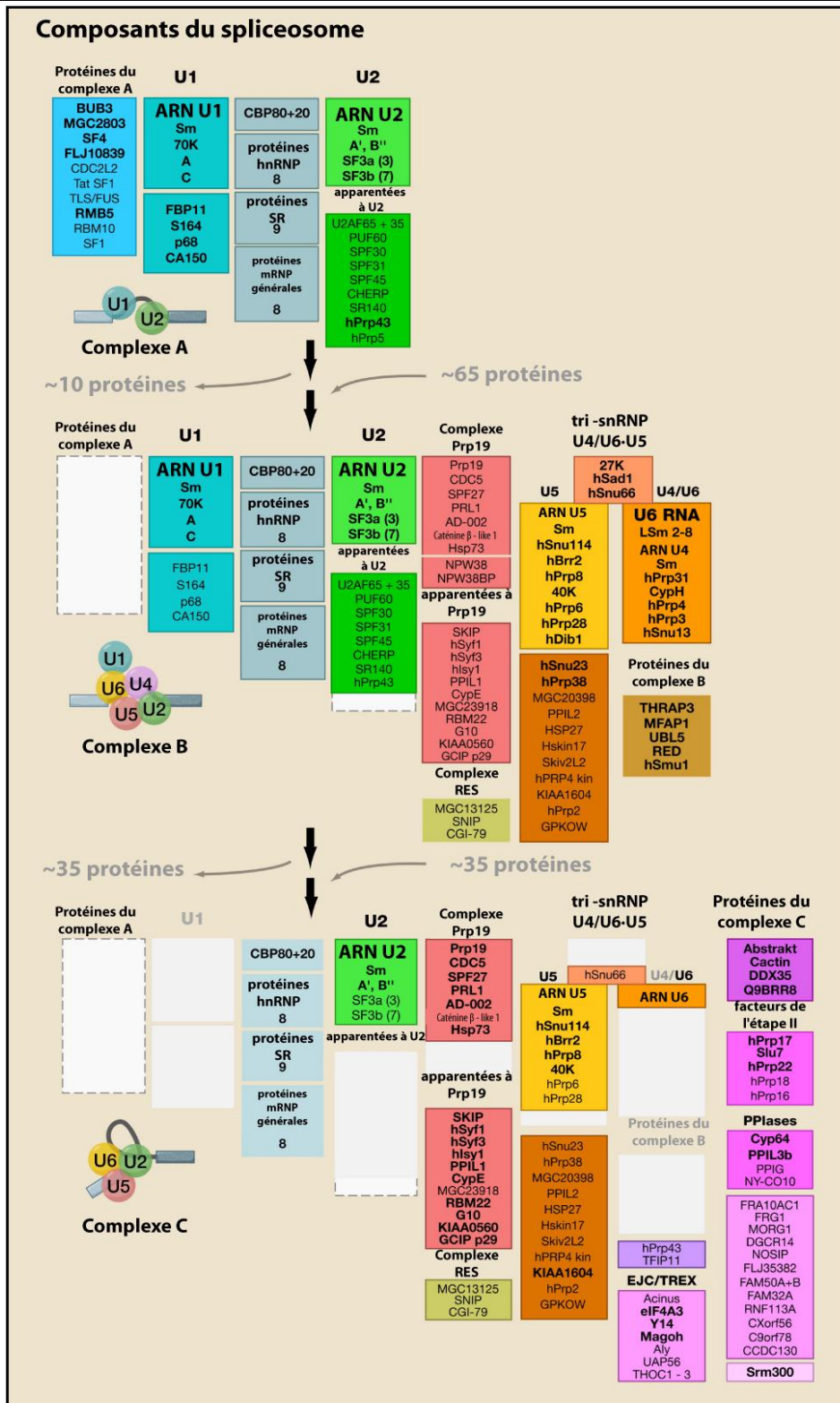


Figure I-4 : Schéma illustrant la complexité et la dynamique de la composition des complexes A, B et C du spliceosome chez l'Homme (tiré de Wahl et al., 2009). Les protéines (nomenclature chez l'Homme) sont regroupées selon leur fonction, leur association stable avec les snRNP, ou leur présence dans un autre complexe spliceosomal, comme indiqué. La composition en protéines des complexes a été déterminée par spectrométrie de masse. Les protéines présentes en quantités substoechiométriques, conformément au nombre de peptides détectés par spectrométrie de masse, sont grisés. Les protéines SR identifiées avec les transcrits utilisés dans ces expériences sont SRSF1 (ASF), SRSF7 (9G8), SRSF3 (SRp20), Tra2-α, SRSF9 (SRp30c), SRSF10 (SRp38), SRSF5 (SRp40), SRSF6 (SRp55) et SRSF4 (SRp75). Les protéines hnRNP identifiées sont les hnRNP A1, A3, C, G, K, M, U, et PCBP1.

L'assemblage des spliceosomes humains nécessite, en plus des séquences des sites d'épissage, la présence de séquences diverses servant à la fixation de ces protéines régulatrices (Lim and Burge, 2001). Chez l'homme, certaines de ces protéines qui s'associent souvent de manière transitoire, ne sont pas recrutées à chaque site d'épissage, et peuvent assurer diverses fonctions régulatrices. Cela rend chaque réaction d'épissage unique et contribue à la complexité des régulations d'épissage. D'autres protéines, comme nous le verrons, servent à faire un lien avec la transcription, ou avec la traduction et l'export des ARNm.

I.2.3. Rôle des protéines SR et SR-like dans la formation des complexes spliceosomaux

Une famille de protéines, les protéines SR, joue un rôle essentiel à la fois dans l'assemblage de la machinerie de base de l'épissage, comme la définition des exons, et dans sa régulation. La définition des protéines SR est basée sur l'observation suivante : la réaction d'épissage n'est possible *in vitro* qu'avec un extrait cytoplasmique, déficients en facteurs d'épissage, que si l'on ajoute la fraction correspondant aux protéines SR. Ces protéines tirent leur nom de leur domaine d'interaction avec les protéines ou domaine RS, riche en dipeptides Arginine-Sérine (R-S), qui permet l'interaction avec d'autres protéines possédant également un domaine RS (Zahler et al., 1992). Elles possèdent également à l'extrémité N-ter un ou deux domaines d'interaction avec l'ARN, de type RRM ("RNA Recognition Motif", aussi appelé RBD pour "RNA Binding Domain"). Dix protéines SR, avec des tailles allant de 20 à 75kDa, ont été identifiées chez l'homme : ASF/SF2, SC35 ou SRp30b, SRp20, SRp30c, SRp40, SRp55, 9G8, SRp75, SRp54 et SRp46) (tableau I-1, § suivant) (Ayane et al., 1991; Cavaloc et al., 1994; Chaudhary et al., 1991; Fu and Maniatis, 1992; Ge et al., 1991; Sreanion et al., 1995; Soret et al., 1998; Zahler et al., 1993). Les protéines SR sont très conservées chez la plupart des espèces eucaryotes, à quelques exceptions près comme par exemple la levure *S. cerevisiae* qui ne possède qu'une protéine apparentée aux protéines SR, et la plante *A. thaliana* chez qui on a retrouvé jusqu'à 18 protéines SR différentes (Barta et al., 2010; Graveley, 2000). Une nouvelle nomenclature des protéines SR humaines, plus cohérente, a été récemment introduite afin de regrouper ces protéines sans ambiguïté sous le même type d'appellations chez l'homme (tableau I-1, § suivant)(Manley and Krainer, 2010). Dans cette nomenclature, deux protéines autrefois considérées comme "SR-like" ou "SR-related protein" (SRrp), c'est à dire semblables aux protéines SR mais avec une organisation légèrement différente, à savoir SRp38 (Srrp40) et Srrp35, ont été incluses dans la famille des protéines SR. La nouvelle nomenclature des protéines SR a été établie en fonction leur organisation. En effet, 51 protéines humaines possèdent un domaine SR et 1 à deux domaines RRM. Parmi celles-ci, seulement 16 ont un domaine RS situé en aval du ou des domaines RRM. Ces 16 protéines peuvent encore être classées en 3 groupes différents selon leur organisation. Le premier groupe est composé des protéines comprenant deux domaines RRM, à savoir ASF/SF2, SRp75,

SRp40, SRp55, SRp30c. Les protéines des deux autres groupes ne contiennent qu'un seul domaine RRM ; le second groupe rassemble les protéines SC35, SRp20, SRp46, 9G8, Srrp35, Srrp40, SRp54. Le troisième groupe renferme les 4 protéines SFRS12, CPSF7, RBMX2 et SR140 qui ne sont pas incluses dans la famille des protéines SR car la séquence de leur domaine RRM est trop divergent par rapport au consensus des domaines RRM des autres protéines SR. D'ailleurs, les protéines CPSF7 et RBMX2 ne sont pas impliquées dans l'épissage (Manley and Krainer, 2010). Les 12 protéines des groupes 1 et 2 sont maintenant dénommées "SRSF" (pour "SR Splicing Factor", facteur d'épissage SR), suivi d'un numéro issu de la nomenclature de leurs gènes et reflétant l'ordre chronologique de leur découverte (tableau I-1, § suivant). Parmi ces protéines on peut noter les rôles particuliers de 3 d'entre elles : la protéine SRp38, qui sera évoquée plus loin dans un chapitre sur le choc thermique, est un répresseur général de l'épissage, activé sous certaines conditions spécifiques, par exemple après un stress cellulaire. Srrp35 et SRp54 sont des régulateurs négatifs et non des activateurs de l'épissage alternatif.

1.2.3.a. Structure des protéines SR

Les deux groupes de protéines SR sont maintenant respectivement dénommées sous famille SRSF2 (SC35) et sous famille SRSF1 (ASF/SF2), selon qu'elles possèdent respectivement 1 ou 2 domaines RRM. Le ou les domaines RRM, qui permettent la fixation à l'ARN sont également présents chez de nombreuses autres protéines, en particulier les protéines hnRNP. Le premier domaine RRM en N-ter des protéines (RRM1), est très structuré, car constitué principalement par un feuillet β de 4 brins antiparallèles, jouant le rôle de plateforme de fixation à l'ARN (Nagai et al., 1990). Ce feuillet est flanqué de part et d'autre par deux très courtes hélices α , et comporte deux séquences signatures RNP-1 (RGFGFVxF) et RNP-2 (LFVGNL), intervenant dans beaucoup de cas dans l'établissement de contacts directs avec l'ARN (Gorlach et al., 1992). Néanmoins selon les protéines SR d'autres éléments de séquences situés dans des boucles flexibles reliant les brins du feuillet β peuvent avoir un rôle très important dans la spécificité de reconnaissance de l'ARN. Le second domaine RRM2, est présent seulement dans les protéines SR de la sous famille de SRSF1 (Tableau I-1). Il présente des séquences signatures RNP-1 et 2 très dégénérées par rapport à celles des domaines RRM1. En revanche, il contient tous les acides aminés conservés qui forment le cœur hydrophobe du domaine RRM1 (Birney et al., 1993). Ce domaine RRM2, aussi appelé pseudo-RRM ou RRMH (pour RRM Homologue), participe aussi à la définition de la spécificité de reconnaissance de l'ARN par SRSF1 (Caceres et al., 1997; Mayeda et al., 1999). La protéine SRSF7 (9G8) a la particularité de posséder un motif doigt à zinc, en plus du domaine RRM, et ce motif contribue lui aussi à la spécificité de fixation à l'ARN. Chaque protéine SR a une capacité large de reconnaissance de l'ARN avec une faible spécificité et présente une forte affinité pour un nombre limité de séquences bien définies, qui ont pu être définies par des expériences de SELEX, comme nous le verrons plus en détails. Ces séquences jouent un rôle majeur dans la régulation de l'épissage.













Symbole du gène	Noms précédents de la protéine	Nouvelle nomenclature	Domaines protéiques
SFRS1	ASF/SF2 ; ASF ; SRp30a	SRSF1	
SFRS2	SC35 ; PR264 ; SRp30b	SRSF2	
SFRS3	SRp20	SRSF3	
SFRS4	SRp75	SRSF4	
SFRS5	SRp40 ; HRS	SRSF5	
SFRS6	SRp55 ; B52	SRSF6	
SFRS7	9G8	SRSF7	
SFRS2B	SRp46	SRSF8	
SFRS9	SRp30c	SRSF9	
SFRS13A	SRp38 ; TASR1 ; Srrp40	SRSF10	
SFRS11	SRp54 ; p54	SRSF11	
SFRS13B	Srrp35	SRSF12	

Tableau I-1 : Nouvelle nomenclature des protéines SR récemment proposée. Adapté de (Manley and Krainer, 2010). RS : domaine riche en arginines et en sérines à l'origine d'interactions protéine-protéine. RRM : "RNA Recognition Motif", domaine d'interaction avec l'ARN. RRMH : RRM Homologue. Zn : motif doigt à zinc.

Le domaine RS en C-ter des protéines SR contient de grands enchaînements de dipeptides R-S. La longueur du domaine RS varie de 24 résidus pour SRSF9 (SRp30c) à 316 résidus pour SRSF4 (SRp75). Ce domaine permet aux protéines SR d'établir des interactions avec les domaines RS d'autres protéines SR, mais également avec les domaines RS d'autres protéines possédant des domaines RS, en particulier des protéines constitutives des UsnRNP. Ces interactions protéine-protéine permettent de faciliter l'assemblage du spliceosome. Le domaine RS des protéines SR peut aussi participer à la reconnaissance générale de faible affinité des ARN par les protéines SR et les interactions stables établies avec le ou les domaines RRM. Ainsi, des expériences de pontage aux UV ont montré que le domaine RS de SRSF1 peut interagir avec le pré-ARNm au niveau de la boîte de branchement (Shen and Green, 2004).

La phosphorylation du domaine RS est indispensable à l'assemblage du spliceosome (Rosigno and Garcia-Blanco, 1995; Xiao and Manley, 1997). En effet, ce domaine est soumis à de nombreux événements de phosphorylation qui permettent de réguler l'activité des protéines SR, mais aussi leur localisation cellulaire (Sanford and Bruzik, 2001). Plusieurs types de protéines kinases agissant sur les protéines SR ont été identifiées : la famille SRPK dont les protéines SRPK1 et SRPK2 (pour "SR protein kinase"), la famille Clk/Sty, Akt, et l'ADN topoisomérase I (Colwill et al., 1996; Patel et al., 2005; Prasad et al., 1999; Rossi et al., 1996). Les protéines SR sont également soumises à l'activité de phosphatases spécifiques PP1A, PP2A et PP2C.

1.2.3.b. Autres protéines à domaines SR

Ces protéines appartiennent à la famille SR-like ou bien SRrp ("SR related proteins") (figure I-5) (pour revue, Graveley, 2000). Cette famille comprend des protéines impliquées directement dans l'épissage, autre que les protéines SR, comme U2AF35 et U2AF65, les protéines U1-70K, U5-100K et [U4/U6.U5]-27K des UsnRNP, des protéines régulatrices de l'épissage (hTra2 α et hTra2 β) et des protéines co-activatrices de l'épissage (SRm160 et SRm300). A la différence des protéines SR, les protéines SRm160 et SRm300 ne possèdent pas de domaine de liaison à l'ARN bien qu'elles possèdent de multiples répétitions RS. Elles s'associent en un complexe SRm160/SRm300, qui favorise l'épissage par l'établissement d'interactions entre les protéines SR et les constituants du spliceosome comme la snRNP U1 (Blencowe et al., 2000; Eldridge et al., 1999).

Certaines protéines SR-like ont des activités ARN hélicases (hPrp16 et HRH1). Les protéines kinases Clk1, 2, 3, capables de phosphoryler le domaine RS des protéines SR, sont aussi des protéines SR-like. Les protéines CASP sont une autre catégorie de protéines SR-like jouent un rôle important dans l'épissage en faisant le lien entre la machinerie de transcription et d'épissage. Les interactions mises en jeu entre les domaines RS sont des interactions ioniques impliquant des arginines chargées positivement et des sérines phosphorylées. C'est pourquoi l'état de phosphorylation du domaine RS influence fortement les interactions des protéines SR avec leurs partenaires protéiques. Il en résulte que l'activité des protéines SR diminue lorsqu'elles sont hyper ou hypo-phosphorylées (Kanopka et al., 1998; Prasad et al., 1999).

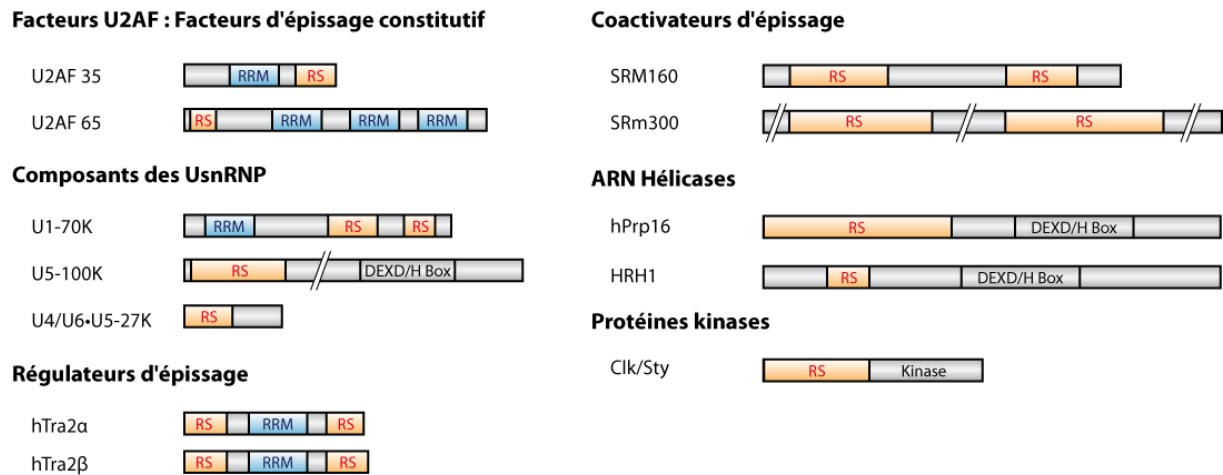


Figure I-5 : Exemples de protéines à domaines RS autres que les protéines SR (Graveley, 2000; Long and Caceres, 2009). RS : domaine riche en arginines et en sérines à l'origine d'interactions protéine-protéine. RRM : "RNA Recognition Motif", domaine d'interaction avec l'ARN. Zn : motif doigt à Zinc. Le motif DEXD/H est caractéristique des protéines à activité hélicase.

1.3.2.c. Interaction des protéines SR avec l'ARN

En plus d'être impliquées dans l'épissage constitutif, les protéines SR peuvent également réguler l'épissage alternatif. En se fixant au niveau de séquences activatrices exoniques (ESE) ou beaucoup plus rarement introniques (ISE), elles contribuent généralement à l'inclusion des exons en favorisant localement l'épissage. Les séquences ARN de haute affinité pour les protéines SR ont été déterminées grâce à des expériences de SELEX (tableau I-2). Ces séquences servent souvent de base à la prédiction de sites régulateurs d'épissage, notamment par les logiciels informatiques. Les éléments ESE peuvent également recruter des co-activateurs ayant un effet similaire, comme les 2 protéines SR-like Tra2α et Tra2β (Blencowe et al., 2000; Eldridge et al., 1999).

1.2.3.d. Rôle des protéines SR dans l'épissage

Par les interactions qu'elles établissent avec le pré-ARNm et différentes protéines tout au long de l'assemblage des complexes spliceosomaux, les protéines SR participent à la stabilisation de l'assemblage des différents composants (Sanford et al., 2005). Au cours des étapes précoces de l'assemblage du spliceosome, elles contribuent à la définition des exons et des introns. En se fixant sur des séquences spécifiques du pré-ARNm, appelées ESE (Exonic Splicing Enhancer) ou ISE (Intronic Splicing Enhancer), elles peuvent favoriser et stabiliser la fixation de la snRNP U1 au site 5' d'épissage et celle du facteur U2AF sur la séquence polypyrimidine. En effet, par leur domaine RS, elles peuvent interagir avec les protéines U1-70K de la UsnRNP U1 et U2AF35 (Tarn and Steitz, 1994; Wu and Maniatis, 1993).

Protéine SR	Site de fixation à l'ARN
SRSF1 (ASF/SF2)	RGAAGAAC (SELEX) AGGACRRAGC (SELEX) SRSASGA (S100) (SELEX fonctionnel) CRSMGW (S100) (SELEX fonctionnel) UGRWG (CLIP)
SRSF2 (SC35)	AGSAGAGUA (SELEX) GUUCGAGUA (SELEX) GRYY C/A G/C YR (S100) (SELEX fonctionnel) UGUUCSAGWU (SELEX) GWUWCCUGCUA (SELEX) GGGUAUGCUG (SELEX) GAGCAGUAGKS (SELEX) AGGAGAU (SELEX) UGCNGYY (EN) (SELEX fonctionnel)
SRSF3 (SRp20)	WCAKCUUUA (SELEX) AUCCNNA (SELEX) WCAUCGAY (SELEX) YWCUUCAU (SELEX) CUWCAAC (SELEX) GGUCCUCUUC (EMSA) WCWWC (Test épissage) CUCKUCY (EN) (SELEX fonctionnel) CCUCGUCC (EN) (SELEX fonctionnel)
SRSF4 (SRp75)	GAAGGA (pontage UV)
SRSF5 (SRp40)	GAGCAGUCGGCUC (SELEX) UGGGAGCRGUYRGCUCGY (SELEX) YRCRKM (S100) (SELEX fonctionnel) ACDGS (S100) (SELEX fonctionnel)
SRSF6 (SRp55)	USCGKM (S100) (SELEX fonctionnel) YYWCWSG (S100) (SELEX fonctionnel) GRUCAACCNGGCGACMG (SELEX)
SRSF7 (9G8)	UCAACA (pontage UV) AGACKACGAY (SELEX) ACGAGAGAY (SELEX) GGACGACGAG (EN) (SELEX fonctionnel)
SRSF9 (SRp30c)	GACGAC (SELEX fonctionnel) AAAGAGCUCGG (SELEX fonctionnel) CUGGAUU (EMSA)
SRSF10 (SRp38)	ACAAAGACAAA (SELEX)
SRSF11 (p54/SRp54)	Riche en C (pontage UV)

Tableau I-2: Séquences reconnues par les protéines SR identifiées par la méthode SELEX (pour revue, Bourgeois et al., 2004; Long and Caceres, 2009). La méthode utilisée pour identifier les séquences SELEX est indiquée (SELEX de liaison ou SELEX fonctionnel). Pour les séquences isolées par expériences de SELEX fonctionnel, il est également indiqué si les expériences ont été réalisées en extrait nucléaire (EN) ou avec un extrait cytoplasmique S100 (S100).

Nucléotides : R: A/G; Y: C/U; S: C/G; K: G/U; W: A/U; M: A/C; D: A/G/U; N: nucléotide quelconque.

De cette manière elles peuvent également faciliter le rapprochement dans l'espace des sites 5' et 3' d'épissage au cours de l'assemblage du spliceosome (figure I-6). Les protéines SR auraient également une fonction dans le respect du bon ordre de ligation des exons au cours de l'épissage d'un transcrit, en empêchant l'exclusion des exons constitutifs (Ibrahim et al., 2005). A l'extrémité 5' de l'ARNm, la reconnaissance du premier exon est facilitée par l'action des protéines CBP20 et CBP80 associées à la coiffe et qui favorisent la fixation de la snRNP U1 sur le site donneur grâce aux protéines SR (Lewis et al., 1996). A l'extrémité 3' de l'ARNm, la reconnaissance du dernier exon est facilitée par la protéine PAP (pour PolyA Polymérase), associée avec le signal de polyadénylation, et qui recrute le facteur U2AF65 sur le site 3' d'épissage situé en amont grâce aux protéines SR (Vagner et al., 2000).

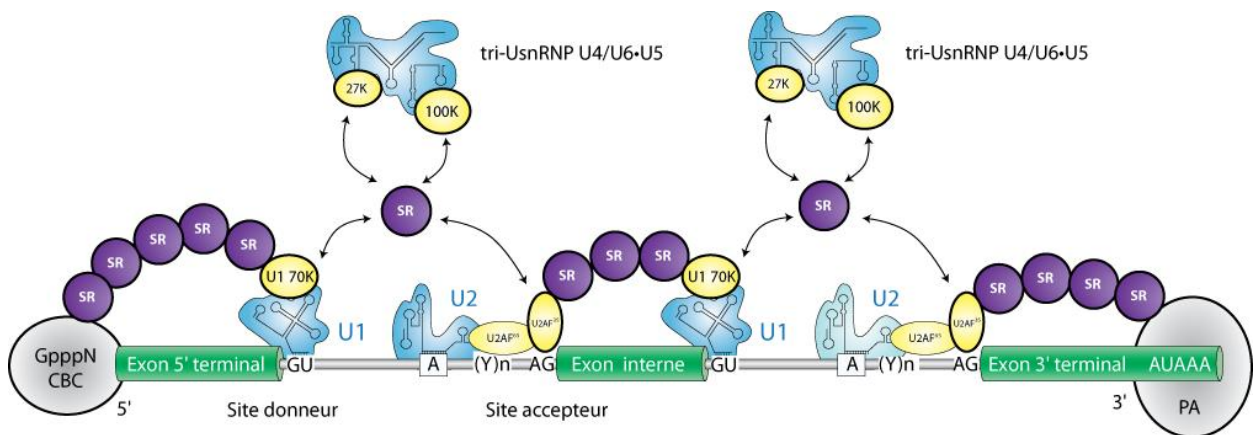


Figure I-6 : Fonctions des protéines SR dans l'épissage constitutif. A. Rôle des protéines SR dans la reconnaissance des exons. Les protéines SR se fixent sur des séquences exoniques activatrices, les ESE (en vert) et facilitent la reconnaissance des sites 5' et 3' d'épissage (activation indiquée par une flèche verte) grâce à des interactions protéine-protéine, en particulier U1-70K et U2AF35, via leurs domaines RS. Ces interactions permettent la définition de l'exon. **B.** Rôle des protéines SR dans la reconnaissance des introns. Les protéines SR facilitent l'assemblage du spliceosome sur l'intron et assurent la cohésion entre les différents partenaires, grâce à leurs interactions avec les protéines possédant un domaine RS.

1.2.3.e. Localisation des protéines SR

Dans les noyaux, les protéines SR sont fortement concentrées dans des régions subnucléaires appelées "speckles", déjà évoquées précédemment, qui servent de site de stockage des UsnRNP et des facteurs d'épissage (Spector and Lamond, 2011). Le domaine RS contient généralement un signal de localisation nucléaire de la protéine, mais celui-ci peut parfois s'étendre jusqu'au domaine de fixation à l'ARN. Le signal de localisation nucléaire renforce leur concentration dans le noyau et permet à ces protéines de rejoindre les speckles (Cazalla et al., 2002; Lamond and Spector, 2003; Spector and Lamond, 2011). L'action de ce signal peut être contrecarrée par l'hyperphosphorylation de leur domaine RS, entraînant leur export dans le cytoplasme. Aussi, certaines protéines SR comme SRSF1, SRSF3 (SRp20), SRSF4 (SRp75), SRSF6 (SRp55) et SRSF7 (9G8), ainsi que le facteur U2AF, sont capables de faire continuellement la navette entre le noyau et le cytoplasme tandis que SRSF2 (SC35) ne fait pas

d'aller-retour dans le cytoplasme (Caceres et al., 1998; Cazalla et al., 2002; Sapra et al., 2009). Les protéines SR peuvent également être exportées sous forme liée à l'ARNm et intervenir dans la régulation de sa traduction (Blencowe et al., 1994; Le Hir et al., 2000a). D'autre part, dans le noyau, les protéines SR peuvent se déplacer entre les speckles et les sites de transcription, qui sont spatialement proches (Misteli and Spector, 1999; Misteli et al., 1997). Ce déplacement serait aussi régulé par le degré de phosphorylation du domaine RS des protéines SR (Misteli and Spector, 1998; Misteli et al., 1998). Les protéines SR déphosphorylées sont associées aux speckles tandis que leur phosphorylation permet leur migration vers les sites de transcription active.

I.2.4. Autres fonctions des protéines SR

Les protéines SR ont plusieurs autres fonctions en dehors de leur rôle dans l'épissage, dont l'une est le transport des ARNm vers le cytoplasme. Par exemple, les protéines SR SRSF7 (9G8) et SRSF3 (SRp20) servent d'adaptateurs entre les ARNm et les protéines TAP, des facteurs d'exports nucléaires majeurs (Hautbergue et al., 2008; Huang et al., 2003). Il a également été montré qu'elles activent l'export dans le cytoplasme de l'ARNm de l'histone H2A, qui est dépourvu d'introns, par fixation sur une séquence spécifique de 22 nt (Huang and Steitz, 2001). Plusieurs études ont montré que, comme il a été évoqué ci-dessus, les protéines SR pouvaient jouer un rôle dans la régulation de la traduction. C'est le cas de la protéine SRSF1, qui peut activer la traduction cap-dépendante (Blaustein et al., 2009; Michlewski et al., 2008). Cette régulation fait intervenir les séquences ESE, le domaine RRM2 et l'état de phosphorylation de la protéine (Sanford et al., 2004, 2005). Il a également été reporté l'implication de SRSF3 (SRp20) dans la traduction médiée par IRES dans les ARN de poliovirus dans les extraits de cellules HeLa, et des protéines SRSF5 (SRp40) et SRSF6 (SRp55) dans l'augmentation de la traduction de la protéine Gag à partir de l'ARN du virus HIV-1 (Bedard et al., 2007; Swanson et al., 2010). En fait, comme nous allons le voir, il existe des liens fonctionnels entre la transcription, l'épissage, l'export des ARNm et leur traduction, ainsi qu'avec d'autres processus comme l'état de condensation de la chromatine dont dépend également la transcription. Des protéines SR et SR-like ont été retrouvées co-purifiées avec l'ARN polymérase II et peuvent s'associer avec les protéines SR-like CASP, elles-mêmes liées de façon directe au domaine CTD de l'ARN polymérase II (Das et al., 2007)(pour revue, Alexander and Beggs, 2010; Corden and Patturajan, 1997). Ces interactions se forment via des domaines RS sur les protéines CASP permettant l'interaction avec les protéines SR phosphorylées. Tous ces phénomènes sont imbriqués, c'est pourquoi les modèles considérant ces processus de manière parfaitement segmentée et distincte sont considérés aujourd'hui comme obsolètes (pour revue, Alexander and Beggs, 2010; Braunschweig et al., 2013).

I.3. Liens existants entre l'épissage et l'état de la chromatine, la transcription et l'export des ARNm

L'épissage peut avoir lieu *in vitro* dans des extraits nucléaires et en absence de transcription, donc peut bien avoir lieu en absence de couplage avec d'autres processus nucléaires. Toutefois, *in vivo*, les complexes spliceosomaux ne s'assemblent que sur les transcrits de l'ARN pol II et le couplage entre transcription et épissage assure l'efficacité nécessaire au processus et permet que des introns mesurant 1500 nt ou plus de 100 kb puissent être éliminés efficacement en moins de 5 min après la transcription de l'exon précédent. *In vitro*, l'épissage est rarement efficace sur des substrats dont la taille est supérieure à 500 nt, et prend beaucoup plus de temps, souvent de l'ordre d'une demi-heure. L'ajout d'ARN pol II dans un extrait d'épissage renforce son efficacité. Ces observations impliquent qu'il existe un lien fonctionnel entre les machineries d'épissage et de transcription, un phénomène connu sous le nom de couplage. *In cellulo*, la reconnaissance des introns et des exons dépend aussi d'autres paramètres comme l'architecture de l'ADN en cours de transcription.

I.3.1. Couplage fonctionnel entre transcription et épissage des ARNm

Les mécanismes d'épissage sont étroitement liés à la transcription, mais aussi la présence de la coiffe en 5' et la polyadénylation des ARNm. La coiffe m7G des transcrits pol II du complexe protéique de la coiffe CBC ("Cap Binding Complex"), est nécessaire à la définition du premier exon. Les protéines fixées sur la coiffe peuvent interagir avec des facteurs d'épissage dont les protéines SR pour participer à la définition du 1er exon (Furger et al., 2002; Lewis and Izaurralde, 1997; Maniatis and Reed, 2002). Des expériences d'épissage *in vitro* ont montré qu'en plus d'améliorer la stabilité des ARN, la présence d'une coiffe augmente fortement l'efficacité de leur épissage (Proudfoot et al., 2002). De la même façon, les protéines du complexe de polyadénylation participent à la définition de l'exon terminal du pré-ARNm.

La coordination entre la transcription, l'épissage et la polyadénylation a lieu principalement grâce aux liens existants entre la machinerie d'épissage et le domaine C-terminal (CTD) de la sous-unité de l'ARN polymérase II. Ce domaine joue un rôle essentiel, puisque sa troncation conduit à des défauts de mise en place de la coiffe, de polyadénylation et de l'épissage (McCracken et al., 1997). Chez les vertébrés, ce domaine est constitué par une cinquantaine de répétitions d'acides aminés YSPTSPS, contenant 3 résidus sérines phosphorylables. Les capacités d'interaction du domaine CTD sont dépendantes du degré de phosphorylation des 3 résidus sérines des répétitions. La transition entre complexe d'initiation et complexe d'élongation par l'ARN pol II est régie par le changement du degré de phosphorylation de ces motifs répétés. Le passage à la phase d'élongation permet la dissociation de l'interaction établie entre l'ARN pol II et les protéines de la coiffe, et permet la mise en place d'interactions avec les facteurs d'épissage. Les résidus sérines sont déphosphorylés en fin d'élongation

lors de l'interaction de l'ARN pol II avec les complexes de polyadénylation. Le domaine CTD agit comme une plate-forme de chargement sur laquelle s'assemblent les complexes multiprotéiques impliqués dans la formation de la coiffe, l'épissage et la polyadénylation des ARNm, mais aussi des protéines intervenant dans la régulation de l'épissage constitutif et alternatif. Un lien direct entre la transcription et l'épissage a été montré au travers du recrutement par le domaine CTD du complexe de facteurs d'épissage lié au sites 3' U2AF65-Prp19 (David and Manley, 2010; David et al., 2011). Le facteur U2AF forme également des interactions avec la protéine CF1 ("Cleavage Factor I") et la poly-A polymérase, et stimule à la fois l'épissage de l'exon terminal et le clivage et la polyadénylation du pré-ARNm (Millevoi et al., 2006). Plusieurs cas sont également répertoriés où des facteurs d'épissage, comme hnRNP H, entrent en compétition avec les facteurs de clivage et de polyadénylation afin de réguler l'utilisation de sites de polyadénylation alternatifs (Evsyukova et al., 2013; Katz et al., 2010).

I.3.2. Couplage cinétique entre transcription et épissage des ARNm

Il a été proposé initialement que ce couplage entre transcription et épissage soit nécessaire pour que les exons soient ligués entre eux dans leur ordre de synthèse au sein du pré-ARNm. Selon ce modèle, chaque exon adjacent à un intron en cours d'épissage reste associé au domaine CTD jusqu'à ce qu'il soit relié à l'exon suivant, de manière à faciliter l'adjonction des exons (Maniatis and Reed, 2002; Zeng and Berget, 2000). La technologie de séquençage haut débit des ARN a permis de confirmer qu'une majorité des introns est épissé de manière co-transcriptionnelle (jusqu'à environ 90% chez la drosophile)(Ameur et al., 2011; Khodor et al., 2011). Ce type d'étude a montré également que ce couplage dépend de la longueur de l'intron et de sa distance par rapport à l'extrémité 5' de l'ARN, et aussi que le clivage en 3' et la polyadénylation des pré-ARNm sont le plus souvent réalisés avant que tous les introns soient épissés, et que ces transcrits restent associés à la chromatine jusqu'à ce que l'épissage soit achevé (Bhatt et al., 2012; Khodor et al., 2012). Ce couplage entre transcription et épissage se fait donc également à un niveau cinétique. En effet, l'élongation des transcrits par l'ARN pol II n'est pas un processus uniforme et régulier. Il fluctue et est soumis à des étapes de pauses, en fonction de la distribution locale des séquences parcourues (Kornblihtt et al., 2004). Ce phénomène est particulièrement important lors de l'épissage alternatif, car la sélection d'un site d'épissage dépend de la vitesse de transcription. Lorsque la vitesse de transcription est faible ou que des pauses de la polymérase ont lieu, un spliceosome peut par exemple s'assembler en utilisant un site 5' d'épissage de faible affinité qui n'aurait pas eu le temps d'être reconnu si la cinétique de synthèse par la pol II était plus rapide. Dans ce cas le site 5' plus fort situé en aval aurait été utilisé (Perales and Bentley, 2009).

Ce principe peut être appliqué de la même façon à la reconnaissance des sites de fixations forts ou faibles d'un facteur de régulation de l'épissage, une transcription lente facilitant la fixation des facteurs présentant une faible affinité. Ainsi, la reconnaissance des sites d'épissage peut être influencée par

plusieurs facteurs affectant la vitesse de transcription, comme le type de promoteur utilisé ou encore la présence de séquences ARN fortement structurées, où encore des facteurs de transcription altérant la processivité de l'ARN pol II (Cramer et al., 1997; Rosonina et al., 2005; Vargas et al., 2011) (pour revue, Braunschweig et al., 2013). Nous allons voir qu'en conséquence, l'état de condensation de la chromatine et d'autres paramètres eux-mêmes reliés à l'activité transcriptionnelle d'un gène, ont également un effet sur le déroulement de l'épissage, et vice-versa.

I.3.3. Liens entre l'organisation de la chromatine et l'épissage

I.3.3.a. Influence de l'état de la chromatine et des modifications épigénétiques sur la transcription

L'état de la chromatine, qui joue un rôle majeur dans l'activité transcriptionnelle des gènes, peut donc également influencer l'épissage des transcrits. En premier lieu, en lien direct avec le couplage entre transcription et épissage, l'état de la chromatine peut influencer directement la vitesse d'élongation de la transcription et donc le déroulement de l'épissage. De nombreux co-activateurs de la transcription provoquent l'ouverture de la chromatine afin de faciliter l'assemblage et la progression de l'ARN polymérase II. Le complexe de remodelage de la chromatine Brahma (SWI/SNF), par exemple, interagit à la fois avec l'ARN pol II et des facteurs d'épissage, et est capable de moduler l'épissage alternatif et la polyadénylation de certains pré-ARNm en modifiant la vitesse d'élongation de l'ARN pol II (Batsché et al., 2006; Waldholm et al., 2011; Zraly and Dingwall, 2012). De même, la déplétion du facteur DBIRD, qui facilite la transcription des séquences riches en A/T, plus difficiles à transcrire par l'ARN pol II, entraîne une diminution de la vitesse d'élongation et de l'utilisation des sites d'épissage situés à proximité de séquences riches en A/T (Close et al., 2012).

D'autre part, des protéines associées aux transcrits naissants, comme les protéines Hu de la famille des hnRNP et qui sont impliquées dans la stabilité de l'ARN, peuvent inhiber l'action d'histones déacétylases HDAC1 et HDAC2, lorsqu'elles sont associées à leur ARN cible, induisant l'acétylation des histones (Zhou et al., 2011). Cela résulte en un accroissement local de la vitesse de transcription, ce qui influence l'inclusion d'exons alternatifs. Pour rappel, les nucléosomes constituent le premier niveau d'organisation de la chromatine. Ils sont formés par association des histones H2A, H2B, H3 et H4 à l'ADN. Les modifications épigénétiques présentes sur ces histones, principalement les acétylations et méthylation, ainsi que la méthylation de l'ADN, affectent l'état de condensation de la chromatine. Ainsi, la délétion d'un site 5' d'épissage fonctionnel proche d'un promoteur peut diminuer le taux de transcrits naissants correspondant (Furger et al., 2002). Ceci est probablement lié au fait que l'ADN correspondant à la première jonction exon-intron est plus favorable à la formation de nucléosomes, qui présentent en plus un taux élevé de modifications des histones H3K4me3 et H3K9ac associées aux promoteurs actifs et à une occupation élevée par l'ARN pol II. En effet, nous allons voir dans le paragraphe suivant que la

position des nucléosomes et les modifications épigénétiques des histones interviennent dans la définition des exons. Dans certains cas, la délétion du premier intron réduit le taux de H3K4me3 et diminue la transcription du pré-ARNm (Bieberstein et al., 2012).

1.3.3.b. L'organisation de la chromatine intervient dans la définition des exons

Ces modifications n'ont pas une distribution uniforme le long du génome et des signatures particulières de modifications sont observées au niveau des promoteurs des gènes, mais également des exons (Smolle and Workman, 2013). Des expériences de type ChIP-seq (ChIP, couplée au séquençage haut débit) ont montré que les exons sont enrichis en nucléosomes par rapport aux introns (Andersson et al., 2009; Spies et al., 2009; Tilgner et al., 2009; Wilhelm et al., 2011). On note la présence fréquente d'un nucléosome dans la région 5' de l'exon. Ce qui peut s'expliquer par un biais de la composition en bases de l'ADN au niveau des exons, les exons étant plus riches en G/C que les introns. Cette haute teneur en G/C favorise l'occupation par les nucléosomes. Par ailleurs, ces nucléosomes sont enrichis en triméthylations du résidu K36 de l'histone 3 (H3K36me3) (Andersson et al., 2009; Spies et al., 2009). Les données actuelles mettent donc en évidence un lien possible entre l'enrichissement en nucléosomes à l'extrémité 5' des exons, l'identité des modifications des histones de ces nucléosomes, et la définition des exons.

1.3.3.c. La chromatine intervient dans le recrutement des facteurs d'épissage

L'état de la chromatine a également un impact direct sur le recrutement des facteurs d'épissage. En effet, des expériences d'immunoprécipitation de la chromatine (ChIP) ont montré le recrutement de facteurs d'épissage au niveau de gènes de manière concomitante à la transcription des sites d'épissages reconnus par ces protéines, chez la levure comme chez l'homme (Lacadie and Rosbash, 2005; Listerman et al., 2006; Sapra et al., 2009). Ce recrutement est supposé faciliter le chargement des facteurs d'épissage sur le transcrit naissant. Il n'a pas lieu sur les gènes dépourvus d'introns, comme des gènes codant des "heat shock proteins", des protéines de réponse au choc thermique. Il a été montré par exemple que les modifications H3K36me3 permettent le recrutement de la protéine Psip1, qui s'associe à son tour avec des facteurs d'épissage comme SRSF1 (Pradeepa et al., 2012). Dans certains cas, ces modifications sont aussi impliquées dans les régulations de l'épissage alternatif : dans le pré-ARNm du récepteur FGFR-2, deux exons mutuellement exclusifs IIIb et IIIc sont flanqués de régions enrichies en modifications des histones H3K36me3. Elles permettent de recruter l'adaptateur MRG15, qui à son tour promeut le recrutement du régulateur d'épissage PTB1 ("Polypyrimidine Tract Binding Protein 1"), ce qui empêche l'inclusion de l'exon IIIb dans les cellules mésenchymateuses (Luco et al., 2010). De la même manière, un nombre significatif d'exons épissés alternativement peuvent être affectés par la déplétion en protéines PTBP1 et MRG15. Certaines expériences montrent que d'autres systèmes

d'adaptateurs de ce type existent, comme CHD1, recruté par H3K4me3, qui interagit avec la snRNP U2, et dont la déplétion provoque une diminution générale de l'épissage et du recrutement de la snRNP U2 au niveau des complexes spliceosomaux (Sims et al., 2007)(pour revue, Luco et al., 2011). De même, les nucléosomes formés dans les régions exoniques sont généralement enrichis en un variant de l'histone H2A, H2A.bbd, qui est capable d'interagir avec de nombreux composants du spliceosome, et dont la déplétion perturbe de façon globale l'épissage, constitutif comme alternatif (Tolstorukov et al., 2012). Enfin, l'état de modification des histones peut affecter la localisation de facteurs d'épissage. En effet, l'induction de l'hyperacétylation des histones ou la déplétion de la protéine de l'hétérochromatine HP1- α induit l'accumulation de facteurs d'épissage tels que U2AF65 dans les "speckles", du fait de la diminution de leur recrutement au niveau des ARN naissants (Schor et al., 2012).

1.3.4. Couplage de l'épissage avec la stabilité et le transport des ARNm

Par son rôle dans la mise en place du complexe protéique de contrôle qualité des ARN nommé complexe EJC (pour "Exon Junction Complex"), l'épissage est également impliqué dans le contrôle qualité et le transport des ARNm. Au cours de l'épissage et de manière épissage-dépendante, ce complexe EJC se fixe sur les ARNm 20 à 24 nt en amont de chaque jonction exon-exon (Le Hir et al., 2000b; Jurica and Moore, 2003; Reichert et al., 2002). Ce complexe n'est dissocié de l'ARNm que par le passage des ribosomes pendant les premiers cycles de traduction (Dostie and Dreyfuss, 2002; Isken et al., 2008). Il contient un coeur tétramérique stable constitué par les protéines Y14, Magoh, eIF4A III et MLN51(Tange et al., 2005). Il contient également les protéines Upf1, Upf2 et Upf3a/b (pour "Upstream Frameshift"), qui participent au système de surveillance des ARNm, appelé NMD (pour "Nonsense Mediated Decay"). Le système NMD induit la dégradation des ARNm dont le codon stop est situé plus de 50 nt en amont de la dernière jonction exon-exon, ceci grâce aux protéines recrutées par le complexe EJC (pour revue, Rebbapragada and Lykke-Andersen, 2009). Le fonctionnement de ce système est basé sur le fait que les séquences des ARNm de mammifères ont évolué de manière à ce que les codons stop soient situés après ou à proximité de la dernière jonction exon-exon. Etant donné que le mécanisme complexe de l'épissage peut conduire par erreur ou du fait de mutations à des ARNm présentant un codon stop prématuré, le système NMD permet d'éliminer l'ARNm aberrant (figure I-7). Lorsque le ribosome rencontre un codon stop situé plus de 50-55 nt en amont d'une jonction exon-exon terminale, les composants de l'EJC recrutent les protéines Upf et déclenchent le relargage du ribosome par interaction avec des facteurs de relargage, les protéines eRF. Ceci conduit à la prise en charge de l'ARNm par les complexes classiques de dégradation des ARNm impliquant l'élimination de la coiffe, la déadénylation et l'action d'exoribonucléases (pour revue, Schoenberg and Maquat, 2012). Ceci évite la production de protéines tronquées pouvant avoir un effet délétère dans les cellules.

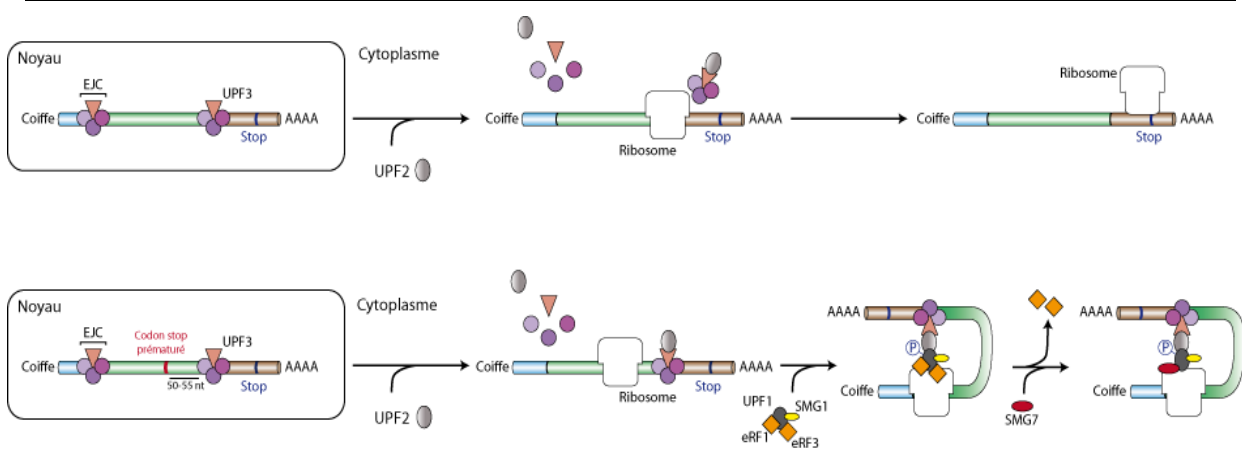


Figure I-7 : Elimination des ARNm présentant un codon stop prématuré par le système NMD (adapté de Carrier et al., 2010). En haut, ARNm ne présentant pas de codon stop prématuré. La présence d'un codon stop prématuré (ARNm du bas, en rouge) provoque le l'arrêt du ribosome et le recrutement des protéines du système NMD si ce codon stop est situé plus de 50-55 nt en amont de la jonction exon-exon, repérée par les protéines du complexe EJC.

L'importance du système NMD est reflétée par le fait que des mutations dans les composants de ce système chez la souris provoquent un arrêt du développement embryonnaire à un stade précoce, et que les mutations de gènes essentiels au NMD chez les mammifères sont létales (Medghalchi et al., 2001). Le système NMD constitue également un moyen de réguler post-transcriptionnellement l'expression de certains gènes. Ainsi, l'inclusion ou l'exclusion d'un exon contenant un codon stop prématuré peut induire soit une diminution (dégradation) soit une augmentation du taux d'ARNm (Hillman et al., 2004a; Lareau et al., 2007; Plocik and Guthrie, 2012).

Le complexe EJC a aussi un rôle dans l'export des ARNm du noyau vers le cytoplasme. Des expériences réalisées dans des oocytes de xénope ont montré que le complexe EJC augmente l'efficacité d'export des ARNm vers le cytoplasme (Zhou et al., 2000). Des protéines du complexe EJC, UAP56, Aly/REF et THO, formant le sous-complexe TREX (pour TRanscription/EXport), interagissent avec la protéine TAP permettant le passage des ARNm au travers des pores nucléaires (Gatfield et al., 2001; Jensen et al., 2001). Les ARNm sans introns disposent pour la plupart de séquences intrinsèques permettant le recrutement direct du complexe TREX/TAP (Lei et al., 2011). Il a été montré récemment que l'interaction directe entre les complexes EJC répartis le long des jonctions exon-exon de l'ARNm avec de nombreuses protéines SR et SR-like permet de former un megacomplexe intervenant dans la compaction et la protection des ARNm et leur export vers le cytoplasme (Singh et al., 2012).

I.4. L'épissage alternatif

Comme il a été évoqué précédemment, chez les vertébrés, les séquences des sites d'épissage sont souvent dégénérées par rapport aux séquences consensus et sont par conséquent sous-optimales. Aussi, chez l'homme, les séquences des sites 5' et 3' ainsi que du point de branchement ne sont généralement pas suffisantes pour définir les introns et les exons des ARNm. Elles sont généralement accompagnées de séquences régulatrices (Lim and Burge, 2001). Cela permet une régulation du taux d'utilisation de ces sites et aussi le choix de sites alternatifs avec production d'isoformes d'ARNm. Les possibilités croissantes d'épissage qui se sont développées au cours de l'évolution ont permis d'augmenter fortement la capacité codante des génomes. L'épissage alternatif permet notamment de moduler l'expression des gènes en fonction du type cellulaire, du stade de développement ou de l'état physiologique d'un organisme (Black and Grabowski, 2003). Environ 80% des événements d'épissage alternatif ont pour conséquence de modifier la séquence codante des protéines et conduisent à la production de protéines distinctes. D'abord estimé comme affectant environ 60% des pré-ARNm au début des années 2000, ce chiffre a plus récemment été porté à environ 90-95% grâce à l'émergence de nouvelles technologies d'analyse comme le séquençage haut débit des ARN (Modrek and Lee, 2002; Nilsen and Graveley, 2010; Pan et al., 2008; Wang et al., 2008a).

I.4.1. Les différents types d'événements d'épissage alternatif

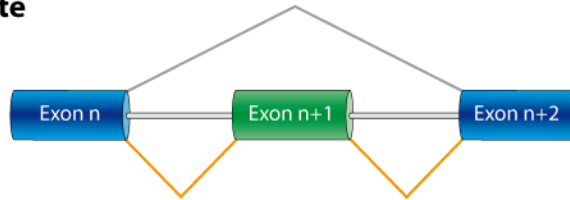
Le processus d'épissage alternatif peut se dérouler selon plusieurs schémas (figure I-8) :

- exons « cassettes » : ces exons alternatifs peuvent être soit inclus, soit exclus de l'ARNm mature.
- exons mutuellement exclusifs : deux exons mutuellement exclusifs qui ne peuvent être présents simultanément dans l'ARNm. Le choix d'un des couples de sites 5' et 3' abolit l'utilisation de l'autre couple et vice versa,
- sites 5' et 3' d'épissage alternatifs : plusieurs sites 5' ou 3' d'épissage sont utilisables pour un même exon et sont donc en compétition les uns avec les autres. Suivant le site utilisé, la taille de l'ARNm produit est différente.
- rétention d'introns : dans de rares cas, l'absence d'épissage entre un site 5' et un site 3' fait qu'un intron peut échapper à l'épissage et rester présent dans l'ARNm mature
- promoteurs ou sites de polyadénylation alternatifs : l'utilisation de promoteurs et de sites de polyadénylation différents influence l'identité des exons terminaux d'un ARNm.

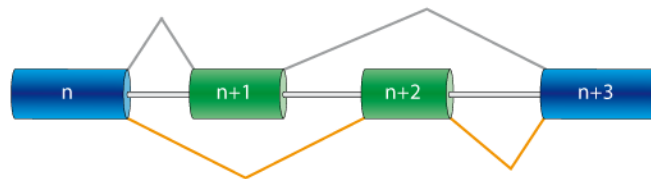
Ces différents cas de figures peuvent être combinés au sein d'un même pré-ARNm pour conduire à différentes isoformes d'ARNm. La production de ces isoformes dépend de l'identité des sites d'épissage et de la présence de séquences régulatrices. La sélection des sites d'épissage est souvent dépendante du

type cellulaire, des conditions environnementales et de l'état physiologique de la cellule, et bien sûr du patrimoine génétique de l'individu (le sexe par exemple, mais aussi des polymorphismes de l'ADN ou de mutations génétiques).

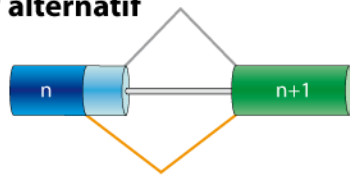
Exon cassette



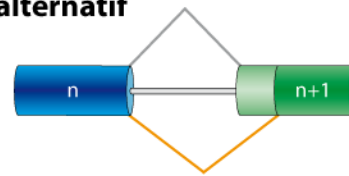
Exons mutuellement exclusifs



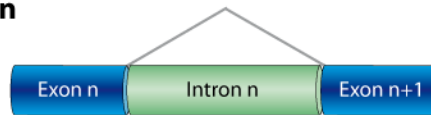
Site 5' alternatif



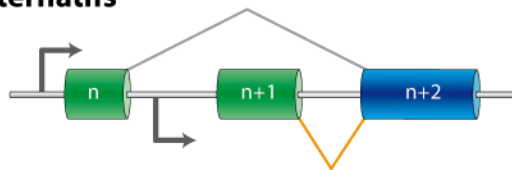
Site 3' alternatif



Rétention d'intron



Promoteurs alternatifs



Sites de polyadénylation alternatifs

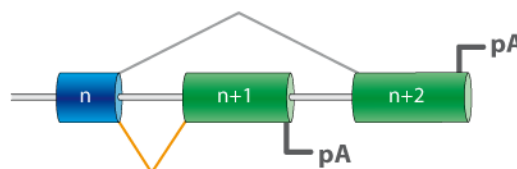


Figure I-8 : Les différents modes d'épissage alternatif des pré-ARNm. Les exons constitutifs sont représentés en bleu et les exons épissés alternativement sont représentés en vert. Les flèches coudées représentent des promoteurs alternatifs et la symbole "pA" représente des sites de polyadénylation alternatifs.

I.4.2. Séquences cis-régulatrices de l'épissage

Lorsque les séquences des sites d'épissage sont très dégénérées, ou que leur accès est rendu difficile par la fixation de protéines générant un encombrement stérique ou encore par une forte structuration de l'ARN, l'assemblage des complexes spliceosomaux nécessite l'association de facteurs protéiques activateurs sur des éléments de séquences agissant *in cis*. Ces séquences activatrices peuvent être exoniques (ESE, pour "Exonic Splicing Enhancer"), ou introniques (ISE, pour "Intronic Splicing Enhancer"). A l'inverse, les éléments de séquences interagissant avec des protéines inhibitrices sont appelés ISS (pour "Intronic Splicing Element"), et ESS (pour "Exonic Splicing Element").

Au niveau des sites d'épissage alternatifs, on retrouve souvent à la fois des séquences inhibitrices et des séquences activatrices. Un certain nombre d'outils bioinformatiques permettent de faciliter l'identification de ces différentes séquences régulatrices (Cartegni et al., 2003; Fairbrother and Chasin, 2000; Voelker and Berglund, 2007; Wang et al., 2004; Zhang and Chasin, 2006). Ces séquences cis-régulatrices agissent par l'intermédiaire de protéines partenaires : les protéines SR (déjà évoquées précédemment) sont majoritairement associées aux éléments activateurs (ISE et ESE), tandis que les protéines hnRNP sont souvent associées aux séquences inhibitrices (ISS et ESS) (pour revue, Chen and Manley, 2009).

I.4.3. Les protéines hnRNP

Les protéines hnRNP (heterogeneous nuclear RiboNucleoProteins) ont été initialement décrites pour leur capacité à s'associer co-transcriptionnellement aux transcrits de l'ARN pol II pour former des complexes ribonucléoprotéiques hétérogènes hnRNP (heterogeneous nuclear RiboNucleoParticles) (pour revue, Dreyfuss et al., 1993). Ces particules sont localisées *in situ* dans des régions du noyau caractérisées en microscopie électronique comme comportant des fibrilles et granules péri-chromatiniennes ("perichromatine fibrils"), qui environnent la chromatine. Très tôt il avait été montré que ces régions contiennent une grande partie des protéines s'associant de façon co-transcriptionnelle à l'ARN naissant, comme les hnRNP, mais également des UsnRNP, des protéines SR et des protéines de l'EJC (Fakan et al., 1984). Des traitements aux RNases ou par des sels de ces complexes hnRNP produisent deux types de particules aux propriétés biochimiques distinctes. Le premier type correspond au composant hnRNP fibrillaire associé à la matrice nucléaire. Il est peu sensible aux traitements aux RNases et n'est pas facilement dissous par les sels. Le deuxième type, constitué par des mono-particules hnRNP d'environ 40S contenant des fragments d'ARN de moins de 1 kb, est plus facilement dissous en présence de sels (Gallinaro and Jacob, 1981; Mähl et al., 1989). Ces mono-particules sont principalement enrichies en protéines hnRNP nommées A1, A2, B1, B2, C1, C2, hnRNP H3 (2H9) et hnRNP M (Beyer et al., 1977; Datar et al., 1993; Mahé et al., 1997). Comme il sera expliqué dans le dernier chapitre de cette introduction, certaines protéines hnRNP voient leurs propriétés d'association avec les pré-ARNm

modifiées dans certaines conditions comme par exemple suite à un choc thermique. Aujourd'hui plus de 25 protéines hnRNP différentes, classées par ordre alphabétique de A (34 kDa) à U (120 kDa) ont été identifiées, principalement grâce à l'immuno-sélection des complexes hnRNP (Beyer et al., 1977; Dreyfuss et al., 1993; Pinol-Roma et al., 1988). Certaines des protéines hnRNP font partie des protéines les plus abondantes dans le noyau. Les protéines hnRNP A1 et C1 sont présentes en plus grandes quantités que les protéines de la snRNP U1, et jouent un rôle général dans le métabolisme des ARNm (pour revue, Dreyfuss et al., 1993; Han et al., 2010a). Les principaux rôles connus des protéines hnRNP concernent la structuration et la maturation des ARN, dont l'épissage alternatif. Elles peuvent également intervenir dans l'activation de la transcription et le remodelage de la chromatine, et sont présentes dans le cytoplasme où elles ont un rôle dans la stabilité, le transport et la traduction des ARN (Björk and Wieslander, 2011; Busch and Hertel, 2012; Dreyfuss et al., 2002; Han et al., 2010a; Kafasla et al., 2012; Ostareck-Lederer and Ostareck, 2012).












1.4.3.a. Domaines d'interaction des protéines hnRNP avec leurs ARN cibles

Les protéines hnRNP contiennent toutes un ou plusieurs domaines d'interaction avec l'ARN : domaines RRM, retrouvés dans les protéines SR, domaines RRM H, ou et domaines KH (pour K Homology) (tableau I-3). Le domaine KH a été mis en évidence en étudiant la protéine hnRNP K (d'où son nom, KH, pour "K Homology"). Il peut être présent en de multiples exemplaires dans la même protéine (3 par exemple dans hnRNP K (Siomi et al., 1993)). La forme basique de ce domaine, d'une cinquantaine d'acides aminés, est un enchaînement de structures $\beta\alpha\beta$, avec au coeur un motif conservé VIGX₂GX₂I (ou X est n'importe quel acide aminé), mais les domaines KH peuvent contenir une extension supplémentaire C-terminale $\beta\alpha$ (ils sont dits de type I, comme c'est le cas pour la protéine hnRNP K) ou N-terminale $\alpha\beta$ (ils sont alors dits de type II) (pour revue, Grishin, 2001)

Les protéines hnRNP peuvent s'associer avec un large panel de séquences sur l'ARN, mais certaines protéines hnRNP ont aussi une forte affinité avec certains types de séquences, comme par exemple les séquences riches en G pour les hnRNP A/B, E, H, H', 2H9, M et U. Les spécificités d'interactions de certaines protéines hnRNP avec l'ARN ont été évaluées par des expériences SELEX (tableau I-3).

Les domaines d'interaction avec l'ARN des protéines hnRNP sont souvent combinés avec un ou plusieurs domaines auxiliaires Gly (riches en résidus glycines), qui participent à la stabilisation des interactions avec l'ARN, et qui sont responsables d'interactions protéine-protéine (tableau I-3). Des résidus phénylalanines et tyrosines, présents tout le long du domaine Gly, sont nécessaires aux interactions hydrophobes avec d'autres protéines. Ces domaines sont également retrouvés dans de nombreuses protéines se fixant à l'ARN (Kiledjian and Dreyfuss, 1992; pour revue, Krecic and Swanson, 1999). Le domaine riche en G de hnRNP A1 stabilise les interactions des domaines RRM et

permet, par le biais d'interaction protéine-protéine, la fixation coopérative de plusieurs molécules de hnRNP A1 sur une même molécule d'ARN (Casas-Finet et al., 1993; Hallay et al., 2006). Cela constitue une propriété importante de hnRNP A1 pour la régulation de l'épissage alternatif, puisque la fixation de hnRNP A1 à un site de haute affinité grâce aux domaines RRM1 et RRM2 peut s'accompagner de la fixation sur des sites adjacents de plus faible affinité par oligomérisation (Hallay et al., 2006). La protéine hnRNP U est dépourvue de domaine RRM, et son interaction avec l'ARN est établie grâce à la présence de motifs RGG ou de domaines riches en Gly.

Nom	Domaines	Séquence d'interaction ARN
A/B		G ₄ (AU ₃) ₄ A (A0) UAGGG, (Py)UAG (A1, A1 ^B) GCCAAGGAGCC (A2, B1) GCCAAGGAGCC(A3)
C1,C2		Riche en U, AGUA(U)SGUGGA
D1 à D4		Riche en A/U, UUAG/G
E1, E2		Poly(C), Poly(G)
F		Poly(G)
G		AAGU
H, H'		GGGA/C, Poly(G)
H3 (2H9)		GGGA
HAP (SAF-B)		
I (PTB)		Riche en C/U
K1, K2		Poly(C), AU(C) _{3/4} (U/A) ₂







Nom	Domaines	Séquence d'interaction ARN
L		Répétitions C/A
M1 à M4		GAGGAGG
P		
Q1 à Q5		
R		
U (SAF-A)		Poly(G), riche en A/U

Tableau I-3 : Organisation structurale des protéines hnRNP (adapté de Han et al., 2010a). Certaines de ces hnRNP étant présentes en plusieurs isoformes différentes, il s'agit ici de représentations génériques pour chaque famille d'isoformes. D'autres hnRNP de la même famille sont issus de gènes distincts, c'est le cas des protéines de la famille hnRNP A/B dont A0, A1, A2/B1, A3 et AB, qui sont issues de l'épissage alternatif des pré-ARNm produits à partir de plusieurs gènes distincts (Han et al., 2010b). Les différents domaines des hnRNP sont indiqués. RRM : "RNA Recognition Motif" ; RRM H : RRM Homologue ; RGG et Gly : domaines riches en répétitions RGG ou en Glycine ; KH : K Homology. En plus des domaines notés ici, les protéines hnRNP présentent également divers autres domaines spécifiques (signalés par une étoile). Les séquences de fixation des hnRNP sur l'ARN établies par SELEX sont indiquées à droite (pour revue, Dreyfuss et al., 2002; Krecic and Swanson, 1999).

De plus, les protéines hnRNP subissent de nombreuses modifications post-traductionnelles. Il s'agit en majorité de phosphorylations, de méthylation des résidus arginines et de sumoylation. Des changements dans ces taux de modifications post-traductionnelles entraînent des modifications de la capacité d'interaction des protéines hnRNP avec l'ARN. Par exemple, suite à des stress cellulaires comme le stress oxydant, la protéine hnRNP K est phosphorylée au niveau de résidus sérines. Cela augmente son affinité pour d'autres protéines comme des protéines kinases ou des répresseurs de la transcription, et diminue *in vitro* sa capacité d'interaction avec l'ARN (Schullery et al., 1999). Les protéines hnRNP C et M sont majoritairement nucléaires, et elles interagissent avec certains ARNm pour les retenir dans le noyau. La sumoylation de ces protéines diminue leur affinité pour leurs ARNm cibles, ce qui faciliterait leur export vers le cytoplasme (Vassileva and Matunis, 2004).

1.4.3.b. Localisation des protéines hnRNP dans la cellule

Les protéines hnRNP ont généralement une distribution diffuse dans le nucléoplasme, mais sont exclues du nucléole. Elles font la navette entre le noyau et le cytoplasme, à quelques exceptions près, et

ce par plusieurs mécanismes différents. Les hnRNP de la famille A1 et I ne contiennent pas de NLS classique, et leur mouvement est couplé à la transcription, tandis que le transport de hnRNP K, qui contient un NLS (pour "Nuclear Localization Signal"), est indépendant de la transcription (Gama-Carvalho and Carmo-Fonseca, 2001). Les protéines hnRNP A1, A2 et B1 contiennent toutes en C-terminal un motif de 38 résidus, appelé M9, riche en glycines et en résidus aromatiques, leur permettant de faire la navette entre le noyau et le cytoplasme grâce à l'interaction avec la transportine (Pollard et al., 1996; Siomi and Dreyfuss, 1995; Weighardt et al., 1995). La localisation de hnRNP A1 dans la cellule est régulée par son degré de phosphorylation. La phosphorylation d'une séquence située en aval du signal M9, appelé peptide F, empêche l'interaction avec la transportine et diminue l'import nucléaire de hnRNP A1 (Allemand et al., 2005). Il a été montré que la protéine hnRNP A1, ainsi que la protéine hnRNP P, sont exportées vers le cytoplasme après un choc thermique, lors de la formation de granules enrichis en ARNm dont la traduction est inhibée. Ces granules sont appelés granules de stress cytoplasmiques (Andersson et al., 2008; Guil et al., 2006). La déplétion en protéine hnRNP A1 altère la capacité de récupération des cellules après un stress, indiquant un rôle important de hnRNP A1 dans ce processus. Les protéines hnRNP A2 et B1 sont elles impliquées dans le transport des ARNm le long des microtubules pour permettre leur traduction dans des régions particulières de la cellule, en particulier dans les neurones (Sinnamon and Czaplinski, 2011). Les protéines PTB, ou hnRNP I, contiennent un élément NLS correspondant aux 50 premiers résidus et permettant leur import dans le noyau (Ghetti et al., 1992). Bien que ces protéines soient majoritairement localisées dans le nucléoplasme, elles peuvent également être réexportées dans le cytoplasme.

I.4.4. Rôle des protéines hnRNP dans l'épissage alternatif

Grâce à leurs propriétés d'interaction avec l'ARN et avec des protéines partenaires, les protéines hnRNP sont impliquées dans la formation de complexes ribonucléoprotéiques influençant la fixation des facteurs d'épissage et des composants du spliceosome.

La plupart du temps, les protéines hnRNP jouent le rôle d'inhibiteur de l'épissage en empêchant la fixation de la machinerie d'épissage (figure I-9). Les protéines hnRNP interagissent sur l'ARN avec des séquences inhibitrices introniques (ISS, Intrinsic Splicing Sequence) ou exoniques (Exonic Splicing Sequence), et agissent sur les sites donneurs d'épissage aussi bien qu'accepteurs. Ainsi hnRNP A1 par exemple, en se fixant sur des séquences inhibitrices exoniques (ESS) contenant des triplets de nucléotides UAG, empêche l'assemblage du spliceosome au niveau des sites d'épissage du pré-ARNm (Caputi et al., 1999; Hallay et al., 2006; Marchand et al., 2002). Cette inhibition par encombrement stérique peut également avoir lieu au niveau de points de branchement. La propriété de certaines hnRNP à oligomériser sur le pré-ARNm, comme hnRNP A1 ou bien hnRNP I, prend part à ce mécanisme

d'encombrement stérique en leur permettant de recouvrir l'ARN et de gêner l'accès des autres protéines (Hallay et al., 2006).

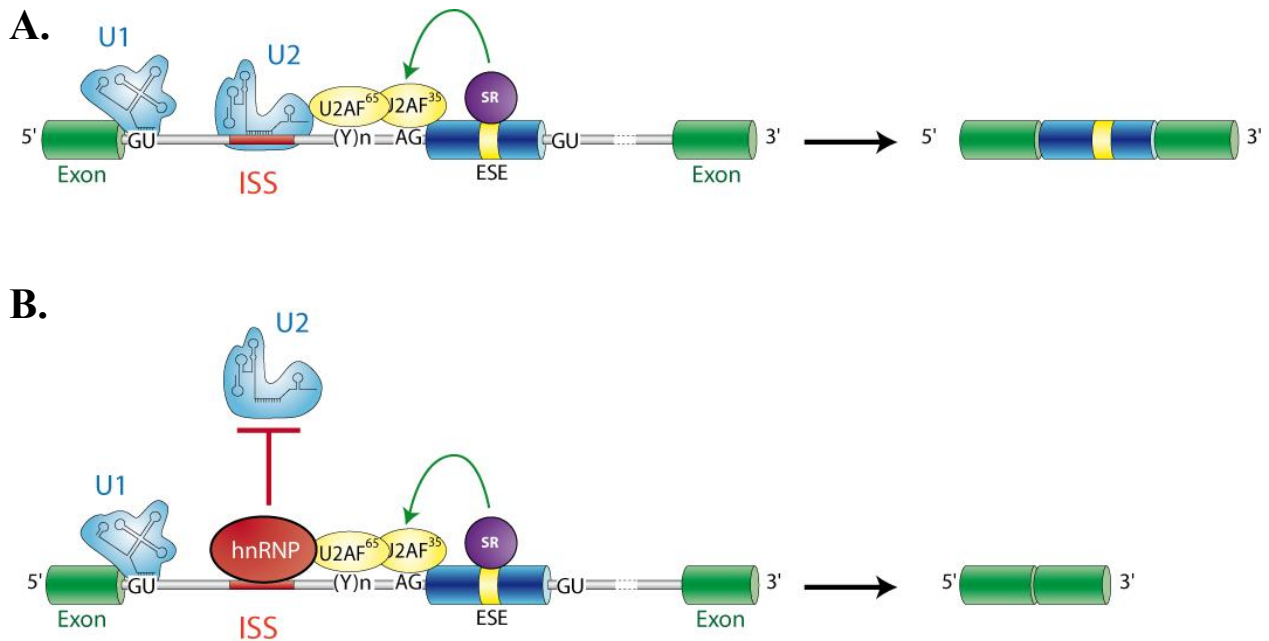


Figure I-9 : Exemple d'inhibition de l'épissage exercée par les hnRNP en limitant l'accessibilité de la machinerie d'épissage aux sites d'épissage. A. Epissage sans l'action de hnRNP. **B.** Epissage avec l'action de hnRNP.

Un autre mode intéressant de régulation par certaines protéines hnRNP est la formation de grandes boucles d'ARN séquestrant les sites d'épissage. C'est par exemple par ce mode de régulation que la protéine hnRNP A1 peut réguler l'épissage de son propre pré-ARNm. L'exon 7B de ce pré-ARNm est bordé par deux introns contenant chacun un site de haute affinité pour hnRNP A1. La fixation de hnRNP A1 à ces sites et l'interaction entre les deux protéines fixées induit la formation d'une boucle dans l'ARN (figure I-10). Il en résulte la séquestration spatiale de l'exon 7B et le rapprochement des sites 5' et 3' d'épissage correspondant aux exons 7 et 8 conduisant à l'exclusion de l'exon 7B et à la formation d'une isoforme spécifique nommée hnRNP A1^B. Ce type de régulation par hnRNP A1 semble assez général, en particulier pour de très larges introns. Les protéines hnRNP H et PTB (hnRNP I) régulent également l'épissage alternatif en formant ce type de boucles (Martinez-Contreras et al., 2006).

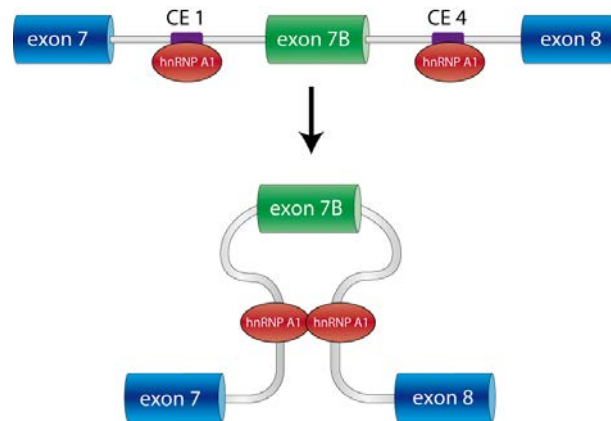


Figure I-10 : Régulation par hnRNP A1 de l'épissage de son propre pré-ARNm. L'interaction distale entre deux molécules de hnRNP A1 (en rouge), fixées sur leurs sites de fixation nommés CE1 et CE4, de part et d'autre de l'exon 7B, conduit à l'exclusion spatiale de cet exon grâce à la formation d'une boucle.

I.4.5. Régulation de l'épissage alternatif par les facteurs activateurs et inhibiteurs de l'épissage

I.4.5.a. Antagonisme entre les protéines SR et hnRNP dans l'épissage alternatif

De la même façon qu'elles limitent l'accessibilité de l'ARNm à la machinerie d'épissage, les protéines hnRNP peuvent interférer avec la fixation des protéines SR et contrecarrer leur effet activateur. C'est le cas, par exemple, lorsqu'il y a juxtaposition entre deux séquences exoniques, une de type ESE, recrutant une protéine SR, et une de type ESS, recrutant une protéine hnRNP (pour revue, Smith and Valcarcel, 2000). Les deux protéines entrent alors en compétition pour la liaison sur le même site, et le taux relatif de ces protéines détermine quelle protéine est fixée en majorité et donc l'effet sur l'épissage. Si le seul effet de la protéine SR est de contrecarrer la fixation de la protéine hnRNP, son ou ses domaines RRM peuvent être suffisants pour cette activité.

Cet antagonisme entre les protéines SR et hnRNP est un mécanisme général de régulation de l'épissage alternatif, et ainsi l'épissage alternatif dépend du ratio SR / hnRNP présent dans les cellules, ratio qui peut varier par exemple de manière tissu-spécifique. Un autre paramètre essentiel dans cette régulation est l'affinité des protéines régulatrices en compétition pour le pré-ARNm.

Par exemple, il a été montré très tôt que si 2 sites 5' d'épissage sont en compétition au niveau d'un intron, que la protéine SRSF1 induit l'utilisation du site 5' proximal, alors que la protéine hnRNP A1 induit l'utilisation du site distal (Mayeda and Krainer, 1992). L'oligomérisation de la hnRNP A1 le long de l'ARNm empêche l'accès des facteurs d'épissage comme la snRNP U1 alors qu'à l'inverse la protéine SRSF1 favorise le recrutement de cette UsnRNP (figure I-11) (Eperon et al., 2000). L'expression de différents isoformes protéiques dépend donc souvent directement du taux relatif entre les

protéines SRSF1 et hnRNP A1 au sein du noyau, qui varient selon les tissus (Caceres et al., 1994; Dirksen et al., 2000; Hanamura et al., 1998; Kamma et al., 1995).

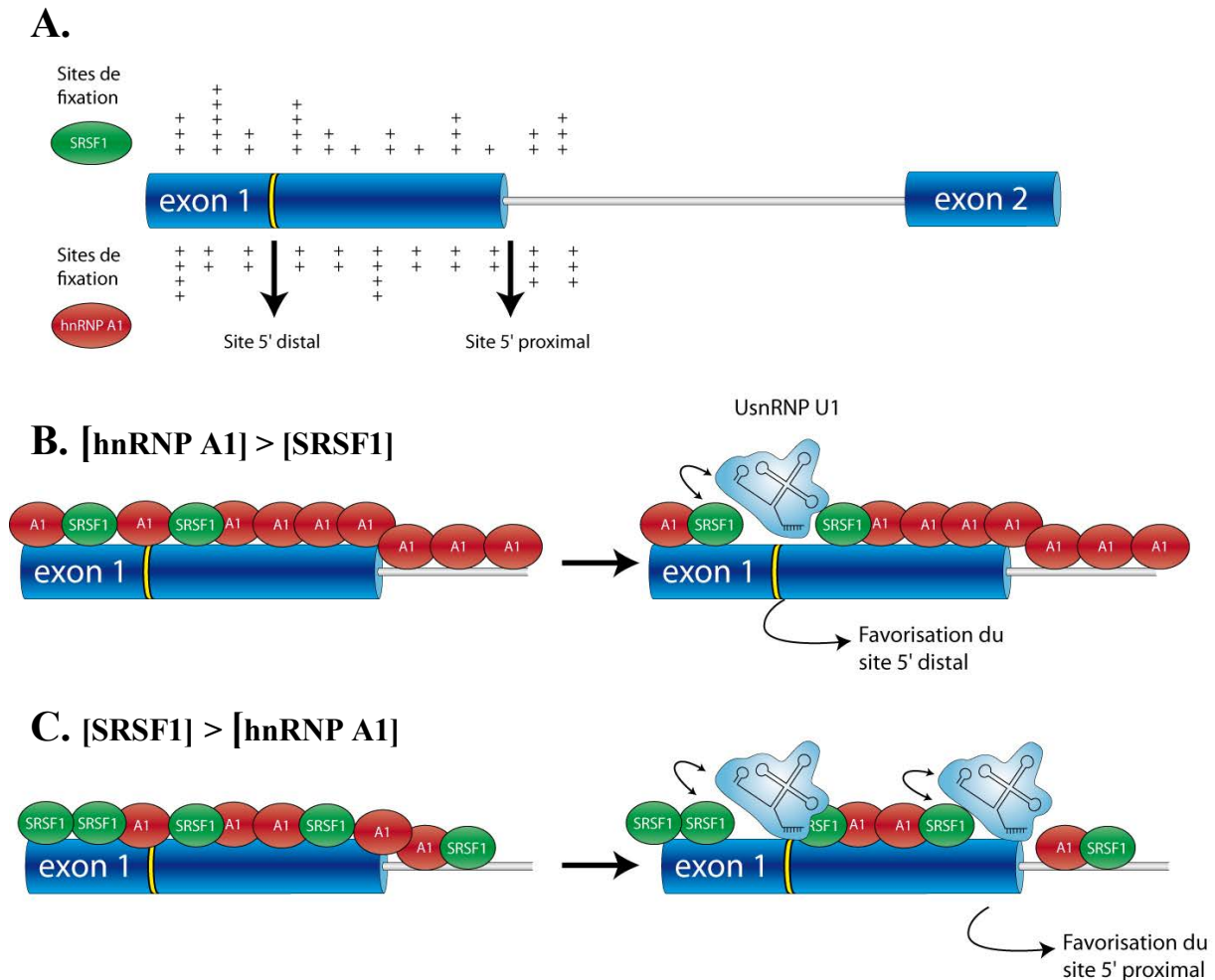


Figure I-11 : Représentation schématique de la compétition entre hnRNP A1 et SRSF1 au niveau de sites d'épissage 5' en compétition (d'après Eperon et al., 2000). **A.** Sites de fixation des deux protéines, d'affinité faible (+) à très élevée (++++), sur un pré-ARNm disposant de deux sites 5' d'épissage, distal et proximal, au niveau du premier exon. Cette distribution prend en compte le fait que ces protéines, bien qu'elles aient des sites de fixation bien définis, sont capables de se répartir sur une grande partie de l'ARN aux concentrations présentes dans les expériences d'épissage alternatif *in vitro* utilisées pour établir ce modèle. **B.** Lorsque la concentration en hnRNP A1 est beaucoup plus élevée que celle de SRSF1 dans un extrait nucléaire, SRSF1 n'occupe que les sites pour lesquels elle a une haute affinité. Dans ce cas, le site 5' distal est favorisé grâce à sa proximité avec un site de fixation de SRSF1, qui améliore la fixation de la snRNP U1. A l'inverse, le site proximal est inhibé par compétition avec la protéine hnRNP A1. **C.** Lorsque la concentration en SRSF1 est plus élevée dans l'extrait nucléaire, cette protéine occupe les sites d'affinité faible, moyenne et forte. Dans ce cas, la snRNP U1 occupe les deux sites 5' d'épissage, ce qui favorise le site proximal.

Il a également été montré au laboratoire que la régulation du site A7, critique dans la production de l'ARNm de la protéine virale Tat du VIH, fait intervenir une fixation coopérative de hnRNPA1, qui peut s'oligomériser et recouvrir toute une partie de l'ARN à proximité de ce site, à partir de seulement 3

sites de régulation de forte affinité. Cela empêche, tout comme dans l'exemple ci-dessus, la fixation d'autres protéines régulatrices de l'épissage comme SRSF1, et inhibe l'utilisation de ce site (Marchand et al., 2002).

La régulation de l'utilisation des sites d'épissages peut faire intervenir des mécanismes complexes d'antagonisme entre les protéines SR et hnRNP, et un grand nombre de partenaires. Ainsi, au laboratoire, il a également été montré que le site d'épissage A3 de l'ARN du VIH1 est régulé par l'intervention d'un grand nombre d'éléments ESS et ESE qui peuvent chacun interagir avec plusieurs régulateurs de l'épissage. Si bien qu'au moins 4 protéines inhibitrices, hnRNP A1, hnRNP H et hnRNP A2 ainsi que la protéine DAZAP1, et les protéines activatrices SRSF1, SRSF2, SRSF5 et SRSF7 sont impliquées dans les différentes régulations possibles de ce site (Hallay et al., 2006; Jacquenet et al., 2001; Ropers et al., 2004; Thèse G. Khoury "Étude des mécanismes moléculaires régulant l'expression de la protéine tat du virus de l'immunodéficience humaine, au niveau de la production de ses ARNm et de leur traduction"). D'autre part, en plus de faire intervenir les protéines hnRNP A1 et SRSF1, la régulation de l'épissage au site A7 fait intervenir la protéine hnRNP K qui inhibe l'utilisation de ce site et favorise la fixation de la protéine hnRNP A1 (Marchand et al., 2011).

Enfin, il peut être bon de mentionner que les protéines hnRNP, bien que généralement inhibitrices de l'épissage, peuvent dans de rares cas avoir un rôle d'activateur de l'épissage, tandis que les protéines SR bien que généralement activatrices peuvent avoir un rôle d'inhibiteur de l'épissage. C'est le cas par exemple de la protéine hnRNP H, qui active l'inclusion de l'exon 19 dans l'ARNm d'un récepteur au glutamate, GRIN1 (Han et al., 2005).

1.4.5.b. Interdépendance complexe entre les différents facteurs de l'épissage pour la régulation de l'épissage alternatif

Les protéines SR et hnRNP ne sont pas les seules protéines intervenant dans la régulation de l'épissage alternatif. Il existe en effet de nombreuses autres protéines régulatrices de l'épissage qui n'ont pas été présentées ici et qui interviennent dans des cas plus spécifiques. C'est le cas des régulateurs Nova-1 et Nova-2, qui ne sont exprimés que dans les cellules du système nerveux central (Buckanovich et al., 1993 ; (Buckanovich et al., 1993; Park et al., 2011). D'autres protéines, comme les protéines Fox-1 et Fox-2, les protéines de la famille CELF (pour "CUG-BP and ETR-3-like factor"), les protéines MBNL-1, MBNL-2 et MBNL-3, les protéines TDP-43 et RBMY ou encore le facteur YB-1, n'interagissent qu'avec des séquences d'ARN spécifiques (Cooper, 1998; Dredge and Darnell, 2003; Ladd et al., 2001; Ponthier et al., 2006; Ryan and Cooper, 1996). YB-1, par exemple, interagit spécifiquement avec des séquences activatrices d'épissage riches en nucléotides A et C, appelées ACE (Stickeler et al., 2001). L'épissage alternatif d'un pré-ARNm donné implique donc une combinaison complexe de facteurs d'épissage, parmi tous les facteurs qui ont été mentionnés au cours de ce chapitre.

Enfin, un dernier type d'élément vient encore s'ajouter à l'ensemble déjà complexe des régulations affectant l'épissage, à savoir la structure secondaire de l'ARN. Les pré-ARNm contiennent en effet des séquences permettant de faciliter ou au contraire de bloquer l'accès du spliceosome et/ou des facteurs de régulation de l'épissage à l'ARN.

I.4.6. Influence de la structure secondaire des ARN

I.4.6.a. Implication de la structure secondaire dans l'inhibition de l'épissage

Comme nous l'avons vu au travers de l'exemple du pré-ARNm de la protéine hnRNP A1, la présence de sites de liaison de hnRNP A1 de part et d'autre de l'exon 7B favorise la formation au sein de l'ARN d'une boucle qui, tout en réduisant la distance physique entre le site 5' distal et le site 3', séquestre l'exon 7B et empêche son inclusion. De façon intéressante, le même effet peut-être obtenu grâce à des séquences inversées complémentaires situées de part et d'autre de l'exon 7B. Ces séquences complémentaires appariées permettent de reproduire la même boucle grâce à la formation d'une structure tige-boucle longue distance. Il est intéressant de constater que la formation de ces structures tige-boucle peut emprisonner les sites d'épissage et diminuer leur accessibilité aux composants du spliceosome et par conséquent inhiber l'épissage.

C'est le cas pour le pré-ARNm de la β -tropomyosine de poulet (Sirand-Pugnet et al., 1995). Les exons 6A et 6B de ce pré-ARNm sont épissés de façon mutuellement exclusive et de façon tissu-spécifique. L'exon 6B n'est inclus que dans les cellules musculaires squelettiques et est exclu dans toutes les autres lignées cellulaires en faveur de l'exon 6A. L'épissage de l'exon 6B est régulé par des séquences introniques régulatrices situées en amont avec laquelle interagit notamment la protéine PTB (Saulière et al., 2006). L'étude expérimentale de la structure secondaire de l'ARN révèle que l'exon 6B et les introns qui le flanquent forment une structure en tige-boucle nécessaire *in vitro* et *in vivo* pour empêcher son inclusion dans les cellules non musculaires (figure I-12) (Clouet d'Orval et al., 1991a, 1991b; Libri et al., 1991; Sirand-Pugnet et al., 1995).

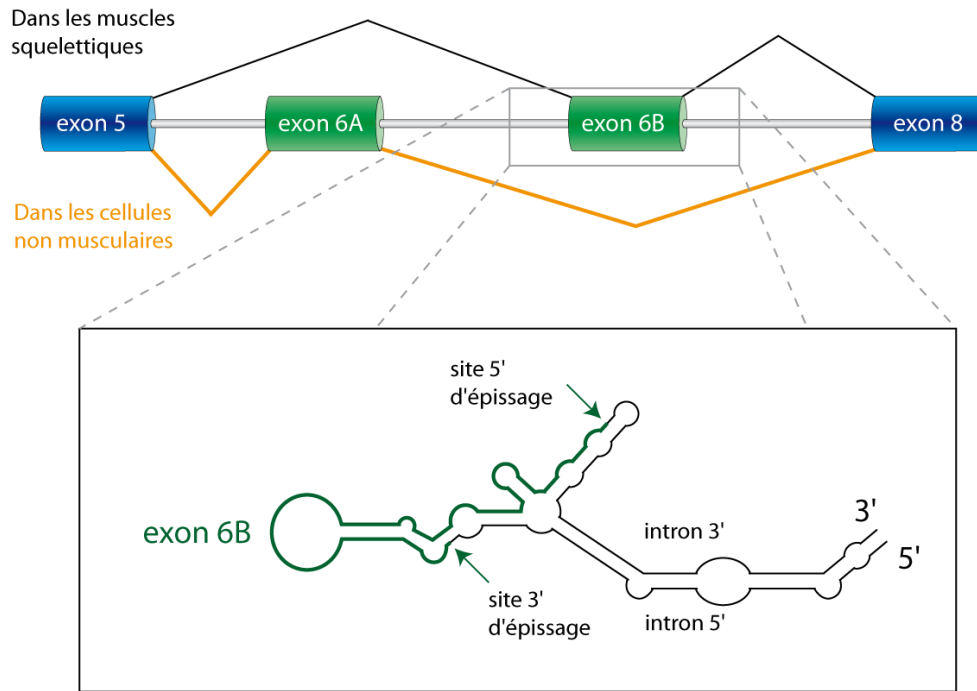


Figure I-12 : Exemple de l'inhibition de l'épissage de l'exon 6B du pré-ARNm de la β -tropomyosine dans les cellules non musculaires. L'emprisonnement de l'exon 6B dans une structure tige-boucle complexe (indiquée dans l'encadré) contribue à l'inhibition de l'épissage de cet exon dans les cellules non musculaires.

Un autre exemple est donné par l'ARNm de la protéine tau, qui est impliquée dans la polymérisation et la stabilisation des microtubules et joue un rôle dans le transport neuronal. Le pré-ARNm de cette protéine présente plusieurs exons épissés alternativement dont l'exon 10 (figure I-13) (pour revue, Andreadis, 2005). Le site 5' de l'exon 10 est enfermé au sein d'une structure tige-boucle, ce qui limite son utilisation lors de l'épissage (pour revue, Andreadis, 2005; Buratti and Baralle, 2004). Un déséquilibre entre les quantités d'ARNm contenant ou non cet exon 10 est à l'origine de l'agrégation de protéines tau insolubles menant à des maladies neurodégénératives. Ce type de déséquilibre peut notamment être causé par des mutations ponctuelles qui perturbent ou renforcent la structure secondaire de l'exon 10, ce qui affecte la fixation de la snRNP U1 au niveau du site donneur et donc le taux d'inclusion de cet exon.

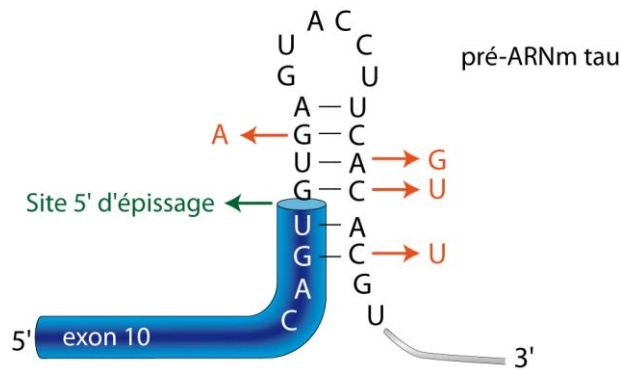


Figure I-13 : Inhibition de l'utilisation du site 5' d'épissage de l'exon 10 du pré-ARNm de la protéine tau. La protéine tau est une protéine importante dans les neurones, et l'exon 10 est situé dans le domaine d'interaction de la protéine tau avec les microtubules qui est important pour la polymérisation de ces derniers. La formation d'une structure secondaire tige-boucle localement au niveau du site 5' d'épissage de l'exon 10 du pré-ARNm tau contribue à l'inhibition de l'utilisation de ce site. Les mutations indiquées en orange sont des mutations retrouvées chez des patients souffrant de maladies neurodégénératives qui perturbent cette structure en tige-boucle, ce qui affecte l'inclusion de l'exon 10 dans l'ARNm de la protéine tau et donc la fonction de la protéine.

1.4.6.b. Implication de la structure secondaire dans l'activation de l'épissage

La structure secondaire peut également participer à l'activation de l'épissage des pré-ARNm. En effet, dans certaines conformations d'ARN les sites d'épissage ou les sites régulateurs peuvent être plus accessibles aux composants spliceosomaux ou aux facteurs de régulation. C'est le cas notamment pour l'exon alternatif EDA du pré-ARNm de la fibronectine de rat. Dans cet ARN, l'exon EDA est exclu dans les hépatocytes et est inclus dans toutes les autres lignées cellulaires. Cet exon contient un élément ESE, placé dans une boucle terminale, ce qui le rend accessible à la protéine SRSF1 et conduit à l'inclusion de l'exon EDA (figure I-14) (Buratti and Baralle, 2004; Lavigueur et al., 1993; Muro et al., 1999). Des mutations dans la séquence ESS, située en aval de l'élément ESE provoquent un changement de conformation de l'ARN. L'élément ESE se retrouve alors dans une structure en double-brin et est devenu inaccessible aux protéines SR. Il en résulte l'exclusion de l'exon EDA (Muro et al., 1999). Les structures secondaires peuvent aussi être impliquées dans le rapprochement dans l'espace de séquences régulatrices. Le pré-ARNm *FGFR-2* contient deux séquences introniques, ISE-2 et ISAR, distantes de 735 nt mais s'appariant sur 18 nt. Ces deux séquences sont situées entre les exons mutuellement exclusifs IIIb et IIIc et leur appariement favorise l'inclusion de l'exon IIIb dans les cellules épithéliales, chez l'homme comme chez le rat (figure I-14). En effet, toute mutation affectant cette structure secondaire affecte l'épissage, qui peut être restauré par des mutations compensatoires restaurant la structure (Muh et al., 2002). De plus, le remplacement des séquences ISE-2 et ISAR par deux séquences complémentaires ne perturbe pas cette régulation. Cette structure permet en effet le rapprochement d'un élément de séquence ESE nommé GCAUG, et d'une séquence ISS (nommée DISS), qui est inhibée par l'élément GCAUG (figure I-14).

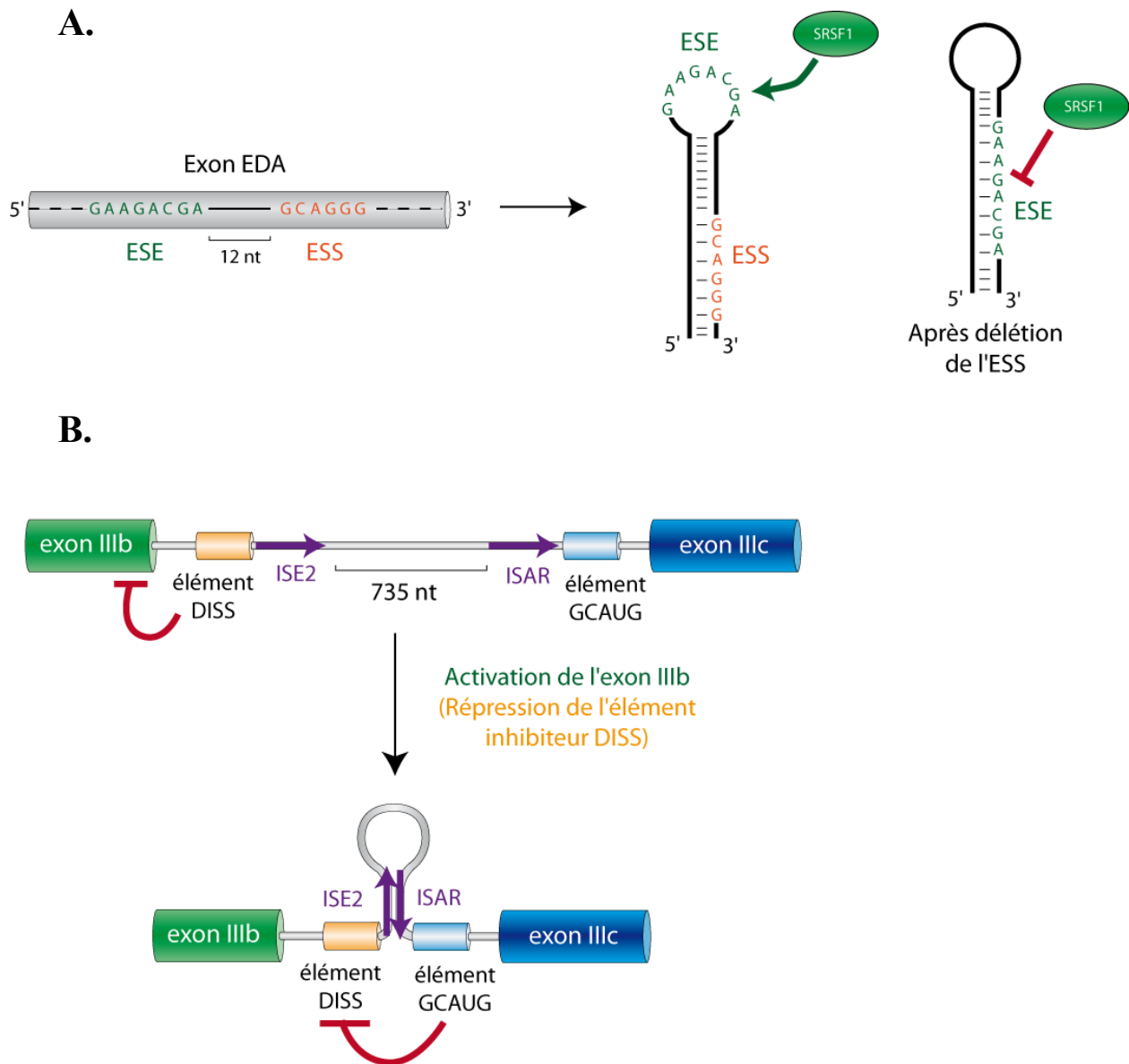


Figure I-14 : Exemples de l'effet de la modification de la structure secondaire de l'ARN sur l'activation de l'épissage. A. Exemple de l'exon EDA du pré-ARNm de la fibronectine de rat. Au sein de l'exon EDA, un élément ESE jouxte un élément ESS (à une distance de 12 nt). L'élément ESE est compris dans une boucle, ce qui le rend accessible à l'activateur d'épissage SRSF1 (ASF/SF2), ce qui permet l'inclusion de l'exon EDA. Lorsque l'élément ESS est déléte (à droite), l'élément ESE se retrouve emprisonné au sein de la structure tige-boucle, ce qui le rend inaccessible à la protéine SRSF1 et provoque l'exclusion de l'exon EDA. **B.** Exemple de l'exon IIIb du pré-ARNm FGFR-2. Les exons IIIb et IIIc indiqués sont mutuellement exclusifs. C'est l'appariement des éléments ISE2 et ISAR présents dans l'intron situé entre ces deux exons, qui va permettre l'inclusion de l'exon IIIb dans les cellules épithéliales. En effet, cet appariement permet d'inhiber l'élément régulateur DISS par le rapprochement de l'élément GCAUG.

I.5. Épissage alternatif et pathologies

I.5.1. Maladies génétiques liées à des mutations ponctuelles affectant directement l'épissage

On estime aujourd'hui qu'environ 20% à 30% des maladies génétiques sont liées à des mutations affectant l'épissage (Corcos and Solier, 2005; Singh and Cooper, 2012; Tazi et al., 2009). Pendant longtemps, les mutations silencieuses n'ont pas été prises en compte comme cause potentielle de maladies génétiques car elles ne modifiaient pas la séquence codante. Cependant il est clairement établi aujourd'hui que de nombreuses mutations silencieuses modifient l'épissage alternatif des pré-ARNm et donc affectent la séquence de la protéine résultante bien que la nature de l'acide aminé encodé ne soit pas modifiée (Raponi and Baralle, 2010). Par ailleurs, l'impact de mutations dans les introns pouvant affecter l'épissage alternatif a aussi longtemps été sous-estimé et non pris en compte dans l'étude des causes des maladies génétiques.

Les mutations à l'origine de pathologies génétiques liées à l'épissage peuvent être de plusieurs types. D'une part, ces mutations peuvent affecter directement les séquences de régulation de l'épissage ou les sites d'épissage d'un pré-ARNm, qu'ils soient constitutifs ou alternatifs. Ces mutations peuvent entraîner la synthèse d'isoformes aberrants d'ARNm conduisant à la production de protéines non fonctionnelles, ou avec une activité modifiée du fait de la perte ou l'ajout d'exons correspondant à des domaines protéiques. Ces mutations peuvent aussi affecter l'épissage de séquences 5' ou 3' UTR ayant un rôle important, ou encore mener à la dégradation de l'ARNm par le système NMD (pour revue, Cooper et al., 2009). A titre d'exemple, quelques-uns des gènes présentant des mutations affectant directement l'épissage au niveau de leur pré-ARNm (mutations dites « en cis ») sont présentées dans le tableau ci-dessous (Tableau I-4), avec les pathologies correspondantes :

Gène muté	Pathologie correspondante	Fonction de la protéine	OMIM
<i>PBGD</i>	Porphyrie intermittente aiguë	Synthèse de l'hème	17600
<i>BRCA1</i>	Cancer du sein et de l'ovaire	Prolifération cellulaire, composant de l'ARN pol II	113705
<i>PMM2</i>	Syndrome de Jaeken	Synthèse de GDP-mannose	212065
<i>CFTR</i>	Fibrose kystique	Canal chlore	602421
<i>Lysyl-hydroxylase</i>	Syndrome de Ehlers-Danlos du collagène	Formation d'hydroxylysines	153454
<i>FANCG</i>	Anémie de Fanconi	Réparation de l'ADN	227650
<i>HPRT</i>	Syndrome de Lesch-Nyhan	Métabolisme de l'ADN	308000
<i>Pyruvate déshydrogénase E1α</i>	Encéphalomyélopathie de Leigh	Métabolisme du pyruvate	312170
<i>Fibrilline-1</i>	Syndrome de Marfan	Constituant des microfibrilles	134797

Gène muté	Pathologie correspondante	Fonction de la protéine	OMIM
<i>NFI</i>	Neurofibromatose de type 1	Inhibiteur de la voie Ras	162200
<i>OCT</i>	Hyperammoniémie	Cycle de l'urée	300461
<i>UPC</i>	Porphyrie cutanée	Métabolisme des porphyrines	176000
<i>Hexosaminidase</i>	Maladie de Sandhoff	Métabolisme des gangliosides	606869
<i>Adénosine désaminase</i>	Immunodéficience combinée sévère	Métabolisme de l'adénosine	102700
<i>FAH</i>	Tyrosinémie de type 1	Métabolisme de la tyrosine	276700
<i>LMNA</i>	Progéria	Architecture du noyau	176670

Tableau I-4 : Exemples de maladies génétiques provoquées par des mutations affectant en cis l'épissage des pré-ARNm. Le nom des gènes portant la mutation est indiqué (tiré de Corcos and Solier, 2005).

Ces mutations peuvent être classées dans différentes catégories selon qu'elles affectent la fixation des complexes spliceosomaux, la fixation des facteurs de régulation de l'épissage ou bien créent de nouveaux sites d'épissage au sein du pré-ARNm.

La première catégorie de mutation affecte directement les sites 5' ou 3' d'épissage. La plupart du temps, ces mutations abolissent la reconnaissance de ces sites par les complexes spliceosomaux. Ces mutations conduisent généralement à l'élimination de l'exon possédant un site d'épissage muté, phénomène qui est appelé "exon skipping" (figure I-15). Dans d'autres cas, elles favorisent la reconnaissance du site d'épissage et augmentent l'inclusion de l'exon.

La seconde catégorie de mutation affecte les séquences régulatrices d'épissage introniques ou exoniques de type "splicing enhancer" ou "splicing silencer". La mutation d'une séquence activatrice de l'épissage peut mener à l'exclusion d'un exon, tandis que la mutation d'une séquence inhibitrice peut au contraire mener à l'inclusion d'un exon (figure I-15).

Ces deux catégories de mutations peuvent aussi affecter l'équilibre d'utilisation de deux sites alternatifs préexistants en compétition l'un avec l'autre, en favorisant un des deux sites. Par exemple, en favorisant un site 5' d'épissage faiblement utilisé en temps normal, au détriment d'un deuxième site 5' dont l'utilisation est diminuée en présence de la mutation. Dans le même ordre d'idée, ces mutations peuvent aussi mener à l'activation d'un site d'épissage préexistant, avec une séquence proche du consensus des sites d'épissage, mais qui n'était pas du tout utilisé dans un contexte sauvage. On nomme alors ce site un site cryptique.

Une troisième catégorie de mutation génétique peut conduire à la création d'un site d'épissage alternatif *de novo*, à partir d'une séquence à l'origine trop éloignée du consensus d'un site d'épissage et qui ne permettait pas le recrutement des complexes spliceosomaux (figure I-15).

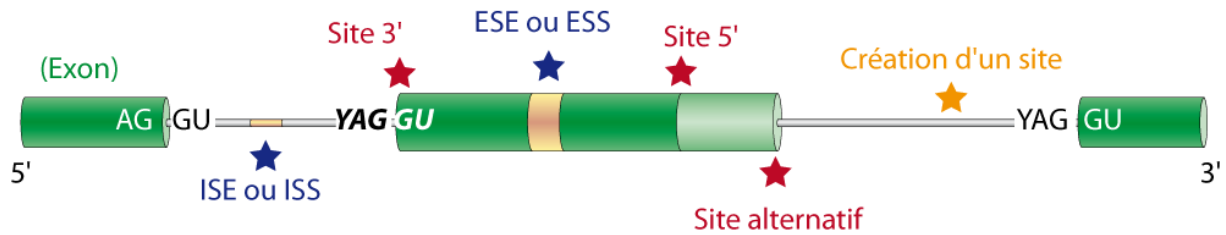


Figure I-15 : Représentation schématique des mutations ponctuelles à l'origine de maladies génétiques affectant en *cis* l'épissage. Chaque étoile indique un type de mutation possible. Elles affectent au choix l'un des éléments suivants : soit directement un site d'épissage préexistant (en rouge), ou bien un élément de séquence de régulation de l'épissage (en bleu). Une dernière catégorie de mutation permet la création *de novo* d'un site d'épissage (en orange). On peut noter qu'une mutation affectant un site d'épissage alternatif peut affecter un site alternatif 3' ou 5', même si ce n'est pas précisé ici.

I.5.2. Maladies génétiques affectant des facteurs d'épissage ou des protéines impliquées dans la biogenèse des spliceosomes

Certaines mutations *trans* peuvent également affecter les protéines de la machinerie d'épissage ou des facteurs régulateurs d'épissage. Cela perturbe alors la régulation de l'épissage de façon plus globale, et concerne généralement de nombreux pré-ARNm simultanément (pour revue, Faustino and Cooper, 2003). Un aperçu non exhaustif des pathologies résultant de ce type de mutations est fourni dans le tableau ci-dessous (tableau I-5) :

Protéine d'épissage affectée	Maladie	Numéro OMIM de la protéine
MBNL1, -2, -3	Dystrophies myotoniques	606516, 607237, 300413
FUSIP1	Leucémie, sarcomes	605221
FUS-TLS	Liposarcome, leucémie aiguë myéloblastique	137070
GSK-3 β	Alzheimer	605004
HMGA1 α	Alzheimer	600701
HCC1	Carcinome hépatocellulaire	604739
NOVA1	Syndrome paranéoplasique	602157
PRPF3, -8, -31	Rétinite pigmentaire	607301, 606419, 607300
RBMY1A1	Azoospermie	400006
SFPQ	Carcinome papillaire rénal	605199
TP73L	Syndrome de Hay-Wells	603273

Tableau I-5 : Pathologies résultant de mutations affectant directement des facteurs d'épissage. tiré de (Corcos and Solier, 2005).

Au niveau de l'organisme, l'ensemble de ces mutations peut donc se traduire par des conséquences dramatiques voire létales, en fonction du rôle de la ou des protéines affectées. Ces

mutations affectent en général directement la séquence codante des protéines, mais certaines maladies peuvent affecter leur localisation. C'est le cas de la dystrophie myotonique de type I (OMIM 160900) ou maladie de Steinert, liée à la perte partielle de fonction du facteur d'épissage MBNL1. C'est une maladie autosomale dominante caractérisée par un affaiblissement progressif et une atrophie musculaires ainsi qu'une myotonie. Les dangers proviennent notamment des risques cardiovasculaires provenant de la myotonie du muscle cardiaque. Cette maladie est provoquée par un phénomène d'amplification de triplets CTG présents dans la région du gène *DMPK* correspondant à la partie 3'UTR de l'ARNm *DMPK* (pour Dystrophie Myotonique Protéine Kinase). Le phénotype et l'âge d'apparition des symptômes sont dépendants du nombre de répétitions CTG présents sur le gène *DMPK*, à savoir en dessous de 30 pour les individus sains, de 100 à 1000 pour la forme classique, et plus de 2000 pour la forme congénitale, la plus grave, où les symptômes apparaissent dès la naissance (Brook et al., 1992). Ces répétitions retrouvées sous la forme CUG dans le pré-ARNm *DMPK*, s'organisent alors en de longues structures tige-boucle qui provoquent le recrutement et la séquestration de la protéine d'épissage MBNL1 (Muscleblind 1), un facteur régulant l'épissage alternatif essentiellement exprimé dans le cœur, le cerveau et les muscles squelettiques. Cette séquestration a lieu dans le noyau au niveau de *foci* situés à la périphérie des speckles (Holt et al., 2007), et conduit à une perte de fonction de MBNL1, ainsi qu'au gain de stabilité et à l'hyperphosphorylation de la protéine CUGBP1, un antagoniste de MBNL1 (Savkur et al., 2001). La perte de fonction de MBNL1 et le gain d'activité de son antagoniste CUGBP1 conduisent à la dérégulation de l'épissage alternatif de plusieurs pré-ARNm, tels que ceux codant la troponine T cardiaque (hcTNT), le récepteur à l'insuline (IR), de la protéine tau, ou d'un canal à chlore. Cette dérégulation participe au développement des symptômes cardiovasculaires, de la dégénérescence musculaire, ainsi qu'à d'autres symptômes non liés à cette dégénérescence comme la résistance à l'insuline.

Les travaux de ma thèse ont porté sur une maladie génétique rare, résultant d'une mutation ponctuelle affectant directement un site d'épissage. Cette maladie, appelée progéria, est présentée dans le chapitre suivant.

II. La progéria, une maladie génétique rare affectant l'épissage du pré-ARNm de la lamine A

II.1. Le syndrome de Hutchinson-Gilford ou progéria

La progéria (OMIM 176670), nommée d'après le terme grec « geras », qui veut dire « vieux », est une maladie génétique rare, caractérisée par un ensemble de symptômes qui s'apparentent pour la plupart à un vieillissement accéléré. Elle a été décrite dès 1886 par Jonathan Hutchinson et en 1887 par Hastings Gilford, c'est pourquoi le syndrome de la progéria est aussi désignée sous la forme abrégé HGPS pour Hutchinson-Gilford Progeria Syndrome (cf. Gilford, 1897; Hutchinson, 1886). Cette maladie, diagnostiquée pour environ 1 individu sur 4 à 8 millions, apparaît de façon sporadique et seule une centaine de cas ont été décrits depuis 1886. Les symptômes apparaissent entre la naissance et l'âge d'un an, âge où le diagnostic peut généralement être posé. Ces symptômes comprennent alopecie (perte des cheveux), perte des graisses sous-cutanées, retard de la dentition, hypoplasie mandibulaire ou micrognathie (atrophie de la mâchoire), disproportion craniofaciale, retard de croissance, arthrite et problèmes cardiovasculaires (pour revue, Pollex and Hegele, 2004). Les enfants atteints de cette maladie ont rapidement une apparence de personne âgée. Ils ont notamment des veines saillantes, des yeux proéminents, et une stature réduite et courbée. Cela s'accompagne d'hypertension et d'athérosclérose, liées à la dégradation rapide de leurs systèmes respiratoire et cardiovasculaire. Par conséquent les individus atteints ont une espérance de vie d'en moyenne 13,5 ans et décèdent le plus souvent des suites d'accidents cardiovasculaires. Bien que les patients atteints de progéria présentent des symptômes s'apparentant à un vieillissement accéléré, ils ne développent pas tous les signes possibles du vieillissement, comme la cataracte ou la dégénérescence cognitive (c'est pourquoi le HGPS est parfois qualifié de syndrome progéroïde "segmentaire"). De plus, contrairement à certaines autres maladies progéroïdes, la progéria ne semble pas augmenter les risques d'apparition de tumeurs. Ce dernier point est encore sujet à discussion, puisque 2 cas référencés de HGPS sont associés à l'apparition d'un ostéosarcome développé aux âges de 13 et 9 ans (King et al., 1978; Shalev et al., 2007). Le manque de données et la faible espérance de vie des patients ne permettent cependant pas de déterminer s'il s'agit ou non d'une coïncidence. Comme il sera décrit dans les chapitres suivants, la progéria peut en effet être associée à une certaine instabilité génétique. En effet, cette maladie est causée par des mutations dans le gène *LMNA*, qui code la protéine lamine A qui appartient à une structure tapissant l'intérieur du noyau appelée lamine nucléaire ou *lamina* et qui est importante pour l'organisation du génome.

II.2. Présentation de la lamine nucléaire ou *lamina*

II.2.1. Composition, structure et organisation

II.2.1.a. Description de la lamine nucléaire

La lamine nucléaire est une structure composite localisée sur la face interne de l'enveloppe nucléaire des cellules de la plupart des métazoaires. Elle est composée majoritairement d'un assemblage de protéines filamenteuses appelées lamines, associées à un ensemble complexe de protéines intégrées totalement ou partiellement à la membrane nucléaire, ainsi qu'à des facteurs protéiques associés à la chromatine. Elle constitue une interface fonctionnelle entre l'enveloppe nucléaire et la chromatine et peut d'ailleurs interagir directement avec la chromatine, notamment au niveau de séquences conservées riches en résidus A et T (Meuleman et al., 2013). Il est difficile de visualiser la lamine nucléaire en tant que structure définie et individualisée. Elle peut apparaître en microscopie électronique comme une couche dense et continue de 40 à 300 nm selon le type cellulaire (pour revue, Bridger et al., 2007). Cependant, après traitement des noyaux avec des nucléases et solubilisation des lipides membranaires par un détergent non ionique, cette structure qui a globalement conservé sa forme, est observable indépendamment comme une fine couche protéique fibreuse d'environ 15 nm même dans des cellules où elle n'est à l'origine pas visible sans ce traitement (Aaronson and Blobel, 1975). Elle est également observable comme une fine couche fibrillaire associée à la membrane interne de l'enveloppe nucléaire après traitement des cellules pour éliminer la chromatine (Dwyer and Blobel, 1976). La lamine nucléaire est un réseau protéique présentant une organisation géométrique régulière qui est localisé à la surface de l'enveloppe nucléaire et qui est associé aux pores nucléaires (ou NPC, pour « Nuclear Pore Complex »), et est à l'origine de la répartition organisée et régulière de ces derniers au sein de l'enveloppe nucléaire (Aaronson and Blobel, 1975; Aebi et al., 1986).

L'organisation en figures tétraogales de cette structure protéique (voir figure II-1) est due aux caractéristiques particulières d'assemblage de ses constituants principaux, les lamines, qui jouent un rôle majeur au sein de la *lamina* et lui confèrent la plupart de ses propriétés physico-chimiques et biologiques.

II.2.1.b. Les différents types de lamines

Les lamines appartiennent à la famille des protéines appelée filaments intermédiaires (FI) (ou "Intermediate Filaments") de type V. Les FI sont des protéines qui polymérisent en de longues structures fibreuses par l'interaction de grands domaines centraux hélicoïdaux. Elles ont un rôle structural ou matriciel du fait de leurs propriétés de polymérisation particulières. A la différence des lamines, les FI de type I à IV sont toutes des protéines cytoplasmiques. Les FI de type I sont des kératines à forte prédominance de résidus amino-acyl dites acides, les FI de type II sont dites kératines basiques.

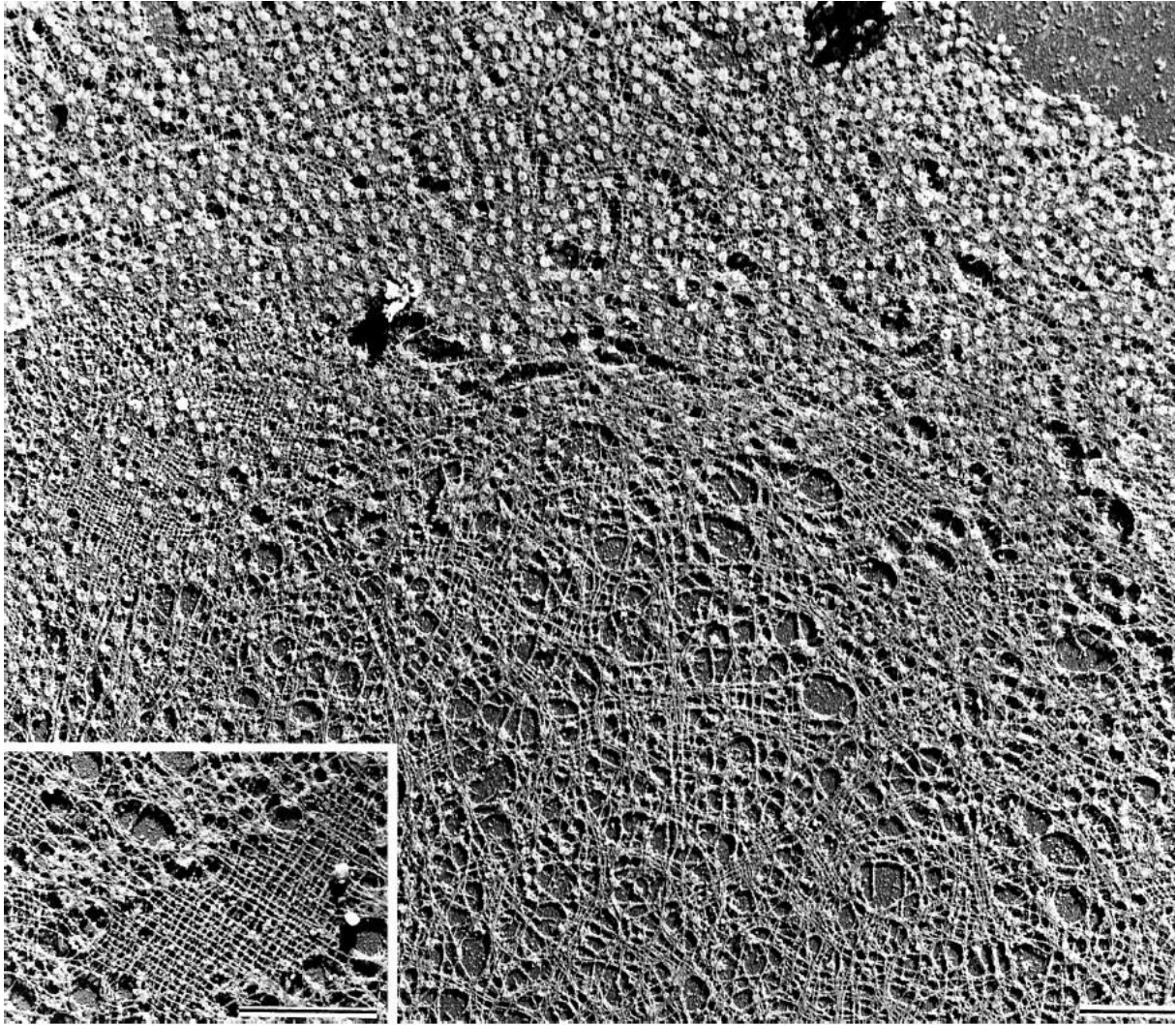


Figure II-1 : Visualisation par microscopie électronique d'une préparation d'un extrait nucléaire d'ovocyte de *Xenopus laevis* traité au détergent Triton X-100. La vue générale présente le réseau protéique de la lamine nucléaire partiellement couverte de NPC, qui apparaissent sous la forme de points blancs, organisés de façon régulière sur le réseau (haut de la section). L'encadré en bas à gauche montre l'agrandissement d'une section mieux préservée de cet extrait présentant un réseau protéique en quadrillage du aux propriétés d'assemblage des lamines. Barre : 1 μ m. Tiré de (Aebi et al., 1986).

Les FI de type III et IV sont désignées respectivement comme protéines desmin-like et FI neuronales. Les lamines ont été classées dans les FI car elles partagent une certaine homologie de séquence et de structure avec les FI de type I à IV, et des propriétés similaires de polymérisation, qui seront décrites dans le paragraphe suivant (Aebi et al., 1986; Fisher et al., 1986).

L'expression d'au moins un type de lamine dans tous les types cellulaires est indispensable à la survie cellulaire. Ce qui n'est pas le cas des protéines FI cytoplasmiques dont l'expression est dépendante du type cellulaire ou tissu spécifique. Chez les espèces des *phyla* d'invertébrés de la branche des métazoaires, comme chez les nématodes, il n'existe qu'un seul type de lamine dans la *lamina* (souvent appelée Célamine chez *C. elegans* par exemple), qui constitue les premières lamines apparues au cours de l'évolution. Plusieurs éléments indiquent d'ailleurs que les différents types de lamine et les protéines

FI cytoplasmiques dérivent de ce gène ancestral de lamine, notamment parce que tous les organismes connus possédant des FI cytoplasmiques possèdent également des lamines, alors que la réciproque n'est pas vraie (Harborth et al., 2001).

Chez les mammifères, il existe deux types de lamines, les lamines de type A (principalement lamines A, C et AΔ10) et de type B (lamines B1 et B2). Elles diffèrent par leurs caractéristiques et leurs propriétés biochimiques (Gerace and Blobel, 1980). Chez l'homme, les lamines de type A ont un point isoélectrique élevé, et sont produites par épissage alternatif des pré-mRNAs issus du gène *LMNA*. Les lamines B1 et B2 ont un point isoélectrique bas et sont exprimées à partir de deux gènes distincts *LMNB1* et *LMNB2*. Elles présentent une plus forte homologie de séquence (*LMNB1* en particulier) avec les lamines primordiales d'invertébrés de type Célamine qu'avec les lamines de type A. Chez les organismes exprimant les deux types de lamines, comme les mammifères, les lamines B sont essentielles à la survie de l'organisme et sont exprimées constitutivement dans toutes les cellules somatiques (Harborth et al., 2001). Au contraire, les lamines de type A dont les 2 formes principales sont les lamines A et C, ont un profil d'expression temporel particulier puisqu'elles sont majoritairement exprimées dans les cellules différenciées. Enfin, d'autres lamines particulières, telles que les lamines C2 (de type A) et B3 (de type B), qui sont respectivement issues de l'épissage alternatif des pré-ARNm *LMNA* et *LMNB2*, sont uniquement exprimées dans les lignées germinales. La lamine C2 est exprimée uniquement lors de la méiose, tandis que la lamine B3 l'est uniquement dans les lignées germinales mâles (Furukawa and Hotta, 1993; Jahn et al., 2010).

II.2.1.c. Structure, propriétés de polymérisation et de désassemblage des lamines

Les lamines de vertébrés ou d'invertébrés comportent en général 3 domaines : un court domaine N-terminal globulaire dit « tête » (ou « head »), un domaine central fuselé α -hélicoïdal, appelé « helical rod », ainsi qu'un domaine C-terminal appelé « queue » ou « tail », dont le repliement en un feuillet de 9 brins β est proche de celui des immunoglobulines (figure II-2) (Dhe-Paganon, 2002).

Le domaine central des lamines est organisé de manière identique à celui des FI cytoplasmiques : il possède 4 hélices α très conservées, nommées hélices 1A, 1B, 2A et 2B, qui sont séparées par 3 régions « linkers » non hélicoïdales L1, L1-2 et L2 (Figure II-2). Ce domaine fuselé participe à l'établissement de la rigidité fibrillaire de la protéine et permet la formation de dimères stables contenant des structures en « super-hélice » (motif « coiled-coil »). Ces super-hélices sont générées par l'enroulement en orientation parallèle de deux molécules de lamines. La formation de ces structures « coiled-coil » repose sur la présence dans les hélices α de répétitions de 7 acides aminés caractérisées par la présence d'acides aminés hydrophobes en positions 1 et 4, dont les chaînes latérales sont souvent orientées vers l'extérieur de l'hélice (Figure II-2). Pour toutes les protéines FI, lamines comprises, ces

répétitions et en particulier ces résidus 1 et 4 permettent la formation d'interactions hydrophobes entre les hélices de deux molécules distinctes et favorisent donc une dimérisation en orientation parallèle (Strelkov et al., 2004).

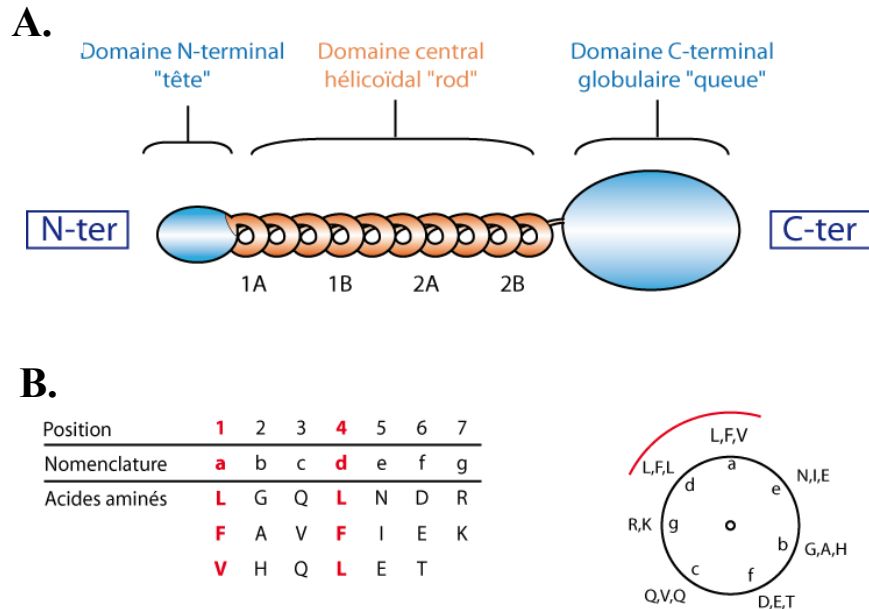


Figure II-2 : Organisation des protéines lamines **A.** Les différents domaines des protéines lamines sont représentés. Les domaines N et C-terminaux sont globulaires et le domaine tubulaire "rod" est majoritairement constitué d'hélices α . Dans le domaine central tubulaire la position de 4 hélices α conservées au sein de la famille des lamines (1A, 1B, 2A, 2B) est indiquée. **B.** Composition des répétitions de 7 acides aminés (notés a à g) présentes dans les domaines tubulaires des lamines. Les acides aminés en position 1 et 4 (a et d) permettant les interactions hydrophobe de type "coiled-coil" entre les dimères de lamines sont indiqués en rouge. A droite est représentée la position de ces acides aminés au sein d'une hélice, dans une coupe longitudinale du domaine tubulaire.

Ces dimères constituent le premier degré d'organisation suivi d'une longue cascade d'étapes d'assemblages, aboutissant en dernier lieu à la formation de macrostructures protéiques, les filaments intermédiaires, d'où ces protéines tirent leur nom. Les dimères s'associent d'abord spontanément de "tête à queue" (« head-to-tail ») pour former un polymère linéaire, c'est-à-dire que les deux extrémités C-ter d'un dimère interagissent avec les deux extrémités N-ter du dimère suivant. Un regroupement de deux polymères linéaires de façon antiparallèle permet la formation de protofilaments de 2 à 3 nm d'épaisseur. L'association de deux de ces protofilaments va donner une protofibrille de 4-5 nm. Une dernière étape de multimérisation de ces protofibrilles va donner un filament intermédiaire, d'une épaisseur d'environ 10 nm (Figure II-3) (pour revue, Dittmer and Misteli, 2011; Fuchs and Weber, 1994). Les polymères formés par les lamines diffèrent de ceux formés par les FI cytoplasmiques car ils sont issus d'une association linéaire de dimères, contrairement aux associations plus compactes en tétramères de deux dimères antiparallèles observés pour les autres FI (Figure II-3).

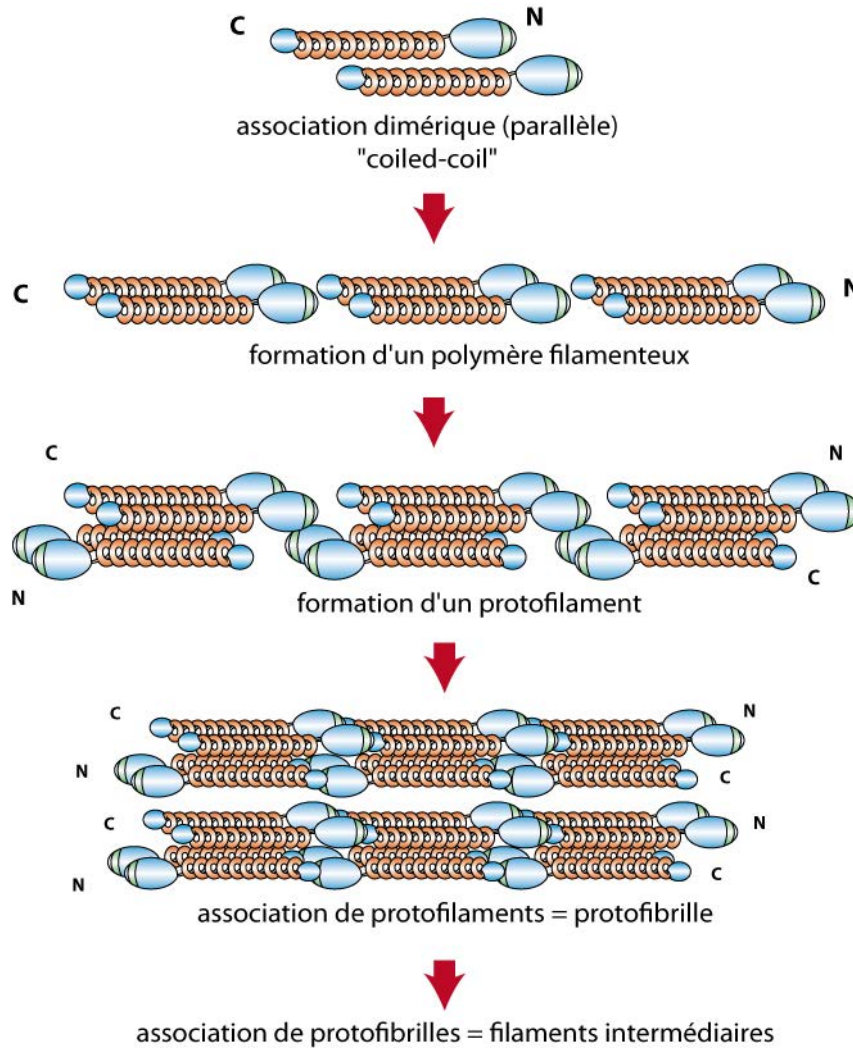


Figure II-3 : Polymérisation des lamines et formation des filaments intermédiaires. Les interactions hydrophobes entre les lamines permettent l'établissement d'interactions "coiled-coil", ce qui donne lieu à la multimérisation des lamines à plus grande échelle. Ce schéma représente les différentes étapes de cet assemblage, en partant du dimère de protéine lamine, formé par deux protéines disposées parallèlement. Les extrémités N-terminales (N) et C-terminales (C) des lamines sont indiquées afin de mieux visualiser les agencements parallèles et antiparallèles des différentes structures.

Le domaine tête N-ter des lamines est essentiel à la formation des interactions « head-to-tail » entre dimères *in vitro* (Stuurman et al., 1996). Cet assemblage « head-to-tail » des dimères de lamines repose en effet sur l'existence d'interactions électrostatiques entre le domaine tête et le début du domaine queue, chargés positivement du fait de la présence de nombreux résidus arginine, avec les parties N- et C-terminales du domaine central « rod » chargées négativement (Strelkov et al., 2004). Le domaine C-terminal « queue », bien que non essentiel pour la polymérisation, est cependant indispensable *in vivo* car il contient un signal de localisation nucléaire (NLS, pour « Nuclear Localisation Signal ») requis pour l'adressage de la protéine néo-synthétisée dans le noyau. Ce domaine C-terminal serait également impliqué dans une interaction directe avec la chromatine, notamment avec les histones H2A et H2B (Goldberg et al., 1999; Taniura et al., 1995).

De plus, les lamines se caractérisent par la présence de 42 aa, absents dans les protéines FI cytoplasmiques, correspondant à 6 motifs heptamériques supplémentaires au sein de l'hélice 1B dans le domaine central. Cette extension comprend de nombreux résidus sérine, dont la phosphorylation pourrait conférer aux lamines des propriétés particulières de liaison à la chromatine et de dépolymérisation que n'ont pas les FI cytoplasmiques (Fisher et al., 1986). En effet, il est établi que la *lamina* subit une dépolymérisation réversible nécessaire au bon déroulement de la mitose, qui implique la phosphorylation de résidus sérine conservés du domaine N-ter « tête » des lamines, par la kinase du cycle cellulaire cdc2 (Gerace and Blobel, 1980; Peter et al., 1990). Au cours du cycle cellulaire, alors que les lamines de type B restent associées à la membrane nucléaire désassemblée, la lamine A est dissociée de la *lamina* et relâchée sous forme soluble (Stuurman et al., 1998)(pour revue, Gruenbaum et al., 2003).

Les rôles propres de chacun des isoformes ou des variants des lamines, ainsi que des modifications post-traductionnelles de ces dernières sont encore mal connus. Néanmoins, au cours de l'évolution, l'accroissement du nombre de lamines et de leurs modifications est probablement lié à la variété et la complexité des types cellulaires dans lesquels elles sont présentes. Cet accroissement est par ailleurs accompagné d'un besoin croissant de diversification des propriétés fonctionnelles des lamines, ainsi que, comme nous allons le voir, de leurs partenaires.

II.2.1.d. Autres composants de la lamine nucléaire

Au sein de la *lamina*, les lamines sont associées à un grand nombre de partenaires protéiques souvent intégrés totalement ou partiellement à la membrane nucléaire. Parmi les principaux partenaires on peut citer : les protéines de la famille LAP (pour « Lamins Associated Proteins »), l'émerine, l'otefine, le récepteur à la lamine B (LBR, pour "Lamin B Receptor"), et le facteur BAF (pour « Barrier to integration Factor ») (figure II-4). On compte aussi parmi ces partenaires le complexe protéique LINC, contenant les protéines Sun, intégrées à la membrane et interagissant avec les lamines A, ainsi que les protéines nesprines, qui font le lien entre la *lamina*, les filaments d'actine et protéines du cytosquelette (pour revue, Burke and Roux, 2009). La plupart des partenaires des lamines sont des protéines intégrées à l'enveloppe nucléaire qui sont importées dans le noyau *via* le réticulum endoplasmique. Ainsi, le récepteur à la lamine B, LBR, présent chez les drosophiles et les vertébrés, mais pas chez *C. elegans*, possède 8 domaines transmembranaires putatifs. Il interagit directement avec la lamine B *in vitro* et *in vivo*, mais aussi avec la protéine spécifique de l'hétérochromatine HP-1, et les histones H3 et H4, ainsi qu'avec l'ADN (Duband-Goulet and Courvalin, 2000; Holmer et al., 1998; Polioudaki et al., 2001; Ye et al., 1997). La protéine LAP2, qui interagit avec la lamine B, est présente chez les vertébrés mais pas chez la drosophile (Furukawa and Kondo, 1998). Elle possède un domaine transmembranaire de 42 aa, nommé LEM, qui est commun à d'autres protéines de la *lamina* comme l'émerine, l'otefine, MAN1, et à plusieurs protéines encore non caractérisées (Cohen et al., 2001). Quatre des isoformes de la protéine

LAP2 (β , γ , δ , ϵ) contiennent le domaine LEM alors que deux autres isoformes (δ et α) ne le contiennent pas. Ce domaine pourrait servir à la mise en place d'interactions avec d'autres protéines de la lamina telle que BAF, comme cela a été montré dans le cas de LAP2 β . L'isoforme LAP2 α peut également interagir avec les lamines de type A, avec lesquelles elle peut co-localiser à l'intérieur du nucléoplasme et non plus à la périphérie du noyau (Dechat et al., 2000) (Figure II-4). LAP2 β forme également un complexe avec les protéines LBR, HA95 (une protéine de la matrice nucléaire et de la chromatine), et avec l'émerine (Martins et al., 2000). Ainsi, les protéines de la *lamina* forment un réseau d'interactions complexes à la fois entre elles et avec la chromatine, ce qui participe notamment à l'une des fonctions principales de la *lamina*, à savoir le maintien de l'architecture du noyau (Furukawa et al., 1998)(pour revue, Gruenbaum et al., 2003).

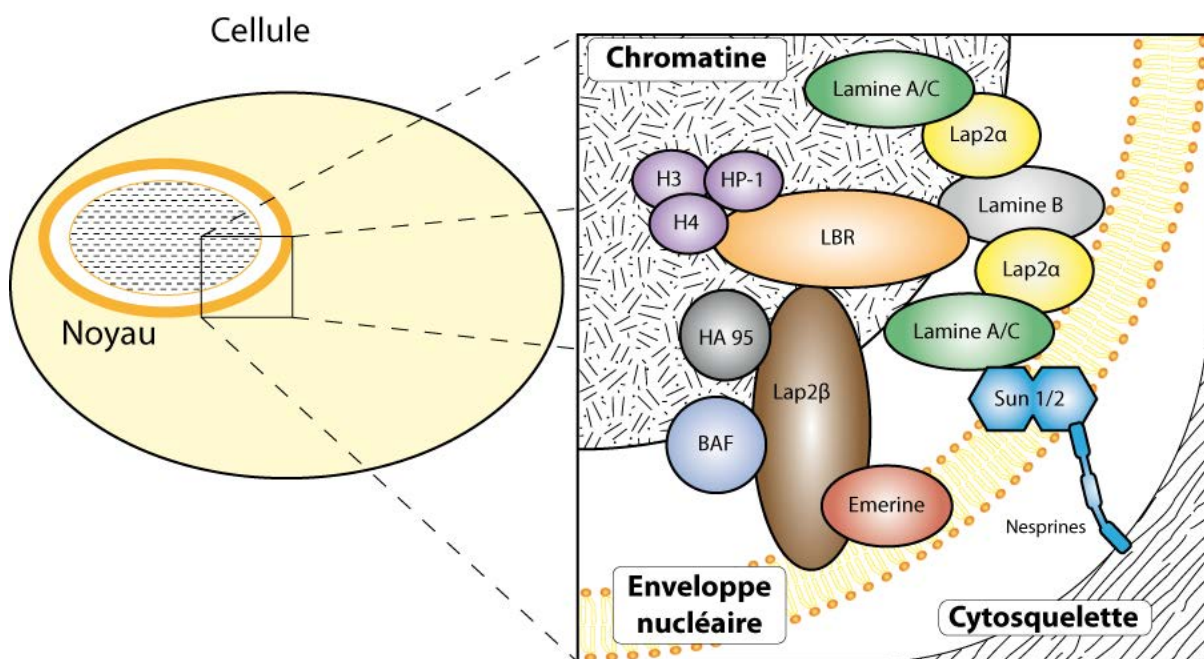


Figure II-4 : Un réseau complexe d'interactions protéine-protéine et protéine-ADN assure la cohésion de la *lamina*. L'encadré de droite est un agrandissement schématisé de la région de l'enveloppe nucléaire où se situe la *lamina*, montrant une représentation non exhaustive des protéines présentes au sein de la *lamina* et des interactions qu'elles sont susceptibles d'établir (pour revue, Burke and Roux, 2009; Gruenbaum et al., 2003). Lap2 β , LBR et l'émerine sont des protéines membranaires, HA95, HP-1, H3 et H4 sont des protéines de la chromatine. LBR, BAF, Lap2 α et β , ainsi que la lamine A et la lamine C sont capables d'établir des interactions avec la chromatine. Lap2 α , la lamine A et la lamine C peuvent également interagir avec la chromatine sans être associées à la *lamina*. Les protéines Sun et Nesprines forment le complexe LINC qui relie structurellement et fonctionnellement la *lamina* au cytosquelette.

L'importance de la *lamina* au niveau de l'organisme et de son développement transparait au travers de l'existence d'une catégorie de pathologies graves affectant la *lamina* et son organisation. Elles sont regroupées sous le terme laminopathies. Bien que cela ne soit pas toujours le cas, ces pathologies,

auxquelles appartient la progéria, sont souvent associées à des ensembles de symptômes s'apparentant plus ou moins directement au vieillissement qu'on appelle syndromes « progéroïdes ».

II.2.2. Laminopathies et syndromes progéroïdes

Comme mentionné précédemment, les syndromes progéroïdes représentent un ensemble de pathologies caractérisées par des signes de vieillissement prématuré. Du point de vue physiopathologique, les syndromes progéroïdes peuvent être classés en deux catégories majeures. La première concerne les syndromes progéroïdes causés par des mutations dans des gènes codants des protéines impliquées dans les mécanismes de réparation de l'ADN. Ainsi, le syndrome de Werner, provoquant des symptômes très proches du vieillissement naturel, est causé par des mutations de l'hélicase RecQL2/WNR, dont le rôle est de préserver l'intégrité du génome en évitant les événements de recombinaison délétère (Navarro et al., 2004). De même, le syndrome de Cockayne, qui s'accompagne en plus de troubles du développement des neurones, est causé par des mutations d'enzymes du système de réparation de l'ADN par excision de bases (Andressoo and Hoeijmakers, 2005). La seconde catégorie de syndromes progéroïdes regroupe les syndromes associés à des défauts primaires ou secondaires des lamines de type A et fait partie intégrante des pathologies appelées laminopathies. En effet, les laminopathies représentent une large classe de maladies génétiques, affectant les protéines qui constituent la lamine nucléaire. Les mutations à l'origine de ces pathologies affectent les lamines de type A et B, ou des protéines impliquées dans leur maturation comme la protéase ZMPSTE24, des protéines partenaires des lamines A/C et B comme l'émerine, LAP2 α , le LBR, MAN1, et la protéine Aladin des NPC.

Pour rester centré sur ce qui a fait l'objet de mon sujet de thèse je décrirai ici les laminopathies provoquées par des mutations dans le gène *LMNA*. Ces mutations peuvent être à l'origine d'au moins 10 syndromes cliniques différents, affectant des tissus très divers. Ces syndromes peuvent être regroupés en 4 classes majeures : (1) les maladies musculaires du type cardiomyopathies, telles que la dystrophie musculaire d'Emery-Dreyfuss (EDMD), (2) les lypodystrophies, dont la dysplasie mandibulo-acrale, (3) une neuropathie périphérique, la maladie de Charcot-Marie-Tooth (de type 2B1) qui affecte les nerfs, et enfin, (4), les maladies du vieillissement que sont le syndrome de Werner atypique, la dermatopathie restrictive létale et le syndrome de Hutchinson-Gilford (pour revue, Worman and Bonne, 2007). Le point commun à la plupart de ces laminopathies est l'observation d'une dégénérescence des cellules qui composent les différents tissus affectés.

Le fait que les mutations d'un seul gène peuvent conduire à autant de phénotypes différents dans des tissus différents reflète les rôles multiples et l'importance des lamines de type A chez les mammifères. Ainsi, la majorité des pathologies des muscles striés (première catégorie) proviennent de mutations faux-sens localisées tout au long du gène *LMNA* (Bonne et al., 1999; Vytopil et al., 2003). Par ailleurs, les souris K.O. pour le gène *LMNA* développent après quelques semaines une cardiomyopathie

sévère et une dystrophie myotonique, et décèdent en moyenne après 8 semaines (Nikolova et al., 2004). Ces symptômes sont retrouvés à différents degrés de sévérité dans les maladies musculaires humaines provenant des mutations faux-sens dans le gène *LMNA*, indiquant qu'ils proviennent vraisemblablement d'une perte de fonction. En outre, le syndrome EDMD est également provoqué par des mutations dans les gènes d'autres composants de la *lamina* comme l'émerine ou les protéines du complexe LINC qui font le lien entre la *lamina* et le cytoplasme, montrant que les liens architecturaux entre protéines de la *lamina* reflètent des liens fonctionnels (Zhang et al., 2007).

Les autres laminopathies peuvent davantage être associées à des mutations affectant spécifiquement certains domaines de la lamine A. C'est le cas des pathologies de type progéroïdes souvent associées à des mutations de l'extrémité C-terminale (pour revue, Worman and Bonne, 2007). En effet, dans certains syndromes de laminopathies progéroïdes, dont la progéria, la dégénérescence des cellules a pu être reliée à des défauts de maturation de la lamine A. La dermatopathie restrictive par exemple est une pathologie provoquée par la mutation d'une enzyme jouant un rôle majeur dans la maturation de la lamine A, la métalloprotéase *ZMPSTE24*. Il s'agit d'une pathologie progéroïde létale parfois dès le stade fœtal ou souvent peu après la naissance, qui est également provoquée par des mutations dans le gène *LMNA* menant à des défauts d'épissage de son pré-ARNm et à la production d'une protéine ayant une extrémité C-terminale aberrante (Navarro et al., 2004; Pendas et al., 2002). Ces syndromes sévères sont en fait dus à l'accumulation d'une forme non mature de la lamine A, appelée progérine, qui est toxique pour la cellule. Nous verrons, après une description plus détaillée de la lamine A et des mutations à l'origine de la progéria, que plusieurs hypothèses permettent de relier la production de progérine à des phénotypes progéroïdes au niveau moléculaire et cellulaire.

II.3. Lamine A et progéria

II.3.1. Organisation du gène *LMNA* et des différentes isoformes protéiques codées par ce gène

Le gène *LMNA* est à l'origine de la production de quatre isoformes de lamines par épissage alternatif. Ces isoformes sont les lamines A et C, déjà décrites précédemment, ainsi que la lamine A Δ 10 et la progérine, qui sont des variantes faiblement produites de la lamine A (figure II-5). La progérine, une isoforme produite de façon minoritaire et sur laquelle nous reviendront, est similaire à la lamine A mais ne contient pas une partie protéique correspondant à la partie 3' de l'exon 11.

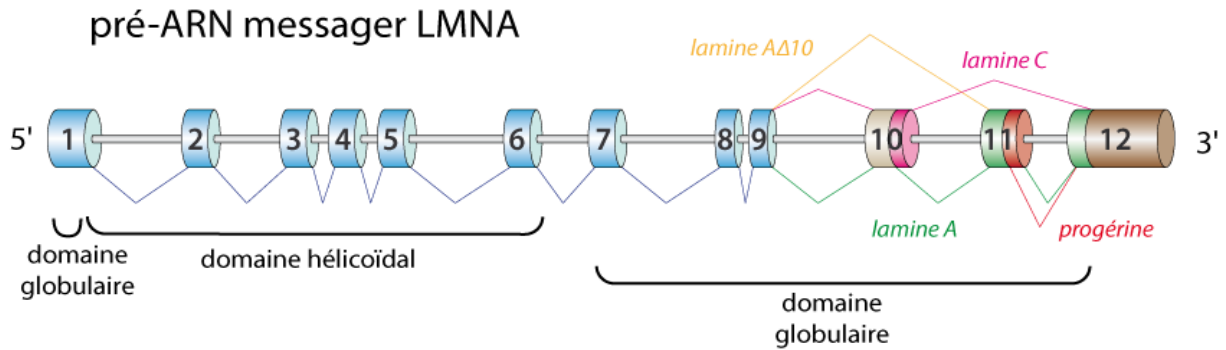













Figure II-5 : Epissage alternatif du transcrite issu du gène LMNA humain. Le gène LMNA comporte 12 exons (indiqués par des cylindres et numérotés) et 11 introns. Les trois derniers exons peuvent être épissés alternativement pour donner des ARNm codant 4 isoformes nommées respectivement lamine A, lamine C, progéline et lamine AΔ10. Les domaines protéiques correspondants aux exons codants sont indiqués. Les évènements d'épissage requis pour produire chacun des ARNm sont indiqués par des lignes brisées joignant les extrémités de chaque exon. Les couleurs utilisées dans cette figure correspondent aux régions et aux profils d'épissage suivants :

	Région exonique commune à tous les transcrits LMNA		Région intronique
	Région exonique commune à la lamine A et la Lamine C		Evènements d'épissage communs à tous les ARNm
	Région exonique spécifique à la lamine C		Evènements d'épissage spécifiques à la lamine A
	Région exonique spécifique à la lamine A / progéline		Evènements d'épissages spécifique à la lamine C
	Région 3' non traduite		Evènement d'épissage spécifique à la lamine AΔ10
	Région supprimée dans la progéline		

Les ARNm de ces 4 isoformes ont en commun les exons 1 à 9 et toute la partie 3' UTR. Les ARNm des lamines A et C ont en commun une partie de l'exon 10, tandis que l'autre partie de cet exon est spécifique de la lamine C. La partie 5' de l'exon 10 et la totalité de l'exon 11, ainsi que le début de l'exon 12, ce qui représente 93 résidus, sont spécifiques de la lamine A. Cette région code une partie du domaine C-terminal globulaire de la lamine A, qui est essentielle dans le processus de maturation de la lamine A et à son ancrage dans la *lamina*. Ce sont des mutations affectant cette région qui sont en grande partie responsables des syndromes progéroïdes et de la plupart des cas de progéria. L'isoforme lamine C n'est donc pas directement impliquée dans ces pathologies, car elle ne possède pas la même séquence C-terminale et ne subit pas les mêmes étapes de maturation que la lamine A (Beck et al., 1990; Fisher et al., 1986; McKeon et al., 1986).

II.3.2. La maturation de la lamine A et son intégration dans la *lamina*

Les lamines A subissent un certain nombre de modifications post-traductionnelles avant de pouvoir exercer leurs fonctions dans le noyau. Après traduction, les pré-lamines A possèdent à leur

extrémité C-terminale une séquence conservée est appelée « boîte CAAX », où A désigne un résidu aliphatique et X un des résidus S, M, C, A, Q (M dans le cas de la lamine A). Le résidu Cys de cette séquence est la cible d'une farnésylation (addition d'un acide gras) catalysée par l'enzyme farnésyl-transférase conduisant à la formation d'une isoprénylcystéine (Vorburger et al., 1989) (figure II-6). Cette modification permet l'ancrage de la pré-lamine A à l'enveloppe nucléaire. Un clivage est ensuite réalisé par la métalloprotéase ZMPSTE24/FACE1, entre le résidu Cys farnésylé et le premier résidu aliphatique de la boîte CAAX (Corrigan et al., 2005; Pendas et al., 2002). Ensuite, le résidu Cys devenu terminal est carboxyméthylé par l'enzyme isoprénylcystéine carboxyméthyl-transférase et la lamine A mature est libérée par un second clivage endoprotéolytique qui a lieu 15 résidus en amont du résidu Cys farnésylé, au niveau d'une séquence conservée RSYLLG (Corrigan et al., 2005; Pendas et al., 2002). L'isoprénylation de la boîte CAAX est une étape nécessaire pour l'adressage et l'intégration de la lamine A mature au sein de la *lamina* (Lutz et al., 1992). C'est pourquoi des mutations du gène *LMNA* mais aussi *ZMPSTE24* perturbant cette maturation, en particulier l'action de la métalloprotéase *ZMPSTE24*, peuvent avoir des conséquences dramatiques pour la *lamina*.

II.3.3. Mutations du gène *LMNA* à l'origine de la progéria

Plusieurs mutations du gène *LMNA* peuvent être à l'origine de la progéria. La majorité de ces mutations sont des mutations ponctuelles sporadiques, autosomales dominantes, apparaissant *de novo*. Le seul cas répertorié de transmission mendélienne consanguine de la maladie est celui d'une famille indienne, où les géniteurs (cousins au 1^{er} degré) sont tous les deux porteurs hétérozygotes d'une mutation autosomale récessive K542N sur le gène *LMNA*, et ne sont atteints d'aucun symptôme. Cinq de leurs 7 enfants ont hérité des deux allèles mutés de leurs parents et ont développé les symptômes classiques de la progéria (Plasilova et al., 2004). De plus, deux cas exceptionnels révèlent une isodisomie uniparentale : c'est-à-dire que les deux copies de la région 1q du chromosome 1, comprenant le gène *LMNA*, ont été héritées du même parent. Il a également été reporté un cas de HGPS pour lequel une région de 6 Mb comprenant le gène *LMNA* avait été délétée dans le chromosome 1 (Eriksson et al., 2003).

Les mutations ponctuelles à l'origine du HGPS affectent une région bien précise et spécifique de la lamine A, et tout particulièrement l'épissage alternatif du pré-ARNm au niveau de cette région. La découverte des premières mutations majeures à l'origine de la progéria a été faite indépendamment et en parallèle par trois équipes lors de l'analyse à grande échelle du génome de patients présentant les symptômes caractéristiques de ce syndrome (Cao and Hegele, 2003; Eriksson et al., 2003; De Sandre-Giovannoli, 2003). La mutation retrouvée le plus fréquemment, soit environ 90 % des cas étudiés, est la mutation silencieuse c.1824C>T (G608G). La récurrence de cette mutation ponctuelle à cet endroit précis pourrait être expliquée par la présence d'une méthylation épigénétique au niveau du résidu C muté, constituant d'un dinucléotide CpG, ou « îlot de méthylation » (Eriksson et al., 2003).

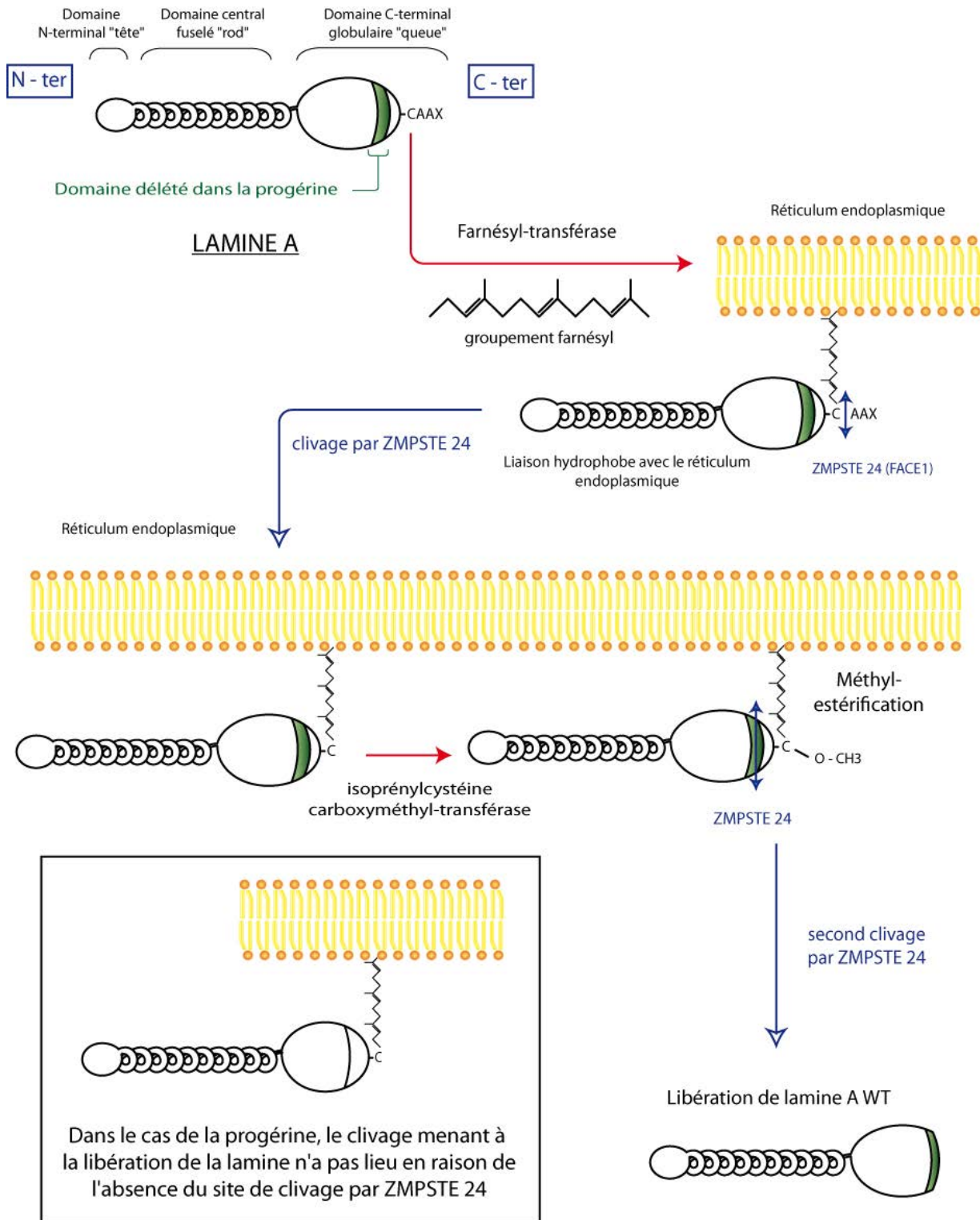


Figure II-6 : Les étapes de maturation de la lamine A et l'effet de l'absence du second clivage par ZMPSTE 24. La première étape de maturation de la lamine A est l'ajout d'un groupement farnésyl sur le résidu Cys de la "boîte CAAX" par la farnésyl-transférase. Ceci permet l'ancrage de la protéine à la membrane du réticulum endoplasmique, où elle subit des étapes de maturation impliquant l'isoprénylcystéine carboxyméthyl-transférase et la protéase ZMPSTE 24. Pour comparaison (voir encadré), le domaine indiqué en vert dans la lamine A est absent dans la progérine. Ce domaine contient le second site de clivage par ZMPSTE 24. La dernière étape de clivage ne peut donc avoir lieu et la progérine ne peut pas être libérée et reste associée à la membrane.

En effet, un nucléotide C méthylé fait partie des bases à l'origine de fréquentes mutations dans le génome des vertébrés, puisque ceux-ci peuvent spontanément être convertis en T par désamination et ainsi conduire à une mutation lors de la réplication de l'ADN génomique. Lors de ces études, la mutation c.1822G>A (G608S), pendant faux-sens de la mutation silencieuse c.1824>T (G608G) évoquée ci-dessus, a également été répertoriée, mais elle est à l'origine d'un nombre de cas beaucoup plus restreints. Une mutation à l'origine d'une forme nettement moins sévère du HGPS a été observée chez un patient japonais. Il s'agit de la mutation c.1868C>G (T623S) (Fukuchi et al., 2004). Le patient a développé tous les symptômes du HGPS avec un retard de plusieurs années et avec une évolution plus lente de la maladie. En effet, le diagnostic a été effectué à 12 ans (au lieu d'un an pour les cas classiques), et le patient a vécu jusqu'à l'âge de 45 ans. La même mutation a été retrouvée en 2007 chez un patient développant la même forme retardée du HGPS, mais il était également atteint d'un ostéosarcome depuis l'âge de 9 ans (Shalev et al., 2007).

Les mutations c.1824C>T, c.1822G>A et c.1868C>G ont en commun d'être toutes les trois à l'origine d'un défaut de l'épissage. Les mutations les plus fréquentes c.1824C>T, c.1822G>A conduisent à l'activation d'un site 5' donneur d'épissage alternatif situé 150 nt en amont du site 5' donneur normalement utilisé de l'exon 11 (figure II-7). Chez un individu sain, ce site alternatif est présent mais est peu utilisé, alors que chez les individus atteints de progéria, il est utilisé dans la grande majorité des transcrits synthétisés à partir de l'allèle muté (Scaffidi and Misteli, 2005, 2006). L'utilisation de ce site alternatif a pour conséquences une délétion d'une séquence spécifique de la lamine A de 150 nt dans l'ARNm, et à la production d'une protéine tronquée de 50 aa, appelée progérine ou LAΔ50. La région protéique éliminée contient le second site de clivage par la métalloprotéase ZMPSTE24, qui, comme nous l'avons vu dans le paragraphe précédent, est indispensable à la dernière étape de maturation de la lamine A. La progérine conserve donc le groupement farnésyl fixé à son extrémité C-terminale et reste alors associée à la face interne de l'enveloppe nucléaire, ce qui, comme nous le verrons par la suite, a de graves conséquences sur la physiologie de la cellule. La mutation c.1868C>G au sein de l'exon 11 conduit à l'apparition d'un autre site 5' d'épissage n'existant pas en absence de la mutation. Ce site est situé 105 nt en amont du site 5' canonique. Son utilisation provoque donc une délétion de 105 nt dans l'ARNm mature (figure II-7), qui se traduit par la production d'une forme alternative de la progérine, appelée LAΔ35, qui, comme la forme LAΔ50, ne peut pas subir le second clivage par ZMPSTE 24. Néanmoins, dans ce cas, l'épissage anormal n'a lieu que dans environ 20% des transcrits synthétisés à partir de l'allèle muté, ce qui peut expliquer les symptômes moins sévères de la maladie (Fukuchi et al., 2004).

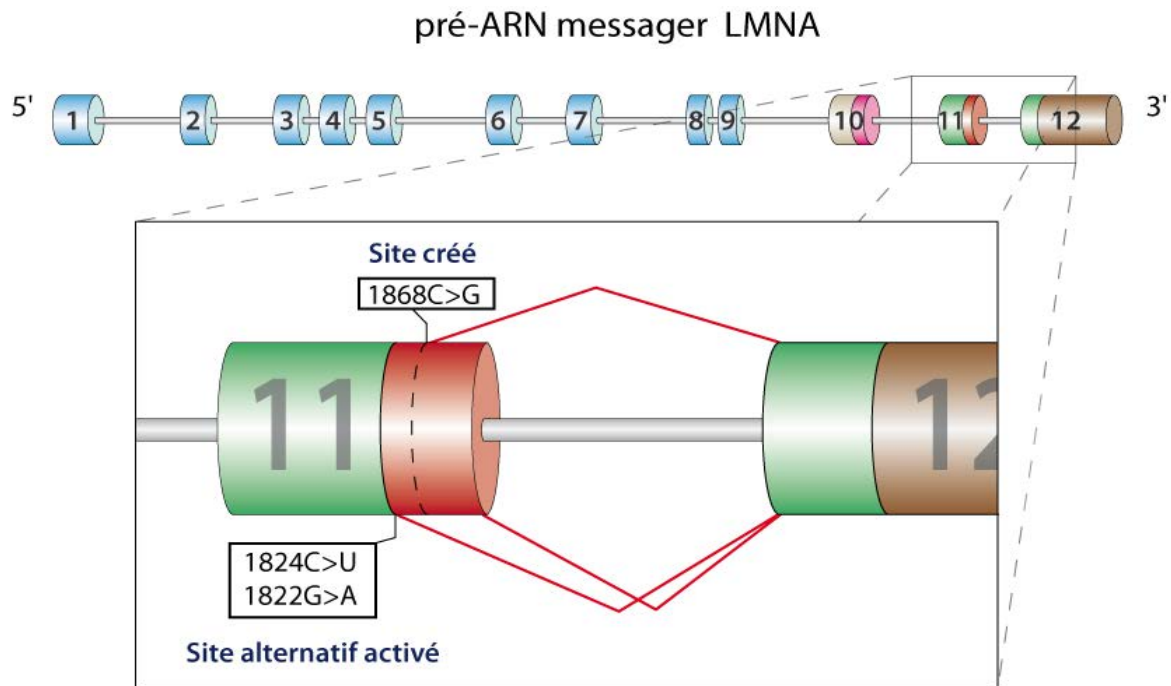








Figure II-7 : Mutations à l'origine des défauts d'épissage de l'exon 11 du pré-ARNm LMNA affectant l'exon 11. En haut, l'organisation des exons (cylindres numérotés de 1 à 12) et des introns du pré-ARNm LMNA est représentée de manière schématique. Un agrandissement de la région contenant l'exon 11 et le début de l'exon 12 est présenté dans l'encadré. Les positions des mutations 1824C>T, 1822G>A et 1868C>G, retrouvées chez plusieurs patients atteints de progéria, sont indiquées. Ces mutations sont présentes dans la région de l'ARN spécifique de la lamine A (indiquée en vert). Chaque mutation provoque l'activation ou la création d'un site 5' donneur d'épissage qui, lorsqu'il est utilisé, provoque la délétion d'une partie de l'ARNm (indiquée en rouge) par rapport à l'ARNm WT. Cette délétion est de 150 nt dans le cas des mutations 1824C>U ou 1822G>A qui activent un site 5' alternatif, et de 105 nt pour la mutation 1868C>G (en pointillés) qui crée un site 5'. Les couleurs utilisées dans cette figure correspondent aux régions exoniques et introniques suivantes :

	Région exonique commune à tous les transcrits LMNA		Région exonique spécifique à la lamine A / progérine
	Région exonique commune à la lamine A et la Lamine C		Région 3' non traduite
	Région exonique spécifique à la lamine C		Segment 3' terminal de l'exon 11 supprimé dans la progérine

II.4. Les fonctions de la lamine A et les perturbations entraînées par l'expression de progérine

II.4.1. Atteintes à l'organisation du noyau et à sa dynamique

Comme mentionné précédemment, le principal rôle de la lamine et de la *lamina* consiste au maintien de l'architecture nucléaire, et elle contribue à la répartition spatiale régulière des pores nucléaires (NPC) au niveau de l'enveloppe nucléaire (Aaronson and Blobel, 1975; Liu et al., 2000; Zhang et al., 1996). Ainsi, la délétion du gène *LMNA* ou la déplétion de son produit chez la souris, ou la délétion du gène de lamine Dm0 chez la drosophile, ou du gène de la Célamine chez *C. elegans*,

provoquent des déformations majeures de l'enveloppe nucléaire et une désorganisation importante des NPC (Lenz-Böhme et al., 1997; Sullivan et al., 1999; Zhang et al., 1996). Une des caractéristiques des cellules de patients atteints de progéria est la présence d'une forte proportion de noyaux adoptant une forme plurilobée aberrante (figure II-8)(Eriksson et al., 2003; Goldman et al., 2004). Ces déformations apparaissent graduellement de façon concomitante à l'accumulation de progérine et s'accompagnent d'un épaissement important de la *lamina*. L'effet de l'accumulation de la progérine sur la forme de l'enveloppe nucléaire a été confirmé par le fait que l'injection de progérine dans des cellules saines a un effet dominant négatif immédiat et entraîne la formation de noyaux multilobés dans ces cellules (Goldman et al., 2004). La formation de ces lobes affecte l'organisation des partenaires de la lamine A, comme les LAP, la lamine B1 et, d'une manière plus générale, des pores nucléaires (NPC) (Aaronson and Blobel, 1975; Goldman et al., 2004). Le caractère toxique de la progérine a également été montré dans des modèles souris progéria *ZMPSTE24* *-/-*, où la maturation de la lamine A ne peut pas être achevée, ce qui conduit à l'accumulation de pré-lamine A, dont les effets s'apparentent à ceux de la progérine (Pendas et al., 2002). Dans ces modèles, la délétion hétérozygote d'un allèle du gène *LMNA* diminue la quantité de progérine produite et donc sa toxicité, phénomène visible au travers de la diminution de la formation de noyaux multilobés (Fong et al., 2004).

Les propriétés de polymérisation des lamines leur confèrent également un rôle dans l'assemblage de l'enveloppe nucléaire à la fin de la mitose ou même *in vitro*, rôle qui est étroitement lié à l'architecture finale du noyau. Ainsi, l'utilisation d'un système d'assemblage des noyaux en extraits a permis de montrer que l'immunodéplétion des lamines empêche l'assemblage de l'enveloppe nucléaire, ou conduit à la formation de noyaux petits et fragiles (Burke and Gerace, 1986; Meier et al., 1991; Ulitzur and Gruenbaum, 1989). Par ailleurs, l'analyse du comportement des noyaux pendant l'interphase par des expériences de FRAP a révélé qu'au-delà d'une concentration critique en progérine, la *lamina* est beaucoup moins dynamique. Cela se traduit d'ailleurs par des difficultés d'extraction biochimique du fait de son interaction plus forte avec la membrane nucléaire via le groupement farnésyl de la progérine (Cao et al., 2007; Dahl et al., 2006; Goldman et al., 2004). Ce rôle négatif du groupement farnésyl est attesté par le fait que des inhibiteurs de farnésyl-transférase permettent le retour à un phénotype normal (voir aussi ci-dessous, paragraphe sur les potentialités thérapeutiques). La progérine, contrairement à la lamine A sauvage, n'est en effet pas libérée de la *lamina* pendant la mitose. Elle reste en grande partie associée à la membrane nucléaire et forme des foci cytoplasmiques aberrants après la dissociation du noyau (Cao et al., 2007) (figure II-8). Elle met ensuite plus de temps à rejoindre la lamina nucléaire néoformée, ce qui a pour effet de prolonger significativement la durée de la mitose, et pourrait affecter la morphogénèse du noyau dans les cellules filles (Cao et al., 2007). Une augmentation du taux d'anomalies de ségrégation chromosomique, telles que la production de cellules binucléées ou la formation de ponts chromosomiques entre deux cellules filles, est observée dans les cellules exprimant la progérine (Cao et al., 2007).

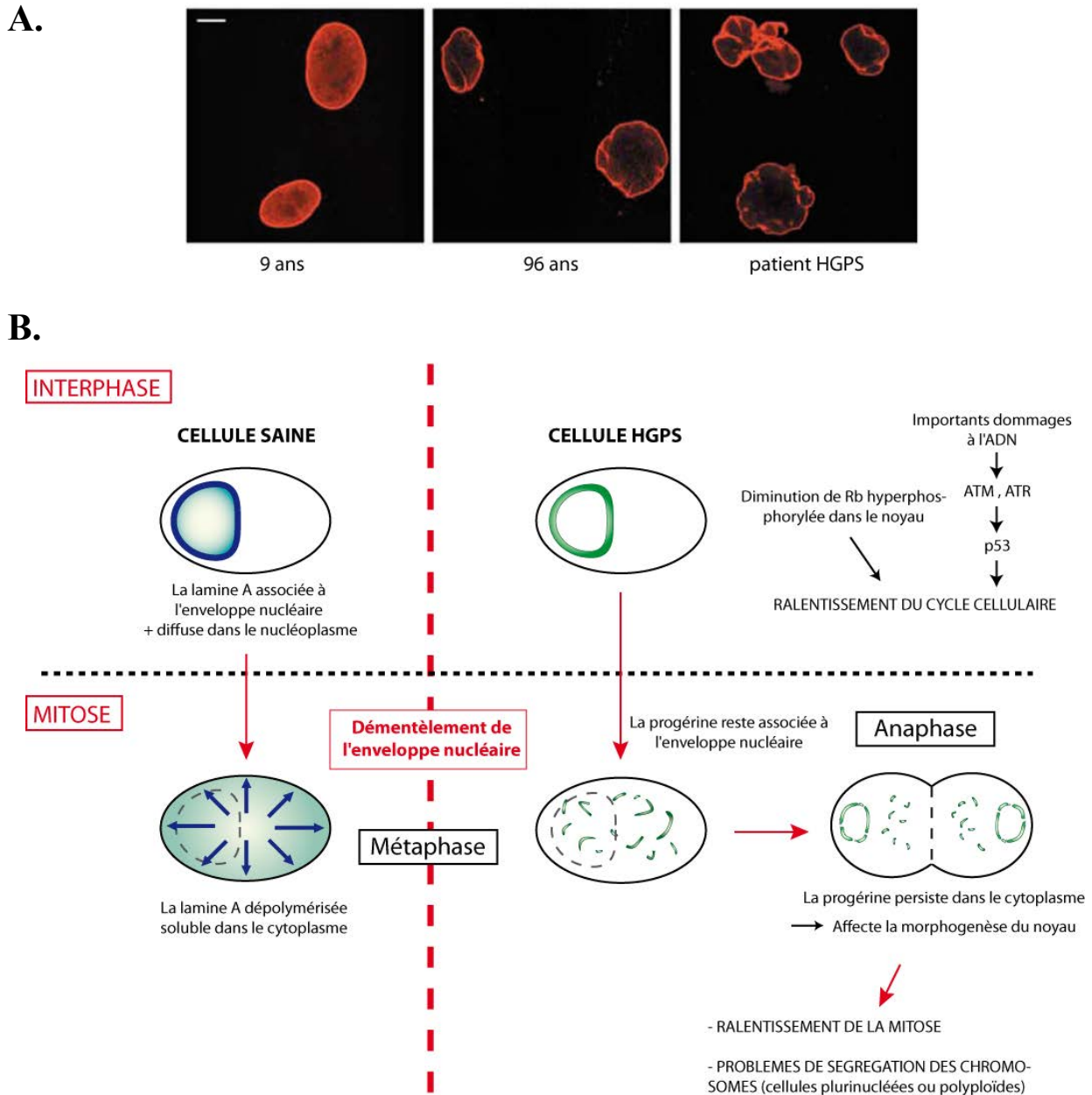


Figure II-8 : Le comportement anormal de la progérine dans le noyau. **A.** Utilisation de la microscopie de fluorescence pour la localisation de la lamine A/C dans des fibroblastes. Les fibroblastes proviennent d'individus âgés de 9 et 96 ans et d'un individu souffrant de progéria. Le noyau de l'individu âgé de 96 ans et celui de l'individu atteint de progéria adoptent une forme plurilobée aberrante, du fait de l'accumulation de progérine au niveau de l'enveloppe nucléaire. La barre blanche représente une distance de 10 μm (tiré de Scaffidi and Misteli, 2006). **B.** Représentation schématique du comportement de la progérine lors de l'interphase (partie supérieure) et de la mitose (partie inférieure). A gauche : comportement de la lamine A WT, représentée en bleu, dans les cellules d'individus sains. A droite, comportement de la progérine, représentée en vert, dans les cellules de patients atteints de progéria. Les événements annexes liés à la présence de progérine et conduisant à un ralentissement du cycle cellulaire sont également indiqués (voir texte).

II.4.2. Perturbation du métabolisme de l'ADN et du cycle cellulaire

L'interaction des filaments de lamine périphériques avec l'ADN sous diverses formes, comme des fibres de chromatine (Belmont et al., 1993; Yuan et al., 1991), les chromosomes lors de la mitose (Glass and Gerace, 1990), avec les histones (Goldberg et al., 1999; Taniura et al., 1995), ou *in vitro* avec des séquences spécifiques de l'ADN comme les "Scaffold/Matrix Attachment Regions" (S/MAR) (Baricheva et al., 1996; Ludérus et al., 1992, 1994; Shoeman and Traub, 1990; Zhao et al., 1996) montre le rôle de ces protéines dans le métabolisme de l'ADN. Ces données suggèrent aussi que les lamines constituent un des liens qui se crée entre la matrice nucléaire et l'ADN, et qu'elle est nécessaire à l'organisation du génome.

II.4.2.a. Perturbation de la réplication de l'ADN

Dans des noyaux assemblés *in vitro* à partir d'extraits interphasiques de xénope, l'ADN ne se réplique pas lorsque ces extraits sont déplétés en lamine (Meier et al., 1991). Les mêmes résultats sont obtenus avec des protéines lamines dominant négatif, perturbant l'assemblage de la *lamina* chez le xénope ou chez l'homme (Ellis et al., 1997; Moir et al., 2000a; Spann et al., 1997). Le mécanisme impliqué n'est pas connu en détails, mais il a été proposé que l'implication des lamines dans la réplication soit en partie liée à la nécessité de maintenir l'enveloppe et la matrice nucléaire, afin de permettre le transport et un positionnement initial efficace des facteurs de réplication. Cette proposition est confortée par l'observation d'une co-localisation des lamines avec le facteur de réplication PCNA dans des foci au cours de la phase S, et par le fait que la localisation de PCNA soit perturbée par l'expression d'un mutant dominant négatif de la lamine A (Moir et al., 1995, 2000a). En effet, en plus de leur présence à la périphérie du noyau, une partie des lamines de type A et B sont présentes dans la partie interne du noyau, d'une part au sein de minuscules foci dont l'apparition dépend de la progression dans le cycle cellulaire, mais également de façon plus diffuse dans le nucléoplasme (Bridger et al., 1993; Goldman et al., 1992; Liu et al., 2000; Moir et al., 2000b; Zhang et al., 1996). Elles peuvent y être associées avec des complexes de réplication de l'ADN (Goldberg et al., 1995; Moir et al., 2000a; Spann et al., 1997), et des protéines de la chromatine comme les histones H2A, H2B, H3 et H4 ou de la protéine HP-1 (Goldberg et al., 1999; Polioudaki et al., 2001; Taniura et al., 1995). La formation de ces foci est perturbée par l'expression de la progérine (Manju et al., 2006).

II.4.2.b. Instabilité génomique

Plusieurs études ont mis en évidence un rôle des lamines dans l'état de la chromatine. Les lamines interagissent en effet spécifiquement avec des régions d'ADN dont l'expression est particulièrement faible voire inexistante, comme les centromères et les télomères, et l'hétérochromatine

intercalaire ou en périphérie du noyau (Franklin and Cande, 1999; Marshall and Sedat, 1999). Une perte de l'hétérochromatine périphérique a été observée chez les patients atteints du HGPS : la présence de progérine est en effet associée à un certain nombre de dérégulations épigénétiques, incluant les modifications des histones qui affectent l'état de condensation de la chromatine (Goldman et al., 2004; Shumaker et al., 2006). Des études ont montré que l'absence de lamine A ou du récepteur de la lamine B (LBR) provoque une perte de la chromatine périphérique du noyau, montrant l'importance de la lamine A dans l'état de condensation de la chromatine. L'absence de lamine A provoque également l'apparition ectopique d'hétérochromatine dans les régions centrales du noyau, ainsi que la formation de chromosomes aberrants fusionnés (Solovei et al., 2013). Une perte de la chromatine périphérique a aussi été observée dans des cellules dépourvues en histones méthyltransférases telles que Suv4-20h2, responsable des méthylations d'histones H4K20me3. Ceci suggère l'importance de ces modifications dans la chromatine périphérique pour l'ancrage à la *lamina* (Hahn et al., 2013).

Il a également été montré, dans des souris modèles de la progéria, que leur durée de vie est augmentée de 60% et le phénotype progéria diminué lorsque ces souris n'expriment pas l'histone méthyltransférase Suv39h1 (Liu et al., 2013). Ceci peut s'expliquer par le fait que cette méthyltransférase interagit avec les lamines A, mais aussi de façon encore plus prononcée avec la progérine. La progérine présente dans les cellules progéria protège Suv-39h1 de la dégradation par le protéasome, ce qui provoque une augmentation des modifications des histones H3K9me3 et perturbe la réparation de l'ADN. Il a en effet été observé une accumulation de dommages à l'ADN, en particulier des cassures double-brins, dans les cellules exprimant la progérine (Liu et al., 2006; Manju et al., 2006). L'instabilité génomique provenant de lésions de ce type est un phénomène observé également lors du vieillissement physiologique (Lombard et al., 2005). Le fait qu'un défaut de recrutement des facteurs de réparation 53BP1 et RAD51 soit observé dans des fibroblastes HGPS ou des fibroblastes sains exprimant une prélamine A ne pouvant être maturée, pourrait expliquer les défauts de réparation observés (Liu et al., 2005).

D'autre part, l'histone déacétylase NAD⁺ dépendante SIRT1, responsable de l'acétylation d'histones comme H3, est capable d'interagir avec la lamine A et la matrice nucléaire dans les cellules souches adultes, et cette interaction est beaucoup plus faible avec la prélamine A ou la progérine (Liu et al., 2012). La présence de progérine dans les modèles de souris progéria K.O. pour le gène *ZMPSTE24* diminue la déacétylation des histones par SIRT1, ce qui a été précédemment associé à une instabilité génétique, mais aussi au niveau de l'organisme à une réduction de la durée de vie, des déficiences cardiaques ou rénales et une perturbation de la gamétogenèse (Cheng et al., 2003; Liu et al., 2012; Wang et al., 2008b).

II.4.2.c. Ralentissement du cycle cellulaire, et entrée en sénescence des cellules

L'accumulation de dommages à l'ADN provoque une perturbation du passage en phase S, voire l'entrée en sénescence des cellules, en empêchant le passage des différents points de contrôle qualité ("checkpoints") du cycle cellulaire liés à la réplication de l'ADN (Liu et al., 2006). En effet, les "checkpoint kinases" ATM et ATR sont recrutées au niveau des lésions et activent la protéine p53, "ange gardien du génome", dans les modèles de souris progéria *ZMPSTE24* *-/-* *LMNA* *+/-* (figure II-8) (Varela et al., 2005). La protéine p53 est un régulateur global, aux propriétés anti-proliférative, dont l'activation peut mener de façon ultime à l'apoptose. C'est pourquoi l'inactivation des kinases ATM et ATR peut partiellement restaurer le passage en phase S des cellules progéroïdes (Liu et al., 2006). De la même façon, la déplétion de Suv-39h1, évoquée un peu plus haut, diminue les modifications H3K9me3, et surtout restaure la réparation de l'ADN ce qui retarde la sénescence des cellules progéroïdes (Liu et al., 2013).

Par ailleurs, une interaction de la protéine du rétinoblastome, Rb, avec la lamine A, a été mise en évidence lorsque Rb est hypophosphorylée (Mancini et al., 1994; Ozaki et al., 1994). Comme p53, la protéine Rb est un suppresseur de tumeur et un régulateur majeur du cycle cellulaire. Sous sa forme hypophosphorylée, elle se lie avec le facteur de transcription E2F requis pour la progression dans le cycle cellulaire et inhibe son activité (pour revue, Giacinti and Giordano, 2006). Pour la progression dans le cycle cellulaire, Rb est hyperphosphorylé par action de kinases cyclines-dépendantes, ce qui provoque le relargage des facteurs de transcription E2F. Or, dans les cellules HGPS, une réduction significative du taux de protéine Rb hyperphosphorylée a été observée. Cette diminution, vraisemblablement liée à la présence de progérine, serait à l'origine du ralentissement voire de l'arrêt du cycle cellulaire (figure II-8) (Dechat et al., 2007). Il est à noter que Rb est également capable de réprimer la transcription en interagissant avec des protéines de remodelage de la chromatine (pour revue, Giacinti and Giordano, 2006), auxquelles l'histone méthyltransférase Suv39h1 évoquée ci-dessus appartient, mais le lien avec la progéria n'a jamais été démontré.

Ainsi, il a été observé que la présence de progérine est reliée à une durée de vie répllicative *in vitro* plus courte, dans les cellules non immortalisées (Cao et al., 2011; Huang et al., 2008). Les cellules humaines d'individus sains exprimant de façon ectopique des dérivés de la progérine sont d'ailleurs exposées à une sénescence prématurée et présentent un phénotype similaire à celui des cellules des patients HGPS (Kudlow et al., 2008). L'inhibition de l'expression de progérine par ARN interférence provoque par conséquent, outre le retour à un phénotype plus proche des cellules wt, une augmentation de la capacité proliférative des cellules et une diminution de leur entrée en sénescence (Huang et al., 2005). Il est connu que le raccourcissement des télomères des chromosomes est un phénomène directement relié à l'entrée en sénescence des cellules. Toutefois, la présence de progérine dans les cellules n'a pas pu être spécifiquement reliée au phénomène de raccourcissement des télomères, bien qu'il soit présumé que la *lamina* intervienne dans le métabolisme des télomères. En effet, la

surexpression de lamine A wt ou de progérine dans des fibroblastes immortalisés induit un raccourcissement des télomères (Huang et al., 2008). A l'inverse, le raccourcissement ou le dysfonctionnement des télomères active la production de progérine et la formation de noyaux multilobés (Cao et al., 2011). C'est probablement pour cette raison que la sénescence associée à l'expression de progérine, ectopique ou non, peut être en partie contrecarrée par l'expression de l'unité catalytique protéique de la télomérase hTERT, même si cela ne diminue pas la proportion de noyau multilobés (Cao et al., 2011; Kudlow et al., 2008).

II.4.3. Influence de la progérine sur l'expression des gènes

Au sein de foci intranucléaires ou à la périphérie du noyau, la lamine A peut également interagir avec des complexes de transcription de l'ARN polymérase II, ainsi qu'avec des facteurs de maturation de l'ARN (Jagatheesan et al., 1999; Spann et al., 2002). La co-localisation intranucléaire des lamines avec ces facteurs suggère un rôle des lamines dans la régulation de l'expression des gènes. Les changements du profil d'expression des différentes lamines sont corrélés avec différents stades du développement et de la différenciation cellulaire. La lamine A, notamment, est exprimée dans les cellules en différenciation et contribuerait à l'expression de gènes spécifiques de la différenciation cellulaire, probablement *via* l'interaction avec la chromatine et le contrôle de la condensation de la chromatine. Ainsi, l'expression ectopique de lamine A dans des cellules myogéniques de poule, ou normalement seule la lamine B est exprimée, induit l'expression de protéines spécifiques des muscles (Lourim and Lin, 1992). Cependant, l'expression ectopique de lamine A n'induit pas la différenciation de cellules embryonnaires cancéreuses (Peter and Nigg, 1991). De façon intéressante, les perturbations de l'organisation de la chromatine induites par la présence de progérine pourraient induire la transcription d'ARN dans les régions péri-centriques (Shumaker et al., 2006). L'exemple le plus impressionnant de ce type d'activation est l'induction de la transcription des ARN satellite III, sur lequel nous reviendrons par la suite. D'autre part, plusieurs études comparatives du transcriptome de fibroblastes sains et HGPS ont été effectuées (Csoka et al., 2004; Ly et al., 2000; Marji et al., 2010; Scaffidi and Misteli, 2008). Ces études mettent en évidence de façon récurrente des variations d'expression des gènes impliqués dans l'état de condensation de la chromatine, la réplication de l'ADN et de façon plus importante des gènes de facteurs de transcription.

Une influence de la progérine sur le transcriptome est également en accord avec le fait que la disruption de l'organisation des lamines, par injection de lamines sans domaine "tête", provoque des changements dans la morphologie des domaines subnucléaires liés à l'expression des gènes. Cette disruption provoque, entre autres, la redistribution des UsnRNP et des facteurs d'épissage dans des speckles de plus grande taille, ainsi que l'agrégation du facteur TBP ("Tata Binding Protein") (Spann et al., 2002). Il en résulte une diminution drastique de l'activité transcriptionnelle de l'ARN polymérase II.

Ces données et celles du paragraphe précédent suggèrent que l'influence des lamines sur la réplication et la transcription de l'ADN découle de leur rôle dans l'organisation du noyau et de la formation d'une matrice, à la périphérie et à l'intérieur du noyau, servant à l'ancrage de facteurs nécessaires à ces processus et à la structuration de la chromatine (Goldman et al., 2002; Hendzel et al., 1991; Tsutsui et al., 2005).

II.4.4. Liens avec le phénotype observé à l'échelle de l'organisme

Les liens étiologiques entre l'expression de progérine et les symptômes observés chez l'homme sont difficiles à définir avec précision, du fait des effets multiples et globaux de cette protéine. Toutefois, des liens ont pu être établis entre les perturbations détectées au niveau cellulaire et les symptômes observés. Le lien le plus évident est le phénomène d'entrée en sénescence des cellules par les mécanismes décrits dans les paragraphes précédents. Le vieillissement est alors provoqué par le peu, voire l'absence, de renouvellement des tissus par mitose. Ce phénomène est sans doute exacerbé par l'augmentation de la durée du cycle cellulaire inhérent aux déficiences des systèmes de réparation de l'ADN et au ralentissement de la dislocation de la lamine contenant de la progérine pendant la mitose (Anckar and Sistonen, 2011; Cao et al., 2007; Dahl et al., 2006). L'étude comparative du transcriptome de cellules souches mésenchymateuses exprimant ou non la progérine a montré que cette dernière affecte le potentiel de différenciation de ces cellules (Scaffidi and Misteli, 2008). Cela suggère que la progérine serait à l'origine d'un vieillissement au niveau de l'organisme non seulement lié à la sénescence des cellules mais également à la perte du potentiel de différenciation des cellules pluripotentes.

Les tissus affectés dans la progéria (en particulier le système cardiovasculaire et les os) et plus globalement dans la plupart des laminopathies, sont principalement issus des cellules souches du mésoderme de l'embryon. L'hypothèse selon laquelle la progéria serait en partie une "maladie du mésoderme" permet ainsi de mieux comprendre les symptômes à l'échelle de l'organisme. Parmi les gènes différenciellement régulés dans la progéria, on retrouve principalement des facteurs de transcription mais aussi des facteurs de synthèse de la Matrice ExtraCellulaire (MEC) (Csoka et al., 2004). Ceci peut s'expliquer par les données d'une autre étude mettant en évidence la perturbation de la signalisation Wnt, dont dépend la synthèse de la MEC, dans les cellules HGPS (Hernandez et al., 2010). Des déficiences dans la synthèse de la MEC pourraient expliquer une partie des symptômes observés puisqu'on sait qu'une telle déficience provoque l'apoptose des cellules des tissus osseux et peut également entraîner l'apoptose des vaisseaux sanguins à proximité du cœur.

D'autre part, la lipodystrophie est un symptôme récurrent des pathologies résultant des mutations du gène *LMNA*, et est un des symptômes de la progéria. Des travaux de l'équipe de J. Tazi et de leurs collaborateurs sur un modèle de souris progéria ont pu montrer qu'une forte expression de progérine est délétère pour les cellules des tissus adipeux, ce qui se manifeste par une perte de la graisse sous cutanée

similaire à celle observée chez les patients atteints de progéria ou les individus âgés (Osorio et al., 2011 et thèse I.C. Lopez-Mejia, “Alternative splicing of LMNA gene: lessons from a new mouse model of Hutchinson-Gilford progeria syndrome”). Par ailleurs, une forte expression de la progérine ou une exposition à long terme à la progérine chez la souris mène à une perturbation du métabolisme énergétique, notamment une augmentation de la sensibilité à l'insuline. Cette perturbation du métabolisme énergétique pourrait être l'un des phénomènes sous-jacent du phénotype progéroïde, la lipodystrophie étant un symptôme récurrent dans ce type de maladies.

Il a également été proposé que certains symptômes cardiovasculaires du HGPS soient liés aux propriétés mécaniques des cellules. Il a en effet été montré par la technique FRAP que l'application de forces mécaniques de pression ou d'étirement dans les cellules exprimant la progérine entraîne une rigidification significative des cellules, due aux changements de propriétés biochimiques et mécaniques de la *lamina* (Dahl et al., 2006). La *lamina* est quasiment incapable de retourner à une forme normale après l'application d'un effet mécanique d'étirement. La mécano-transduction d'un stimulus mécanique pourrait donc être affectée dans les cellules exprimant la progérine, rendant les cellules incapables de répondre correctement à un stress mécanique. Cette rigidité des cellules pourrait être néfaste pour les cellules subissant des contraintes mécaniques continues, comme les cellules musculaires lisses ou encore les cellules endothéliales vasculaires, constamment soumises aux pressions mécaniques résultant du flux sanguin (Davies et al., 2003). Une calcification des tissus vasculaires a aussi été observée récemment chez des souris exprimant la progérine. Elle résulte du dysfonctionnement mitochondrial qui provoque une baisse de la production d'ATP, d'où une augmentation de la concentration extracellulaire en pyrophosphate (PPi) et son appauvrissement au sein des cellules (Villa-Bellosta et al., 2013). Le PPi est un inhibiteur de la calcification vasculaire, un traitement à base de PPi permet d'ailleurs de diminuer cette calcification chez les souris exprimant la progérine, ce qui pourrait constituer une piste thérapeutique valable pour l'avenir.

II.4.5. Potentialités thérapeutiques face à la progéria

A ce jour, il n'existe aucun traitement permettant de guérir la progéria. Toutefois, certaines stratégies thérapeutiques sont en cours d'investigation ou d'expérimentation, comme l'emploi d'inhibiteurs de la farnésyl-transférase (les FTI). *In cellulo*, ces inhibiteurs diminuent clairement la proportion de noyaux multilobés, que ce soit dans des modèles de cellules HeLa progéria ou dans des fibroblastes issus de patients atteints d'HGPS (Capell et al., 2005; Glynn and Glover, 2005; Mallampalli et al., 2005; Toth et al., 2005). Cette stratégie permet un retour des cellules à un phénotype quasi normal en ce qui concerne la dynamique de la *lamina* et la forme du noyau, mais ne réduit pas les dommages subis par l'ADN (Liu et al., 2006 ; Cao et al., 2007). Ces inhibiteurs ont aussi un effet positif sur la durée de vie et provoquent une diminution des accidents cardiovasculaires dans un modèle de souris progéria

ZMPSTE24 ^{-/-} (Varela et al., 2008; Yang et al., 2006). Il existe de nombreux FTI et un rapport préliminaire récent sur les essais cliniques d'une de ces drogues, le lonafarnib, rapporte les effets bénéfiques d'un traitement de 2 ans sur les symptômes du HGPS chez plusieurs patients. Ces effets bénéfiques sont observés tant au niveau cardiovasculaire, avec une forte diminution des attaques cardiaques, qu'au niveau du squelette, du gain de poids, et de l'audition, avec de faibles effets secondaires (Gordon et al., 2012). Certains patients présentant des symptômes neurologiques comme des maux de têtes chroniques ont également vu ces symptômes diminuer (Ullrich et al., 2013). Les FTI sont plus efficaces lorsqu'ils sont utilisés conjointement avec des inhibiteurs de géranyl-géranyl-transférase, comme lors d'un traitement utilisant conjointement les statines et les aminobiphosphates (Varela et al., 2008). En effet, le groupement géranyl-géranyl peut être utilisé pour la prénylation de la lamine A lorsque la farnésylation est bloquée.

Il a également été montré que le résveratrol (surtout connue pour ses propriétés antioxydantes) est capable de restaurer l'activité de l'histone déacétylase SIRT1, dont l'activité est fortement diminuée par la présence de progérine dans les cellules (Liu et al., 2012). Chez les souris progeria ZMPSTE24 ^{-/-}, le résveratrol permet d'une part de retarder le déclin des cellules souches adultes, mais permet également à l'échelle de l'organisme d'améliorer la structure et la densité minérale des os, d'augmenter la durée de vie, et de retarder la perte de poids (Liu et al., 2012).

Toutefois, d'autres études ont montré la toxicité résultant de l'accumulation de pré-lamine A non farnésylée, montrant les limites de la thérapie basée sur l'inhibition de la farnésylation de la lamine A ou de la progérine, mettant en évidence le besoin de modes de thérapie complémentaires (Davies et al., 2010; Le Dour et al., 2011; Lee et al., 2010). D'autres stratégies thérapeutiques qui sont envisagées reposent sur l'inhibition de la production ou la suppression de l'ARNm de la progérine. Une de ces stratégies est l'élimination de l'ARNm de la progérine par ARN interférence grâce à l'emploi de shRNA dirigés spécifiquement contre l'ARNm de la progérine. Ces shRNA ont été introduits dans les cellules de patients HGPS grâce à un vecteur viral et ont permis une diminution d'environ 75% de l'expression de la progérine, menant à une atténuation du phénotype progéroïde des cellules (Huang et al., 2005). Une autre stratégie importante est l'emploi d'oligonucléotides antisens morpholino permettant l'inhibition par gêne stérique de l'utilisation du site 5' d'épissage progérine dans le pré-ARNm LMNA. L'efficacité de cette approche a pu être démontrée *in cellulo* mais aussi dans un modèle de souris progeria, puisqu'elle a mené à l'augmentation de la durée de vie de ces souris (Osorio et al., 2011; Scaffidi and Misteli, 2005). Ces stratégies reposent sur une bonne connaissance des mécanismes régissant l'utilisation des sites 5' d'épissage de l'exon 11.

II.4.6. La progéria en tant que modèle d'étude du vieillissement cellulaire

En plus de son intérêt pour le développement d'une thérapie efficace, l'étude de la progéria a un intérêt plus général. En effet, la pathologie correspond à un vieillissement accéléré et constitue donc un modèle intéressant pour l'étude des phénomènes de vieillissement. Le fait que l'expression de la progérine induise de nombreux phénomènes cellulaires qui apparaissent au cours du vieillissement, en particulier la sénescence, a permis de mettre en évidence un rôle central de la *lamina* dans ces processus. Une autre démonstration d'un lien possible entre progéria et vieillissement normal est liée à la détection de l'utilisation du site alternatif d'épissage de l'exon 11 du pré-ARNm LMNA dans les fibroblastes de donneurs sains et la production de progérine dans ces cellules (McClintock et al., 2007; Scaffidi and Misteli, 2006). Cette progérine produite à faible taux s'accumule au fur et à mesure des divisions cellulaires, et est retrouvée en plus grande quantité dans les noyaux de fibroblastes de personnes âgées, chez qui la proportion de noyaux multilobés augmente également en conséquence (Goldman et al., 2004; McClintock et al., 2007; Scaffidi and Misteli, 2006).

II.5. Liens existant entre la progéria, le vieillissement, et les modèles cellulaires du stress

De nombreuses pistes concernant l'étude du vieillissement ont été fournies par les maladies progéroïdes, notamment l'importance de la *lamina* dans de nombreux processus nucléaires, dont la stabilité génomique. Il est également possible de relier l'impact des maladies progéroïdes avec plusieurs théories formulées pour expliquer le vieillissement cellulaire.

II.5.1. Progéria et stress oxydant

Le stress oxydant est depuis longtemps considéré comme un facteur majeur du vieillissement cellulaire (Harman, 1956). Les dommages causés par le stress oxydant résultent de la formation incontrôlée dans la cellule de molécules oxygénées hyper-réactives, instables (principalement des radicaux libres) pendant les processus respiratoires impliquant le dioxygène (Sena and Chandel, 2012). Ces molécules peuvent endommager la majorité des constituants cellulaires, que ce soit les acides nucléiques, les protéines ou les lipides. Ces dommages augmentent avec l'âge et pourraient contribuer voire être à l'origine du vieillissement de l'organisme, mais cette vision du vieillissement a été révisée au cours des dernières années. Les données sont contradictoires : certaines études montrent une corrélation entre la durée de vie et l'activation des mécanismes de défense face au stress oxydant (pour revue, Landis and Tower, 2005). Par contre d'autres études plus récentes ont mis en évidence que la surexpression de

protéines impliquées dans ces mécanismes de défense a au contraire peu d'effet sur la longévité, ceci dans plusieurs organismes modèles (Gems and Doonan, 2009; Lapointe and Hekimi, 2010; Pérez et al., 2009). Cela est probablement dû au caractère multifactoriel et complexe du vieillissement, qui ne peut se résumer aux seuls effets du stress oxydant. Quoiqu'il en soit, plusieurs liens directs ont été établis entre la progéria et le stress oxydant : premièrement, l'accumulation de dommages aux protéines et à l'ADN par le biais d'espèces oxydantes dans les cellules de patients HGPS a pu être reliée directement à la présence de progérine (Richards et al., 2011; Viteri et al., 2010). Le stress oxydant peut également lui-même générer la formation de noyaux multilobés, du fait de modifications subies par la *lamina*, similaires à ce qui est observé dans la progéria (Barascu et al., 2012). La lamine B1 séquestre au niveau de la *lamina* le facteur de transcription Oct-1, dont les gènes cibles sont impliqués dans la réponse cellulaire au stress oxydant (Malhas et al., 2009). En accord avec ces données, les cellules n'exprimant pas de lamine B1 fonctionnelle sont hypersensibles au stress oxydant.

L'analyse de protéomes cellulaires de souris modèles de la progéria mutées au niveau du gène *LMNA*, a permis de mettre en évidence une diminution des protéines de défense face au stress oxydant ainsi que de la chaîne respiratoire mitochondriale, ce qui conduirait à une dysfonction des mitochondries et une baisse de la production d'ATP dans les fibroblastes et dans plusieurs tissus (Rivera-Torres et al., 2013; Villa-Bellosta et al., 2013).

II.5.2. Progéria et choc thermique

Une autre façon d'appréhender les mécanismes entraînant le vieillissement est de prendre en compte l'apparition au cours du temps d'une détérioration progressive de la capacité des cellules à maintenir l'intégrité de leur protéome, ce qui entraînerait l'apparition de protéines toxiques au cours du temps (protéotoxicité) (Morimoto and Cuervo, 2009). Un des exemples les plus marquants est celui de la maladie d'Alzheimer. Dans certains mutants de différents organismes modèles comme la souris ou *C. elegans*, une extension significative de la durée de vie a pu être reliée à l'augmentation de l'efficacité de mécanismes permettant de limiter la protéotoxicité, en particulier l'action de protéines chaperonnes empêchant l'agrégation et la dénaturation d'autres protéines (Cohen et al., 2006, 2009; Morley et al., 2002). Je décrirai en particulier le cas du choc thermique, qui a fait l'objet de la seconde partie de mes travaux de thèse. Les effets observés de ce stress ont des liens avec ceux observés pour la progéria et lors du vieillissement. Une étude a par exemple montré que soumettre des cellules à des chocs thermiques modérés (41°C) et répétés pendant une période prolongée (répartition de 35 chocs régulièrement sur une période de 140 jours), diminue les malformations nucléaires apparaissant au cours du vieillissement, bien que cela n'augmente pas la durée de vie pendant laquelle les cellules peuvent se répliquer (Rattan, 1998). Une autre étude a de plus montré que les fibroblastes HGPS sont hypersensibles au choc thermique (Paradisi et al., 2005). Un fort taux de mort cellulaire (25%) est observé 24 h après le choc thermique

alors que la population de fibroblastes sains augmente d'environ 40% dans le même temps. De plus, la croissance des cellules HGPS ne reprend qu'environ 48 h après la fin du stress.

Ces liens entre la progéria et le choc thermique ont également été observés au niveau moléculaire. D'une part, l'interaction des composants de la *lamina* avec des protéines du choc thermique (les HSP) a été mise en évidence. L' α B-crystalline et la protéine hsp25, qui sont des HSP de petite taille ("small HSP"), sont impliquées dans la stabilisation des foci intranucléaires contenant la lamine A (Adhikari et al., 2004). D'autre part, bien que la disposition et la localisation des composants de la *lamina* ne semblent pas impactées par le choc thermique, une augmentation de l'expression de la lamine B 40 min après le début du stress et une diminution de l'expression de l'émerine pendant la période de régénération post-stress ont été observées dans des cellules HeLa (Haddad and Paulin-Levasseur, 2008). Ces variations perdurent plusieurs heures après la fin du choc thermique ce qui suggère un rôle de ces protéines dans la régénération des cellules, peut-être au travers de leurs effets sur la chromatine ou sur l'expression des gènes.

Le lien le plus marquant entre la progéria et le stress thermique est la production d'ARN satellite III à partir des régions péricentromériques de certains chromosomes. Ces ARN ne sont transcrits qu'après divers stress, en particulier le stress thermique, et ils forment dans les noyaux des granules qui contiennent des facteurs d'épissage et dont la formation pourrait affecter l'épissage. Des granules contenant de l'ARN sat III, formés probablement à cause de variations épigénétiques, ont été décrits dans les cellules de patients HGPS (Shumaker et al., 2006). L'étude de ces granules fait par ailleurs l'objet de la seconde partie de ma thèse.

Partie 1 - Objectifs

Lorsque j'ai débuté ma thèse on savait que la progéria était due à des défauts d'épissage du gène *LMNA* codant les lamines A et C et plus particulièrement à des défauts entraînant la production de lamines A tronquées présentant des altérations de maturation l'immobilisant au niveau de l'enveloppe nucléaire. On savait aussi que ces défauts d'épissage étaient liés à des mutations ponctuelles dans le gène *LMNA* conduisant à l'utilisation de sites 5' donneurs différents du site 5' donneur normal (5' SS lamine A) lors de l'épissage de l'intron 11. L'utilisation de ces sites 5' situés en amont du site 5' SS lamine A, conduit à la production d'ARNm tronqués et donc à la synthèse de protéines tronquées : la progérine pour les mutations 1824C>U et 1822G>A (utilisation du site 5' SS progérine) et la lamine A Δ 35 pour la mutation 1868C>G (utilisation du site 5' SS LA Δ 35).

Comme nous l'avons mentionné, différentes stratégies pouvant corriger des défauts d'épissage sont en cours d'investigation dans différents laboratoires et entreprises dans une perspective d'application thérapeutique. Ces stratégies reposent sur une connaissance approfondie des mécanismes régulant l'utilisation du site normal et des sites créés, ou dont l'utilisation est renforcée suite à une mutation du génome. En vue de pouvoir essayer d'appliquer de telles stratégies à la progéria, il était donc nécessaire de commencer par caractériser les propriétés du site 5' SS lamine A dont l'utilisation est diminuée dans le HGPS, et les propriétés des sites dont l'utilisation augmente.

Les données de l'équipe de Misteli, obtenues sur des fibroblastes d'individus sains et de patients souffrant de progéria, indiquaient que le site 5' donneur, dont l'utilisation est renforcée en présence des mutations 1824C>U ou 1822G>A (site 5' progérine), pouvait être utilisé à un taux très faible par rapport au site 5' SS lamine A (50 fois moins environ) (Scaffidi and Misteli, 2006). Ceci suggérait qu'il s'agit d'un site 5' donneur alternatif de l'intron 11. Vu le faible taux d'expression observé, il était néanmoins important de le vérifier par des approches *in vitro* et *in cellulo*, ce que nous avons fait.

D'une manière générale, lorsque j'ai débuté ma thèse, aucune étude n'avait été réalisée sur les mécanismes assurant la régulation du taux d'utilisation du site 5' SS lamine A, ni sur ceux pouvant s'exercer sur les sites 5' SS progérine et 5' SS LA Δ 35, aussi bien dans un contexte sauvage qu'en présence des mutations. Nous avons abordé cette question sous deux angles. D'une part, nous avons testé l'effet possible de la structuration de l'ARN au niveau des différents sites 5' donneurs de l'intron 11 en présence ou en absence des mutations. D'autre part, nous avons étudié les effets possibles des facteurs régulant l'épissage (protéines SR et hnRNP) sur ces sites 5' d'épissage, là encore en présence ou en absence des mutations.

Ce travail a été réalisé en collaboration avec l'équipe J. Stévenin (IGBMC Strasbourg) et J. Tazi (IGMM Montpellier). Il a constitué un des volets d'un projet ANR commun, ainsi qu'un des projets développés dans le cadre du réseau européen EURASNET dirigé par R. Lührmann et auquel

appartenaient les trois équipes. Par ailleurs, l'équipe de J. Tazi avait en parallèle développé une collaboration avec l'équipe N. Levy à Marseille, spécialiste de la pathologie progéria. Les deux équipes ont construit ensemble des modèles souris de la progéria qui ont permis de tester chez l'animal les hypothèses formulées sur la base des données que nous avons obtenues *in vitro* et *in cellulo*.

PARTIE 1

PROGERIA ET EPISSAGE

(Résultats)

III. Etude expérimentale des mécanismes à l'origine des régulations et dérégulations de l'épissage du transcrit LMNA en lien avec la progéria

III.1. Etude des mécanismes induisant l'activation du site 5' SS progérine en présence des mutations 1824C>U

Les données des travaux concernant cette partie de mes résultats sont en grande partie rassemblés dans un article publié dans le journal Human Molecular Genetics dont je suis second auteur. Cet article est intégré à la fin de ce chapitre. Dans les paragraphes suivants je présenterai la stratégie de l'étude réalisée dans cet article ainsi que ma contribution aux résultats publiés.

III.1.1. Constat initial

Nous avons constaté que le simple examen du niveau de complémentarité entre le snRNA U1 et la séquence du site 5' SS progérine (progerin 5' Splice Site ou 5'SS dans la publication n°1) ne permettait pas d'expliquer l'utilisation prédominante de ce site dans le contexte de la mutation 1824C>U ou 1822G>A à l'origine de la progéria. L'appariement entre la région 5' terminale du snRNA U1 et la séquence en 5' des introns est un évènement précoce dans la définition des exons et le niveau de complémentarité de ces deux séquences est un paramètre important dans la définition de la force d'un site 5' d'épissage. La séquence du site 5' lamine A (lamin A 5' Splice Site ou 5'SS dans la publication n°1) est identique à la séquence consensus des sites 5' d'épissage des vertébrés. De ce fait, ce site 5' est prédit comme ayant une forte efficacité d'après les estimations faites par logiciel (un score de 12,2 pour SScoreCalculation et 10,67 pour MaxEnt, ces scores reflétant l'efficacité intrinsèque d'un site d'épissage) (Zhanglab ; Yeo and Burge, 2004) (figure 1 de la publication n°1, p. 2). Des scores très inférieurs sont obtenus pour le site 5' SS progérine (scores de 7,1 et 8,07) ceci du fait d'une divergence à trois positions (+3, +4 et +6 dans l'intron) par rapport à la séquence consensus. Ces différences expliquent sans doute en partie la très faible utilisation du site 5' SS progérine en absence des mutations causant la pathologie. Ces mutations réduisent au nombre de deux les différences par rapport à la séquence consensus. La mutation 1824C>U permet de remplacer la paire de bases A-C à l'extrémité de l'appariement par une paire de bases A-U plus stable. La mutation 1822G>A convertit une paire G•U en une paire A-U plus stable (figure 1 de la publication n°1, p. 2). Nous avons aussi examiné la complémentarité des sites 5' WT et muté avec le snRNA U6. En effet, lors de la formation des complexes catalytiques à partir des pré-spliceosomes, le snRNA U1 est relargué et son appariement avec l'extrémité 5' de l'intron est remplacé par un

appariement avec le snRNA U6. Alors que la mutation 1824C>U renforce la complémentarité avec le dinucléotide AC du snRNA U6, la mutation 1822G>A ne semble pas augmenter cette complémentarité.

Clairement les variations de complémentarité de la séquence du site 5'SS progérine avec aussi bien le snRNA U1 que le snRNA U6 en présence des mutations ne permettaient pas de comprendre l'utilisation prédominante de ce site lorsqu'il est muté, ses scores restant toujours nettement inférieurs à ceux du site 5' SS lamine A. D'autres effets de ces mutations devaient aussi avoir un impact sur l'utilisation relative des deux sites 5'. Comme nous l'avons vu en introduction, deux autres paramètres peuvent jouer un rôle très important dans les choix alternatifs des sites d'épissage : la structure secondaire du pré-ARNm et les facteurs régulateurs de l'épissage. Nous avons donc décidé de tester si les mutations pouvaient modifier la structuration du pré-ARNm et/ou si elle pouvait détruire un élément inhibiteur ISS/ESS ou créer un élément activateur ISE/ESE au voisinage du site 5' SS progérine.

III.1.2. Mise au point d'un système d'étude *in vitro* de l'épissage de l'intron 11 du transcrit LMNA

Pour pouvoir décrypter les mécanismes intervenant dans les régulations et dérégulations de l'utilisation des sites 5'SS lamine A et progérine, il était important de pouvoir disposer d'un système d'étude *in vitro*, ceci aussi bien pour tester l'effet de la structure secondaire du pré-ARNm que pour tester l'effet de facteurs régulant l'épissage, en particulier les protéines SR.

Pour cela, nous avons utilisé une construction produite par l'équipe de J. Tazi pour la réalisation de tests d'épissage *in cellulo* (figure III-1). On savait qu'*in cellulo*, la génération de la mutation 1824C>U dans cette construction conduisait, comme c'est le cas dans les cellules de patients atteints de progéria, à l'utilisation majoritaire du site 5'SS progérine (Figure 4B, p16 de la publication n°1). Nous avons donc décidé de l'adapter à la production du transcrits dont l'épissage pouvait être étudiés *in vitro* en extrait nucléaire.

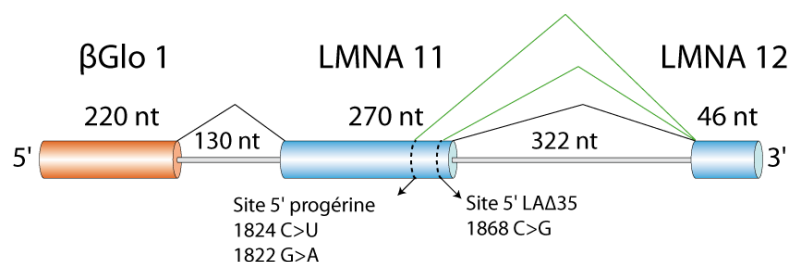


Figure III-1 : Représentation schématique de l'ARN utilisé pour les expériences d'épissage *in vitro*. Cet ARN contient de 5' vers 3' : le premier exon (en rouge) et le premier intron de la β -globine, puis l'exon 11, l'intron 11 et l'exon 12 de l'ARN LMNA. Les tailles des différents segments de l'ARN sont indiquées, ainsi que les mutations testées. Les lignes brisées représentent les sites potentiellement utilisés au sein de ce transcrit. Les lignes vertes représentent les événements d'épissage impliquant le site 5'SS progérine ou LA Δ 35.

Cette construction insérée dans le plasmide pCND3/V5-His TOPO produit par linéarisation par l'enzyme *NotI* et par transcription avec l'ARN polymérase du phage T7 un pré-ARNm appelé *βglo*-LMNA contenant l'exon 11, l'intron 11 et la partie codante de l'exon 12. Pour être épissé efficacement, le pré-ARNm doit porter une coiffe m⁷G en 5' et pour le suivi des produits d'épissage, il doit être uniformément marqué. La transcription *in vitro* de l'ARN polymérase du phage T7 est donc effectuée en présence d'[α-³²P]-UTP et de m⁷GpppGp.

Les premiers essais d'épissage *in vitro* de l'ARN *βglo*-LMNA réalisés par l'équipe de J. Stévenin avaient montré que les conditions classiquement utilisées pour l'épissage *in vitro* ne permettent pas d'observer un épissage au niveau des sites 5' de la partie LMNA de l'ARN *βglo*-LMNA. J'ai donc dû dans un premier temps trouver des conditions expérimentales adaptées à l'épissage *in vitro* de l'ARN produit à partir de cette construction. Les différents tests réalisés (variations de la quantité de Mg²⁺, de la quantité d'ATP, du volume d'extrait nucléaire utilisé et du temps d'incubation) ont permis d'établir que l'ARN *βglo*-LMNA étudié nécessitait une concentration en Mg²⁺ deux fois plus faible que la normale (1,6 μM au lieu de 3,2 μM) afin être épissé efficacement au niveau des sites 5' LMNA. Pour identifier les produits d'épissage séparés par électrophorèse en gel de polyacrylamide dénaturant à 5% nous avons déposé en parallèle un marqueur de taille radiomarké. L'estimation relative du taux des différents produits d'épissage a été réalisée par mesure de la quantité de radioactivité présente dans chaque bande du gel en utilisant le logiciel Image Quant.

Une fois les conditions d'épissage mises au point, nous avons effectué une cinétique d'épissage avec des l'ARN *βglo*-LMNA WT et des ARN *βglo*-LMNA portant la mutation 1824C>U avec pour comparaison une cinétique réalisée avec un ARN portant la mutation 1819G>A détruisant le site 5'SS progérine. Les produits d'épissage obtenus après 30, 60, 90, 120 et 150 min d'incubation en extrait nucléaire de cellule HeLa ont été analysés par électrophorèse (figure III-2). Les expériences réalisées avec 2h d'incubation à 30°C ont été répétées 3 fois. La quantification des résultats obtenus atteste de la reproductibilité des données.

La réaction d'épissage du transcrit *βglo*-LMNA WT conduit dès 30 min d'incubation à l'apparition d'un produit d'épissage de grande taille (environ 900 nt). D'après la taille des différents introns potentiels de l'ARN *βglo*-LMNA, il correspond à l'élimination de l'intron amont situé entre l'exon *β*globine et l'exon 11 LMNA. La rapidité de ce premier événement d'épissage est sans doute liée à la forte efficacité du site 5' situé en aval de l'exon *β*-globine. Ensuite à partir de 60 min d'incubation, un second produit d'épissage d'environ 540 nt apparaît dans le cas de l'ARN WT. D'après sa taille, il est épissé à la fois au niveau du premier intron et au niveau du second intron, ceci avec utilisation du site 5'SS lamine A. En accord avec cette identification, après 120 min d'incubation, une part importante de l'ARNm seulement épissé au niveau du premier intron est convertie en ARNm doublement épissé avec utilisation du site 5'SS lamine A.

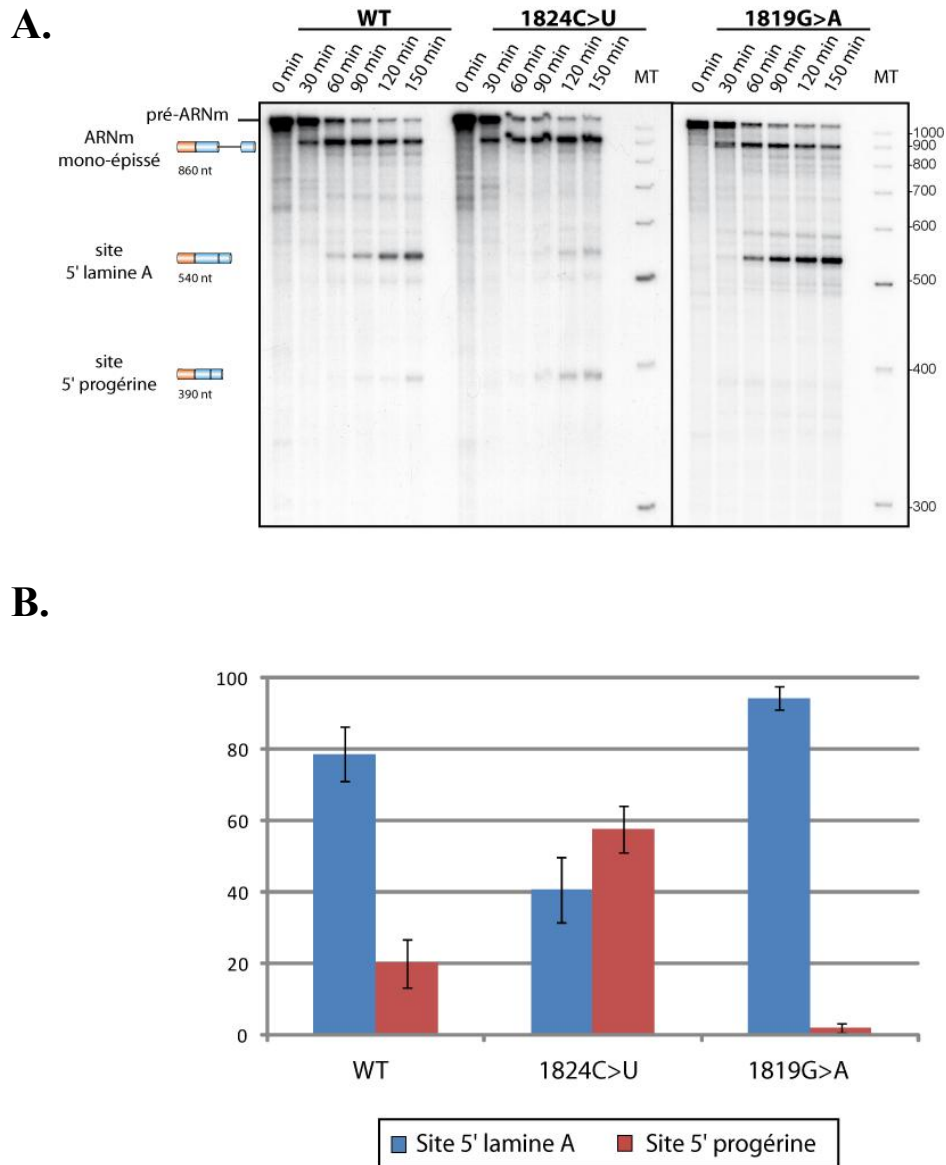


Figure III-2 : Cinétiques d'épissage *in vitro* de l'ARN β Glo-LMNA **A.** Fractionnement en gel de polyacrylamide dénaturant à 5% des produits ARN issus de l'épissage des pré-ARNm β Glo-LMNA WT, 1824C>U et 1819G>A. La réaction d'épissage a été réalisée à 30°C pendant 0 ; 30 ; 60 ; 90 ; 120 ou 150 min dans un extrait nucléaire de cellules HeLa contenant 1,2 mM d'ATP. Les différents produits formés sont indiqués à gauche du gel, ainsi que l'identité du site 5' de l'exon LMNA 11 utilisé. Le poids moléculaire des ADN utilisés comme marqueur de taille (MT) sont indiqués à droite. **B.** Quantifications des différents produits d'épissage formés après 120 min de réaction, montrant les quantités relatives (en %) de chacun des produits doublement épissés par rapport à la quantité totale de produits résultant d'un double épissage. Cela reflète le taux relatif d'utilisation de chacun des deux sites 5' de l'exon 11 dans les conditions d'épissage utilisées. L'expérience a été répétée plusieurs fois ($n>3$), les barres d'erreur représentent l'écart-type obtenu sur la moyenne des quantifications.

Au cours de l'incubation, on a donc épissage de l'intron 1, puis au sein de cet ARN épissé, épissage de l'intron 2. On note une faible bande d'une taille d'environ 670 nt qui pourrait correspondre à un ARNm dans lequel l'intron 2 aurait été éliminé sans élimination de l'intron 1, mais son taux reste très faible probablement du fait de l'élimination rapide de l'intron 1. Enfin, on note à partir de 30 min

d'incubation l'apparition d'une bande très faible qui reste faible même après 120 min d'incubation et qui correspond à un ARNm d'environ 390 nt, c'est-à-dire à un ARN épissé au niveau du premier intron et au niveau du second intron mais cette fois en utilisant le site 5'SS progérine (délétion de 150 nt). Conformément à ce qui était attendu, aucune trace d'ARNm pouvant résulter de l'utilisation du site 5'SS LAΔ35 (créé par la mutation 1868C>G retrouvée chez un patient) n'était détectée. Ces données montraient donc que le site 5'SS progérine peut être utilisé dans un extrait nucléaire, sans qu'il y ait de mutation, ce qui n'est pas le cas du site 5'SS LAΔ35.

En accord avec les données obtenues chez les patients, la substitution 1824C>U conduit à une utilisation renforcée du site 5'SS progérine par rapport au site 5'SS lamine A (figure III-2). A l'inverse, non seulement la substitution 1819G>A abolissait comme attendu l'utilisation du site 5'SS progérine mais en plus, elle renforçait fortement l'utilisation du site 5'SS lamine A, conduisant après 120 et 150 min d'incubation à un taux d'ARN doublement épissé plus élevé que dans le cas de l'ARN sauvage (figure III-2).

Nous disposions donc d'un test *in vitro* permettant d'étudier les effets de la structuration de la région du pré-ARNm LMNA contenant les sites 5'SS lamine A et progérine, ainsi que l'effet des protéines SR *in vitro* sur l'efficacité d'utilisation de ces sites. Par ailleurs, nous confirmons clairement *in vitro* la différence entre le site 5' SS progérine et le site 5'SS lamine AΔ35 à savoir que le premier est un site alternatif, alors que le second n'est utilisé qu'après mutation 1868C>G.

III.1.3. Etude de la structure secondaire en solution d'une grande région du transcrit LMNA contenant les sites 5'SS lamine A et progérine

III.1.3.a. Stratégie d'analyse de la structure secondaire

Dans un premier temps, afin de tester si la présence de la mutation 1824C>U modifiait la structure secondaire du pré-ARNm au niveau de l'un des deux sites d'épissage, nous avons réalisé une construction génétique permettant de produire un grand fragment (638 nt) du pré-ARNm LMNA (LMNA 638) sous contrôle de l'ARN polymérase du phage T7. Cet ARN comprend l'exon 11 entier, l'intron 11 entier et la partie de l'exon 12 située en amont du codon de terminaison de la traduction. Ce fragment contenait donc les deux sites 5' SS alternatifs de progérine et de la lamine A, ainsi que le site 3' d'épissage auquel ils peuvent être couplés. J'ai produit des versions mutées de cette construction permettant d'exprimer des fragments d'ARN équivalents avec soit la mutation 1824C>U, soit la mutation 1822 G>A.

Ces différents transcrits ont été renaturés après transcription selon le protocole décrit dans la partie Matériels et Méthodes. Puis ils ont été soumis à des hydrolyses ménagées par les RNAses T1, T2 et V1 dans le tampon utilisé pour la réalisation d'épissage *in vitro* en extrait nucléaire (tampon D).

- La RNase T1 hydrolyse spécifiquement les ARN en 3' des résidus guanine et préférentiellement dans les segments non appariés en conditions limitantes (figure III-3).

- La RNase T2 hydrolyse aussi préférentiellement les régions en simple brin. Cette enzyme présente peu de spécificité vis-à-vis des nucléotides encadrant la liaison phosphodiester clivée et libre, comme la RNase T1, des fragments ARN possédant des extrémités 3' phosphate et 5'-OH (figure III-3).

- La RNase V1 hydrolyse préférentiellement les régions en double-brin et les régions structurées par empilement de bases lorsqu'elle est en conditions limitantes, ceci avec peu de spécificité vis-à-vis de l'identité des résidus. Cette enzyme libère des fragments avec des extrémités 3'-OH et 5'-phosphate (figure III-3).

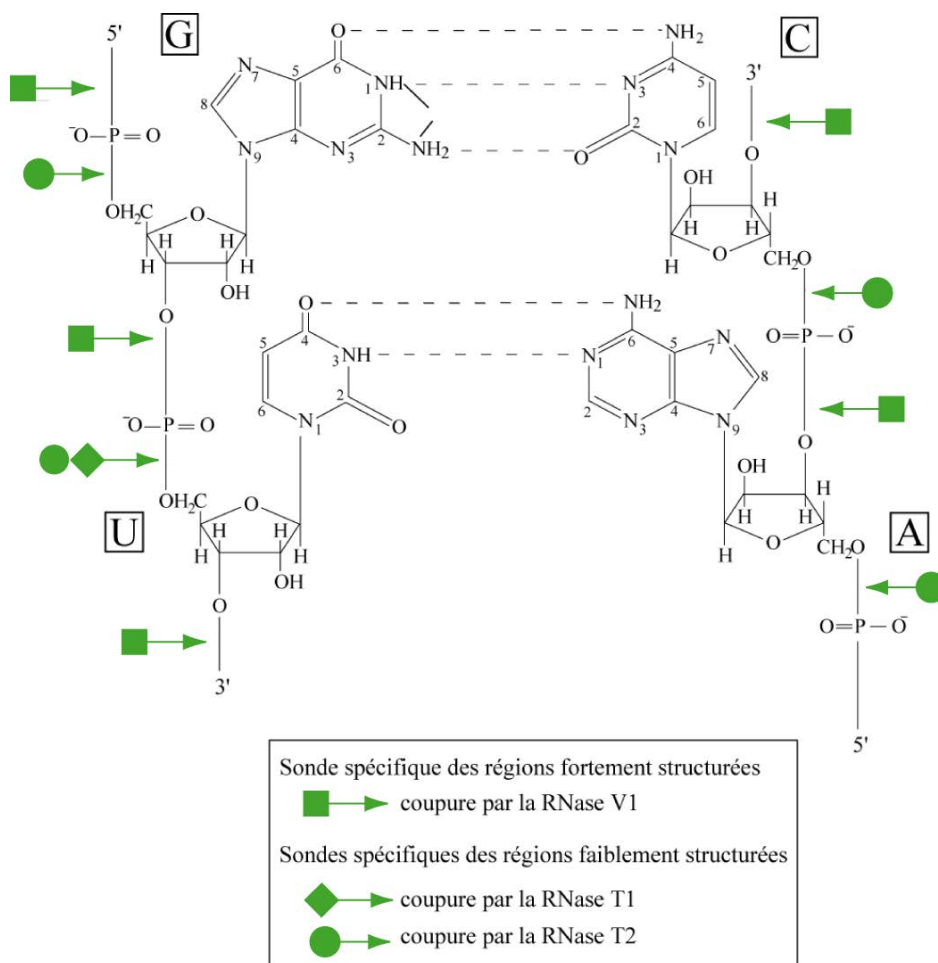


Figure III-3 : Représentation schématique des sites de clivages par les sondes enzymatiques RNases T1, T2, et V1 au sein des ARN.

Pour que l'analyse de structure secondaire en solution soit optimale, il faut que le nombre de clivages par molécule réalisés par ces enzymes soit limité mais que statistiquement, au sein du mélange de molécules d'ARN, des clivages aient lieu dans toute la région d'ARN étudiée. Des conditions

d'utilisation statistiques de ces enzymes avaient été mises au point depuis longtemps au laboratoire. Néanmoins, une adaptation est nécessaire pour chaque nouvel ARN étudié. J'ai donc dû adapter ces conditions à l'analyse de l'ARN LMNA 638. Pour cela chacune des RNases a été utilisée à trois concentrations différentes et la position des clivages a été identifiée par rétrotranscription des produits d'hydrolyse en utilisant des oligonucléotides amorce radiomarqués. Les ADNc produits ont été fractionnés par électrophorèse en conditions dénaturantes, en parallèle avec les produits d'une réaction de séquençage réalisée avec le même oligonucléotide amorce. Une réaction d'extension d'amorce a également été réalisée sur un ARN non hydrolysé afin d'identifier les pauses naturelles de la transcriptase inverse. La nature de l'oligonucléotide, la quantité d'ARN utilisée et les conditions d'hybridation ont été optimisées et deux oligonucléotides donnant des profils ADNc satisfaisants, 5478 et 5480, ont été retenus. L'oligonucléotide 5478 s'hybride des positions positions 320 à 300 de l'ARN LMNA 638. Cela correspond aux positions 1968+50 à 1968+30 selon la numérotation de séquence de l'ARNm de la lamine A chez l'homme, adoptée dans la figure 2, p.4 de la publication n°1 (1968 est la position du site 5' SS lamine A dans l'ARNm, les positions suivantes sont situées dans l'intron 11). Il a permis l'analyse de la région contenant le site 5' SS lamine A, tandis que l'oligonucléotide 5480 qui s'hybride des positions 213 à 189 de l'ARN LMNA 638 (soit 1911 à 1887 sur l'ARNm de la lamine A) a été utilisé pour l'analyse de la région contenant le site 5' SS progérine. L'utilisation de ces 2 oligonucléotides a permis d'analyser la structure secondaire de la région allant des positions 1 à 340 de l'ARN LMNA 638 (soit 1699 à 1968+64 du pré-ARNm de la lamine A). Les expériences ont été répétées trois fois sur des échantillons d'ARN différents.

III.1.3.b. Structure secondaire adoptée par l'ARN et influence des mutations sur cette structure

Des résultats reproductibles ont été obtenus, permettant de distinguer clairement des régions en simple brin et des régions en double brin. Les données obtenues ont ensuite été utilisées pour réaliser une prédiction de la structure 2D de la région étudiée contenant le site 5'SS lamine A et le site 5'SS progérine. Pour cela, j'ai utilisé le logiciel M-Fold (<http://mfold.rna.albany.edu/?q=mfold>) et introduit l'ensemble des données d'hydrolyses enzymatiques en tant que contraintes. Pour certains des segments, un accord particulièrement satisfaisant a pu être obtenu entre la structuration ayant la meilleure stabilité thermodynamique calculée et les données d'accessibilité, notamment au niveau de la région contenant le site 5'SS progérine. Ces données ont révélé que ce site 5' est situé dans une structure tige-boucle (structure tige-boucle IV) comportant une boucle terminale peu sensible à l'action des RNases, et que la région complémentaire au snRNA U1 est emprisonnée dans l'hélice (Figure 2, p.4 de la publication n°1). A l'inverse, le site 5'SS lamine A est situé dans un segment en simple-brin fortement accessible aux nucléases et ce dans toutes les prédictions de structure réalisées. Ces résultats suggèrent donc très

fortement l'existence d'un contexte structural défavorable pour l'utilisation du site 5' SS progérine au sein du pré-ARNm LMNA sauvage en comparaison du site 5' SS lamine A très accessible.

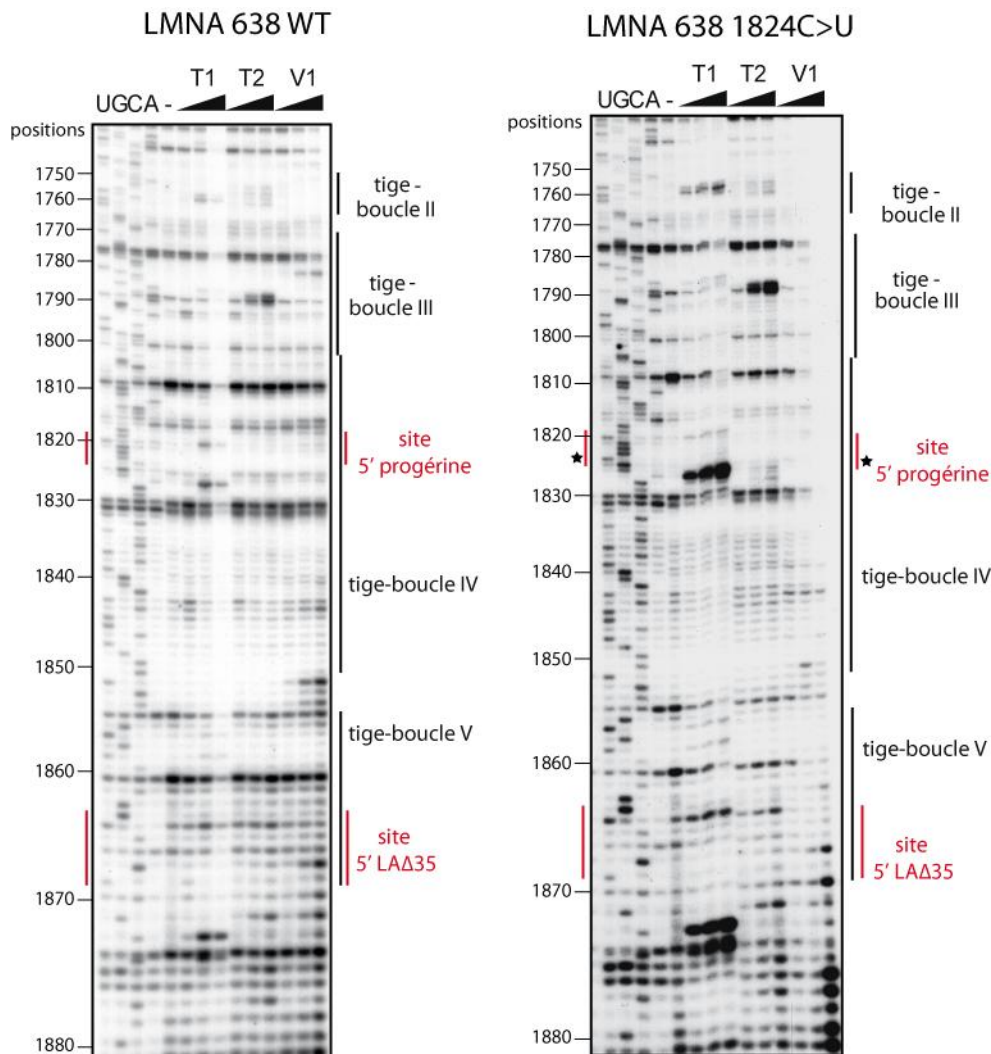


Figure III-4 : Comparaison des profils d'hydrolyses enzymatiques de la région contenant le site 5' SS progérine des ARN LMNA 638 WT et 1824C>U. Après renaturation dans les conditions décrites dans la partie Matériel et Méthodes, l'ARN a été soumis à des hydrolyses ménagées par les RNases T1, T2 et V1. Les RNases T1 et T2 ont été utilisées à des concentrations croissantes allant de 0,1 à 1 U et la RNase V1 à des concentrations croissantes allant de 10^{-4} à 10^{-3} U. Les positions des coupures ont été identifiées par extension de l'amorce 5480 radiomarquée en 5'. Un séquençage du transcrit a été réalisé avec la même amorce (pistes U,G,C,A). La colonne marquée (-) correspond à une extension d'amorce contrôle réalisée sur un ARN incubé en absence d'enzyme. Les différentes structures secondaires identifiées dans la publication n°1 sont annotées sur la partie droite de chaque gel. La même numérotation correspondant à l'ARNm LMNA adoptée dans la publication n°1 et est indiquée à gauche. Le site 5' SS progérine est indiqué en rouge, et la mutation 1824C>U est indiquée par une étoile.

J'ai répété les mêmes analyses structurales sur l'ARN LMNA 638 portant la substitution C en U en position 126, correspondant à la mutation 1824C>U trouvée chez les patients. D'une manière générale, nous n'avons pas observé de changements profonds des sites d'action des différentes RNases utilisées. Les variations observées étaient des variations des taux relatifs des différentes coupures, suivant

l'ARN utilisé. Nous détectons une augmentation de l'accessibilité aussi bien à la RNase V1, qu'aux RNases T1 et T2 du segment 1778 à 1883 comportant le site 5' SS progérine (figure III-4). Ceci suggérerait fortement un changement de compaction de cette région de l'ARN par disparition d'interactions de type tertiaire.

Nous avons noté tout particulièrement une augmentation de la sensibilité à l'hydrolyse par les RNases T1 et T2 de la boucle terminale de la structure tige-boucle IV, qui contient le site 5'SS de la progérine. I. Behm-Ansmant a tiré les mêmes conclusions en comparant l'accessibilité de l'ARN LMNA 638 WT et mutés à des sondes chimiques n'agissant que sur les segments accessibles et en simple brin des ARN. L'augmentation de l'accessibilité aux RNases et aux agents chimiques de la boucle terminale de la structure tige-boucle IV pouvait refléter soit la destruction d'une structuration interne de cette boucle, soit la disparition d'interactions tertiaires qu'elle contracte avec d'autres segments de l'ARN LMNA 638.

Pour trancher entre ces deux hypothèses, après publication de l'article dans *Human Molecular Genetics*, j'ai étudié si la variation d'accessibilité de la boucle terminale de la structure tige-boucle IV en présence de la mutation 1824C>U était détectable dans des sous-fragments de l'ARN LMNA 638. Pour cela, j'ai produit par transcription *in vitro* des ARN LMNA 48 WT et LMNA 48 portants la mutation 1824C>U correspondant à la structure tige boucle IV (positions 1804 à 1852) (figure III-5). J'ai également produit dans le même but des ARN LMNA 103 WT ou 103 porteur de la mutation 1824C>U correspondant aux tiges boucles III, IV et V (positions 1766 à 1869). J'ai analysé leur structure comme ci-dessus pour l'ARN LMNA 638, en employant des sondes enzymatiques. Contrairement à ce qui est observé pour l'ARN LMNA 638, la substitution 1824C>U ne modifie pas l'accessibilité de la boucle terminale de la structure tige-boucle IV, aussi bien pour l'ARN LMNA 48 que pour l'ARN LMNA 103 (figure III-5). Ceci peut s'expliquer par le fait qu'elle présente déjà une forte accessibilité dans les deux ARN WT. Dans les ARN LMNA 48 et 103 WT les sites de coupures observés sont en parfait accord avec la formation de la structure tige-boucle IV mais, en comparaison avec le profil d'hydrolyse obtenu pour l'ARN LMNA 638 WT, on note à la fois une plus forte accessibilité à la RNase T2 des résidus 1824 à 1827 dans la boucle apicale et 1836 à 1839 dans la boucle latérale. Dans l'ARN LMNA 48 1824C>U, les boucles, latérale comme apicale, présentent un niveau d'accessibilité à la RNase T2 très similaire à celui observé dans l'ARN WT.

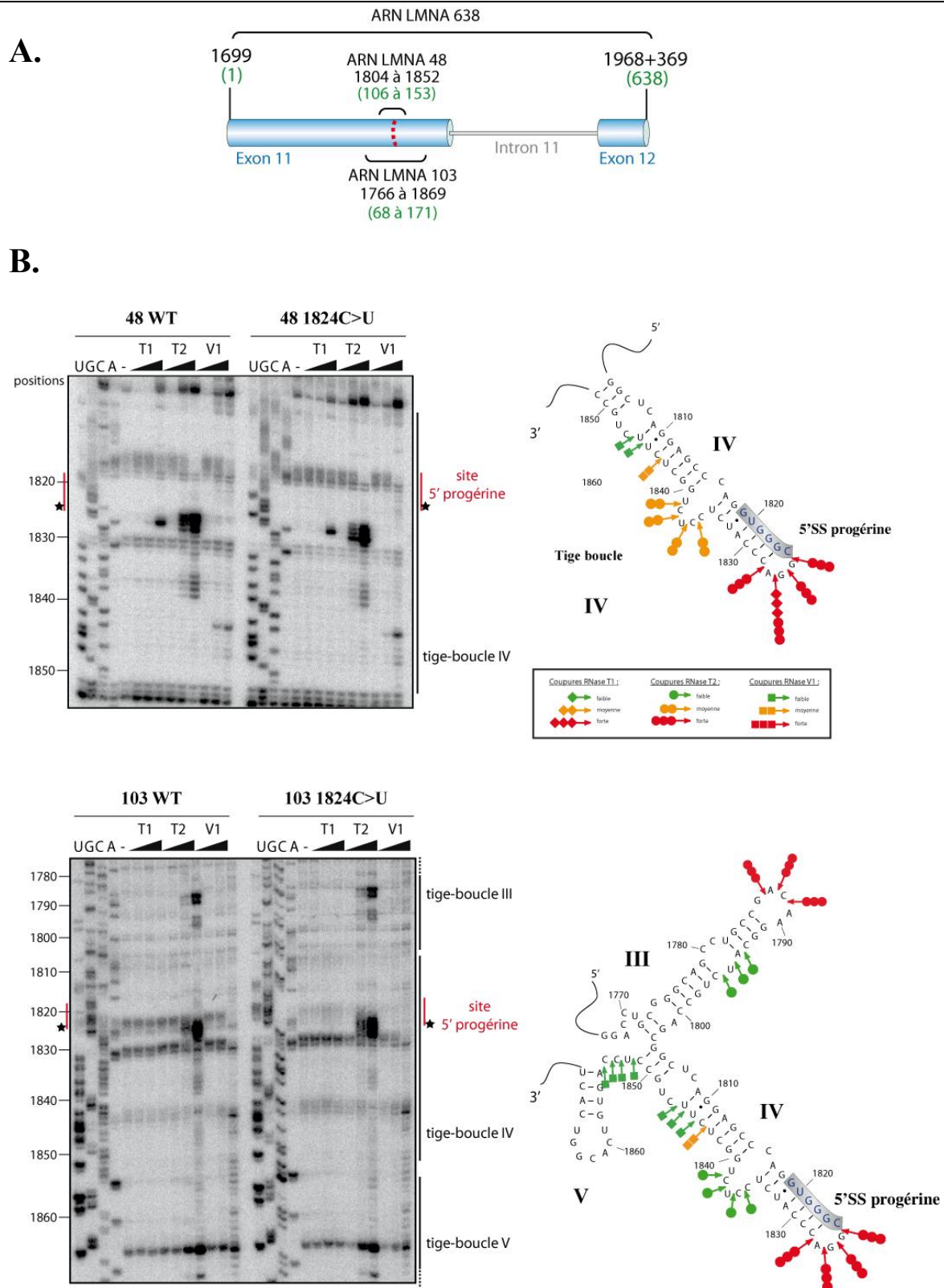


Figure III-5 : La mutation 1824C>U n'a pas ou peu d'effet sur l'accessibilité du site 5'SS progéline dans les ARN LMNA 48 et 103. A. Situation des ARN LMNA 48 et 103 sur l'ARN LMNA 638. Le site 5'SS progéline est indiqué par des pointillés rouges. Les positions indiquées sont celles de l'ARNm LMNA 638 adoptées dans la publication n°1 (en noir), ou celles dans l'ARN LMNA 638 (en vert). **B.** Analyse de la structure secondaire des ARN LMNA 48 et 103 WT ou en présence de la mutation 1824C>U, réalisée dans les mêmes conditions que pour l'ARN LMNA 638 précédemment.

Seules des variations d'intensité relatives des clivages au niveau des résidus de la boucle apicale sont détectées, pouvant être liées au changement de la séquence de la boucle. Des résultats identiques sont observés avec l'ARN LMNA 103, pour lequel les hydrolyses enzymatiques concordent également parfaitement avec la structure modélisée pour l'ARN LMNA 638 (figure III-5). Nous en avons donc conclu que la faible accessibilité de la boucle terminale de la structure tige-boucle IV observée dans l'ARN LMNA 638 devait être due à l'existence d'une interaction tertiaire de cette structure tige-boucle avec une partie du fragment 638 absente dans le fragment et la partie apicale de la structure tige-boucle IV contenant le site 5'SS progérine. Cette interaction tertiaire doit nécessiter la présence du résidu C1824 puisqu'elle semble disparaître lorsqu'il est muté en résidu U. Elle doit aussi nécessiter la présence du résidu G1822, puisque nous avons aussi constaté que la mutation 1822G>A a le même effet sur la structuration de la partie apicale de la structure tige-boucle IV que la mutation 1824C>U. Elle renforce de la même manière l'accessibilité à la RNase T2 de la boucle apicale (données supplémentaires de la publication n°1, figure 2). Cette augmentation de l'accessibilité de la boucle apicale de la structure tige-boucle IV accompagne le renforcement de l'utilisation du site 5' SS progérine situé dans cette boucle, suggérant fortement un rôle de la structure secondaire dans l'utilisation de ce site.

Pour apporter un argument supplémentaire en faveur de cette hypothèse, j'ai produit différentes mutations dans la boucle apicale de la structure tige-boucle IV et j'ai pu montrer un lien direct entre les mutations augmentant l'accessibilité de cette boucle au sein de l'ARN LMNA 638 et leur capacité à renforcer l'utilisation du site 5'SS progérine. Aussi, comme nous l'avons décrit dans l'article, la mutation 1825G>A qui provoque une augmentation de l'accessibilité de la boucle apicale de la structure tige-boucle IV, renforce l'utilisation du site 5'SS progérine dans des tests d'épissage *in vitro* sans que la complémentarité avec le snRNA U1 soit renforcée. Ainsi, les deux mutations 1824C>U et 1822G>A auraient deux effets bien distincts, à la fois améliorer la complémentarité du site 5'SS progérine avec le snRNA U1, et augmenter l'accessibilité de ce site, ce qui favoriserait la fixation de la snRNP U1.

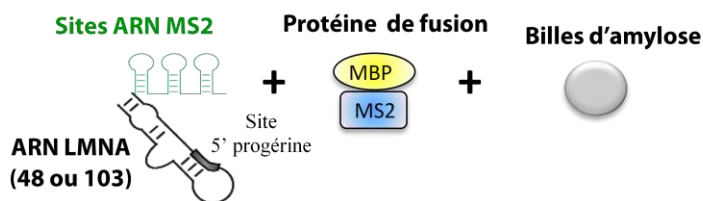
III.1.4. Etude *in vitro* de l'effet des protéines SR sur l'activité des sites 5'SS lamine A et 5'SS progérine

III.1.4.a. Etude de l'association des protéines contenues dans un extrait nucléaire de cellules HeLa avec le site 5'SS progérine

Avant mon arrivée au laboratoire, I. Behm-Ansmant et F. Freyermuth (stagiaire M1) avaient entrepris de purifier les complexes formés entre des fragments du pré-ARNm LMNA et les protéines d'un extrait nucléaire de cellules HeLa compétent pour l'épissage. Pour cette étude, les ARN LMNA 48 et 103 précédemment décrits, WT ou porteurs de la mutation 1824C>U, ont été utilisés (voir aussi figure III-5, A). Chacun de ces 4 ARN avait été produit par transcription *in vitro* sous forme fusionnée avec 3 sites de fixation de la protéine de l'enveloppe du phage MS2. Le domaine de fixation à l'ARN de cette

protéine d'enveloppe avait été produit sous forme de fusion avec la Maltose Binding Protein (MBP) afin de retenir la protéine de fusion MBP/MS2 et par voie de conséquence les ARN de fusion associés à cette protéine sur des billes d'amylose (figure III-6). Des complexes ribonucléoprotéiques avaient ensuite été formés par incubation de l'ARN fixé sur les billes avec un extrait nucléaire de cellules HeLa. Après lavages les complexes retenus avaient été élués par ajout de maltose. Environ 50% des protéines contenues dans ces complexes avaient été fractionnées par électrophorèse en gel SDS-PAGE et après coloration au bleu colloïdal, les bandes du gel avaient été découpées et analysées par spectrométrie de masse par l'équipe d'A. Van Dorssealer à Strasbourg.

A.



B.

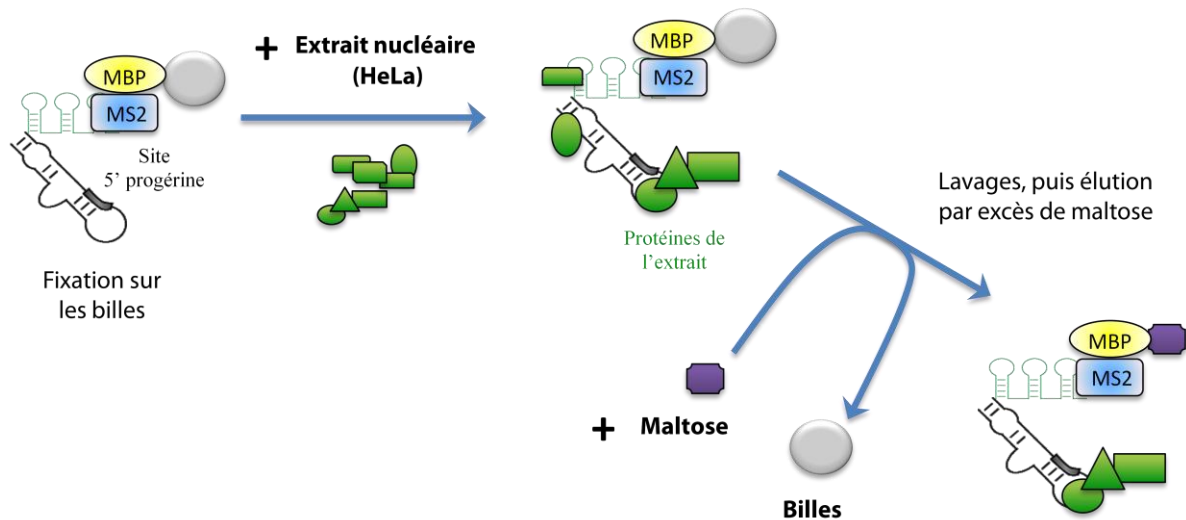


Figure III-6 : Principe de la purification des complexes ribonucléoprotéiques par chromatographie d'affinité basée sur l'emploi de la protéine de l'enveloppe du phage MS2 et de la protéine MBP. A. L'ARN LMNA d'intérêt fusionné co-transcriptionnellement à 3 motifs MS2 est incubé en présence de la protéine de fusion MS2/MBP (Protéine de l'enveloppe du phage MS2/Maltose Binding Protein), puis retenu sur des billes d'amylose **B.** Les fragments du pré-ARNm LMNA fixés sur les billes d'amylose sont ensuite incubés avec un extrait nucléaire de cellules HeLa. Les protéines n'interagissant pas de manière spécifiques sont éliminées du culot de billes par des lavages, et les protéines partenaires de l'ARN sont ensuite récupérées par élution par un excès de maltose entrant en compétition avec la protéine de fusion MBP-MS2. Environ 50% des protéines contenues dans ces complexes sont fractionnées par électrophorèse en gel SDS-PAGE. Après coloration au bleu colloïdal, les bandes du gel sont découpées et analysées par spectrométrie de masse.

Un ARN ne contenant que les sites de fixation de la protéine de l'enveloppe du phage MS2 (ARN MS2) avait également été utilisé en tant que contrôle négatif. L'expérience de purification de complexes avait été réalisée à 2 reprises et avait permis l'identification d'un nombre considérable de protéines associées aux différents ARN étudiés. Néanmoins, un handicap majeur de ce type d'analyse était que les résultats n'étaient pas quantitatifs. De ce fait, des différences de taux de fixation d'une protéine sur deux ARN distincts ne pouvaient pas être appréciées et les données obtenues ne permettaient pas de détecter de différences marquantes entre les protéines associées aux ARN sauvages et aux ARN mutés. C'est pourquoi j'ai réalisé des expériences de Western Blot à l'aide d'anticorps spécifiques dirigés contre les protéines identifiées. Ces expériences ont été réalisées sur les fractions protéiques isolées à partir des complexes purifiés, afin de détecter des variations de taux de fixation des protéines au sein des complexes. Parmi les nombreuses protéines identifiées par spectrométrie de masse, nous avons examiné des protéines connues pour avoir un lien avec les régulations de l'épissage, et tout particulièrement les protéines SR et hnRNP. Les résultats obtenus sont brièvement décrits dans la figure 3 des données supplémentaires de la publication n°1. Ils ont permis notamment de mettre en évidence l'association des protéines SRSF1 et SRSF6 comme se fixant sur des fragments du pré-ARNm LMNA contenant le site 5' progérine et d'observer une association moins forte de ces protéines en présence de la mutation 1824C>U. Cette observation nous a encouragé à étudier l'influence de ces protéines sur la régulation de l'épissage du pré-ARNm LMNA.

III.1.4.c. Test de l'effet des protéines SRSF1 et SRSF6 sur l'utilisation des sites 5' progérine et lamine A in vitro

Ayant mis au point des conditions d'analyse par épissage *in vitro* permettant d'étudier de façon satisfaisante l'utilisation relative des sites 5'SS lamine A et progérine, j'ai alors testé l'influence de l'ajout des protéines SR SRSF1 et SRSF6 sur l'épissage *in vitro* des ARN *βglo*-LMNA WT et *βglo*-LMNA 1824C>U. Les protéines SRSF1 et SRSF6 recombinantes nécessaires à ces expériences ont été produites et purifiées à partir de baculovirus et nous ont été fournies par les équipes de J. Stévenin (SRSF1) et J. Tazi (SRSF6). Ces protéines SR purifiées sont phosphorylées et actives pour l'épissage. Nous avons observé que l'ajout des protéines SRSF1 ou SRSF6 recombinantes (200 ou 400 ng) dans le milieu réactionnel provoquait une diminution drastique de l'utilisation du site 5'SS progérine au profit du site 5'SS lamine A, avec un effet moindre toutefois de SRSF1 (Figure 4D, publication n°1). Afin de déterminer si l'impact des protéines SRSF1 et SRSF6 sur l'utilisation du site 5'SS progérine était dû à une inhibition de l'utilisation de ce site 5' ou une activation du site 5'SS lamine A, j'ai répété ces expériences en utilisant l'ARN *βglo*-LMNA 1969G>A dans lequel le site 5'SS lamine A était inactivé par la mutation 1969G>A. Dans ce cas seule l'influence inhibitrice de SRSF6 sur l'utilisation du site 5'SS progérine était toujours observée, y compris en présence de la mutation 1824C>U. L'effet de SRSF1, en revanche, était moindre, voir quasiment inexistant sur le double mutant 1969G>A + 1824C>U (Figure

4E, publication n°1). Cela suggérait que SRSF6 avait une réelle activité inhibitrice vis-à-vis de l'utilisation du site 5'SS progérine, tandis que la protéine SRSF1 agissait davantage en activant l'utilisation du site 5'SS lamine A.

Les données que j'avais obtenues *in vitro* ont été complétées par des tests d'épissage *in cellulo*, qui ont été réalisés par I. Lopez-Mejia dans l'équipe J. Tazi.

Afin de tester l'effet répresseur de SRSF6 sur le site 5' SS progérine et l'action possible de SRSF1 elle a traité des cellules HeLa transfectées par la construction β Glo-LMNA WT ou mutée par des siRNA dirigés contre SRSF6 ou contre SRSF1. Dans ces expériences, alors que la diminution du taux de SRSF6 intracellulaire augmentait l'utilisation du site 5'SS progérine, la diminution du taux de SRSF1 avait un effet opposé. I. Lopez-Mejia a réalisé le même type d'expérience dans des fibroblastes de patients HGPS. Dans ces cellules, la diminution du taux de SRSF6 augmentait aussi de façon marquée l'utilisation du site 5'SS progérine. Alors que la diminution du taux de SRSF1 conduisait à une légère augmentation de l'utilisation du site 5'SS lamine A, le taux d'utilisation du site 5'SS progérine étant inchangé.

Ces données et nos résultats ont conduit à proposer le modèle suivant : en absence de la mutation 1824C>U, le site 5'SS progérine est localisé dans une région très structurée du pré-ARNm, il est peu accessible et sa séquence de fixation au snRNA U1 est sous-optimale. Par ailleurs, son activité est inhibée par la protéine SRSF6. En se liant au voisinage de ce site, cette protéine pourrait renforcer la stabilité de la structure dans laquelle ce site est impliquée. Des expériences complémentaires, comme des expériences d'empreinte ou d'étude de la structure par RMN, seront nécessaires pour tester cette hypothèse.

La présence d'un site 5'SS progérine en compétition avec le site 5'SS lamine A n'est pas une particularité du gène *LMNA* humain. Cette organisation avec deux sites en compétition est conservée chez les vertébrés en particulier chez les souris. L'une des différences majeures en terme de séquence entre les pré-ARNm *LMNA* humain et murin au niveau du site 5'SS lamine A est la présence en position 1828 d'une substitution C>U. Afin de tester si la structure secondaire de l'ARN pouvait aussi influencer le taux d'utilisation du site 5'SS progérine chez la souris, j'ai produit un ARN *LMNA* 638 porteur de cette substitution 1828C>U. L'analyse de la structure secondaire de cet ARN m'a permis de montrer qu'il formait bien une structure tige-boucle similaire à la structure tige-boucle IV de l'ARN humain. La boucle terminale de cette structure tige-boucle était comme chez l'homme peu accessible mais la mutation 1824C>U augmentait son accessibilité.

Les travaux réalisés par l'équipe de J. Tazi sur des cellules MEFs isolées à partir du modèle progéria de souris développé par N. Levy et J. Tazi ont permis de montrer que l'effet de la protéine SRSF6 est également conservé. Cette conservation renforçait encore l'idée d'une importance fonctionnelle marquante de la présence du couple des sites 5'SS progérine et 5'SS lamine A et de leur

régulation. Par ailleurs l'équipe de J. Tazi sur la base de la comparaison des modèles de souris hétérozygote et homozygote pour la mutation équivalente à la mutation 1824C>T a pu montrer une anti-corrélation entre le nombre d'allèles portant la mutation et l'espérance de vie des souris. Ainsi les souris homozygotes pour cette mutation vivaient moins longtemps que les souris hétérozygotes pour cette mutation qui elles même avaient une espérance de vie plus courte que des souris contrôles. Ces données suggèrent que la présence du couple des sites 5'SS progérine et 5'SS lamine A puisse participer à la définition de la durée de vie d'un organisme.

L'ensemble de ces travaux est décrit dans l'article qui suit.

PUBLICATION N°1

A conserved splicing mechanism of the LMNA gene controls premature aging.

Isabel C. LOPEZ-MEJIA, Valentin VAUTROT, Marion DE TOLEDO, Isabelle BEHM-ANSMANT, Cyril F. BOURGEOIS, Claire L. NAVARRO, Fernando G. Osorio, Jose M. P. FREIJE, James STEVENIN, Annachiara DE SANDRE-GIOVANNOLI, Carlos LOPEZ-OTIN, Nicolas LEVY, Christiane BRANLANT and Jamal TAZI.

Human Molecular Genetics, **2011**, Vol. 20, No. 23, **4540-4555**.

doi:10.1093/hmg/ddr385

A conserved splicing mechanism of the *LMNA* gene controls premature aging

Isabel C. Lopez-Mejia¹, Valentin Vautrot², Marion De Toledo¹, Isabelle Behm-Ansmant², Cyril F. Bourgeois³, Claire L. Navarro⁵, Fernando G. Osorio⁴, José M. P. Freije⁴, James Stévenin³, Annachiara De Sandre-Giovannoli⁵, Carlos Lopez-Otin⁴, Nicolas Lévy⁵, Christiane Branlant² and Jamal Tazi^{1,*}

¹CNRS, UMR 5535, University of Montpellier, Institut de Génétique Moléculaire de Montpellier, 1919 Route de Mende, Montpellier 34293, France, ²CNRS UMR 7567, Biopôle, Faculté de médecine, Université Henri Poincaré-Nancy I, 9 Avenue de la Forêt de Haye, 54506 Vandoeuvre-les-Nancy, France, ³IGBMC Department of Functional Genomics and Cancer, CNRS UMR 7104, INSERM U 964, University of Strasbourg, 1 Rue Laurent Fries, 67404 Illkirch Cedex, France, ⁴Departamento de Bioquímica y Biología Molecular, Facultad de Medicina, Instituto Universitario de Oncología, Universidad de Oviedo, 33006-Oviedo, Spain and ⁵INSERM UMR_S910, Faculté de Médecine, Université de la Méditerranée, Boulevard Jean Moulin, 13385 Marseille Cedex 05, France

Received June 28, 2011; Revised August 12, 2011; Accepted August 24, 2011

Hutchinson–Gilford progeria syndrome (HGPS) is a rare genetic disorder phenotypically characterized by many features of premature aging. Most cases of HGPS are due to a heterozygous silent mutation (c.1824C>T; p.Gly608Gly) that enhances the use of an internal 5' splice site (5'SS) in exon 11 of the *LMNA* pre-mRNA and leads to the production of a truncated protein (progerin) with a dominant negative effect. Here we show that HGPS mutation changes the accessibility of the 5'SS of *LMNA* exon 11 which is sequestered in a conserved RNA structure. Our results also reveal a regulatory role of a subset of serine–arginine (SR)-rich proteins, including serine–arginine rich splicing factor 1 (SRSF1) and SRSF6, on utilization of the 5'SS leading to lamin A or progerin production and a modulation of this regulation in the presence of the c.1824C>T mutation is shown directly on HGPS patient cells. Mutant mice carrying the equivalent mutation in the *LMNA* gene (c.1827C>T) also accumulate progerin and phenocopy the main cellular alterations and clinical defects of HGPS patients. RNAi-induced depletion of SRSF1 in the HGPS-like mouse embryonic fibroblasts (MEFs) allowed progerin reduction and dysmorphic nuclei phenotype correction, whereas SRSF6 depletion aggravated the HGPS-like MEF's phenotype. We demonstrate that changes in the splicing ratio between lamin A and progerin are key factors for lifespan since heterozygous mice harboring the mutation lived longer than homozygous littermates but less than the wild-type. Genetic and biochemical data together favor the view that physiological progerin production is under tight control of a conserved splicing mechanism to avoid precocious aging.

INTRODUCTION

Different proteins are produced by alternative splicing and or polyadenylation of the *LMNA* gene, including lamin A, lamin C and progerin (Fig. 1A) (1–5). Lamin A and lamin C are major components of the nuclear lamina, a complex molecular interface located between the inner membrane of the nuclear

envelope and chromatin (6). Lamins A and C are also distributed throughout the nucleoplasm and are involved in crucial functions, including DNA replication, transcription, chromatin organization, nuclear positioning and shape, as well as the assembly/disassembly of the nucleus during cell division (6–8).

Hutchinson–Gilford progeria syndrome (HGPS) is a dominant disease usually caused by heterozygous *de novo* mutations in the

*To whom correspondence should be addressed at: IGMM-CNRS, 1919 Route de Mende, 34293 Montpellier. Tel: +33 34359685; Fax: +33 34359634; Email: jamal.tazi@igmm.cnrs.fr

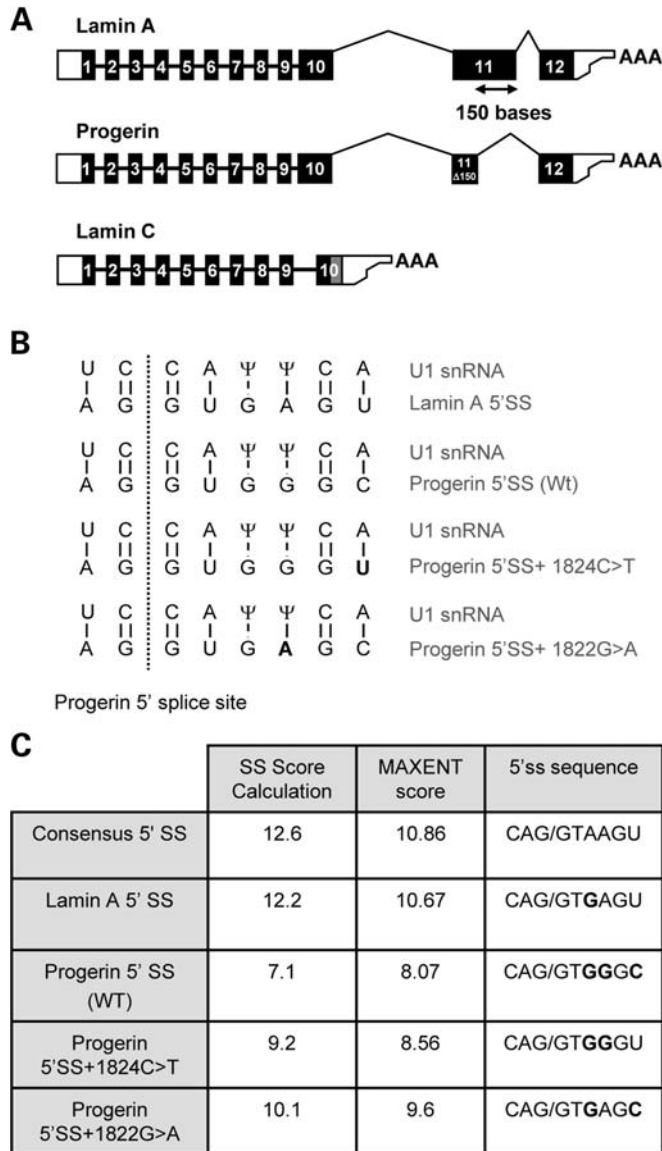


Figure 1. LMNA isoforms and properties of the progerin and lamin 5'SSs. (A) Three different proteins are produced from the *LMNA* gene. Cartoon showing the exon structure of the three different forms of *LMNA* mRNA showing the position of the alternative splice site used in progerin and alternative polyadenylation site that is specific for lamin C. Exons are numbered and shown as boxes (coding regions are solid). Introns are shown as lines. The site of the HGPS mutation is shown in bold. (B) The HGPS progeria mutations increase complementarity of the progerin splice site for U1 snRNA. The wild type and mutant progerin 5'SSs are shown based paired to U1 snRNA sequence. The mutation creates an AU base pair at the plus 6 position shown in bold. (C) Predicted bioinformatic scores of the *LMNA* splice sites. Bioinformatic scores for a consensus splice site (consensus 5'SS), lamin A 5'SS, the progerin 5'SS and the mutants c.1824C>T and c.1822G>A progerin 5'SS are shown using two different programs. Note the near-perfect score of the lamin 5'SS and the moderate improvement in the mutated progerin 5'SS+ 1824C>T compared with WT progerin 5'SS.

LMNA gene that leads to accumulation of a truncated protein, commonly called progerin (4,5). Progerin is hardly detectable in wild-type (WT) cells but a single nucleotide substitution c.1824C>T, recurrent in the majority of typical HGPS patients, promotes activation of a 5' splice site (5'SS) in exon 11, causing a large increase

in the levels of transcripts encoding progerin compared with the other *LMNA* isoforms. Due to a 150 nucleotides deletion from exon 11, progerin retains the C-terminal CAAX box of lamin A (encoded by exon 12), which undergoes methyl esterification and farnesylation, but it lacks 50 internal amino acids essential for a specific cleavage step involved in lamin A maturation (4–11). Progerin is therefore permanently farnesylated and carboxymethylated (12,13) which leads to its abnormal anchoring to the nuclear membrane throughout the cell cycle (13,14). The production of progerin interferes with the integrity of the nuclear lamina, causing misshapen nuclei, and also adversely affects other important cellular processes, such as interphase chromatin, mitosis and cell proliferation (4,5,10,11,14). The finding that progerin is expressed at detectable levels in healthy individuals and seems to accumulate with age has led to the proposal that progerin could be involved in physiological aging (10,15–17). However, to date very little is known about the underlying splicing mechanism leading to usage of the internal 5'SS in exon 11 and how this selection might alter lifespan.

Major players in both splice-site recognition and intron excision are five small nuclear RNAs (U1, U2, U4, U5 and U6) which have dynamic specific interactions via base pairing with the pre-mRNA (18), and notable among these are the base pairing between U1 snRNA and the 5'SS. However, the information contained in the nucleotide sequence in splice sites is not sufficient for regulation and it has been shown that the RNA–protein complexes which form on the pre-mRNA also help the recognition of exons (19). The majority of splicing regulatory proteins in these complexes belongs to the heterogeneous ribonucleoprotein (hnRNP) and serine–arginine (SR)-protein families. hnRNPs were initially defined as the major proteins binding to RNA polymerase II transcripts in the course of transcription. SR proteins containing arginine-serine rich domains play essential roles in the splicing reaction (19–22). hnRNPs and SR proteins contain both RNA-binding and protein–protein interaction domains. They bind with low specificity to accessible parts of the pre-mRNA and RNA secondary structures strongly influence their binding efficiency (23–27).

Here we use *in vitro* and *in vivo* approaches to explore the sharp differences in splice-site usage in progeria patients. By probing the conserved structure of the alternatively spliced RNA as well as the proteins that bind to it, we demonstrate that the c.1824C>T mutation increases the accessibility of the exon 11 internal 5'SS which should facilitate U1 snRNP binding and discover a previously undescribed repression by the SR protein serine–arginine rich splicing factor 6 (SRSF6), whereas SRSF1 has an opposite action. Thus RNA-binding proteins and RNA structure should cooperate with enhanced U1 complementarity in the regulation of production of the progerin. Additionally, mouse embryonic fibroblasts (MEFs) that harbor a synonymous C>T (GGC>GGT, Gly609Gly) single-base substitution of the mouse *Lmna* gene demonstrate that this splicing mechanism is conserved from the mouse to human and is determinant in the control of lifespan.

RESULTS

Conserved RNA structure at the progerin 5'SS

The synonymous c.1824 C>T (GGC>GGT, G608G) single-base substitution in exon 11 of the *LMNA* gene that potentiates

usage of the exonic 5'SS (hereafter referred to as the progerin 5'SS) is not expected to have a strong effect on splicing because it occurs at the final, highly variable (+6) position of the 5'SS consensus (25,26,28–30). Furthermore, computational analyses to score the strength of either the WT or the mutated (c.1824C>T) form of the progerin 5'SS in comparison with the downstream 'authentic' 5'SS that is used to produce lamin A (hereafter referred to as the lamin A 5'SS) revealed that the scores of the progerin 5'SS are lower than the lamin A 5'SS (Fig. 1A–C). While the c.1824C>T mutation augments the score of the progerin 5'SS, it is not sufficient to make it comparable to the lamin A 5'SS which has a close to perfect score (Fig. 1B and C). However, the mutation results in the production of more progerin mRNA than lamin A mRNA (see below), suggesting that the mutation likely enhances recognition of the progerin 5'SS by other mechanisms than simply by enhancing base pairing with U1 snRNP. Given the fact that RNA structure may be essential for the interaction of a target pre-mRNA and spliceosomal components, and thereby may influence splice-site selection (20,31,32), we decided to assay for disruption or occurrence of RNA structures surrounding the c.1824C>T mutation in exon 11 of the *LMNA* gene. To this end, radiolabelled RNAs were transcribed, denatured and refolded and then digested with different ribonucleases specific either for accessible single-stranded regions (T1 and T2 RNases) or for double-stranded or stacked RNA regions (V1 RNase). To avoid border effects, the sites of cleavage were analyzed by primer extension on a series of overlapping fragments of different sizes. The most stable secondary structure that fits the best with the experimental data was predicted with the M-fold software by introducing experimental data as constraints. In all RNA fragments studied, the segment from positions 1804–1851, which contains the progerin 5'SS was always folded into a compact 2-D structure (Fig. 2A) with the progerin 5'SS located in the stem. This part of stem-loop (SL) structure IV will now be denoted as the progerin stem-loop (progerin SL). Its terminal tetra-loop is expected to be structured because of its limited accessibility to RNases. In contrast, the lamin A 5'SS was found to be located in a highly accessible single-stranded region in all the RNA fragments studied (Fig. 2A and Supplementary Material, Fig. S1). Thus, in contrast to the lamin A 5'SS, recognition of the progerin 5'SS in the WT gene is normally disfavored due to its sequestration in a structured region.

Interestingly, when repeating the experiment on transcripts containing the c.1824C>U mutation, we found that the progeria mutation induces some opening of the progerin SL terminal loop which flanks the progerin 5'SS, as shown by the strong accessibility to T1 RNase (Fig. 2B, right panel). Although the mutation is not expected to affect the base pairing of the stem structure, the stability of the terminal G-C pair may be also slightly decreased based on the appearance of some T2 RNase cleavage 3' to residue C1828 (Fig. 2B, right panel). The overall conformational change of the progerin SL might therefore increase the accessibility of the progerin 5'SS to U1 snRNA and explain, at least in part, the increased usage of this site in HGPS.

To further analyze the unexpected accessibility of the loop of the progerin SL after the c.1824C>U mutation, we decided to assess two novel mutations, one located in the stem and

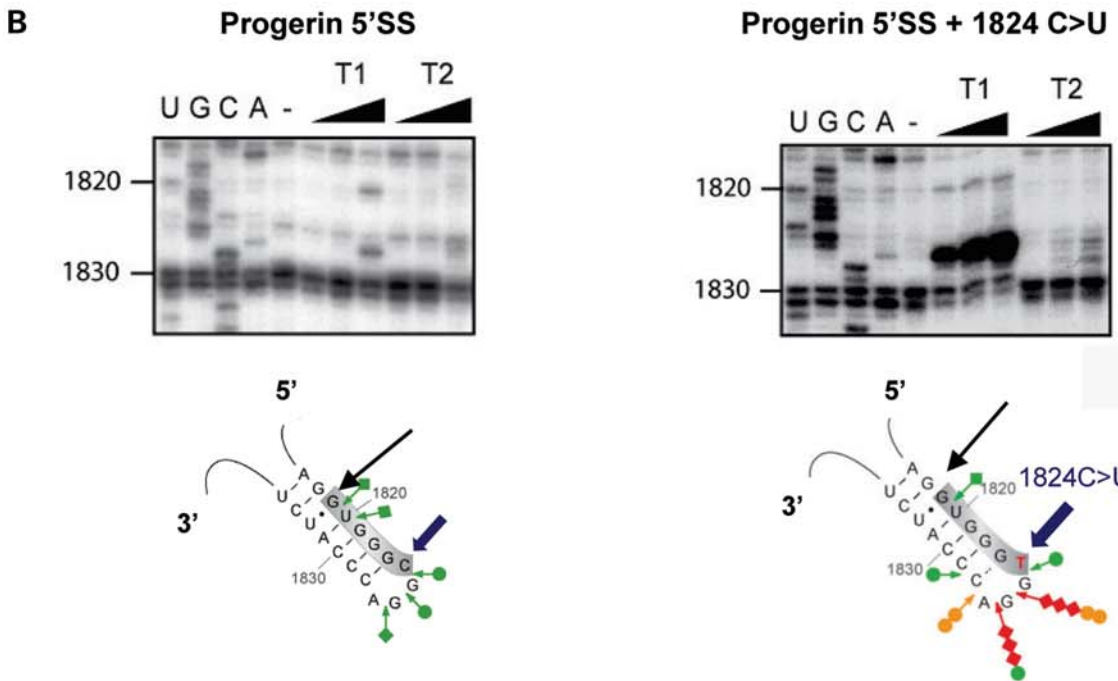
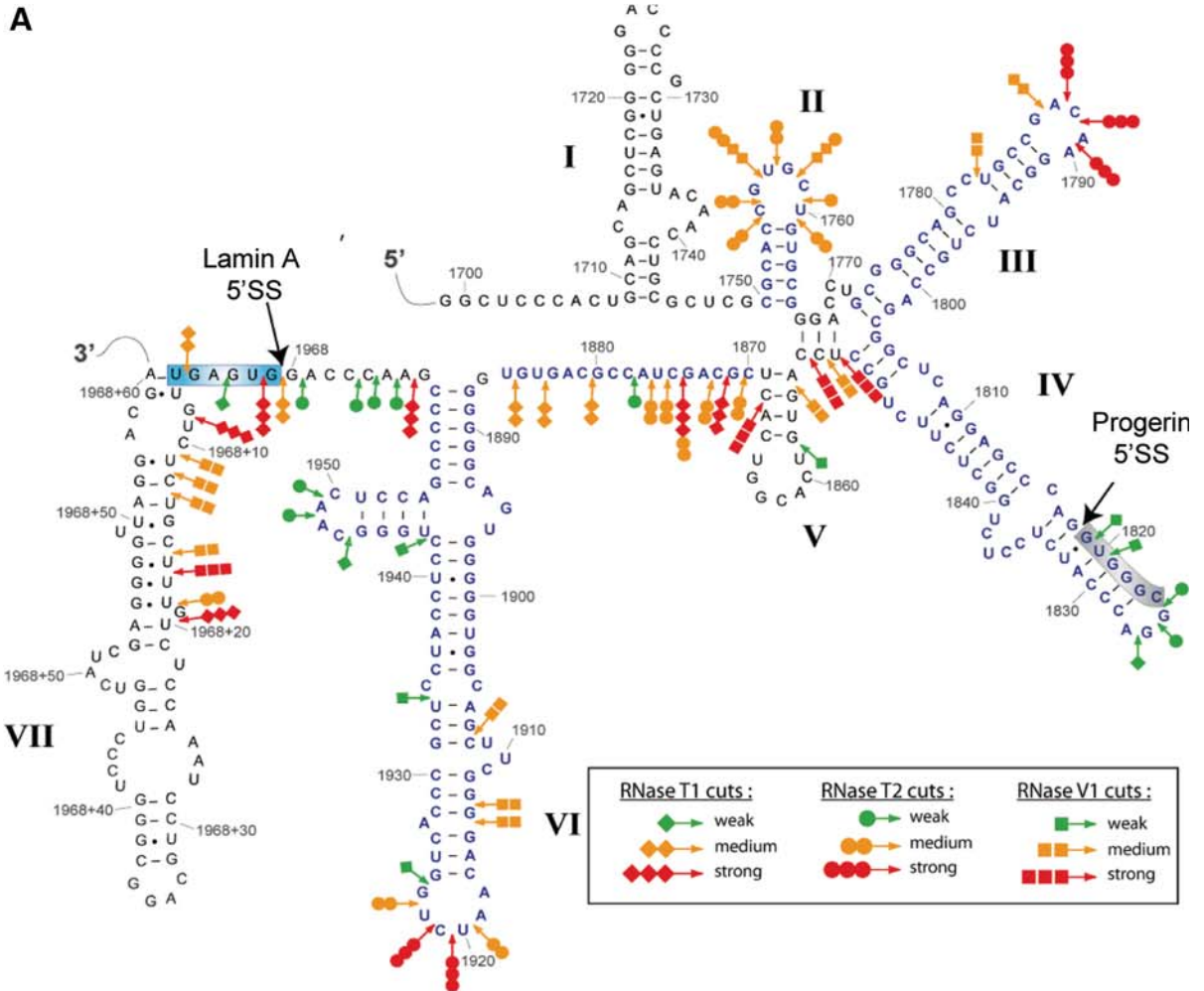
expected to alter RNA structure of the region and a second in the loop of the progerin SL, as the original mutation. The first is a HGPS mutation, which significantly increased the complementarity of the progerin 5'SS with U1 snRNA (Fig. 1), and was also found to increase utilization of the progerin 5'SS (4,5). By probing the structure of the mutated pre-mRNA region with T1 and T2 RNases, we observed an opening of the terminal loop in the progerin SL (Supplementary Material, Fig. S2a). The G>A mutation disrupts a G–C base pair close to the WT terminal loop and as a consequence several G residues are cleaved by T1 RNase within the opened terminal loop. A better accessibility of the 5'SS region, together with an increased complementarity to the U1 snRNA, may explain the strong utilization of the progerin 5'SS in the β -globin-LMNA reporter RNA incubated in HeLa nuclear extract (see below and Supplementary Material, Fig. S2b).

For the second mutation, we replaced the first residue of the loop, residue G1825 which is located 3' to the U1 binding site into an A residue and tested for the effect on both RNA structure and utilization of the progerin 5'SS (Supplementary Material, Fig. S2a and b). Interestingly, this mutation increased both the accessibility of the loop to RNases (especially RNase T2) and the utilization of the progerin 5'SS (Supplementary Material, Fig. S2a and b). Altogether with the mutation c.1824C>U data, these data show a possible correlation between the accessibility of the terminal loop of the progerin SL and the efficiency of the progerin 5'SS utilization.

Recognition of the progerin 5'SS by the SR proteins SRSF1 and SRSF6

Mutations responsible for splicing defects at the origin of human diseases are either found in splice sites, or in regulatory sequences involved in the control of splicing (25,28,33–35). Therefore, we used various approaches to investigate whether the c.1824C>T mutation might modify the binding of some splicing regulatory factors at or in the vicinity of the progerin 5'SS, in addition to its effect on the RNA structure.

First, we identified nuclear proteins able to bind to the progerin SL structure by fusing three MS2 coat protein RNA-binding sites at the 3' extremity of the progerin SL RNA (Supplementary Material, Fig. S3a) and purifying the complexes formed by incubation of the tagged RNA in HeLa nuclear extract. The complexes were affinity purified using an MS2 coat protein fused to the Maltose Binding Protein bound to amylose beads, as previously described (36,37). Both mass spectrometry analysis of protein bands purified from sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) and western blot analysis (Supplementary Material, Fig. S3b) of the bound proteins revealed the presence of the SR proteins SRSF1 and SRSF6. The association of both proteins was also observed when using a longer RNA called progerin ESL (Extended Stem Loop) which contains stem loops III, IV and V in addition to progerin SL, and association of SRSF6 was stronger with this longer RNA (see Fig. 1 and Supplementary Material, Fig. S3a). Interestingly, experiments done with SL or ESL RNAs containing the c.1824 C>U mutation revealed that SRSF1 association with each of the mutated progerin SL and ESL RNAs was strongly diminished,



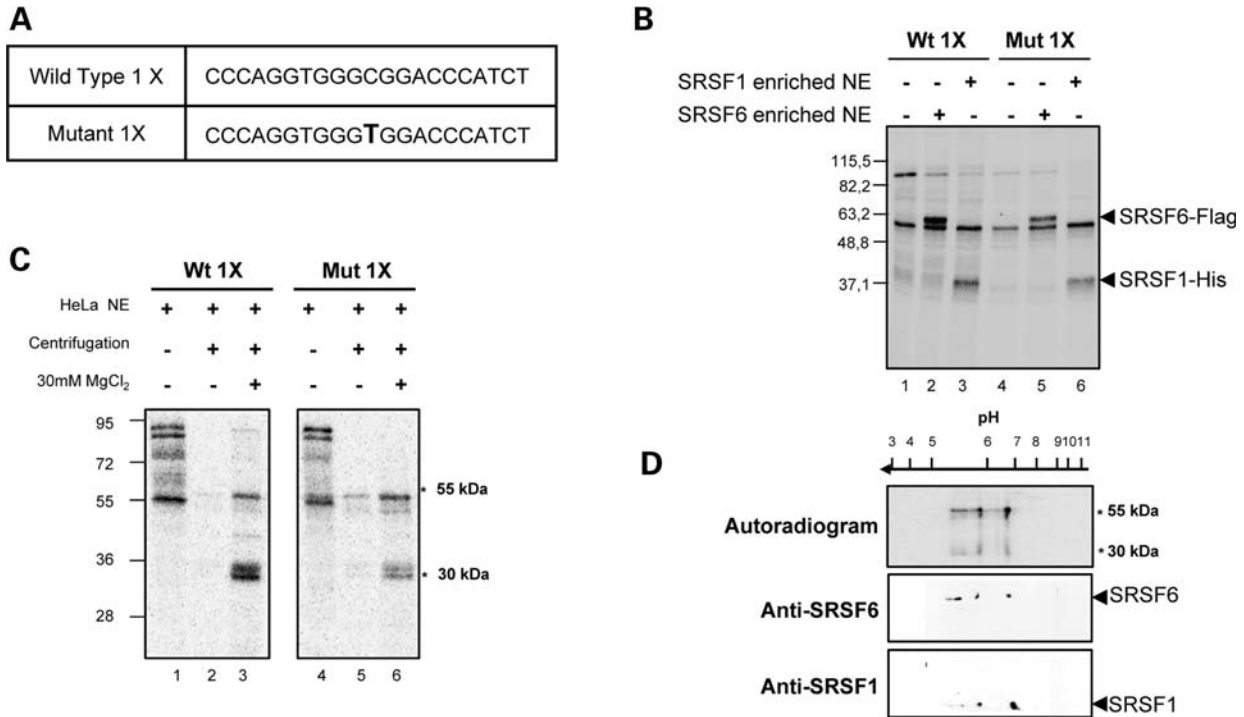


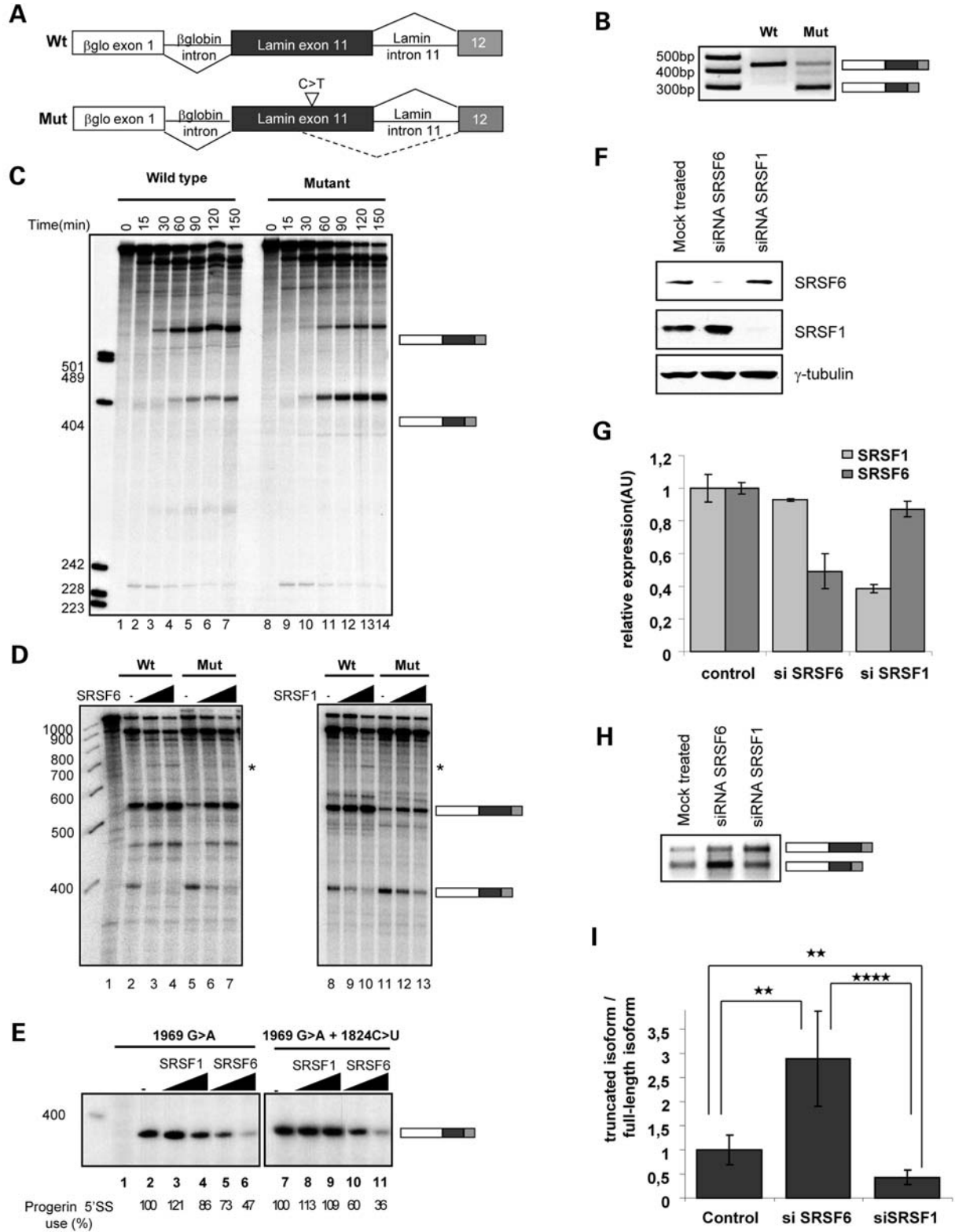
Figure 3. The SR proteins SRSF1 and SRSF6 cross-link to the WT and the c.1824C>U mutant progerin SL. (A) Sequences of the RNA constructs used for cross-linking in nuclear extracts. The sequences shown were embedded within a 40 mer RNA with invariable flanking sequences. The different RNA probes comprise monomers (1X) of the progerin 5' splice site (SS) contained in the SL structure in the wild-type (WT 1X) or mutated versions (Mut 1X). (B) Cross-linking assay in nuclear extract of untransfected cells (lanes 1 and 4) or transfected cells expressing flag-tagged SRSF6 (lanes 2 and 5) or His-tagged SRSF1 (lanes 3 and 6). The constructs (detailed in A) were cross-linked in different nuclear extracts and after RNase cleavage bound proteins were separated by SDS-PAGE and visualized by autoradiography. The size of the protein markers is shown on the left. (C) The cross-linking experiment (detailed in B) was repeated but with HeLa nuclear extract (HeLa NE, lanes 1 and 4) or with supernatant of HeLa nuclear extract treated with 30 mM MgCl₂ and centrifugation (lanes 2 and 5) or with the 30 mM MgCl₂ precipitation-enriched nuclear extract (lanes 3 and 6). (D) 2D gel analysis of UV-cross-links using the mutated version of the progerin SL probe and HeLa nuclear fraction that precipitates at 30 mM MgCl₂. This gel was transferred to a nitrocellulose membrane and submitted to autoradiography and then probed with antibodies specific for SRSF6 and SRSF1.

whereas association of SRSF6 with the SL RNA was diminished to a lower degree, but its association with the ESL RNA was also more extensively reduced (Supplementary Material, Fig. S3b). Altogether, this revealed another property of the mutated RNA which may influence progerin 5' splice site utilization. However, based on gel shift experiments performed with recombinant SRSF1 and an *in vitro* transcribed progerin SL structure, the affinity of the isolated protein SRSF1 for the WT RNA is low, suggesting that its association in the extract is mediated or reinforced by the presence of some other proteins (Supplementary Material, Fig. S3c).

Therefore, to identify proteins that bind directly to the segment containing the progerin 5' splice site, we performed UV cross-linking experiments using ³²P-labeled RNA probes. Probe WT of 21 nucleotides contains the progerin 5' splice site and corresponded

to the upper part of the progerin SL (Figs 2A and 3A). The mutated probe (Mut) included the c.1824C>T mutation. Probes were cross-linked in 293E and HeLa cells nuclear extracts, or in nuclear extract fraction precipitated with magnesium which is one of the steps used in SR protein purification (38). Labeled proteins were analyzed by electrophoresis on SDS-polyacrylamide gels after RNase treatment and visualized by autoradiography. Using whole nuclear extracts, WT probe predominantly cross-links to proteins with apparent molecular weights of 30, 55 and 120 kDa, whereas only 55 and 120 kDa protein bands are revealed with the Mut progerin probe (Fig. 3B, lanes 1 and 4). However, only two main bands at 55 and 30 kDa were detected when UV-cross-linking was performed with the MgCl₂-enriched SR protein fraction (Fig. 3C, lanes 3 and 6). According to their electrophoretic mobility, the

Figure 2. Secondary structure around LMNA exon 11 splice sites. (A) The secondary structure model proposed for exon 11 and part of intron 11 of the human LMNA pre-mRNA. The model is proposed based on thermodynamic considerations and on the results of enzymatic digestions shown in Supplementary Material, Figure S1. V1, T1 and T2 RNase cleavages are represented by arrows surmounted by squares, dots and diamonds, respectively. Red, orange and green symbols indicate a high, medium or low extent of cleavage, respectively. The structure representation shows all the bases used and the position of the splice sites are indicated with arrows. Different predicted SL structures are numbered with roman numerals. Note that the lamin A 5' splice site is in a single-stranded conformation, whereas the Progerin 5' splice site is engaged in a secondary structure (referred to in the text as progerin SL). (B) The progeria mutation opens RNA structure around the progerin 5' splice site. Autoradiographs show primer extension analyses of the WT progerin 5' splice site and the mutant c.1824C>U progerin 5' splice site cleaved by increased amount of T1 and T2 RNases. Lanes U, G, C and A correspond to the sequencing ladder. The dotted line at the level of the terminal G-C pair indicates that the stability of this terminal pair may be slightly decreased in the presence of the progeria mutation based on the appearance of some T2 RNase cleavage 3' to residue C1828.



proteins cross-linked to the progerin probes might be the SR proteins SRSF6 and one of the SRp30 proteins. Since the labeling of the 30 kDa band was decreased with the mutated RNA probe, it might represent SRSF1, which binding to the mutated progerin SL also decreased (Supplementary Material, Fig. S3). However, binding to other SRp30 proteins migrating at the same position of the gel (notably SRSF2 or SRSF7) cannot be completely rule out.

This identification was further confirmed by UV-cross-linking experiments performed with extracts from cells expressing tagged SRSF1 (His-SRSF1) or SRSF6 (Flag-SRSF6). Both tagged proteins were efficiently cross-linked, consistent with their expected apparent mobility (Fig. 3B, lanes 2, 3, 5 and 6) and were immunoprecipitated with the corresponding anti-tag antibodies (data not shown). Another confirmation was obtained by analyzing the cross-linked products with the WT RNA probe by two-D gel electrophoresis followed by western blotting. The 55 and 30 kDa cross-link signals coincided with different phosphorylation isoforms revealed by specific antibodies directed against SRSF6 and SRSF1, respectively (Fig. 3D, compare upper panel with SRSF1 and SRSF6 panels, respectively). To further document the specificity of the cross-linking, we tested two control probes under the same conditions: the RSF1 BS probe, which is a high-affinity binding site for RSF1 (RNA splicing repressor 1) known to bind hnRNP sequences (39), and the SRSF1 probe which is a high-affinity binding site for SRSF1 and other SR proteins (39). None of the probes demonstrated a strong signal at 55 kDa similar to the progerin SL probe (data not shown).

All the data of Figure 3 establish that the progerin SL and ESL RNAs have the ability to bind to SRSF1 and SRSF6. In these experiments, when protein SRSF1 was over-expressed, no significant decrease in the SRSF1 cross-link was detected in the presence of the mutation (Fig. 3B). However, in the presence of a more limited amount of SRSF1 (Fig. 3C and Supplementary Material, Fig. S3), the mutation was found to limit binding of SRSF1 to the RNA.

The SR protein SRSF6 represses splicing to the progerin 5'SS *in vitro*

To assess the contribution of the SRSF1 and SRSF6 splicing factors to progerin production, we cloned the *LMNA*

genomic sequence from exons 11 through 12 with or without the HGPS mutation in a reporter construct that was suitable for both cell transfection and *in vitro* splicing assays (Fig. 4A). In the presence of the HGPS mutation, transient transfections of these constructs in HeLa cells gave rise to a 150 nucleotide deletion in the *LMNA* splicing product due to the predominant use of the progerin 5'SS, whereas only the lamin A 5'SS was selected when using the WT *LMNA* reporter (Fig. 4B). In contrast, when the same constructs were used in *in vitro* splicing experiments (Fig. 4C), both the WT and mutated *LMNA* substrates led to the production of mRNAs arising from the utilization of both lamin A and progerin 5'SSs, implying that the progerin 5'SS can be recognized in the extract as an authentic 5'SS, even in the absence of the HGPS mutation. However, splicing in RNAs transcribed from the mutated *LMNA* construct predominantly used the progerin 5'SS, whereas in RNA of the WT *LMNA* construct splicing mostly proceeded through the lamin A 5'SS (Fig. 4C, compare lanes 4–7 and lanes 11–14). Depletion of the U1 snRNP from the extract abolished usage of both progerin and lamin A 5'SSs and led to exon 11 skipping (data not shown), implying that both splice sites are equally dependent on U1 snRNP. This result also indicate that exon 12 of *LMNA* may assemble a cross exon complex able to subsequently bind the globin 5'SS to generate a B-like spliceosomal complex to perform splicing independently of U1 as has recently been suggested by Schneider *et al.* (40).

To test whether SRSF1 or SRSF6 might be involved in activation of the utilization of the progerin 5'SS, *in vitro* splicing assays were performed with HeLa cell nuclear extracts supplemented with recombinant SRSF1 or SRSF6 proteins (Fig. 4D). An expected effect of the addition of a SR protein is to increase the global splicing activity of the nuclear extract. However, SRSF6 addition strongly inhibited the use of the progerin 5'SS, whereas the selection of the lamin A 5'SS was activated (Fig. 4D, compare lane 2 with lanes 3–4). Addition of SRSF1 to the extract also resulted in less progerin and more lamin A mRNA production from the WT substrate. But this effect was milder than that observed with SRSF6 (Fig. 4D, lanes 9–10). Similar effects of both SR proteins were obtained with the mutated *LMNA* transcript (Fig. 4D, lanes 6–7 and 12–13, respectively). However, SRSF6 had a very strong inhibition activity on the use of the mutated progerin 5'SS in the mutated RNA, whereas

Figure 4. SRSF1 and SRSF6 play opposite roles in *LMNA* splicing. (A) Schematic representation of the minigene reporters β globin-*LMNA* used for the study of *LMNA* splicing. *LMNA* sequences, carrying (Mut) or not carrying (Wt) the mutation, were inserted downstream of sequences derived from the human β globin gene. (B) These minigenes were transfected into HeLa cells and the splicing products were analyzed by RT-PCR of RNA extracted from transfected cells using primers either side of the alternatively spliced regions that give different sized products for the long and short forms. (C) Time course of *in vitro* splicing of WT and Mut reporters showing differential but significant use of both lamin A 5'SS and progerin 5'SS in HeLa nuclear extracts. RNA size markers are indicated on the left. Different times of incubation in minutes are indicated on the top of each panel. (D) Splicing reactions with Wt (lanes 1–4 and 8–10) and Mut (lanes 5–7 and 11–13) reporters in HeLa nuclear extracts alone (lanes 2, 5, 8 and 11) or supplemented with either SRSF6 (lanes 3–4 and 6–7) or SRSF1 (lanes 9–10 and 12–13) recombinant proteins [200 ng (lanes 3, 6, 9 and 12) or 400 ng (lanes 4, 7, 10 and 13)]. (Star) Note that a minor mRNA species of 450 nt long which accumulates in the presence of SRSF6 corresponds to the use of a cryptic 5'SS. (E) Splicing reactions with WT (lanes 1–6) and Mut (lanes 7–11) reporters mutated at lamin 5'SS in HeLa nuclear extracts alone (lanes 2 and 7, respectively) or supplemented with either SRSF6 (lanes 5–6 and 10–11) or SRSF1 (lanes 3–4 and 7–8) recombinant proteins [200 ng (lanes 3, 5, 8 and 10) or 400 ng (lanes 4, 6, 9 and 11)]. The relative efficiency of splicing is indicated below each panel, with a value of 100 for the assays without added SR protein. (F) Depletion of SRSF1 or SRSF6 in HeLa cells using siRNAs was assessed by western blotting and (G) resulting mRNA levels assessed by RT-qPCR. (H) The splicing pattern of the mutant β globin-*LMNA* reporter in treated cells was determined by RT-PCR using primers either side of the alternative spliced region as in (B) and (I) quantification of the ratio between truncated and full-length isoforms from three independent experiments of SRSF1 and SRSF6 siRNA knockdown. The value of 1 is arbitrarily assigned to control siRNA samples.

SRSF1 was unable to inhibit it completely. Also, SRSF6 highly activated the use of the lamin A 5'SS (Fig. 4D, lanes 6–7).

As these experiments could not discern whether SRSF6 acts as a repressor of the progerin 5'SS and/or as an activator of the lamin A 5'SS, we mutated the first G of the lamin A 5'SS in the WT and mutated substrates and tested the effect of SRSF6. Addition of SRSF6 to the extract inhibited strongly progerin mRNA production in the WT and the mutant context (Fig. 4E, compare lanes 2 and 7 with lanes 5–6 and lanes 10–11, respectively). In contrast, addition of SRSF1 had a mild effect in the WT context, and almost no effect in the presence of the c.1824C>T mutation (Fig. 4E, compare lanes 2 and 7 with lanes 3–4 and lanes 8–9, respectively), which is in agreement with its lower association with the mutated progerin SL RNA in HeLa nuclear extract. Taken together, the data suggest that the SRSF1 activity mainly consists in the activation of the lamin A 5'SS, while SRSF6 acts as a repressor of progerin production, most probably by binding to the progerin SL sequence and may be also by an independent activation of the lamin A 5'SS.

The SR protein SRSF6 represses the usage of the progerin 5'SS in transfected cells

If SRSF6 plays an important role in repressing the progerin 5'SS *in vivo*, it should be possible to reverse this phenomenon by decreasing SRSF6 levels. This possibility was next investigated by using small-interfering RNAs (siRNAs) to reduce the amount of SRSF6 in HeLa cells transfected with the mutant *LMNA* minigene, for which the selection of the progerin 5'SS is predominant. Treatment of HeLa cells with siRNAs targeting SRSF6 resulted in a large and specific reduction in the SRSF6 level, as observed 96 h post-transfection by western blotting (Fig. 4F). Under the same conditions, no significant changes were observed in the expression of SRSF1 and γ -tubulin, used as internal controls (Fig. 4F). Quantitative polymerase chain reaction (PCR) analysis also confirmed that SRSF6 expression was reduced at the RNA level (Fig. 4G). The effect of SRSF6 depletion on the splicing profile of the HGPS mutant *LMNA* minigene was analyzed by reverse transcription polymerase chain reaction (RT-PCR), using primers allowing discrimination between mRNAs spliced at the lamin A and the progerin 5'SSs, respectively (Fig. 4H). Results from three independent experiments showed that the siRNA-mediated reduction in the SRSF6 protein level greatly enhanced usage of the progerin 5'SS (Fig. 4H and I), whereas utilization of the lamin A 5'SS remained unchanged.

Surprisingly, siRNA knockdown of SRSF1 to comparable levels of SRSF6 knockdown (Fig. 4F and G) leads to the opposite effect on *LMNA* splicing, as it reduces usage of the progerin 5'SS (Fig. 4H and I). Since SRSF1 has only a limited affinity for the progerin SL, and since it did not activate the progerin 5'SS *in vitro*, the effect observed upon siRNA knockdown of SRSF1 might result from an indirect mechanism. Given that *in vivo* transcription and splicing are coupled (41), it is possible that SRSF1 depletion alters the transcription rate of the RNA pol II (42) allowing better recognition of the progerin 5'SS.

The SR proteins SRSF6 and SRSF1 have opposing effect on the usage of the progerin 5'SS in primary patients' fibroblasts

To directly demonstrate the involvement of SRSF1 and SRSF6 in the regulation of the progerin 5'SS selection, we used fibroblasts isolated from HGPS patients. The levels of SRSF1 or SRSF6 in these fibroblasts were reduced by RNAi-mediated knockdown and the splicing profile of endogenous *LMNA* was evaluated by RT-PCR using primers located in exon 8 and exon 12 to amplify both progerin and lamin A cDNAs. Consistent with the results obtained with transfected reporter constructs, the siRNA-mediated reduction in SRSF6 (Fig. 5B, panel SRSF6) greatly enhanced usage of the progerin 5'SS (Fig. 5A) with concomitant reduction in the lamin A 5'SS utilization. In contrast, SRSF1 knockdown (Fig. 5B, panel SRSF1) allowed slightly more lamin A mRNA production, whereas progerin mRNA levels remained unchanged (Fig. 5A). HGPS fibroblasts treated with siRNA targeting SRSF6 accumulated slightly more progerin than HGPS fibroblasts treated with control siRNA, as judged by western blotting with a lamin A/C antibody (Fig. 5B, panel lamin A/C). However, in keeping with the RNA analysis, the effect of SRSF1 knockdown on progerin expression is less visible. Taken together, these observations established conclusively that SRSF6 is directly responsible for repression of the progerin 5'SS in cells, in agreement with the effects of SRSF6 observed in *in vitro* splicing assays.

The progerin 5'SS splicing mechanism is conserved in the mouse

The mouse *LMNA* gene has 83% identity in the open reading frame when compared with human. Furthermore, the 20 nt long sequences surrounding the progerin 5'SS in the mouse (position 1817–1837) and human (position 1814–1834) mRNAs only show one nucleotide difference (U 1831 in the mouse instead of a C 1828 in human). Therefore, sequestration of the putative progerin 5'SS in a secondary structure in the mouse similar to that in human could be predicted (Fig. 6A), and we verified this hypothesis experimentally (Fig. 6B). We also showed that introduction of the human HGPS mutation at the same position in the mouse as in the human progerin SL (c.1827 C>U) strongly increased the accessibility of the terminal loop as found for the human pre-mRNA, implying that this conserved structure likely has a functional importance (Fig. 6B) and that the progerin 5'SS could be used in mice.

To determine whether regulation of the progerin 5'SS was similar in the two species, the HGPS C>T mutation was introduced 6 nt downstream from the progerin 5'SS (c.1827 C>T) in the *Lmna* mouse gene and mutated mice were generated by homologous recombination (Osorio *et al.*, submitted for publication). To evaluate the splicing profile of endogenous *LMNA* transcripts, RNA extracted from different mouse tissues of WT homozygous and heterozygous knock-in mice were submitted to RT-PCR using primers located in exon 8 and exon 12 to amplify both progerin and lamin A cDNAs (Supplementary Material, Fig. S4a). This analysis was performed with 4-month-old mice because all the homozygous knock-in mice *Lmna*^{G609/G609} died at this age (Osorio *et al.*, submitted

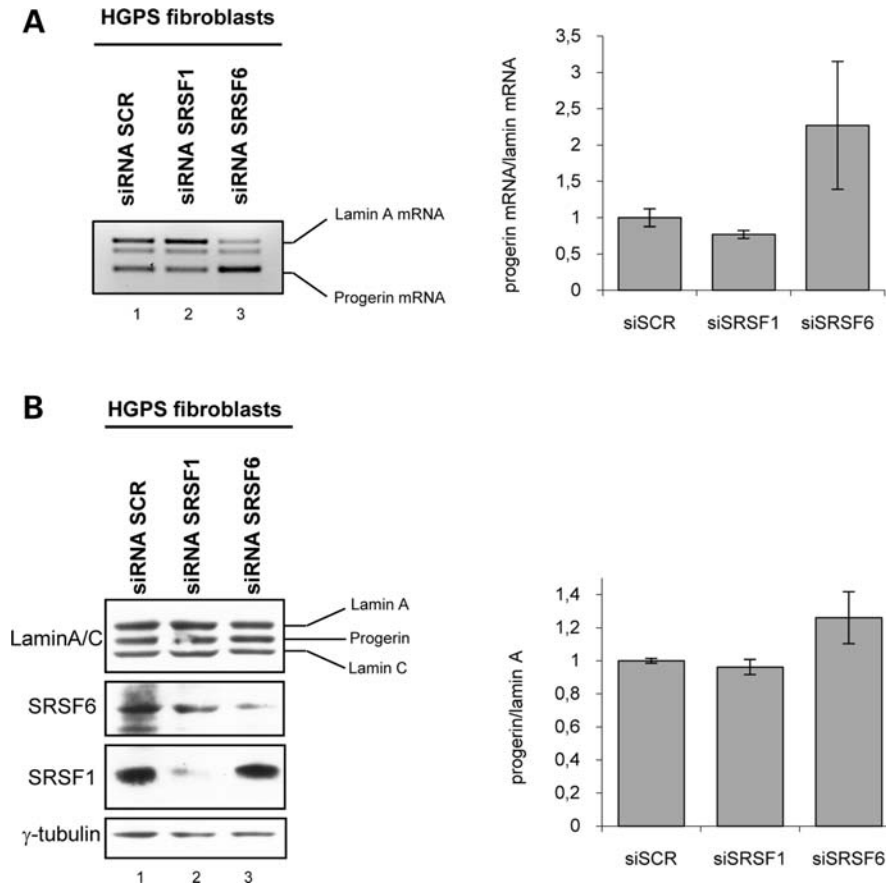


Figure 5. SRSF6 prevents usage of progerin 5'SS in fibroblasts from progeria patients. (A) The levels of lamin A/C and progerin mRNA in treated cells with scrambled siRNAs (lane 1) and with siRNA targeting either SRSF1 (lane 2) or SRSF6 (lane 3) were determined by RT-PCR using primers located in exon 9 and exon 12. Right panel quantification of the ratio between truncated and full-length isoforms from two independent experiments of SRSF1 and SRSF6 siRNA knockdown. (B) Depletions of SRSF1 or SRSF6 in fibroblasts from progeria patients using siRNAs (panels SRSF1 and SRSF6, respectively), and their effects on lamin A, lamin C and progerin expression (panel Lamin A/C), as assessed by western blotting. Right panel quantification of the ratio between progerin and lamin A from two independent experiments of SRSF1 and SRSF6 siRNA knockdown. The value of 1 is arbitrarily assigned to control siRNA samples.

for publication). As expected, both heterozygous and homozygous knock-in but not WT mice expressed progerin mRNA in all tissues (Supplementary Material, Fig. S4a). Both RT-PCR and western blot analysis of extracts from various tissues of 4-month-old homozygous knock-in mice (*Lmna*^{G609G/G609G}) confirmed that the steady-state levels of progerin are very high compared with the lamin A level which is hardly detectable in most tissues (Supplementary Material, Fig. S4a and b, G609G/G609G), implying that the expression of progerin from the mutant allele is predominant over lamin A in most tissues. Consistently, in heterozygous mice where both alleles are expressed, the levels of lamin A and progerin are comparable (Supplementary Material, Fig. S4a and b, G609G/+), whereas in WT mice only lamin A is expressed (Supplementary Material, Fig. S4a and b, +/+). *Lmna*^{G609G/G609G} mice also showed markedly reduced survival compared with WT littermates, since all the homozygous mice died 4 months after birth (Osorio *et al.*, submitted for publication). Strikingly, *Lmna*^{G609G/+} mice were also short lived but survived longer than homozygous mice (8 months on average) (Osorio *et al.*, submitted for publication), suggesting that *Lmna* splicing that favors progerin production can lead to reduced growth and lifespan.

In order to study the contribution of SR protein to the regulation of progerin production, we generated MEFs. MEFs isolated from WT, heterozygous (*Lmna*^{G609G/+}) and homozygous (*Lmna*^{G609G/G609G}) c.1827 C>T mutant embryos at embryonic day 13.5 were harvested to analyze the RNA splicing ratio and protein levels using RT-PCR and western blots, respectively (Fig. 6C and D). As expected, both heterozygous and homozygous but not WT MEFs expressed progerin mRNA (Fig. 6C) and protein (Fig. 6D), demonstrating that the c.1827 C>T mutation, like in human fibroblasts from progeria patients, activates the progerin 5'SS. In keeping with the PCR results, more progerin is produced in the homozygous compared with the heterozygous MEFs (Fig. 6D). However, while the c.1827 C>T mutation induces a large decrease in the expression of the lamin A, it did not abolish its expression in the homozygous MEFs, implying that alternative splicing is controlling the balance between lamin A and progerin expression. Consistent with previous results, immunofluorescence microscopy revealed that heterozygous and homozygous MEFs had more misshapen nuclei with nuclear blebs than WT MEFs (Fig. 6E and F). Thus, the c.1827 C>T mutation of the mouse *Lmna* gene reproduces the splicing-dependent phenotype of fibroblasts from progeria patients.

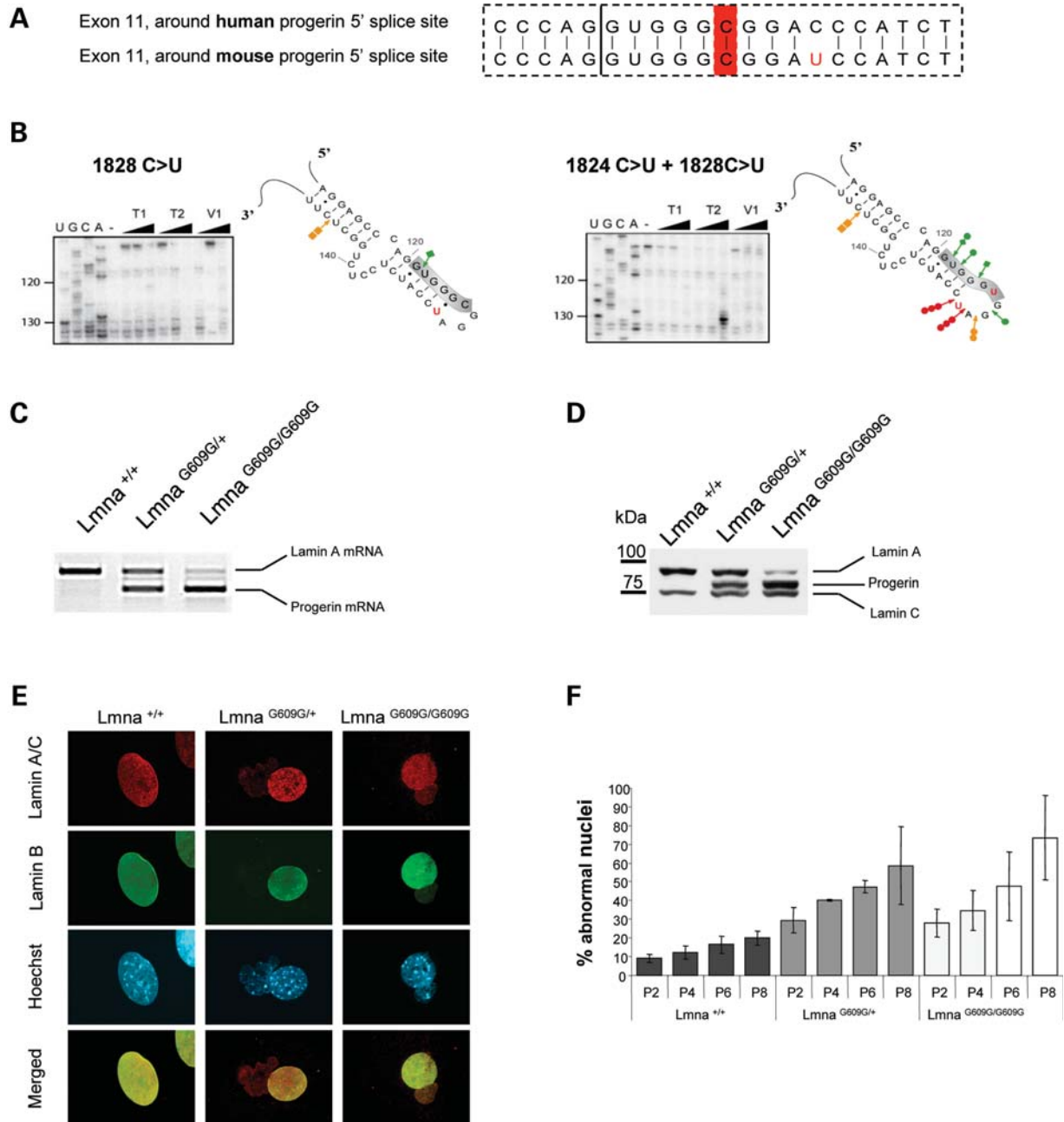


Figure 6. The use of progerin 5'SS is conserved from the mouse to human. (A) Sequence comparison between the mouse (position 1817–1837) and human (position 1814–1834) mRNAs corresponding to the progerin SL. (B) The secondary structure model proposed for WT (left) and mutated (right) mouse progerin SL. The model was proposed based on thermodynamic considerations and on the results of the enzymatic digestions shown. T1, T2 and V1 RNase cleavages are symbolized as in Figure 2. The structure representation shows all the bases used and the position of the progerin 5'SS is indicated. Autoradiographs show primer extension analyses of the WT and the mutant c.1827C>U mouse progerin SL cleaved by increasing amounts of T1, T2 and V1 RNases. Lanes U, G, C and A correspond to the sequencing ladder. Note good accessibility to T1 and T2 cleavage around the mutation. (C) Heterozygous (*Lmna*^{G609G/+}) and homozygous (*Lmna*^{G609G/G609G}) MEFs carrying the equivalent splicing mutation in endogenous *LMNA* express increased amounts of progerin as shown at the RNA level by RT-PCR using primers either side of the mutation and (D) at the protein level as shown by western blotting revealed with lamin A/C antibody. (E) Expression of mouse progerin leads to nuclear abnormalities, as shown by immunofluorescence with anti-lamin A/C and Hoechst staining. Scale bars correspond to 10 μ m. (F) Quantification of nuclear abnormalities in WT (*Lmna*^{+/+}), *Lmna*^{G609G/+} and *Lmna*^{G609G/G609G} mutated fibroblasts at the indicated cell passages.

Next, to test whether SRSF1 and/or SRSF6 modulate the progerin 5'SS *in vivo*, we reduced their levels with specific siRNAs in *Lmna*^{G609G/+} MEFs. *Lmna*^{G609G/+} MEFs were either mock-treated or transfected with siRNAs directed against SRSF1 or SRSF6 and the corresponding proteins

levels were assessed by western blot analysis 5 days after transfection. A reduction in both protein levels was observed with their corresponding siRNAs (Fig. 7A, upper and lower panels). To test the effect of SR protein depletion on the splicing profile of *LMNA* primary transcripts expressed in the

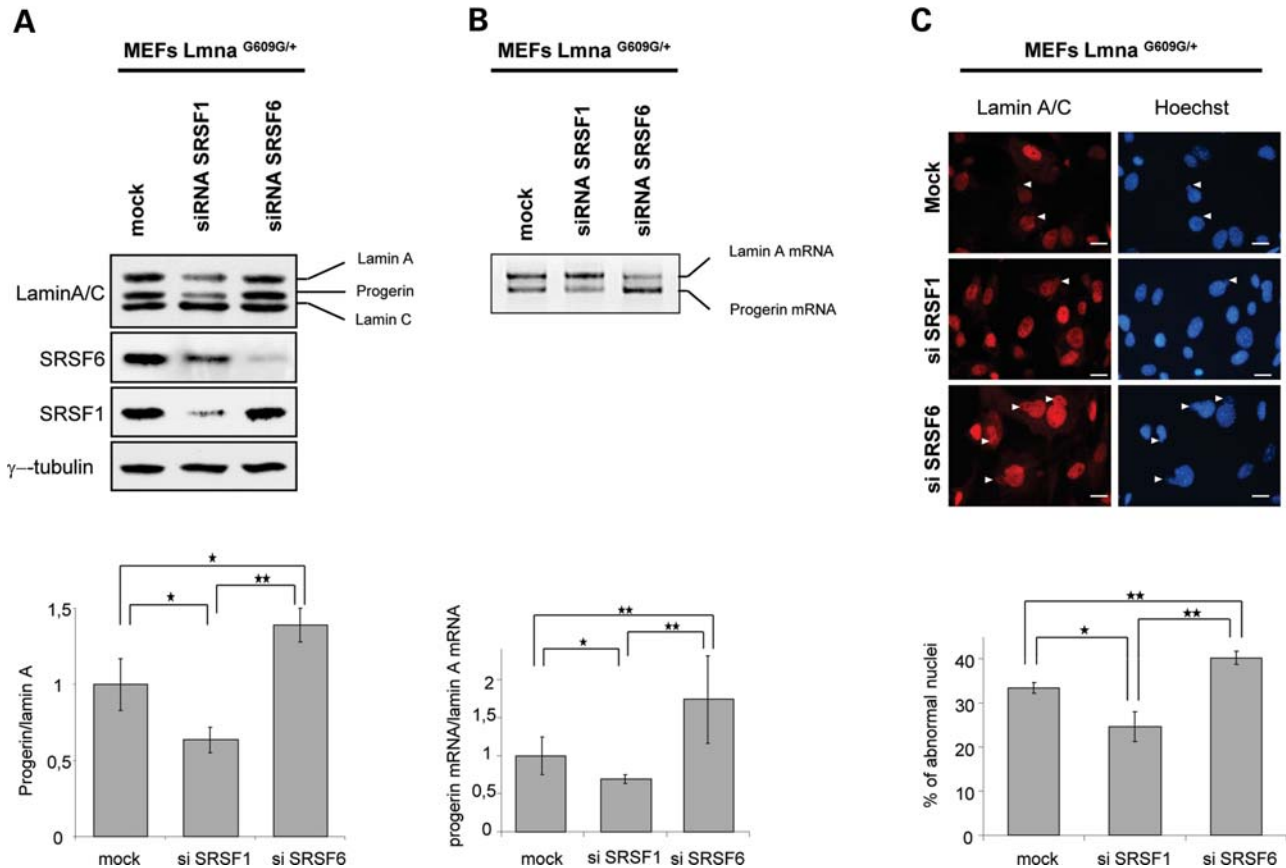


Figure 7. *LMNA* splicing regulation by the SR proteins SRSF1 and SRSF6 is conserved between the mouse and human. SRSF1 or SRSF6 was depleted in *Lmna*^{G609G/+}MEFs using siRNAs. Below each section, the quantifications are shown as bar charts. (A) Western blot showing depletion of SRSF1 or SRSF6 and its effect on the ratio of lamin A/C and progerin. (B) The splicing pattern of the *LMNA* gene in SRSF1 or SRSF6 depleted cells was determined by RT-PCR and analysed as in Figure 5C. (C) Nuclear abnormalities in *Lmna*^{G609G/+}MEFs after SRSF1 or SRSF6 depletion, as shown by immunofluorescence with anti-lamin A/C and Hoechst staining. Scale bars correspond to 30 μ m. Percentage of nuclear abnormalities in siRNA-treated versus control cells is shown.

corresponding fibroblasts, RT-PCR experiments were performed with an oligonucleotide pair allowing amplification of both lamin A and progerin isoforms. The results established that siRNA-mediated reduction in SRSF6 protein levels is correlated to an effective increase in progerin both at the protein (Fig. 7A, upper and lower panels) and mRNA levels (Fig. 7B, upper and lower panels), whereas siRNA-mediated reduction in SRSF1 leads to opposite effect (Fig. 7A and B, upper and lower panels). Given that progerin accumulation was expected to change the nuclear envelope of MEFs, we checked the effect of these SR proteins depletion on nuclear morphology. As predicted, depletion of SRSF6 was associated with more misshapen nuclei, whereas SRSF1 depletion reduced the number of misshapen nuclei. Taken together, these observations establish for the first time that SRSF6 and SRSF1 have opposing effect on *Lmna* gene splicing.

DISCUSSION

In 2003, it was discovered that the silent mutation in exon 11 of the *LMNA* gene responsible for human progeria syndrome allows production of progerin by inducing utilization of a 5'SS in the exon 11 (4,5). Although this phenomenon was

described long ago, the molecular basis for the age-associated splicing alteration of the *LMNA* gene has not been elucidated. Here we demonstrate that the 5'SS in the exon 11 of the *LMNA* gene is engaged in a stable RNA structure which is probably preventing its efficient recognition by the splicing machinery. The progeria mutation induces some opening of the terminal loop which may favor accessibility of the internal 5'SS and facilitate U1 snRNP association. In addition, regulation of the competition between the progerin 5'SS and lamin A 5'SS by the SR proteins SRSF1 and SRSF6 is also modulated by the mutation. The demonstration that these two characteristic features are also conserved in the mouse further underlies the importance of splicing repression to minimize progerin production during aging.

It is indeed striking that the progerin 5'SS is conserved in mammals over millions of years and that the single-point mutation at the normally uninfluential +6 position leads to the same dramatic changes in splicing. Our *in vitro* splicing analyses of the WT and mutated *LMNA* pre-mRNA, performed for the first time, showed that WT progerin 5'SS is a *bona fide* 5'SS, since it is moderately recognized from the WT transcript, and strongly activated from the mutated transcripts, as expected (Fig. 4). Interestingly, our observation of the utilization of the WT progerin 5'SS in the absence of the mutation

is in agreement with several reports indicating that progerin may be detected at low level during physiological aging (10,15–17). Furthermore, the fact that the splicing at the progerin 5'SS can be also induced in WT fibroblasts using anti-sense oligonucleotides against exon 11 sequences (43), and that the splicing of WT and mutated *LMNA* transcripts is modulated by SR proteins strongly supports the idea that the progerin 5'SS is an alternative 5'SS. Therefore, the mutation 1824C>T reinforces a naturally weak 5'SS rather than create a *de novo* 5'SS.

Interestingly, the progerin 5'SS in the WT exon 11 is engaged in a stable SL structure (progerin SL), with a compact terminal tetra-loop. This tight structuration likely prevents its efficient recognition by the splicing machinery. In contrast, the lamin A 5'SS is in a favorable non-structured region (Fig. 2). The progeria mutation markedly increases the accessibility of the terminal loop of the progerin SL, which together with the slight increase in the intrinsic strength of the progerin 5'SS by the C>U substitution is expected to favor the recognition of the 5'SS by U1 snRNP. Indeed, it is well known that RNA–RNA interaction preferentially initiates through loop–loop or single strand–loop interactions (44). Interestingly, we observed that the other mutation 1822 G>A in the human *LMNA* gene, which is associated with HGPS (4,5), also leads both to an opening of the progerin SL structure and an increased progerin 5'SS utilization. It will be interesting to identify the kind of structuration of the progerin SL terminal loop by structural analysis such as NMR 3D structure analysis, since another single mutation in this loop, 1825 G>A, was found to induce both an opening of the loop structure and an increase in the progerin 5'SS utilization *in vitro*. We predict that the occurrence of this mutation in the human *LMNA* gene may generate HGPS. In support to the proposed role of the RNA structure in regulation of the progerin 5'SS, it should be pointed out that there are already well-documented examples for the involvement of RNA secondary structure in the control of alternative splicing (31,45,46). For example, the stability of a predicted SL structure at the 5'SS of the key Alzheimer's protein tau regulates usage of exon 10 (47,48) and the alternative exon of the key adhesion gene fibronectin is highly influenced by secondary RNA structure (31). Bioinformatics analyses indicate that the splicing regulatory sequences are preferentially in a single-stranded conformation, as this allows these sequences to interact with RNA-binding proteins that mostly have a preference for single-strand RNA (45). At present, it is unclear whether SR proteins SRSF6 and SRSF1 that regulate the progerin 5'SS utilization are involved in the stabilization or disruption of the identified RNA structure. According to the present data, SRSF6 might stabilize the RNA structure around the progerin 5'SS, thereby, preventing SRSF1 from enhancing the recruitment of the U1 snRNP. As transcription and splicing are connected, the kinetics of elongating polymerase could also influence the formation of RNA structures in connection with SR protein concentrations (26,41).

Like other splicing mutations that are involved in the etiology of human diseases, the HGPS mutation does not inactivate *LMNA* gene expression completely, but rather gives rise to a mixture of normal and aberrant isoforms. Transitions between splicing isoforms are dynamic and depend on a

balance between factors that maintain a silent alternative splice site and those that promote its activation, possibly in a tissue-specific manner (19). Increased production of progerin in the homozygous and heterozygous mice harboring the splicing mutation triggers many of the features that appear early on in humans with HGPS, such as slow growth, osteolytic lesions in bone, osteoporosis, micrognathia and loss of adipose tissue (Osorio *et al.*, submitted for publication). In contrast, consistent with previous reports (49) (Osorio *et al.*, submitted for publication), no functional abnormalities were associated with mice lacking lamin A and/or progerin expression, indicating that progerin, even at trace levels, could be detrimental for lifespan. Given that progerin production is under the control of both transcription and splicing of the *LMNA* gene, the importance of splicing has escaped analysis in previously described models (50,51). If changes in the RNA structure and/or SR proteins concentrations occur preferentially in some tissues but not others, one would expect that progerin accumulation will also be different between tissues. Our initial analysis of progerin expression in various tissues of the mouse shows that this is indeed the case (Supplementary Material, Fig. S4). In WT mice, where only lamin C and lamin A are detected, it can be clearly seen that while the brain contains roughly equivalent amount of both proteins, in all other tissues the lamin A content is higher than lamin C. Also, the brain accumulated less progerin than lamin C in both homozygous and heterozygous knock-in mice. Consistent with the RNA analysis, in heterozygous knock-in mice, it can be appreciated that less progerin is expressed in both the brain and spleen where the lamin A mRNA was higher than progerin mRNA. In contrast, the heart of heterozygous mice accumulated more progerin than any other tissue.

The detailed knowledge of the factors that are critical for the activation described here and how they cooperate with other spliceosomal components in a given tissue in the HGPS-like mouse model will help the design of new therapeutic agents. The basis for targeting mutated RNA sequences is straightforward and highly specific for a given gene because it relies on antisense molecules base-paired to pre-mRNA to prevent access of the splicing machinery to splice sites (52,53). Combination of the RNA structure and binding of SR proteins described here helps explain past successes and failures of anti-sense oligonucleotides (ASOs) to modulate alternative splicing within *LMNA* exon 11 (43,53). Conversely, we show here that the region 50 nucleotides downstream of the alternative donor site is fully accessible and may serve as binding site for splicing regulators. Interestingly, targeting of ASOs to that region was effective at triggering progerin production from WT *LMNA* pre-mRNA (43,53). It has been postulated that this enhancement might be due to the ability of ASOs to hide a binding site for SRSF5. Our *in vitro* complementation assay is in keeping with this postulate because SRSF5 is shown to enhance the use of the *LMNA* 5'SS at the expenses of progerin 5'SS (43) (data not shown). Our finding that two SR proteins might have opposing effects on progerin 5'SS utilization indicates that the efficiency of ASOs to prevent splicing at this site will depend on the cellular concentration of the two SR proteins, making the outcome of the effect of the ASOs less predictable. This does indeed seem to be the case since in one study a morpholino ASO

hybridizing to the progerin 5'SS was shown to be effective at reducing the aberrant splicing, whereas another study using a 2'O methyl ASO had a negligible effect (43,54).

The new mouse model (Osorio *et al.*, submitted for publication) will facilitate testing the outcomes of ASO treatment in different tissues; it could also lead to development of a therapeutic strategy involving small molecules that target splicing factors (53,55). This can be achieved either by lowering their cellular levels or by interfering with their splicing enhancing/repressing activity. Our data establish a proof of principle that SR proteins could be therapeutic targets. RNAi-mediated reduction in the SR protein SRSF1 levels led to a reduction in progerin expression in heterozygous *Lmna*^{G609G/+} MEFs, whereas SRSF6 knockdown aggravated the MEFs phenotype. As SR protein knockdown is expected to have a pleiotropic effect on gene expression and not only affect progerin expression, targeting of the activity of SR proteins on specific substrates could be more attractive. Indeed, screening of chemical libraries with ESE-dependent splicing substrates resulted in the identification of indole derivatives as a class of small molecule inhibitors of alternative splicing (56). Rather than generally inhibiting splicing, these molecules proved highly selective for splicing events mediated by different classes of ESE sequences and are therefore less toxic. Their selectivity is thought to depend on their interaction with each SR protein (56,57). Further studies are now ongoing to confirm the potency and lack of deleterious side-effects of indole derivatives in the HGPS model.

The control of lamin A/progerin splicing switch of the *LMNA* gene in the various tissues of the mouse will have an important impact on the lifespan. Balancing this output in critical tissues will be a major determinant not only for therapeutic action but also for establishing the longevity of mammalian organisms.

MATERIALS AND METHODS

Ethics statement

All animal procedures were conducted in strict adherence with the European Community Council Directive of 24 November 1986 (86-609/EEC).

Plasmids constructs

LMNA sequences (exon 11, intron 11 and exon 12) were PCR amplified from WT and patient's cell genomic DNA with specific forward and reverse primers. PCR fragments were purified with Concert Rapid PCR purification system (Invitrogen) and subcloned at the *Bam*HI and *Eco*RI restriction sites of the β glo3S plasmid (39) to give the β Glo-*LMNA* constructs used for *in vitro* splicing and transient transfection experiments.

Cell lines, culture and transfection conditions

HeLa cells were grown in Dulbecco's modified eagle medium media (Gibco) complemented with 10% fetal calf serum and antibiotics. They were transfected in six-well plates, at 50% confluency (30% for siRNA experiments), using Dreamfect

(OZ biosciences) for plasmid DNA or oligofectamine (Invitrogen) for siRNAs, according to the manufacturer's instructions. For siRNA depletion of SR proteins, cells were treated for 72 h with siRNAs against hASF/SF2 (hSRSF1) or 96 h with siRNAs against hSRSF6 (with two successive applications carried out at 24 h interval). The mutant β Glo-*LMNA* plasmid was transfected and incubated for 24 h before recovering the cells. The siRNA sequences are available upon request. MEFs were derived from embryos dissected 13.5 days after detection of vaginal plugs according to Nagy *et al.* (58). Fetal livers and/or yolk sacs were used for PCR genotyping. MEFs were transfected in six-well plates with siRNA using oligofectamine (Invitrogen), according to the manufacturer's instructions. For siRNA depletion of SR proteins, cells were treated for 72 h with siRNAs against mSRSF1 or 96 h with siRNAs against mSRSF6 (with two successive applications carried out at 24 h interval). The siRNA sequences are available upon request.

Human HGPS fibroblasts AG01972 were purchased from Coriell Cell Repository. They were cultured at 37°C in DMEM containing GlutaMAX, non-essential amino acids, sodium pyruvate, 0.1 mM β -mercaptoethanol and 10% fetal bovine serum. For siRNA depletion of SR proteins, cells were treated for 72 h with siRNAs against SRSF1 or SRSF6. The siRNA sequences are available upon request.

Expression of tagged SR proteins SRSF6-Flag and SRSF1-His was performed after transfection of the corresponding plasmid in the human embryonic kidney 293 cell line expressing Epstein-Barr virus EBNA1 protein (293E) according to the established protocol (59).

RNA and protein analysis

RNAs were extracted using TRI Reagent (Sigma-Aldrich). The samples were treated with RNase-free DNase (RQ1, Promega). Cellular RNA concentrations were quantified by measuring optical absorption at 260 nm with a nanodrop. 1.5 μ g of RNA was reversed transcribed using 'First strand cDNA synthesis' kit (GE Healthcare biosciences) with random hexamers. PCRs were carried out with 0.1 mM dNTP mix, 1 mM MgCl₂, 1 \times PCR buffer, 1.25 U of Taq DNA polymerase (Invitrogen) and 0.2 μ M of each of the sense and antisense primers. Denaturation, annealing and extension were performed at 94, 64 and 72°C, respectively. Each step lasted 30 s. PCRs were performed on a Master cycler Gradient 96 thermocycler (Eppendorf) with 28 cycles. These cycles were preceded by a 3 min denaturation at 94°C and terminated by a 2 min extension at 72°C. Amplification products were analyzed by 2% agarose gel electrophoresis and visualized by ethidium bromide staining. The intensity of the bands corresponding to the two mRNAs was quantified using GeneSnap acquisition software and GeneTools analysis software (Syngene). The ratio lamin A 5'SS/progerin 5'SS from three independent transfection experiments was normalized using the control sample of each series as a reference, which gave a score for the choice of progerin 5'SS.

HeLa cell proteins were extracted in a buffer containing 50 mM Tris-HCl pH 6.8, 20 mM ethylenediaminetetraacetic acid (EDTA), 5% sodium dodecyl sulfate (SDS), sonicated briefly on ice and analysed by SDS-PAGE and western

blotting. Proteins from MEFs and mouse tissues were extracted in a buffer containing 63 mM Tris-HCl pH 7.5, 2% SDS, 5% β -mercaptoethanol and 8 M urea. Samples were ground in a glass tissue grinder for additional 2 min and then sonicated and centrifuged at 13 200g for 10 min at 4°C. The supernatant was analyzed by western blotting using 4–12% polyacrylamide BisTris gels (Invitrogen) for electrophoresis and fractionated proteins were then transferred to a nitrocellulose membrane. For western blotting, we used the following antibodies with the corresponding dilutions: anti γ -tubuline (sigma) (1/1000); anti-ASF/SF2 (Invitrogen/Zymed, catalog #32-4500); anti-SRSF6 (kind gift of Adrian Krainer) (mAb 8-1-28 culture supernatant); anti-lamin A/C (Santa Cruz, sc-20681); anti-mouse HRP (GE Healthcare) and anti-rabbit HRP (GE Healthcare).

Immunofluorescence

MEFs were plated on glass cover slips at a confluence of 50% before fixation in 3.7% formaldehyde in phosphate-buffered saline (PBS) followed by a 5 min permeabilization in 0.1% Triton X-100 in PBS. Cells were then incubated in PBS containing 3% bovine serum albumin (BSA) before staining with primary antibodies as follows: rabbit polyclonal anti-lamin A/C (N-18) (sc-6215 from Santa Cruz Biotechnology, Inc.) at 4 μ g/ml in PBS–0.1% BSA, goat polyclonal anti-lamin B (M-20) at 4 μ g/ml in PBS–0.1% BSA (sc-6217 from Santa Cruz Biotechnology, Inc.). Primary antibodies were revealed with either an AlexaFluor546- or -488-conjugated anti-rabbit or anti-goat antibody (Invitrogen). Nuclei were stained with Hoechst 33 342 dye (1 μ g/ml) (Sigma B2261). Cells were then washed in PBS and mounted on glass slides in 'ProLong Gold Antifade' Reagent (Invitrogen). Sixteen-bit fluorescent images were captured with a MetaMorph-driven microscope (Leica DM6000) using a Leica 40 HCX PL APO 1.25-0.75 oil immersion objective (Leica) and a 'CoolSNAP HQ2' camera (Photometrics). Individual images were further processed using Adobe Photoshop software.

Recombinant proteins and *in vitro* and *ex vivo* splicing assays

Recombinant SRSF1 and SRSF6 were produced and purified from baculovirus-infected Sf9 cells as previously reported (60). For *in vitro* splicing, radiolabelled RNAs were synthesized by *in vitro* transcription in the presence of 50 U of T7 RNA polymerase (NEB), 2 μ g of the suitable linearized plasmids and 50 μ Ci [α -³²P] UTP (800 Ci/mmol) in 40 μ l reactions according to the manufacturer conditions. After DNase (RQ1, PROMEGA) digestion, *in vitro* transcripts were purified with sigma columns (Sigma, RTN10-1K) and quantified by Cerenkov counting. Splicing reactions with HeLa nuclear extracts (CilBiotech) and 20 fmol of radiolabelled transcripts were performed under standard conditions as described previously (61). Spliced products were cleaned by proteinase K-treatment and phenol extraction, and analyzed by 6% polyacrylamide gel electrophoresis.

UV cross-linking

For UV cross-linking, radiolabelled RNAs were synthesized by *in vitro* transcription in the presence of 20 U of SP6 RNA polymerase (NEB), 1 μ g of the suitable linearized plasmids and 60 μ Ci [α -³²P] GTP (800 Ci/mmol) in 20 μ l reactions according to the manufacturer's instructions. *In vitro* transcripts were purified by phenol-chloroform extraction, verified by denaturing polyacrylamide-urea gel electrophoresis and quantified by Cerenkov counting. HeLa nuclear extracts were pre-incubated for 15 min at 37°C in a 10 μ l solution containing 10 mM triethylamine (TEA) pH 8, 50 mM KCL, 0.1 mM EDTA, 0.25 mM DTT, 10% glycerol, 7.5 mM ATP, 25 mM creatine phosphate, 1 mM MgCl₂ and 0.025 μ g/ μ l of yeast tRNA. One hundred femtomole of indicated radiolabelled RNA was then added into 10 μ l of interaction buffer (10 mM TEA pH 8, 50 mM KCL, 0.1 mM EDTA, 2.25 mM DTT, 12.6% glycerol, 0.75 mM ATP, 25 mM creatine phosphate, 1 mM MgCl₂, 0.2% NP40, 32 U of RNasin (PROMEGA) and 0.06 μ g/ μ l of BSA). Mixture reactions were then incubated 15 min at 30°C, and subsequently transferred to 96-well plates. Samples were irradiated for 20 min on ice with UV light (254 nm) at a distance of 5 cm. After RNase treatment, 20 μ l of 5 \times SDS gel loading buffer was added in each sample and cross-linked proteins were separated by 12% SDS-PAGE. Gels were fixed (30 min in 50% methanol, 10% acetic acid, then 5 min in 7% methanol, 7% acetic acid, 1% glycerol), dried and revealed by autoradiography.

For two-dimensional gel analysis of SR proteins, UV cross-linked proteins were diluted in a buffer containing 7 M urea, 2 M thiourea, and 4% CHAPS, 4% IPG buffer and 10 mM DTT. Samples were subjected to two-dimensional analysis first with Immobiline Dry strips (pH 3-11 NL GE healthcare) using an Ettan IPGphor3 IEF system (GE healthcare life sciences) according to the manufacturer's instructions. After the second dimension electrophoresis, proteins were transferred onto nitrocellulose membranes. The membrane was first revealed by autoradiography and then SRSF1 and SRSF6 were detected with the above indicated antibodies, followed by ECL staining.

Enzymatic and chemical probing of RNA secondary structure

RNA 2-D structures in solution were probed as follows; 200 ng of transcripts dissolved at a 80 nM concentration in buffer D (20 mM Hepes-KOH, pH 7.9, 100 mM KCl, 0.2 mM EDTA pH 8.0, 0.5 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride, 20% (vol/vol) glycerol) were renatured by 10 min heating at 65°C, followed by slow cooling at room temperature with the addition of 1 μ l of 62.5 mM MgCl₂ to a final concentration of 3.25 mM MgCl₂. After a 10 min pre-incubation at room temperature, RNase T1 (0.02 or 0.0375 U/ μ l) or T2 (0.025 or 0.0375 U/ μ l) was added under conditions conducive to cleavage of single-stranded segments. V1 RNase (2.5 \times 10⁻³ or 5 \times 10⁻³ U/ μ l) was used to cleave double-stranded and stacked residues. DMS [1 μ l of a 1/4 or 1/8 (V/V) DMS/EtOH solution] was employed to modify single-stranded A and C residues and CMCT was used (4 or 5 μ l of a 180 mg/ml solution) to modify single-stranded U and to a lesser extent

G residues. Reactions were stopped as previously described (23). Cleavage and modification positions were identified by primer extension (23). Stable secondary structures having the best fit with experimental data were identified with the Mfold software, version 8.1. Probing data were introduced as a constraint in the search.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We thank A. Krainer for the gifts of antibodies against SRSF6. Thanks to members of the Tazi laboratory for helpful discussions. Special thanks to J. Venables for critical reading of the manuscript.

Conflict of Interest statement. None declared.

FUNDING

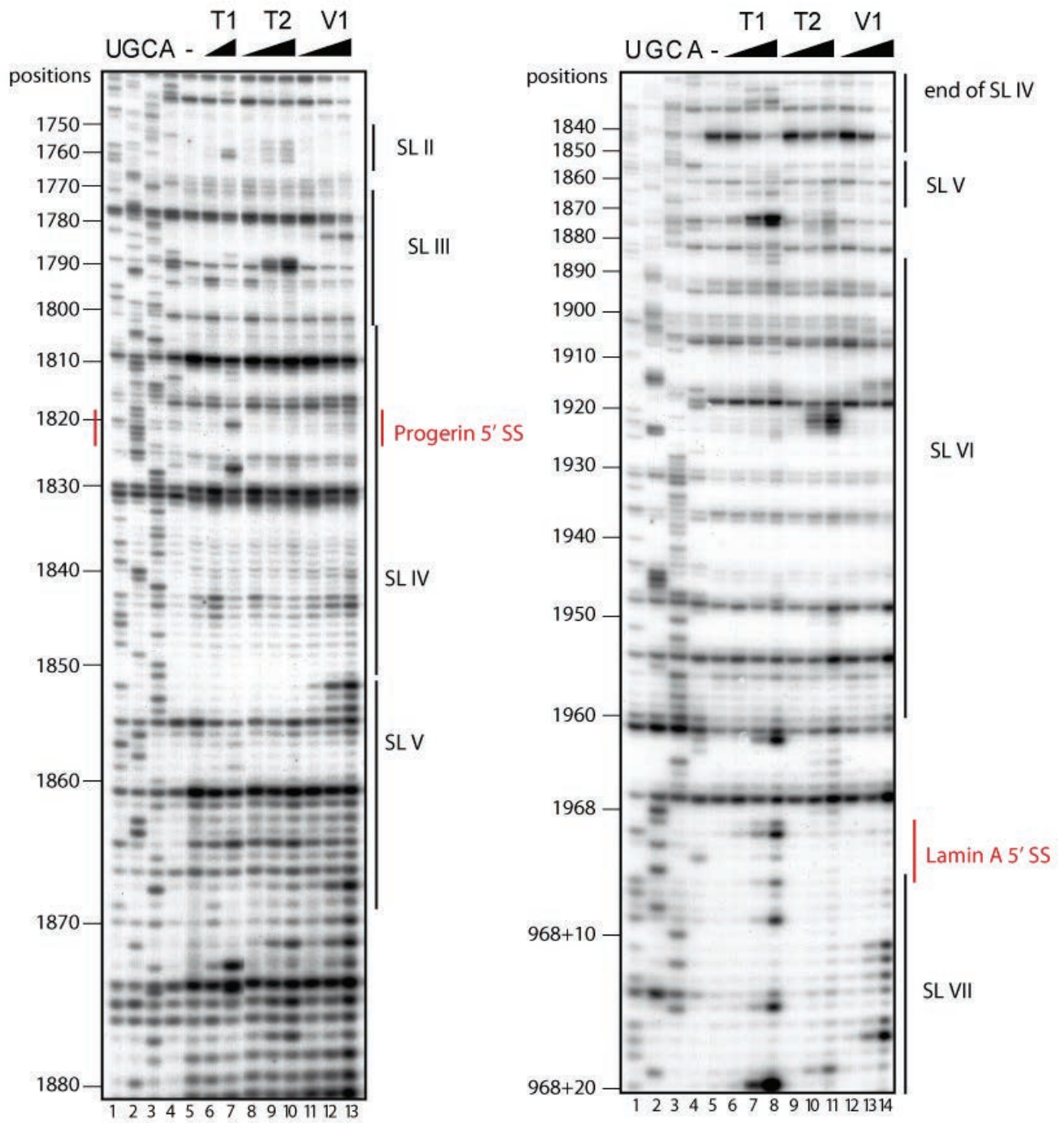
I.C.L.-M. and V.V. were supported by a graduate fellowships from the Ministère Délégué à la Recherche et aux Technologies. M.T. was supported by a post-doctoral fellowship of the Association Française contre les Myopathies (AFM) and CNRS. This work was supported by grants from the Agence Nationale de la Recherche (ANR-05-BLAN-0261-01) and the European Alternative Splicing Network of Excellence (EURASNET, FP6 life sciences, genomics and biotechnology for health). Ministerio de Ciencia e Innovación-Spain, PCTI-FICYT Asturias, European Union (FP7 MicroEnviMet), and Fundación M. Botin.

REFERENCES

- Lin, F. and Worman, H.J. (1993) Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C. *J. Biol. Chem.*, **268**, 16321–16326.
- Nakajima, N. and Abe, K. (1995) Genomic structure of the mouse A-type lamin gene locus encoding somatic and germ cell-specific lamins. *FEBS Lett.*, **365**, 108–114.
- Machiels, B.M., Zorenc, A.H., Endert, J.M., Kuijpers, H.J., van Eys, G.J., Ramaekers, F.C. and Broers, J.L. (1996) An alternative splicing product of the lamin A/C gene lacks exon 10. *J. Biol. Chem.*, **271**, 9249–9253.
- Eriksson, M., Brown, W.T., Gordon, L.B., Glynn, M.W., Singer, J., Scott, L., Erdos, M.R., Robbins, C.M., Moses, T.Y., Berglund, P. *et al.* (2003) Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature*, **423**, 293–298.
- De Sandre-Giovannoli, A., Bernard, R., Cau, P., Navarro, C., Amiel, J., Boccaccio, I., Lyonnet, S., Stewart, C.L., Munnich, A., Le, M.M. and Levy, N. (2003) Lamin A truncation in Hutchinson-Gilford progeria. *Science*, **300**, 2055.
- Dechat, T., Pfleghaar, K., Sengupta, K., Shimi, T., Shumaker, D.K., Limando, L. and Goldman, R.D. (2008) Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes Dev.*, **22**, 832–853.
- Worman, H.J., Fong, L.G., Muchir, A. and Young, S.G. (2009) Laminopathies and the long strange trip from basic cell biology to therapy. *J. Clin. Invest.*, **119**, 1825–1836.
- Worman, H.J. and Bonne, G. (2007) “Laminopathies”: a wide spectrum of human diseases. *Exp. Cell Res.*, **313**, 2121–2133.
- Varela, I., Cadinanos, J., Pendas, A.M., Gutierrez-Fernandez, A., Folgueras, A.R., Sanchez, L.M., Zhou, Z., Rodriguez, F.J., Stewart, C.L., Vega, J.A. *et al.* (2005) Accelerated ageing in mice deficient in Zmpste24 protease is linked to p53 signalling activation. *Nature*, **437**, 564–568.
- Scaffidi, P. and Misteli, T. (2008) Lamin A-dependent misregulation of adult stem cells associated with accelerated ageing. *Nat. Cell Biol.*, **10**, 452–459.
- Shumaker, D.K., Dechat, T., Kohlmaier, A., Adam, S.A., Bozovsky, M.R., Erdos, M.R., Eriksson, M., Goldman, A.E., Khuon, S., Collins, F.S. *et al.* (2006) Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc. Natl Acad. Sci. USA*, **103**, 8703–8708.
- Glynn, M.W. and Glover, T.W. (2005) Incomplete processing of mutant lamin A in Hutchinson-Gilford progeria leads to nuclear abnormalities, which are reversed by farnesyltransferase inhibition. *Hum. Mol. Genet.*, **14**, 2959–2969.
- Dechat, T., Shimi, T., Adam, S.A., Rusinol, A.E., Andres, D.A., Spielmann, H.P., Sinensky, M.S. and Goldman, R.D. (2007) Alterations in mitosis and cell cycle progression caused by a mutant lamin A known to accelerate human aging. *Proc. Natl Acad. Sci. USA*, **104**, 4955–4960.
- Cao, K., Capell, B.C., Erdos, M.R., Djabali, K. and Collins, F.S. (2007) A lamin A protein isoform overexpressed in Hutchinson-Gilford progeria syndrome interferes with mitosis in progeria and normal cells. *Proc. Natl Acad. Sci. USA*, **104**, 4949–4954.
- McClintock, D., Ratner, D., Lokuge, M., Owens, D.M., Gordon, L.B., Collins, F.S. and Djabali, K. (2007) The mutant form of lamin A that causes Hutchinson-Gilford progeria is a biomarker of cellular aging in human skin. *PLoS ONE*, **2**, e1269.
- Haithecock, E., Dayani, Y., Neufeld, E., Zahand, A.J., Feinstein, N., Mattout, A., Gruenbaum, Y. and Liu, J. (2005) Age-related changes of nuclear architecture in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA*, **102**, 16690–16695.
- Rodriguez, S., Coppede, F., Sagelius, H. and Eriksson, M. (2009) Increased expression of the Hutchinson-Gilford progeria syndrome truncated lamin A transcript during cell aging. *Eur. J. Hum. Genet.*, **17**, 928–937.
- Will, C.L. and Luhrmann, R. (2001) Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.*, **13**, 290–301.
- Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.*, **72**, 291–336.
- Soret, J. and Tazi, J. (2003) Phosphorylation-dependent control of the pre-mRNA splicing machinery. *Prog. Mol. Subcell Biol.*, **31**, 89–126.
- Bourgeois, C.F., Lejeune, F. and Stevenin, J. (2004) Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **78**, 37–88.
- Lin, S. and Fu, X.D. (2007) SR proteins and related factors in alternative splicing. *Adv. Exp. Med. Biol.*, **623**, 107–122.
- Hallay, H., Locker, N., Ayadi, L., Ropers, D., Guitte, E. and Branlant, C. (2006) Biochemical and NMR study on the competition between proteins SC35, SRp40, and heterogeneous nuclear ribonucleoprotein A1 at the HIV-1 Tat exon 2 splicing site. *J. Biol. Chem.*, **281**, 37159–37174.
- Marchand, V., Mereau, A., Jacquenet, S., Thomas, D., Mouglin, A., Gattoni, R., Stevenin, J. and Branlant, C. (2002) A Janus splicing regulatory element modulates HIV-1 tat and rev mRNA production by coordination of hnRNP A1 cooperative binding. *J. Mol. Biol.*, **323**, 629–652.
- Tazi, J., Bakkour, N. and Stamm, S. (2009) Alternative splicing and disease. *Biochim. Biophys. Acta*, **1792**, 14–26.
- Pandit, S., Wang, D. and Fu, X.D. (2008) Functional integration of transcriptional and RNA processing machineries. *Curr. Opin. Cell Biol.*, **20**, 260–265.
- Buratti, E., Muro, A.F., Giombi, M., Gherbassi, D., Iaconcig, A. and Baralle, F.E. (2004) RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin ED4 exon. *Mol. Cell Biol.*, **24**, 1387–1400.
- Caceres, J.F. and Kornbliht, A.R. (2002) Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet.*, **18**, 186–193.
- Ast, G. (2004) How did alternative splicing evolve? *Nat. Rev. Genet.*, **5**, 773–782.
- Lear, A.L., Eperon, L.P., Wheatley, I.M. and Eperon, I.C. (1990) Hierarchy for 5' splice site preference determined in vivo. *J. Mol. Biol.*, **211**, 103–115.
- Buratti, E. and Baralle, F.E. (2004) Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol. Cell Biol.*, **24**, 10505–10514.

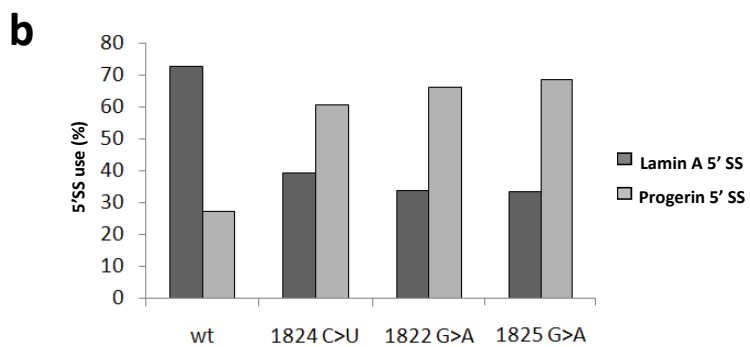
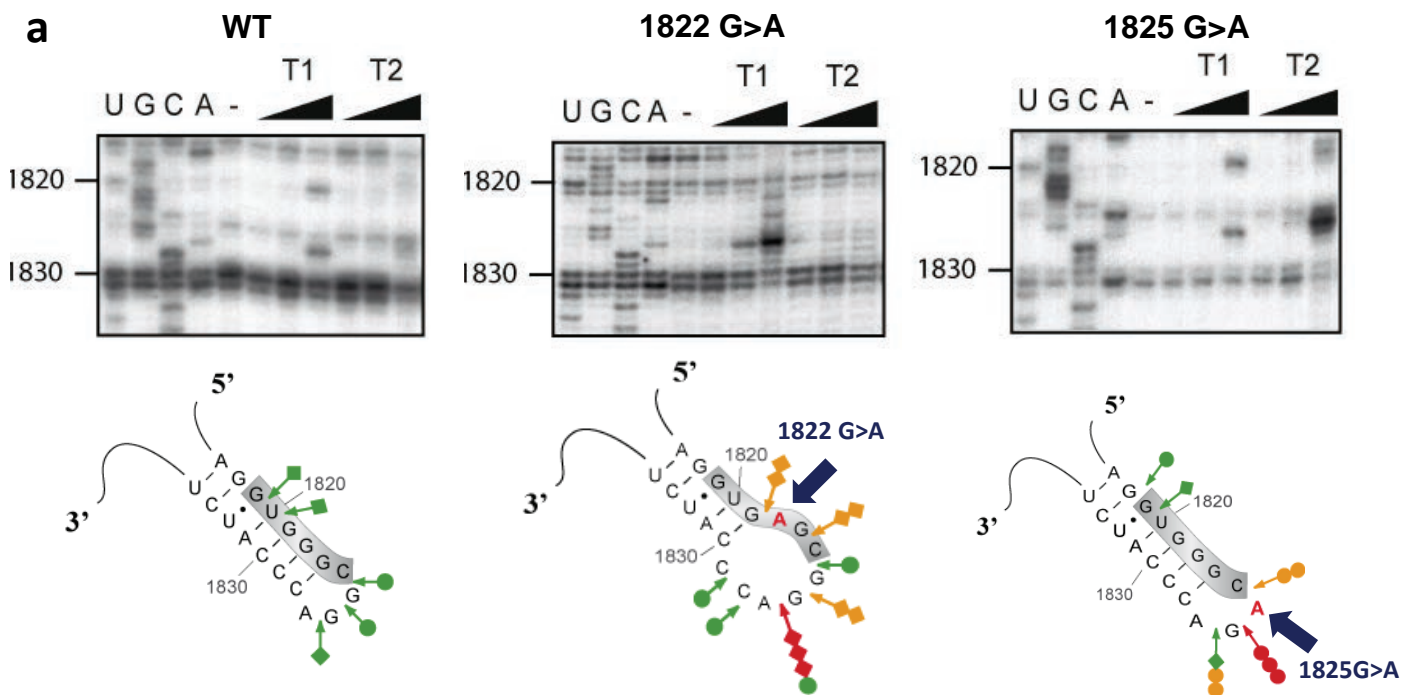
32. Saliou, J.M., Bourgeois, C.F., Ayadi-Ben, M.L., Ropers, D., Jacquenet, S., Marchand, V., Stevenin, J. and Branlant, C. (2009) Role of RNA structure and protein factors in the control of HIV-1 splicing. *Front Biosci.*, **14**, 2714–2729.
33. Cooper, T.A. and Mattox, W. (1997) The regulation of splice-site selection, and its role in human disease. *Am. J. Hum. Genet.*, **61**, 259–266.
34. Baralle, D. and Baralle, M. (2005) Splicing in action: assessing disease causing sequence changes. *J. Med. Genet.*, **42**, 737–748.
35. Gabut, M., Mine, M., Marsac, C., Brivet, M., Tazi, J. and Soret, J. (2005) The SR protein SC35 is responsible for aberrant splicing of the Elalpha pyruvate dehydrogenase mRNA in a case of mental retardation with lactic acidosis. *Mol. Cell Biol.*, **25**, 3286–3294.
36. Marchand, V., Santerre, M., Aigueperse, C., Fouillen, L., Saliou, J.M., Van Dorsselaer, A., Sanglier-Cianferani, S., Branlant, C. and Motorin, Y. (2011) Identification of protein partners of the human immunodeficiency virus 1 tat/rev exon 3 leads to the discovery of a new HIV-1 splicing regulator, protein hnRNP K. *RNA Biol.*, **8**, 325–342.
37. Maenner, S., Blaud, M., Fouillen, L., Savoye, A., Marchand, V., Dubois, A., Sanglier-Cianferani, S., Van, D.A., Clerc, P., Avner, P. *et al.* (2010) 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS Biol.*, **8**, e1000276.
38. Zahler, A.M., Lane, W.S., Stolk, J.A. and Roth, M.B. (1992) SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.*, **6**, 837–847.
39. Labourier, E., Allemand, E., Brand, S., Fostier, M., Tazi, J. and Bourbon, H.M. (1999) Recognition of exonic splicing enhancer sequences by the Drosophila splicing repressor RSF1. *Nucleic Acids Res.*, **27**, 2377–2386.
40. Schneider, M., Will, C.L., Anokhina, M., Tazi, J., Urlaub, H. and Luhrmann, R. (2010) Exon definition complexes contain the tri-snRNP and can be directly converted into B-like pre-catalytic splicing complexes. *Mol. Cell*, **38**, 223–235.
41. Kornblihtt, A.R. (2007) Coupling transcription and alternative splicing. *Adv. Exp. Med. Biol.*, **623**, 175–189.
42. Tuduri, S., Crabbe, L., Conti, C., Tourriere, H., Holtgreve-Grez, H., Jauch, A., Pantescio, V., De Vos, J., Thomas, A., Theillet, C. *et al.* (2009) Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nat. Cell Biol.*, **11**, 1315–1324.
43. Fong, L.G., Vickers, T.A., Farber, E.A., Choi, C., Yun, U.J., Hu, Y., Yang, S.H., Coffinier, C., Lee, R., Yin, L. *et al.* (2009) Activating the synthesis of progerin, the mutant prelamin A in Hutchinson-Gilford progeria syndrome, with antisense oligonucleotides. *Hum. Mol. Genet.*, **18**, 2462–2471.
44. Brunel, C., Marquet, R., Romby, P. and Ehresmann, C. (2002) RNA loop-loop interactions as dynamic functional motifs. *Biochimie*, **84**, 925–944.
45. Hiller, M., Zhang, Z., Backofen, R. and Stamm, S. (2007) Pre-mRNA secondary structures influence exon recognition. *PLoS Genet.*, **3**, e204.
46. Shepard, P.J. and Hertel, K.J. (2008) Conserved RNA secondary structures promote alternative splicing. *RNA*, **14**, 1463–1469.
47. Varani, L., Hasegawa, M., Spillantini, M.G., Smith, M.J., Murrell, J.R., Ghetti, B., Klug, A., Goedert, M. and Varani, G. (1999) Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc. Natl Acad. Sci. USA*, **96**, 8229–8234.
48. Grover, A., Houlden, H., Baker, M., Adamson, J., Lewis, J., Prihar, G., Pickering-Brown, S., Duff, K. and Hutton, M. (1999) 5' splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10. *J. Biol. Chem.*, **274**, 15134–15143.
49. Fong, L.G., Ng, J.K., Lammerding, J., Vickers, T.A., Meta, M., Cote, N., Gavino, B., Qiao, X., Chang, S.Y., Young, S.R. *et al.* (2006) Prelamin A and lamin A appear to be dispensable in the nuclear lamina. *J. Clin. Invest.*, **116**, 743–752.
50. Yang, S.H., Qiao, X., Farber, E., Chang, S.Y., Fong, L.G. and Young, S.G. (2008) Eliminating the synthesis of mature lamin A reduces disease phenotypes in mice carrying a Hutchinson-Gilford progeria syndrome allele. *J. Biol. Chem.*, **283**, 7094–7099.
51. Yang, S.H., Bergo, M.O., Toth, J.I., Qiao, X., Hu, Y., Sandoval, S., Meta, M., Bendale, P., Gelb, M.H., Young, S.G. and Fong, L.G. (2005) Blocking protein farnesyltransferase improves nuclear blebbing in mouse fibroblasts with a targeted Hutchinson-Gilford progeria syndrome mutation. *Proc. Natl Acad. Sci. USA*, **102**, 10291–10296.
52. Hua, Y., Sahashi, K., Hung, G., Rigo, F., Passini, M.A., Bennett, C.F. and Kraimer, A.R. (2010) Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes Dev.*, **24**, 1634–1644.
53. Tazi, J., Durand, S. and Jeanteur, P. (2005) The spliceosome: a novel multi-faceted target for therapy. *Trends Biochem. Sci.*, **30**, 469–478.
54. Scaffidi, P. and Misteli, T. (2005) Reversal of the cellular phenotype in the premature aging disease Hutchinson-Gilford progeria syndrome. *Nat. Med.*, **11**, 440–445.
55. Soret, J., Gabut, M. and Tazi, J. (2006) SR proteins as potential targets for therapy. *Prog. Mol. Subcell. Biol.*, **44**, 65–87.
56. Soret, J., Bakkour, N., Maire, S., Durand, S., Zekri, L., Gabut, M., Fic, W., Divita, G., Rivalle, C., Dauzonne, D. *et al.* (2005) Selective modification of alternative splicing by indole derivatives that target serine-arginine-rich protein splicing factors. *Proc. Natl Acad. Sci. USA*, **102**, 8764–8769.
57. Bakkour, N., Lin, Y.L., Maire, S., Ayadi, L., Mahuteau-Betzer, F., Nguyen, C.H., Mettling, C., Portales, P., Grierson, D., Chabot, B. *et al.* (2007) Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance. *PLoS Pathog.*, **3**, 1530–1539.
58. Nagy, A., Gertsenstein, M., Vintersten, K. and Behringer, R. (2006) Preparing mouse embryo fibroblasts. *Cold Spring Harb. Protoc.*, doi:10.1101/pdb.prot4398.
59. Durocher, Y., Perret, S. and Kamen, A. (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res.*, **30**, E9.
60. Allemand, E., Gattoni, R., Bourbon, H.M., Stevenin, J., Caceres, J.F., Soret, J. and Tazi, J. (2001) Distinctive features of Drosophila alternative splicing factor RS domain: implication for specific phosphorylation, shuttling, and splicing activation. *Mol. Cell Biol.*, **21**, 1345–1359.
61. Damier, L., Domenjoud, L. and Branlant, C. (1997) The D1-A2 and D2-A2 pairs of splice sites from human immunodeficiency virus type 1 are highly efficient in vitro, in spite of an unusual branch site. *Biochem. Biophys. Res. Commun.*, **237**, 182–187.

Supplemental Figure 1



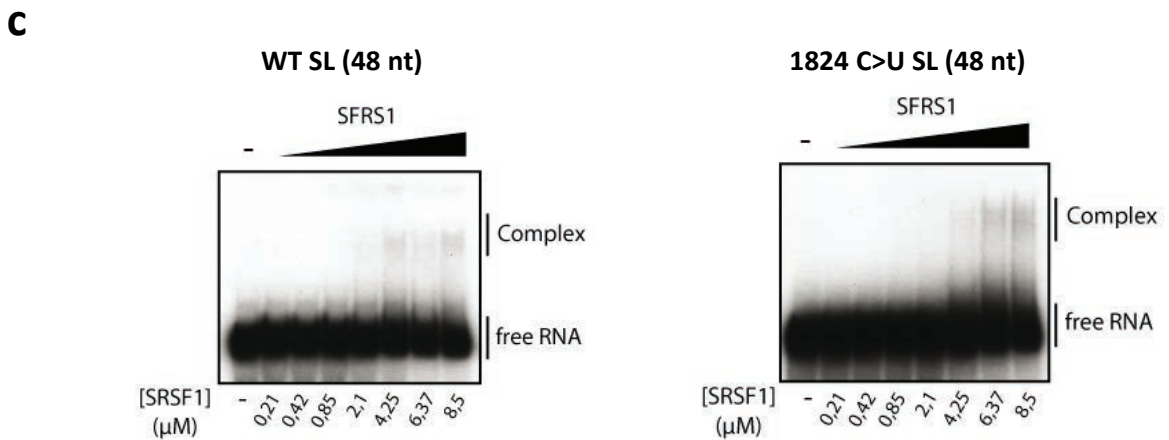
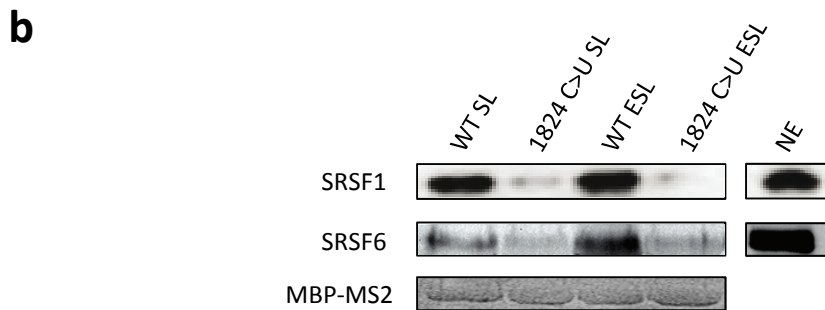
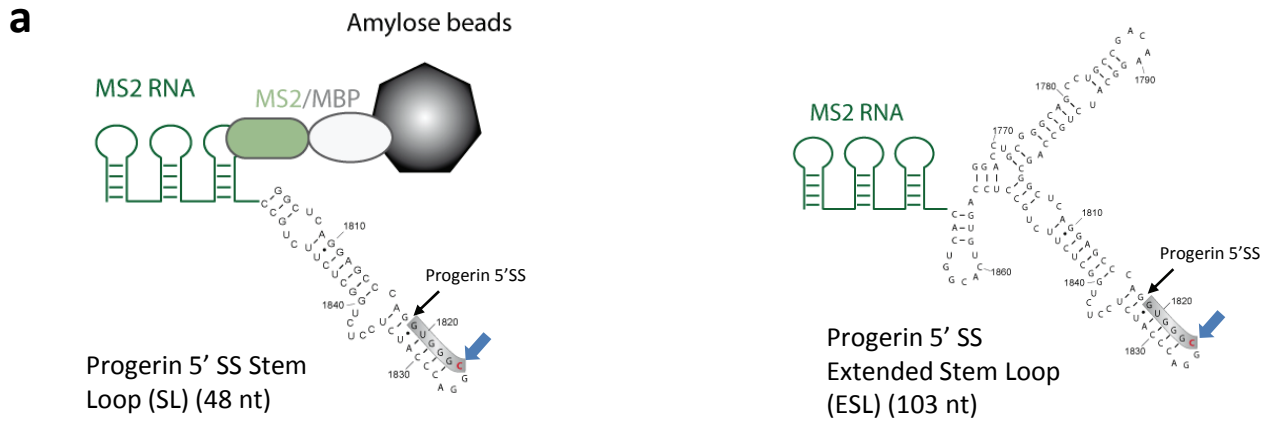
Legend to supplementary Figure 1: Probing of the structure of the WT LMNA pre-mRNA in the region containing exon 11 5' splice sites. RNAs of 638 nts were used to determine the structure around LMNA exon 11 alternative 5' SSs. These RNAs contain exon 11, intron 11 and the coding sequence of exon 12. Digestions with RNases T1 (lanes 6–8), T2 (lanes 9–11) or V1 (lanes 12–14) were carried out as described in the 'Materials and methods' section. As a control, undigested RNA was fractionated in parallel (lanes 5). Position of the cleavages were identified by reverse transcription of digested RNAs. The two primers used for this study allowed the analysis of the secondary structure from positions 1750 to 1968+20 of the LMNA RNA (see Fig. 2). Splice Site (SS) and Stem Loop structures (SL) identified as in the structure model in Fig. 2 are indicated on the right side of the gel. Positions in the LMNA RNA are indicated on the left side.

Supplemental Figure 2



Legend to supplementary Figure 2: The 1822 G>A and 1825G>A mutations both open the terminal stem-loop structure close to the progerin 5' SS and lead to the predominant use of the progerin 5' SS. **(a)** Secondary structures proposed for the WT and mutated human progerin SL RNA based on thermodynamic considerations and positions and yields of enzymatic cleavages. T1, T2 and V1 RNase cleavages are represented by arrows surmounted by squares, dots and diamonds respectively. Red, orange and green symbols indicate a strong, medium or weak cleavage, respectively. The position of the progerin 5' SS is indicated. Autoradiographs show primer extension analyses of WT and mutated c.1822G>A and c.1825G>A human progerin SL RNA cleaved by increasing amounts of T1, T2 and V1 RNases. Lanes U, G, C, and A correspond to the sequencing ladder. **(b)** Relative ratios of lamin A *versus* progerin 5'-splice sites utilization in WT c.1824C>U, c.1822G>A and c.1825G>A b-globin-LMNA reporter RNAs in HeLa nuclear extract, expressed as percentages of the spliced pre-mRNAs.

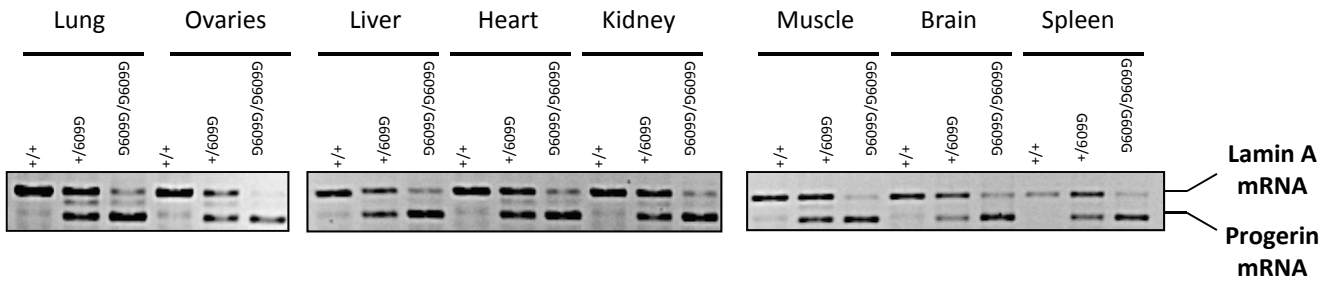
Supplemental Figure 3



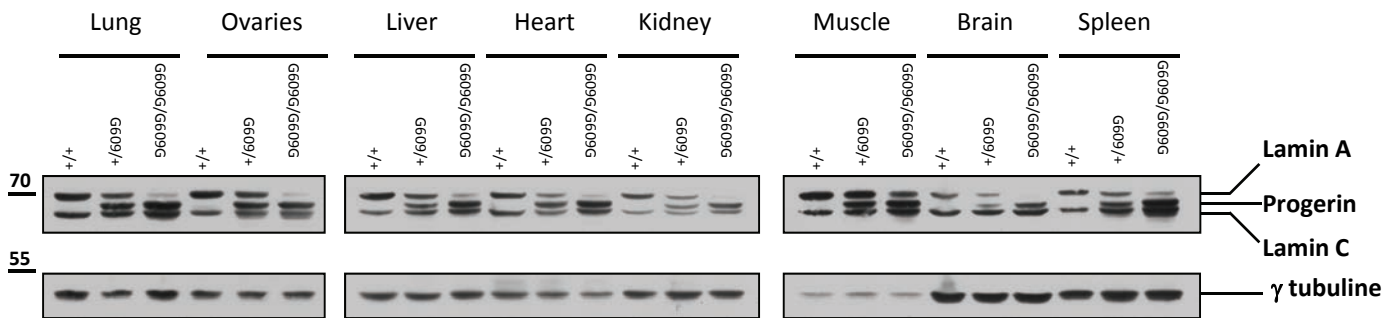
Legend to supplementary Figure 3: Association of SRSF6 and more particularly SRSF1 with the SL and ESL RNAs are decreased in the presence of the c.1824C>U mutation. **(a)** Scheme of the hybrid RNAs used for the purification of proteins associated with Progerin Stem Loop in HeLa cell nuclear extracts. The Progerin Stem Loop (SL) and Extended Stem Loop (ESL) RNAs, both contain the Progerin 5' SS (WT or mutated). These RNAs were fused with three MS2 coat protein binding sites (MS2). The retention of hybrid RNAs on amylose beads was mediated by the MS2/MBP fusion protein. The experiment was performed with WT and mutated Progerin SL RNAs, WT and mutated ESL RNAs and with the MS2 RNA alone. The 1824C>U mutation is indicated by a wide blue arrow. **(b)** Western blot analysis of the eluted proteins. Equal amounts of eluted fractions were separated on SDS PAGE 10%, transferred on nitrocellulose membrane and hybridized using an antibodies specific for the SRSF1 and SRSF6 proteins. Ponceau staining of the nitrocellulose membrane allowed the detection of the MBP-MS2 protein and was used to check for equal loading. NE lane: assay for the presence of SRSF1 and SRSF6 proteins in the HeLa Nuclear extract used to perform the purification. **(c)** Test for the direct binding of SF1 on the WT and mutated Progerin SL RNAs by EMSA. Ribonucleoprotein complexes were formed by the incubation of 47 fmoles of 5'-labeled WT or 1824 C>U SL RNA in the presence of increasing amount of recombinant SRSF1. The final concentration of the recombinant SRSF1 protein is indicated below the gels. Lanes (-) are control experiments in the absence of protein. The complexes were formed at room temperature in the presence of 2 μ g of yeast tRNAs under the conditions described in the "Materials and Methods" section. Complexes were fractionated by electrophoresis in a native 8% polyacrylamide gel

Supplemental Figure 4

a



b



Legend to Supplementary Figure 4: (a) Analysis of RNA by RT-PCR to determine the relative expression levels of lamin a and progerin mRNAs in various tissues from wild (+/+), heterozygous *Lmna*G069G/+ and homozygous *Lmna*G069G/G609G mice. (b) Western blot of extracts from indicated tissues revealed with an anti-lamin A/C antibody. The bottom panels show the loading control (γ - tubulin).

III.2. Etude de l'influence des mutations 1868C>G et 1822G>A sur les taux relatifs d'utilisation des sites 5'SS du pré-ARNm LMNA : présentation de la publication n°2.

Ayant réalisé une étude approfondie des mécanismes de régulation de l'utilisation du site 5'SS progéine en présence de la mutation la plus fréquente chez les patients atteints de la progéria (1824C>U), nous avons entrepris des travaux similaires sur l'effet des deux autres mutations 1822G>A et 1868C>G, minoritaires mais qui conduisent respectivement à la production de progéine et de lamine A Δ 35.

III.2.1. Analyse de l'effet sur la structure secondaire du pré-ARNm LMNA des mutations 1822G>A et 1868C>G

Etant donné que nous avons observé un effet marquant sur l'accessibilité du site 5'SS progéria des mutations 1824C>U et 1822G>A (publication n°1), il nous a paru important de savoir si la mutation 1868C>G modifiait aussi fortement la structuration et donc l'accessibilité du site 5'SS LA Δ 35. Nous avons donc entrepris une étude expérimentale de la structure secondaire adoptée par l'ARN LMNA 638 portant la mutation 1868C>G. L'analyse structurale réalisée sur l'ARN LMNA 638 montrait qu'en l'absence de la mutation 1868C>G, le site 5'LA Δ 35, qui n'est pas utilisé, est situé en amont d'un segment simple brin très accessible aux RNases T1 et T2 spécifiques des régions en simple brin, et que ce site est lui-même peu accessible à ces enzymes et donc vraisemblablement situé dans un élément de structure secondaire. Une petite structure tige-boucle incluant ce site peut en effet être proposée (tige boucle V) (figure III-7). En présence de la mutation 1868C>G, on note une augmentation de l'accessibilité à la RNase T1 de deux résidus G de cette structure tige-boucle, reflétant une déstabilisation partielle. L'effet est néanmoins moins marquant que pour les mutations 1824C>U et 1822G>A.

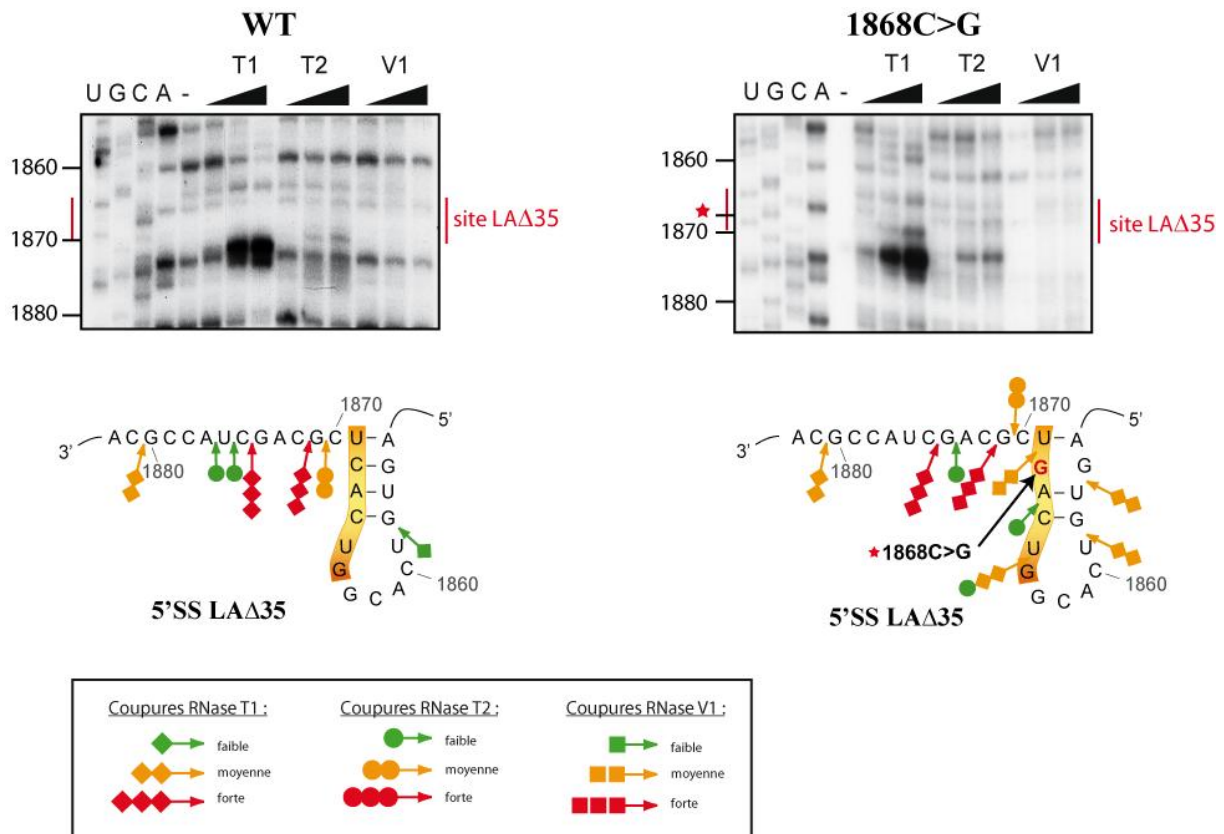


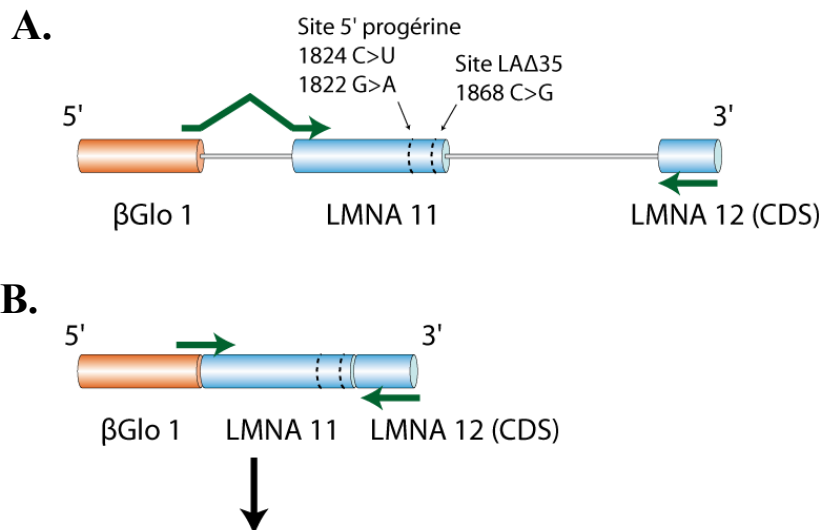
Figure III-7 : Analyse expérimentale de la structure secondaire de la région contenant le site 5'SS LAΔ35. L'analyse a été réalisée par hydrolyse enzymatique (RNase T1, T2 et V1) des ARN LMNA 638 WT et 1868C>G en conditions natives dans les mêmes conditions que pour l'ARN LMNA 638 présentées précédemment (figure III-5). La position de la mutation 1868C>G est indiquée par une étoile. La région présentée dans l'autoradiogramme contient le site 5'SS LAΔ35, qui est indiqué en rouge. La numérotation adoptée est celle de l'ARNm LMNA. Les modèles de structure secondaire correspondant à chaque autoradiogramme sont présentés en dessous des autoradiographies. La partie intronique du site 5'SS LAΔ35 est surlignée en orange, et la mutation 1868C>G est indiquée par une flèche noire. Les hydrolyses enzymatiques sont représentées par des icônes comme indiqué dans l'encadré.

III.2.2. Identification des protéines SR impliquées dans la régulation de l'utilisation des sites 5' progérine et LAΔ35 en absence et en présence des mutations 1822G>A et 1868C>G

Ayant montré dans la publication n°1 que les protéines SR pouvaient moduler le taux d'utilisation du site 5'SS progérine en présence de la mutation 1824C>U, il nous a paru important de tester l'effet des protéines SR sur le taux d'utilisation des sites 5'SS progérine et 5'SS LAΔ35 en présence des mutations 1822G>A et 1868C>G. Pour cela, nous avons utilisé deux approches : (1) l'étude *in cellulo* de l'effet de la surexpression des différentes protéines SR sur le taux d'utilisation des 3 sites 5'SS lamine A, progérine et LAΔ35 dans la construction *βGlo*-LMNA en présence ou en absence de ces mutations et (2) l'étude de l'ajout de protéines SR recombinantes dans des expériences d'épissage *in vitro*

du transcrit β Glo-LMNA en présence ou en absence de ces mutations. Dans ce but, j'ai produit les constructions β Glo-LMNA 1822G>A et 1868C>G par mutagenèse dirigée.

Les constructions β Glo-LMNA, WT ou porteuse d'une des deux mutations, ont été co-transfectées dans des cellules HeLa avec un plasmide pXJ41 permettant de surexprimer une protéine SR, ou le plasmide pXJ41 vide en tant que contrôle. Nous avons ainsi testé l'effet des protéines SR SRSF1 et SRSF5. Les ARN totaux ont ensuite été extraits après 24 h de culture et les taux relatifs d'utilisation des différents sites 5'SS lamine A, progérine et LA Δ 35 ont été évalués par RT-PCR semi-quantitative. L'étape de PCR a été réalisée grâce à un oligonucléotide sens chevauchant le premier exon de la β -globine et l'exon 11 LMNA, et un oligonucléotide antisens s'hybridant au niveau de la partie codante de l'exon 12 (figure III-8). L'utilisation de ce couple d'oligonucléotides permettait la détection simultanée des produits d'épissage produits par utilisation des sites 5'SS progérine, LA Δ 35 et lamine A.



Taille des produits PCR attendus :

BGlo LMNA WT : 304 pb
 BGlo LMNA 1824C>U et 1822G>A : 154 pb
 BGlo LMNA 1868C>G : 199 pb

Figure III-8 : Localisation des sites d'hybridation des oligonucléotides utilisés pour l'analyse par RT-PCR des ARNm issus du minigène β Glo-LMNA obtenus *in cellulo*. Représentation de l'hybridation des oligonucléotides amorce utilisés lors de l'étape de PCR sur la construction β Glo-LMNA (A.) ou sur l'ARNm mature (B.).

Ces expériences ont été réalisées avec l'aide de C. Aigueperse et les résultats obtenus sont décrits dans le manuscrit qui suit. Cette approche nous a permis de montrer qu'*in cellulo*, les deux protéines SRSF1 et SRSF5 induisent une augmentation forte du taux d'utilisation du site 5'SS lamine A aux dépens de l'utilisation des sites 5'SS progérine et LA Δ 35. Afin de savoir si les effets observés *in cellulo* étaient

directs et ne résultaient pas d'une cascade d'événements cellulaires, nous avons testé l'effet de ces protéines en épissage *in vitro*. Les données obtenues *in vitro* confirment un effet activateur direct de la protéine SRSF5 sur le site 5'SS lamine A. Nous avons donc testé l'effet de la surexpression de la protéine SRSF5 dans des fibroblastes de patients HGPS (mutation 1824C>U). Les résultats préliminaires obtenus suggèrent que cette surexpression permet de limiter l'utilisation du site 5'SS progérine. Des expériences complémentaires sont en cours.

L'ensemble des données concernant la régulation des sites 5'SS progérine et LA Δ 35 en présence des mutations 1822G>A et 1868C>G fait l'objet d'un manuscrit en préparation pour soumission à la revue *European Journal of Human Genetics* présenté ci-après, dont je suis premier auteur.

MANUSCRIT PUBLICATION N°2

Identification of SRSF5 as a regulator of LMNA
exon 11 splicing in the context of progeria
inducing mutations

Valentin VAUTROT, Christelle AIGUEPERSE, James STEVENIN, Christiane BRANLANT and Isabelle BEHM-ANSMANT

In preparation for European Journal of Human Genetics

Identification of SRSF5 as a regulator of LMNA exon 11 splicing in the context of progeria inducing mutations

Valentin VAUTROT¹, Christelle AIGUEPERSE¹, James STEVENIN², Christiane BRANLANT^{1*} and Isabelle BEHM-ANSMANT^{1*}

¹UMR 7365 CNRS-UL IMoPA (Ingénierie Moléculaire et Physiopathologie Articulaire), Biopôle de l'Université de Lorraine, 9 Avenue de la forêt de Haye, 54505 Vandoeuvre-lès-Nancy

²University of Strasbourg, IGBMC Department of Functional Genomics and Cancer, CNRS UMR 7104, INSERM U 964, 1 Rue Laurent Fries, 67404 Illkirch Cedex, France

Running title: SRSF5 regulates LMNA splicing

* To whom correspondence should be addressed:

isabelle.behm@univ-lorraine.fr

christiane.branlant@univ-lorraine.fr

ABSTRACT

The Hutchinson Gilford Progeria Syndrom (HGPS), also called progeria, is a rare genetic disease leading to accelerated aging. So far three different mutations leading to HGPS have been identified within exon 11 of the LMNA gene: the two most frequent ones corresponding to the silent 1824C>T (G608G) and the 1822G>A (G608S) enhance the use of an internal alternative 5' splice site used at low level in normal human cells and lead to the production of a truncated dominant negative lamin A protein (progerin or LA Δ 50). The third mutation corresponds to the 1868C>G (T623S) mutation. It creates an internal 5'-splice site located and induces the production of an alternative form of progerin called LA Δ 35 whose accumulation leads to an attenuate form of progeria. We previously studied the molecular bases underlying splicing alteration in the presence of the most frequent 1824C>T mutation (¹⁵). In this study we investigated splicing regulation of the LMNA pre-mRNA in the presence of the less prominent progeria 1822G>A and 1868C>G mutations. By *in cellulo* assays we identified the SR protein SRSF5 as a common regulator of both the alternative and cryptic 5'-splice sites. We demonstrated SRSF5 direct implication in LMNA exon 11 splicing regulation by *in vitro* splicing assays. Moreover we showed that SRSF5 overexpression in HGPS patient cells allows the recovery of a nearly WT splicing pattern and reduces progerin production. This general role of SRSF5 in LMNA exon 11 splicing regulation opens new therapeutical perspectives for premature aging.

INTRODUCTION

The Hutchinson Gilford Progeria Syndrom (HGPS), also called progeria, is a rare genetic disease leading to accelerated aging. Progeria is caused by mutations in the LMNA gene, coding lamin A, a protein of the nuclear envelope substructure called lamina. The most frequent mutation responsible for progeria is the silent c.1824C>T (G608G) mutation, leading to the excision of intron 11 by utilization of an alternative 5'-splice site used at low level in normal human cells (^{1, 2}). As this site is located 150 nts upstream of the canonic one, a truncated version of the lamin A protein is produced (progerin or LA Δ 50). Progerin retains the C-terminal CAAX box of Lamin A, which undergoes methyl esterification and farnesylation, but it lacks amino acids essential for a specific cleavage step involved in lamin A maturation (^{1, 2, 3}). Progerin is therefore permanently farnesylated and carboxymethylated which leads to its abnormal anchoring to the nuclear membrane throughout the cell cycle (^{4, 3, 5}). The nuclear accumulation of progerin interferes with the integrity of the nuclear lamina, causing misshapen nuclei, and also adversely affects other important cellular processes, such as interphase chromatin, mitosis and cell proliferation (^{2, 1, 6, 7, 5}). Two other mutations within the LMNA gene also lead to progeria: the c.1822G>A (G608S) mutation activates the same alternative 5'-splice site and induces progerin production, whereas the c.1868C>G (T623S) mutation activates a cryptic 5'-splice site located 105 nts upstream from the canonic 5'-splice site and induces the production of an alternative form of progerin called LA Δ 35 (^{2, 8}). The accumulation of LA Δ 35 leads to an attenuate form of progeria (⁸). This form was described only once and the patient presented all the progeria symptoms when he died at the age of 45 but the onset of the disease was late (age of 12) and it took a longer time to develop full-blown HGPS (⁸).

The spliceosome is a dynamic machinery responsible for removing the vast majority of pre-mRNA introns (⁹). The U1, U2, U4/U6, and U5 snRNPs are the main building blocks of this machinery. Each snRNP consists of an snRNA (or two in the case of U4/U6) and a variable number of complex-specific proteins. The five small nuclear RNAs (U1, U2, U4, U5, and U6) have dynamic specific interactions via base pairing with the pre-mRNA (¹⁰), and notable among these are the base pairing between U1 snRNA and the 5' splice site. Indeed, the hydrogen bonding pattern between the 5' splice site and U1 snRNA determines the intrinsic strength of the 5' splice site and thus contributes to its recognition and frequency of usage. However, the information contained in splice sites is not sufficient for splicing regulation and RNA-protein complexes which form on pre-mRNAs also help exon recognition (¹¹). The majority of splicing regulatory proteins in these complexes belongs to the heterogeneous ribonucleoprotein (hnRNP) and serine-arginine (SR)-protein families. HnRNPs were initially defined as the major proteins binding to RNA polymerase II nascent RNA transcripts in the form of ribonucleoprotein particles. SR-proteins containing arginine-serine rich domains play essential roles in the splicing reaction (^{11, 12, 13, 14}). HnRNPs and SR-proteins contain both RNA-binding and protein-protein interaction domains and bind with low specificity to accessible parts of the pre-mRNA.

In a previous study we deciphered the molecular basis for the alternative 5'-splice site reinforced utilization in the presence of the progeria c.1824C>T mutation (¹⁵). In this study we investigated the molecular basis for splicing alterations in the presence of the progeria c.1822G>A and c.1868C>G mutations. By *in vitro* and *in cellulo* assays we identified the SR protein SRSF5 as a common and direct regulator of both the alternative and cryptic 5'-splice sites. We also show that its overexpression in HGPS patient fibroblasts allows the recovery of a nearly WT splicing pattern and strongly reduces progerin production

MATERIALS AND METHODS

Plasmids constructs

pCDNA3.1/V5-His-TOPO plasmids containing the WT or c.1824C>T *βGlo*-LMNA construct used for splicing experiments were previously described (¹⁵). The pCDNA3.1/V5-His-TOPO plasmids containing the 1822G>A and 1868C>G *βGlo*-LMNA constructs were produced by site-directed mutagenesis (oligonucleotide sequences are available on request). The pXJ41 plasmids used for SR protein overexpression were described previously (¹⁶).

Cell lines, culture and transfection

Hela S3 cells were grown in Dulbecco's modified eagle medium media (Gibco) complemented with 10% fetal calf serum and antibiotics. Cells from a female HGPS patient (AG11513) (Coriell Cell Repositories, Camden, NJ) were grown in MEM (Gibco) complemented with 15% fetal calf serum. Cells were plated in six-well plates at ~200 000 cells per well, and after 24h co-transfected with 1,5 μg of plasmid expressing *βGlo*-LMNA constructs and 50 ng of plasmid expressing SR proteins using JetPEI (Polyplus) according to the manufacturer's instructions. Cells were collected 36h later.

Recombinant proteins and *in vitro* and *ex vivo* splicing assays

Recombinant SR proteins were produced and purified from baculovirus-infected Sf9 cells as previously reported (¹⁷). *In vitro* splicing reactions were realized as previously described (¹⁵). Spliced products were analyzed by 6% polyacrylamide gel electrophoresis, and radioactive bands were quantified using the ImageQuant software.

For *ex vivo* splicing, RNA were extracted using Trizol reagent (Ambion/Life technologies), then treated with RQ1 DNase (Promega). 1,5 μg of RNA were reverse transcribed using random hexamer oligonucleotides and MMLV retrotranscriptase (Promega) according to the

manufacturer's instructions. PCR were carried out with 25 μ M dNTP mix and 2,5U DreamTaq DNA pol (Thermo scientific) (oligonucleotides sequences are available on request). Denaturation, annealing and extension steps were performed for 30 s at 94, 65 and 72°C, respectively, for 35 cycles. The splicing products were fractionated on 2% agarose gel and quantified in four different experiments using GeneSnap acquisition and GeneTools analysis softwares (Syngene). HeLa and fibroblasts proteins were extracted in cell lysis buffer containing Tris 50 mM pH 8, NaCl 150 mM, Igepal 1%, EDTA 1 mM and Halt protease and phosphatase inhibitor cocktail EDTA-free (Thermo scientific). The supernatant was recovered for western blotting to verify protein surexpression or the progerin/Lamin A ratio. Proteins were fractionated on 10% SDS-PAGE gels and transferred to nitrocellulose membrane 0,45 μ m (Biorad). For western blotting, the following antibodies were used: anti-SRSF1 (Santa Cruz #sc-73026), anti-phosphoepitope SR proteins (Santa Cruz #sc-13509), anti Lamin A/C (Millipore #MAB3211), anti-progerin (Santa Cruz #sc-81611) and anti- β tubulinI (Sigma # T7816).

RESULTS

The SR proteins SRSF1 and SRSF5 regulate the utilization of both progerin and LA Δ 35 splice sites in HeLa cells

The c.1822G>A (G608S) and c.1868C>G (T623S) single-base substitutions in exon 11 of the LMNA gene potentiate usage of an alternative 5' splice site (hereafter referred to as the progerin 5'SS) and a cryptic 5' splice site (hereafter referred to as the LA Δ 35 5'SS), respectively. Computational analyses to score the strength of either the wild type or the mutated forms of the progerin and LA Δ 35 5'SS in comparison with the downstream canonic 5'SS that is used to produce lamin A (hereafter referred to as the lamin A 5'SS), revealed that the scores of both progerin and LA Δ 35 5'SS are lower than the lamin A 5'SS (Figure 1B). In

particular the score of the wild type form of the LA Δ 35 5'SS is almost null which greatly correlates with the fact that this cryptic site is normally not used in human cells. While the c.1822G>A and c.1868C>G augment the score of the progerin and LA Δ 35 5'SS, it is nevertheless not sufficient to make it comparable to the lamin A 5'SS which has a close to perfect score (Figure 1B). However these mutations result in the production of more progerin mRNAs than lamin A mRNA (see below), suggesting that they likely enhance recognition of 5'SS by other mechanisms than simply by enhancing base pairing with U1 snRNP.

Therefore we investigated whether SR proteins could regulate *in cellulo* the utilization of the progerin c.1822G>A and LA Δ 35 c.1868C>G 5'SS. To this goal we used the *β Glo-LMNA* minigene that we previously described (¹⁵) (Figure 2A). In this minigene the LMNA genomic sequence from exons 11 through 12 has been cloned downstream from the β -globin intron 1 in the β glo3S plasmid (¹⁸). The HGPS c.1822G>A and c.1868C>G mutations were introduced by site-directed mutagenesis into the wild-type *β Glo-LMNA* minigene. As expected, the transient transfection of the 1822G>A *β Glo-LMNA* reporter in HeLa cells gave rise to a 150 nucleotide deletion in the LMNA splicing product due to the predominant use of the progerin 5'SS, whereas only the lamin A 5'SS was selected when using the WT LMNA reporter (Figure 2B-C, lanes 7 versus 1). In agreement with the computational analysis, the progerin 1822G>A 5'SS was slightly stronger than the progerin 1824C>T 5'SS and led to slightly more progerin mRNA production (Figure 2B-C, lanes 7 versus 4). Moreover, the transient transfection of the 1868C>G *β Glo-LMNA* reporter gave rise to a 105 nucleotide deletion in almost half of the LMNA splicing products (Figure 2B-C, lane 10). To test for the impact of SR protein on LMNA splicing, constructs allowing the overexpression of various SR proteins were transiently co-transfected with WT or mutant *β Glo-LMNA* reporters in HeLa cells. The efficiency of the overexpression was checked by western blot using dedicated antibodies (Supplementary Figure 1B). We clearly observed that the overexpression of SRSF5

and SRSF1 greatly inhibits the use of both the progerin 1822G>A and LA Δ 35 1868C>G 5'SS, whereas the selection of the lamin A 5'SS was activated (Figure 2B-C, lanes 7-8 and 10-11). Similar effects of SRSF5 and SRSF1 were also observed when using the 1824C>T *β Glo*-LMNA reporter (Figure 2B-C, lanes 4-6).

To check that the level of overexpression of the SR proteins was in a physiological range, dose effect experiments were conducted in HeLa cells. In these experiments increasing amounts (0, 10, 50 and 250 ng) of the pXJ41-SRSF5 plasmid allowing the overexpression of SRSF5 were co-transfected with either the 1822G>A or 1868C>G *β glo*-LMNA reporter and the impact on the rate of utilization of the mutant 5' splice sites was analyzed. A great correlation between the rate of use of the mutant 5' splice sites and the quantity of pXJ41-SRSF5 plasmid was observed (Figure 2D). The quantity of plasmid (50 ng) used for the experiment shown in Figure 2B-C gave an intermediate effect indicating that we did not saturate the cells by a too high overexpression rate of SRSF5. Similar results were obtained for SRSF1 (data not shown). As a negative control a dose effect experiment was performed using the pXJ41-SRSF7 plasmid which allows the overexpression of the SRSF7 protein that does not regulate the rate of utilization of the mutant splice-sites (Figure 2D).

As these experiments could not discern whether SRSF5 and SRSF1 act as repressors of the progerin and LA Δ 35 5'SS and/or as activators of the lamin A 5'SS, we mutated the first G of the lamin A 5'SS (1969G>A mutation) in the wild type and mutated substrates and tested the impact of SRSF5 and SRSF1 overexpression on the splicing pattern in HeLa cells. The introduction of this mutation in the various *β glo*-LMNA constructs leads to the total abolition of the use of the lamin 5'SS and to the exclusive use of either the progerin or the LA Δ 35 5'SS (Figure 2B-C lane 13 and Supplementary Figure 1A, lanes 1, 4 and 7). In good agreement with the computational analysis the introduction of the 1969G>A mutation in the WT *β glo*-LMNA minigene leads to the exclusive use of the progerin 5'SS and confirms the

nearly null score of the LA Δ 35 5'SS in the absence of the 1868C>G mutation (Figure 2B-C, lane 13). Interestingly in the context of the absence of the lamin A 5'SS the effect of the overexpression of the SRSF5 and SRSF1 proteins was lost (Supplementary Figure 1A) suggesting that the action of SRSF5 and SRSF1 probably mainly consists in the activation of the lamin A 5' SS.

The SR protein SRSF5 represses splicing at the progerin and LA Δ 35 5'SS *in vitro*

In order to know whether the effect of the overexpression of SRSF5 and SRSF1 observed in HeLa cells resulted from direct splicing regulation by these proteins or was the resultant of secondary effects due to a cellular cascade of events, we performed *in vitro* splicing experiments. We took advantage of the fact that the β Glo-LMNA minigene we used for *in cellulo* studies was also suitable for *in vitro* splicing assays (Lopez et al., 2011). We first analyzed the splicing pattern of the 1822G>A and 1868C>G β glo-LMNA constructs that were newly generated. Similarly to what we previously observed for the 1824C>T β glo-LMNA construct (Lopez et al., 2011), the 1822G>A β glo-LMNA substrate led to the production of mRNAs arising from the utilization of both lamin A and progerin 5'SS with a predominant use of the progerin 5' SS (Supplementary Figure 2B, lanes 8-13 and 2C). Concerning the 1868C>G β glo-LMNA substrate, mRNAs arising from the utilization of lamin A, progerin and LA Δ 35 5'SS were detected, with a predominant use of the LA Δ 35 5'SS (Supplementary Figure 2B, lanes 14-20 and 2C). Interestingly the rate of use of the progerin 5'SS observed with this substrate is quite similar to the one observed with the WT β glo-LMNA substrate (compare lanes 1-6 and 15-20 in Supplementary Figure 2B) confirming that, in agreement with the computational analysis, the progerin 5'SS can be recognized in the extract as an authentic 5'SS, even in the absence of HGPS mutations.

The effect of SRSF1, SRSF5 and SRSF7 proteins on the splicing patterns produced from the WT, 1822G>A and 1868C>G *βglo*-LMNA substrates was then tested by adding increasing amounts (200 and 400 ng) of recombinant SR proteins to the nuclear HeLa extract used to perform the splicing reaction. Addition of recombinant SRSF1 to the extract had a mild effect in the WT context and led to a slight increase of the use of the lamin A 5'ss at the expense of the progerin 5'SS (Figure 3A, compare lane 1 with lanes 2-3 and see quantifications). This slight activation of the lamin A 5'SS concomitant to a slight decrease of the use of the progerin 5'SS was also observed when analyzing the splicing pattern of the 1868C>G *βglo*-LMNA substrate (Figure 3A, compare lane 4 with lanes 5-6 and see quantifications). In contrast the rate of use of the LAΔ35 5'SS within this substrate was unaffected by SRSF1 addition. This was further confirmed by the analysis of the changes of the splicing pattern of the 1822G>A *βglo*-LMNA substrate following SRSF1 addition to nuclear extract. Indeed, a complete inversion of the relative use of the progerin and lamin A 5'SS was clearly observed with the highest amount of SRSF1 (Figure 3A, compare lane 8 with lanes 9-10 and see quantifications). Taken together the data suggest that SRSF1 most probably directly regulates the use of the progerin and lamin A 5'SS, but does not seem to be directly involved in the regulation of the use of the LAΔ35 5'SS. By contrast, addition of recombinant SRSF5 to the extract abolished the use of both progerin and LAΔ35 5'SS and led to nearly exclusive RNA proceeding through the lamin A 5'SS in all the tested *βglo*-LMNA substrates (Figure 3B, compare lane 1 to 4-5, lane 6 to 9-10, lane 11 to 14-15 and lane 16 to 19-20). It also slightly increased the global splicing efficiency since more double-spliced RNA species are produced compared to controls. *In vitro* splicing experiments performed in the presence of increasing amounts of recombinant SRSF7 protein were used as negative controls. Interestingly whereas SRSF7 addition had no effect on both lamin A and LAΔ35 5'SS (Figure 3B, compare lane 1 to 2-3, lane 6 to 7-8, lane 11 to 12-13 and lane 16 to 17-18), it

slightly repressed the use of the progerin 5'SS in the mutant contexts (Figure 3B, compare lanes 6-8, lanes 11-13 and lanes 16-18). Taken together obtained data show that SRSF5 directly regulates the use of the progerin, LA Δ 35 and lamin A 5'SS of LMNA exon 11.

The SR proteins SRSF5 represses the use of the progerin 5'SS in primary patients' fibroblasts. (to be confirmed, experiments still in progress)

To directly demonstrate the involvement of SRSF5 in the regulation of the progerin 5'SS selection, we used fibroblasts isolated from HGPS patients. As no primary cell lines from patients bearing either the c.1822G>A or c.1868C>G mutation were available we used primary fibroblasts isolated from a patient bearing the c.1824C>T mutation. The level of SRSF5 in these fibroblasts was increased by transfection of the pXJ41-SRSF5 plasmid which allows SRSF5 overexpression as attested by Western Blot analysis (Figure 4B). The splicing profile of endogenous LMNA was evaluated by RT-PCR using primers located in exon 8 and exon 12 to amplify both progerin and lamin A cDNAs. Consistent with the results obtained *in vitro* and in HeLa cells, SRSF5 overexpression reduced usage of the progerin 5'SS with concomitant increase in the lamin A 5'SS utilization (Fig. 4A). HGPS fibroblasts overexpressing SRSF5 also had slightly reduced level of progerin compared to HGPS fibroblasts transfected with empty pXJ41, as judged by western blotting with a lamin A/C antibody (Fig. 4C, panel lamin A/C). Taken together, these observations established conclusively that SRSF5 is directly responsible for repression of the progerin 5'SS in cells, in agreement with the effects of SRSF5 observed in *in vitro* splicing assays, most probably by activating the use of the lamin A 5'SS.

DISCUSSION writing in progress

ACKNOWLEDGEMENTS

V.V. was supported by a graduate fellowship from the Ministère Délégué à la Recherche et aux Technologies and a 4th year PhD fellowship by the Fondation pour la Recherche Médicale (FRM). This work was supported by grants from the Agence Nationale de la Recherche (ANR-05-BLAN-0261-01) and the European Alternative Splicing Network of Excellence (EURASNET, FP6 life sciences, genomics and biotechnology for health).

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. De Sandre-Giovannoli A, Bernard R, Cau P *et al*: Lamin a truncation in Hutchinson-Gilford progeria. *Science* 2003; 300: 2055.
2. Eriksson M, Brown WT, Gordon LB *et al*: Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature* 2003; 423: 293-298.
3. Dechat T, Shimi T, Adam SA *et al*: Alterations in mitosis and cell cycle progression caused by a mutant lamin A known to accelerate human aging. *Proc Natl Acad Sci U S A* 2007; 104: 4955-4960.
4. Glynn MW, Glover TW: Incomplete processing of mutant lamin A in Hutchinson-Gilford progeria leads to nuclear abnormalities, which are reversed by farnesyltransferase inhibition. *Hum Mol Genet* 2005; 14: 2959-2969.

5. Cao K, Capell BC, Erdos MR, Djabali K, Collins FS: A lamin A protein isoform overexpressed in Hutchinson-Gilford progeria syndrome interferes with mitosis in progeria and normal cells. *Proc Natl Acad Sci U S A* 2007; 104: 4949-4954.
6. Scaffidi P, Misteli T: Lamin A-dependent misregulation of adult stem cells associated with accelerated ageing. *Nat Cell Biol* 2008; 10: 452-459.
7. Shumaker DK, Dechat T, Kohlmaier A *et al*: Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc Natl Acad Sci U S A* 2006; 103: 8703-8708.
8. Fukuchi K, Katsuya T, Sugimoto K *et al*: LMNA mutation in a 45 year old Japanese subject with Hutchinson-Gilford progeria syndrome. *J Med Genet* 2004; 41: e67.
9. Will CL, Luhrmann R: Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* 2001; 13: 290-301.
10. Wahl MC, Will CL, Luhrmann R: The spliceosome: design principles of a dynamic RNP machine. *Cell* 2009; 136: 701-718.
11. Black DL: Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 2003; 72: 291-336.

12. Soret J, Tazi J: Phosphorylation-dependent control of the pre-mRNA splicing machinery. *Prog Mol Subcell Biol* 2003; 31: 89-126.
13. Bourgeois CF, Lejeune F, Stevenin J: Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. *Prog Nucleic Acid Res Mol Biol* 2004; 78: 37-88.
14. Lin S, Fu XD: SR proteins and related factors in alternative splicing. *Adv Exp Med Biol* 2007; 623: 107-122.
15. Lopez-Mejia IC, Vautrot V, De Toledo M *et al*: A conserved splicing mechanism of the LMNA gene controls premature aging. *Hum Mol Genet* 2011; 20: 4540-4555.
16. Bourgeois CF, Popielarz M, Hildwein G, Stevenin J: Identification of a bidirectional splicing enhancer: differential involvement of SR proteins in 5' or 3' splice site activation. *Mol Cell Biol* 1999; 19: 7347-7356.
17. Allemand E, Gattoni R, Bourbon HM *et al*: Distinctive features of Drosophila alternative splicing factor RS domain: implication for specific phosphorylation, shuttling, and splicing activation. *Mol Cell Biol* 2001; 21: 1345-1359.
18. Labourier E, Allemand E, Brand S, Fostier M, Tazi J, Bourbon HM: Recognition of exonic splicing enhancer sequences by the Drosophila splicing repressor RSF1. *Nucleic Acids Res* 1999; 27: 2377-2386.

TITLES AND LEGENDS TO FIGURES

Figure 1: Three HGPS-associated mutations are present in LMNA exon 11 and change LMNA splicing.

(A) The most common HGPS-associated mutation, c.1824C>T, causes 150 nucleotides encoded in exon 11 to be spliced out of the final mRNA and results in a protein, progerin or LA Δ 50, that lacks 50 amino acids. The same deletion is also induced by the c.1822G>A mutation which activates the same internal alternative splice site (progerin 5'SS) within exon 11. A similar but smaller C-terminal deletion (35 amino acids) has been identified in a patient with HGPS who survived more than 30 years longer than the average disease-associated life expectancy. In this case, the causative LMNA mutation, c.1868C>G, activated a cryptic splice donor (LA Δ 35 5'SS) 3' to the site that leads to progerin production from the c.1824C>T allele. (B) The LMNA c.1824C>T, c.1822G>A and c.1868C>G mutations increase the score of the Progerin and LA Δ 35 splice sites used in progeria.

Figure 2: SRSF1 and SRSF5 regulate LMNA splicing *in cellulo*.

(A) Schematic representation of the minigene reporters *β glo-LMNA* used for the study of LMNA splicing. LMNA sequences, carrying (1822G>A and 1868C>G) or not (WT) mutations, were inserted downstream of sequences derived from the human β globin gene. (B) These minigenes were co-transfected into HeLa cells with empty pXJ41, pXJ41-SRSF1 or pXJ41-SRSF5 plasmid and the splicing products were analyzed by RT-PCR of RNA extracted from transfected cells using primers either side of the alternatively spliced regions that give different sized products depending on the identity of the 5'SS used for intron 11 excision. (C) Quantification of the rate of utilization of the laminA, progerin and LA Δ 35 5'SS from four independent LMNA experiments of SRSF1 and SRSF5 overexpression. (D) Quantification of the rate of use of progerin 1822G>A 5'SS (left panel) and LA Δ 35 1868C>G 5' SS (right

panel) in dose effect experiments realized by transfection of increasing amount of pXJ41-SRSF5 or pXJ41-SRSF7 plasmids.

Figure 3: SRSF1 and SRSF5 proteins directly regulate LMNA splicing *in vitro*.

(A) Splicing products resulting from splicing reactions with WT (lanes 1–3), 1868C>G (lanes 4–6) and 1822G>A (lanes 8–10) *βglo*-LMNA reporters in HeLa nuclear extracts alone (lanes 1, 4 and 8) or supplemented with 200 ng (lanes 2, 5 and 9) or 400 ng (lanes 3, 6 and 10) of SRSF1 recombinant proteins were separated on denaturing 6% acrylamide gels. Size markers are indicated on the right of the gels. Quantification of the rate of utilization of the laminaA, progerin and LAΔ35 5'SS is indicated on the right panel. (B) Splicing products resulting from splicing reactions with WT (lanes 1–5), 1824C>U (lanes (6-10); 1822G>A (lanes 11-15) and 1868C>G (lanes 16–20) *βglo*-LMNA reporters in HeLa nuclear extracts alone (lanes 1, 6, 11 and 16) or supplemented with 200 ng (lanes 2, 4, 7, 9, 12, 14, 17 and 19) or 400 ng (lanes 3, 5, 8, 10, 13, 15, 18 and 20) of SRSF7 (lanes 2-3, 7-8, 12-13 and 17-18) or SRSF5 (lanes 4-5, 9-10, 14-15 and 19-20) recombinant proteins were separated on denaturing 6% acrylamide gels. Size markers are indicated on the right of the gel.

Figure 4: Overexpression of SRSF5 in HGPS patient fibroblasts reduces progerin 5'SS utilization and progerin production.

(A) The rate of use of Lamin A and progerin 5'SS in fibroblasts from progeria patients transfected with either empty pXJ41 or pXJ41-SRSF5 plasmids was determined by RT-PCR using primers located in exon 8 and exon 12. The quantifications of two independent experiments are shown. (B). Overexpression of SRSF5 in fibroblasts from progeria patients and its effect on lamin A, lamin C and progerin expression, as assessed by Western blotting. Right panel quantification of the ratio between progerin and lamin A from two independent

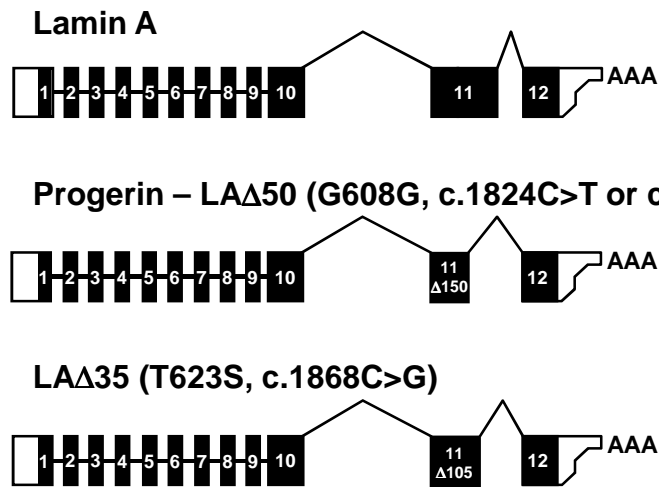
experiments of SRSF5 overexpression.

Supplementary Figure 1:

(A) The 1969G>A+1824C>T, 1969G>A+1822G>A and 1969G>A+1868C>G *β*glo-LMNA minigenes were co-transfected into HeLa cells with empty pXJ41, pXJ41-SRSF1 or pXJ41-SRSF5 plasmid and the splicing products were analyzed by RT-PCR of RNA extracted from transfected cells using primers either side of the alternatively spliced regions that give different sized products depending on the identity of the 5'SS used for intron 11 excision. (B) The overexpression of SRSF5 or SRSF1 in HeLa cells was assessed by Western blotting using anti-phosphoepitope SR proteins, anti-SRSF1 and anti-βtubulin antibodies.

Supplementary Figure 2:

(A) Schematic representation of the *β*glo-LMNA substrates used for the study of LMNA splicing *in vitro*. The size of the expected splicing products is indicated. (B) Time course analysis of *in vitro* splicing of WT, 1822G>A and 1868C>G *β*glo-LMNA reporters showing differential but significant use of both lamin A, progerin and LAΔ35 5'SS in HeLa nuclear extracts. Size markers are indicated on the left. Different times of incubation in minutes are indicated on the top of each panel. (C) Quantification, from at least three independent experiments, of the rate of utilization of the laminA, progerin and LAΔ35 5'SS after 120 min of splicing reaction.

A**B**

	Maxent score	5'ss sequence
Consensus 5'SS	10.86	CAG/GTAAGT
Lamin A 5'SS	10.67	CAG/GT G AGT
Progerin 5'SS	8.07	CAG/GT GGG C
Progerin 5'SS+ 1824C>T	8.56	CAG/GT GG GT
Progerin 5'SS + 1822G>A	9.60	CAG/GT G AGC
LA Δ 35 5'SS	0.44	ACG /GTC ACT
LA Δ 35 5'SS + 1868C>G	8.02	ACG /GT C AGT

Figure 1. Vautrot *et al*

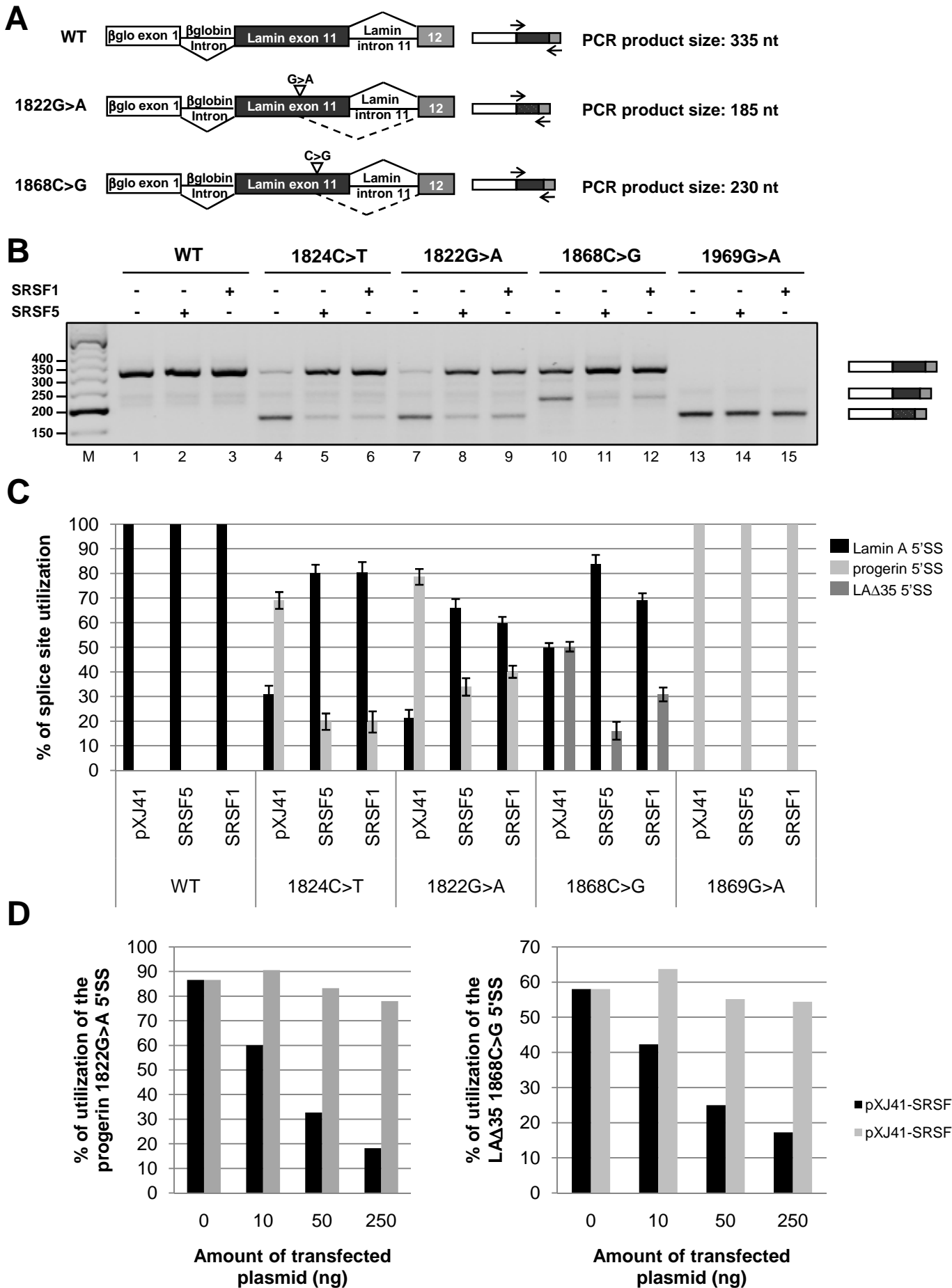


Figure 2. Vautrot *et al*

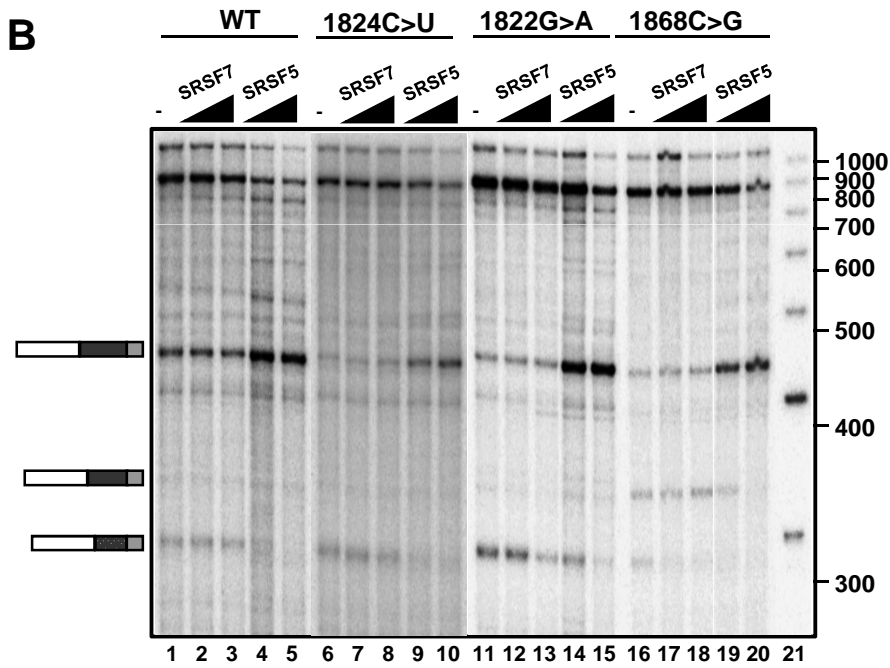
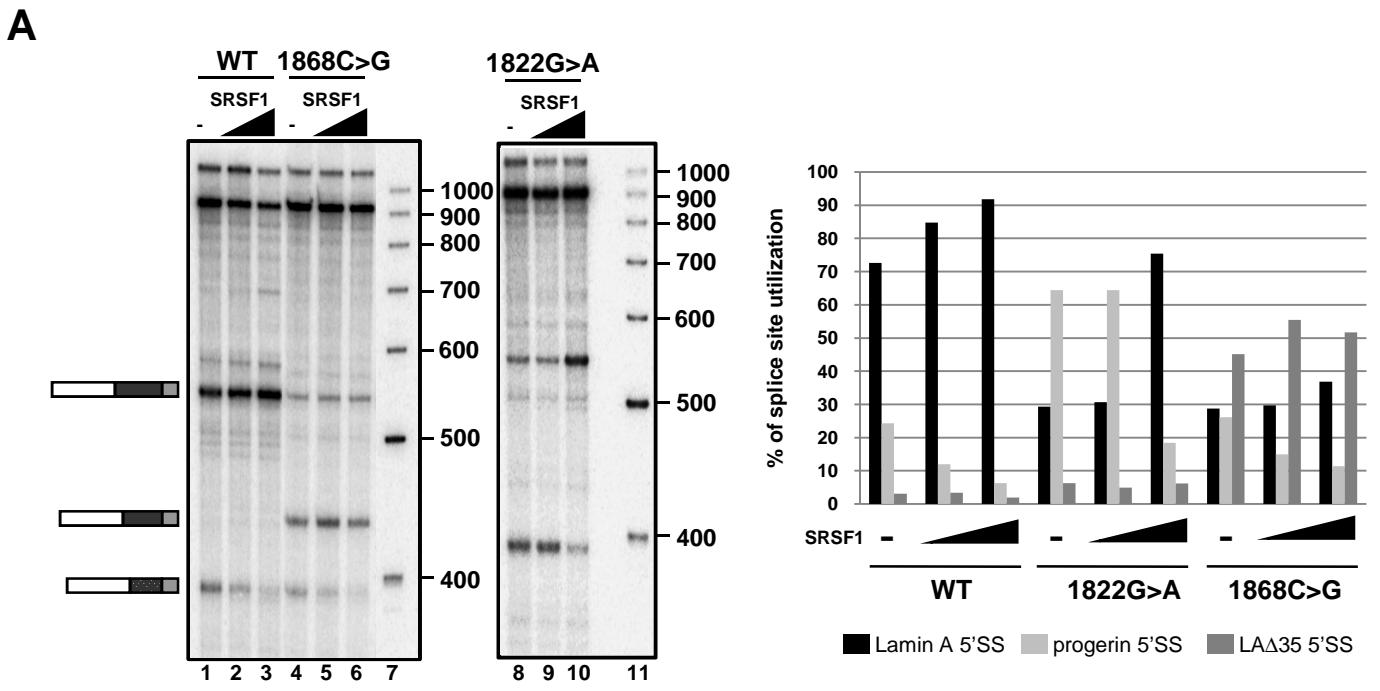
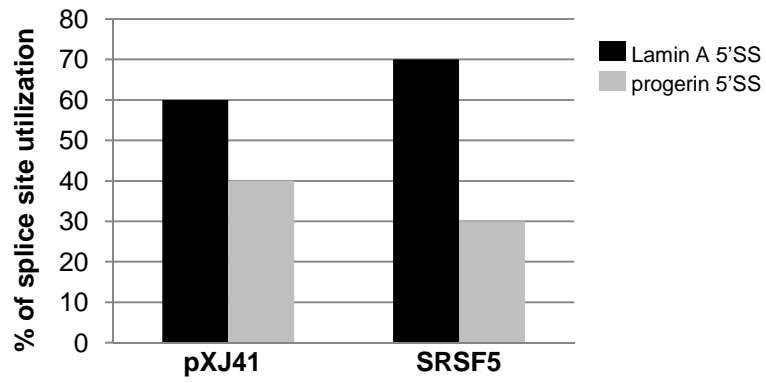


Figure 3. Vautrot *et al*

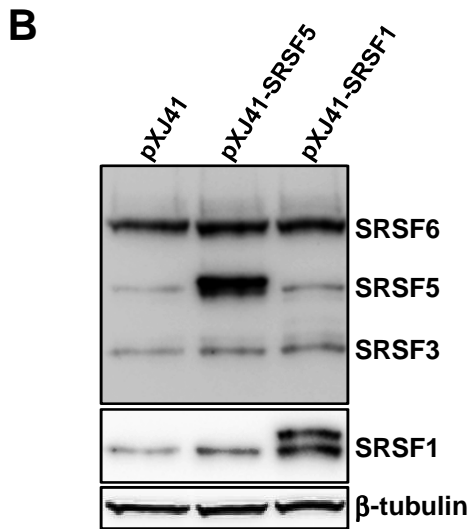
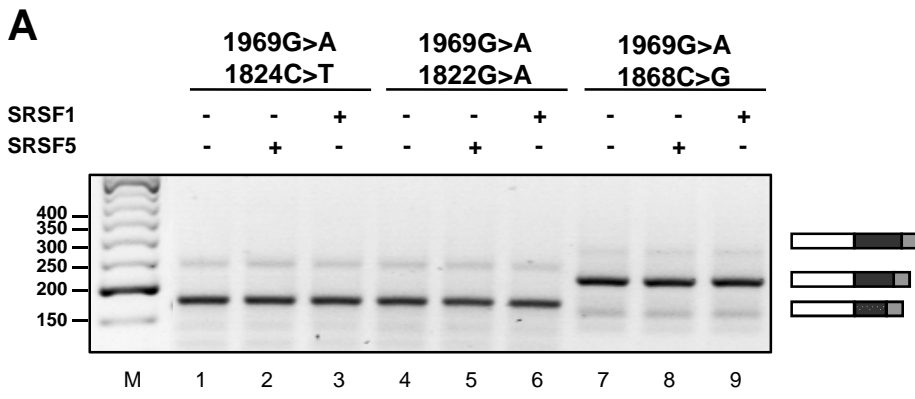
A



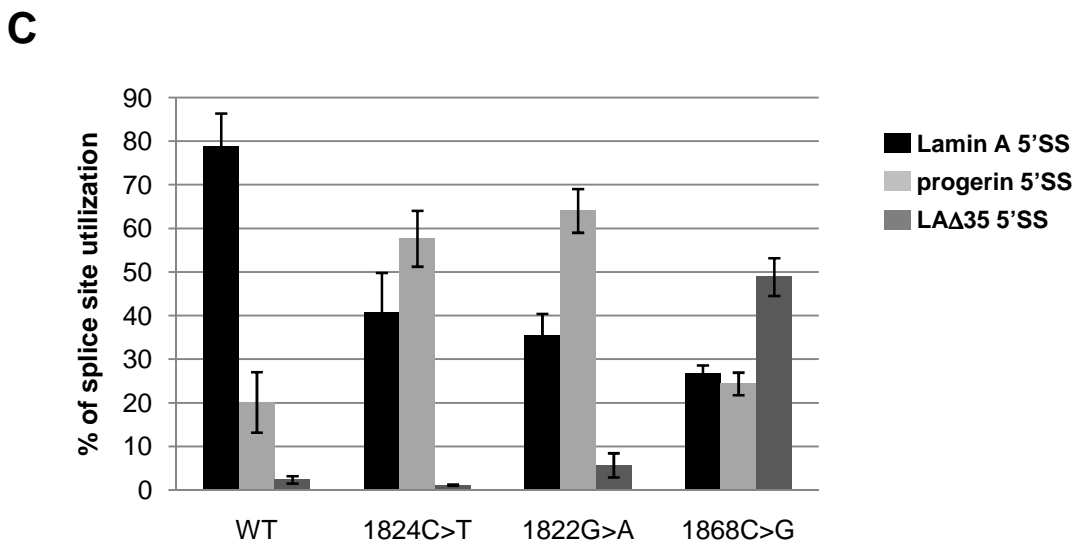
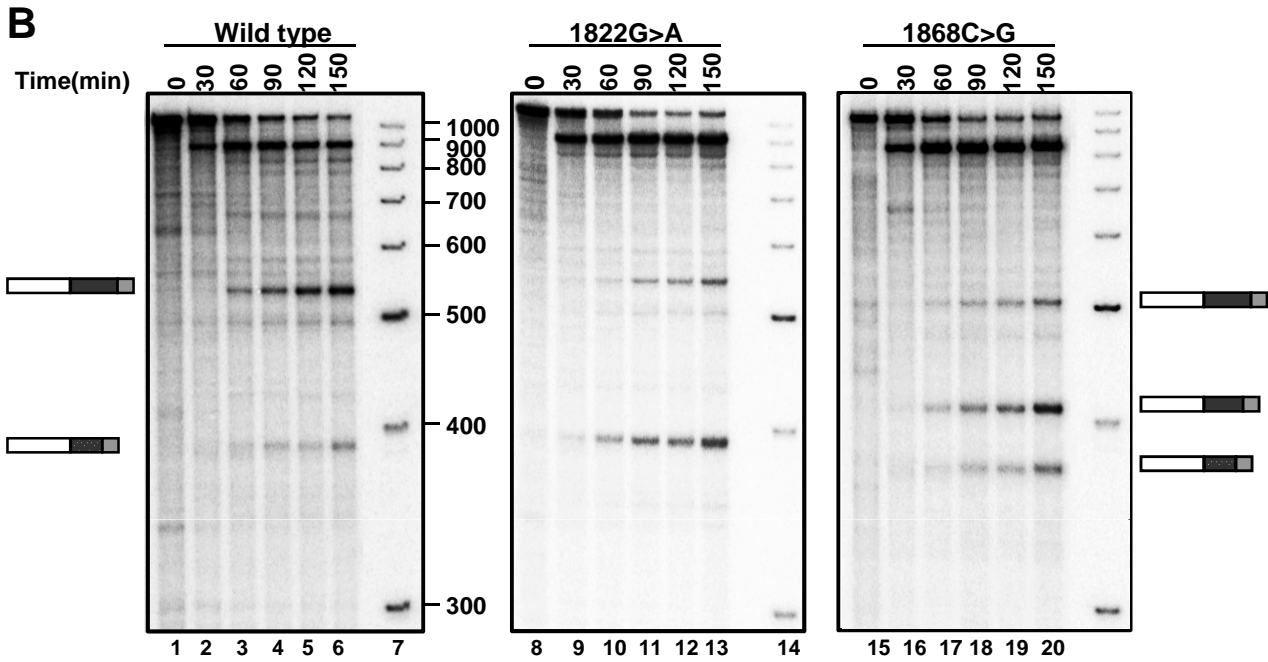
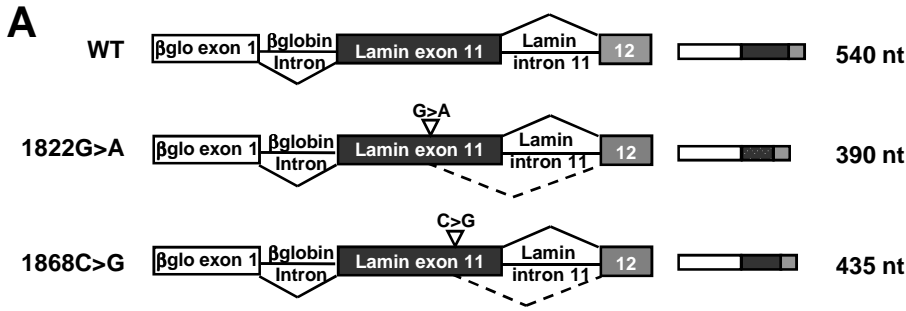
B

b) analysis of progerin/laminA ratio by western blot
+ WB to check the overexpression

Figure 4. Vautrot et al



Supplementary Figure 1. Vautrot *et al*



Supplementary Figure 2. Vautrot *et al*

III.3. Conclusions et perspectives

L'objectif de cette partie de ma thèse était d'étudier plus en détails le mécanisme par lequel les mutations 1824C>U, 1822G>A et 1868C>G favorisent l'utilisation des sites 5'SS progérine et LAΔ35 de l'exon 11 de l'ARN LMNA aux dépens de l'utilisation du site canonique 5'SS lamine A.

L'étude de la structure secondaire de l'ARN dans la région contenant ces sites a révélé qu'un des paramètres impliqué dans l'augmentation de l'utilisation du site 5'SS progérine chez les patients portant les mutations 1824C>U et 1822G>A est la structure secondaire compacte de l'ARN au niveau de ce site qui est relâchée en présence de la mutation. L'ensemble des données structurales obtenues suggère que ce changement conformationnel résulterait de la perte d'une interaction à longue distance impliquant la boucle terminale de la structure tige-boucle IV, dont l'hélice implique le site 5'SS progérine. Nos résultats montrent que les tige-boucle III, IV et V ne sont pas suffisantes pour former cette interaction, ce qui implique que la séquence ARN nécessaire n'est pas comprise dans ces structures. Il n'y pas de séquence parfaitement complémentaire des résidus de la boucle terminale de la structure tige-boucle IV dans l'ARN LMNA 638 utilisé dans les expériences de structure. Cela suggère que les interactions mises en œuvre ne sont pas des appariements Watson-Crick. Il pourrait par exemple s'agir d'une structure tridimensionnelle particulière adoptée par l'ARN résultant en la diminution de l'accessibilité du site 5'SS progérine, et dont la formation dépend de la présence des résidus 1822 et 1824 mutés chez les patients progéria. Ce type de repliement nécessiterait probablement une séquence de grande longueur, ce qui est cohérent avec nos résultats. Des expériences visant à identifier les segments ARN impliqués dans la formation de cette structuration tertiaire sont actuellement en cours et seront poursuivies. Notre but ultime est de déterminer par RMN la structure tridimensionnelle de la structure tige-boucle IV sauvage ou mutée (mutation 1824C>U ou 1822G>A), ceci en présence et en absence du segment d'ARN formant cette interaction longue distance. En effet, la caractérisation du changement conformationnel de l'ARN conduisant à l'augmentation de l'accessibilité du site 5'SS progérine en présence des mutations 1824C>U et 1822G>A peut permettre le design d'approches thérapeutiques ciblées.

Par ailleurs, au cours de notre étude, le développement de systèmes d'études *in vitro* et *in cellulo* de l'épissage de l'intron 11 du gène *LMNA* a permis de confirmer que le site 5'SS progérine est bien un site alternatif d'épissage qui est faiblement utilisé dans un contexte normal et dont l'utilisation est activée en présence des mutations 1824C>U et 1822G>A, alors que le site 5'SS LAΔ35 est un site créé par la mutation 1868C>G. Ces résultats sont intéressants car l'équipe de J. Tazi avec laquelle nous avons collaboré à Montpellier a mis en évidence un lien direct entre l'intensité des symptômes progéroïdes observés dans un modèle de souris progéria et la quantité de progérine produite, en particulier dans les tissus adipeux. Ainsi, les souris hétérozygotes pour la mutation menant à la production de progérine (*LMNA*^{G609G/+}) ont une durée de vie supérieure aux souris homozygotes pour cette mutation (*LMNA*^{G609G/G609G}), mais inférieure à celle des souris WT. Au contraire, les souris génétiquement

modifiées pour ne pas exprimer les exons 11 et 12 du gène *LMNA*, et qui ne produisent donc que la protéine lamine C et pas de lamine A ni de progérine, ont une durée de vie augmentée par rapport aux souris WT (mais développent un plus grand nombre de tumeurs) (Osorio et al., 2011 et thèse I.C. Lopez-Mejia, “Alternative splicing of LMNA gene: lessons from a new mouse model of Hutchinson-Gilford progeria syndrome”). De plus, le fait que nous observons que la proportion d'isoforme d'ARNm LA Δ 35 produite en présence de la mutation 1868C>G soit systématiquement plus faible *in vitro* et *in cellulo* que la quantité d'isoforme d'ARNm progérine en présence de la mutation 1824C>U (respectivement environ 40% contre 60% *in vitro*, et 60% contre 70% *in cellulo*) expliquerait également pourquoi les symptômes observés sont moins importants et que la durée de vie du patient présentant la mutation 1868C>G soit augmentée (Fukuchi et al., 2004).

Au cours de cette étude nous avons également mis en évidence l'implication des protéines SRSF1, SRSF5 et SRSF6 dans la régulation des différents sites 5'SS de l'exon 11 du gène *LMNA*. Les protéines SRF1, SRSF5 et SRSF6 favorisent l'utilisation du site 5'SS lamine A et inhibent l'utilisation du site 5'SS progérine. Les expériences de siRNA réalisées *in cellulo* par l'équipe de J. Tazi ont montré que la déplétion de SRSF1 provoquait, contrairement à ce qui était attendu par rapport aux résultats *in vitro*, non pas l'augmentation de l'utilisation du site 5'SS progérine mais l'augmentation de l'utilisation du site 5'SS lamine A. La contradiction entre le comportement de SRSF1 *in vitro* et *in cellulo* avec les siRNA nous empêchait de conclure précisément sur les mécanismes mis en jeu. Les expériences d'épissage réalisées *in cellulo* avec la construction β Glo-LMNA en transfectant les cellules avec des plasmides permettant la surexpression des protéines SRSF1 ou SRSF6 ont montré que la protéine SRSF1 lorsqu'elle est surexprimée *in cellulo* a en réalité un effet identique à l'effet constaté *in vitro*. Au final, les protéines SRSF1, SRSF5 et SRSF6 ont donc des effets très similaires sur l'utilisation des différents sites 5' *in vitro* ou *in cellulo*. L'effet observé avec les siRNA dirigés contre les ARNm SRSF1 est difficile à expliquer, mais il pourrait provenir d'un effet indirect provoqué par la déplétion de la protéine SRSF1, empêchant l'action d'un autre facteur nécessaire à la régulation des différents sites 5' d'épissage.

Concernant la fixation des protéines, la purification des protéines associées à ce site *in vitro* a montré que dans un extrait nucléaire les protéines SRSF1 et SRSF5 étaient associées efficacement avec l'extrémité de la structure tige-boucle contenant site alternatif en absence de mutation, et que la mutation 1824C>U provoque une diminution de cette fixation (voir partie matériel supplémentaire de la publication n°1). Cette association est observée avec les ARN LMNA 48 et 103, restreints à une petite zone autour du site alternatif. L'effet de la mutation 1824C>U sur l'ouverture de la structure n'étant pas observé avec les ARN LMNA 48 et 103, nous pouvons en déduire que l'interaction de l'ARN LMNA avec SRSF1 et SRSF5 est un phénomène indépendant de l'ouverture locale de la structure. Les expériences de retard sur gel que nous avons ensuite réalisées *in vitro* montrent cependant que la protéine SRSF1 seule n'a qu'une très faible affinité pour la région contenant le site 5'SS progérine (voir partie matériel supplémentaire de la publication n°1). Cela suggère qu'un ou plusieurs facteurs présents dans les

extraits permettent cette interaction. Afin de mieux comprendre comment les protéines SRSF1, SRSF5 et SRSF6 modulent l'utilisation des différents sites 5' d'épissage, il serait nécessaire de réaliser des expériences d'empreintes ("RNA footprinting") pour localiser les sites d'interactions de ces protéines avec l'ARN LMNA 638.

L'ensemble des données que nous avons rassemblées pourrait permettre d'envisager de nouvelles stratégies thérapeutiques. En effet, plusieurs pistes de thérapies géniques ciblant l'épissage de l'ARN LMNA ont été envisagées ces dernières années, comme l'utilisation d'oligonucléotides antisens empêchant stériquement l'utilisation du site alternatif. Toutefois ce type de thérapie est encore trop difficile à mettre en place, et les protéines SR constituent une cible supplémentaire pour la conception de molécules à visée thérapeutique. Notamment, l'utilisation de drogues connues pour réguler l'activité des protéines SR pourrait avoir un impact sur l'utilisation des sites d'épissage.

Enfin, l'influence des protéines SR sur l'épissage de l'exon 11 dépend également de la disponibilité de celles-ci dans le cytoplasme. Des données récentes montrent la formation de granules nucléaires dans les fibroblastes de patients progéria, qui, comme il a été mentionné dans l'introduction, contiennent des ARN satellites III (Shumaker et al., 2006). Ces granules peuvent également être formés dans les cellules de patients sains soumises à des stress tels qu'un choc thermique (Chiodi et al., 2000; Jolly et al., 2004; Rizzi et al., 2004). Dans les cellules soumises à un stress, ces granules, nommés granules de stress nucléaires, contiennent également de nombreux facteurs d'épissage (Denegri et al., 2001). Il est possible que ces facteurs d'épissage soient séquestrés dans les granules de stress nucléaires. Cette séquestration diminuerait leur disponibilité dans le nucléoplasme et pourrait moduler les régulations de l'épissage alternatif. Un des facteurs retrouvés dans les granules de stress nucléaires est SRSF1, et au vu de nos résultats sa séquestration pourrait favoriser l'utilisation du site alternatif menant à la production de progérine lors de l'épissage du pré-ARNm LMNA dans les cellules de patients progéria. De plus, il est possible que les protéines SRSF5 et SRSF6 soient également incluses dans les granules de stress nucléaires. Ceci nous a conduits à nous intéresser à ces granules de stress et à l'influence qu'ils pourraient avoir sur l'épissage alternatif lors de stress.

PARTIE 2

STRESS THERMIQUE ET EPISSAGE

(Introduction)

I. Stress thermique et épissage

I.1. Définition et origines d'un stress thermique

On considère qu'une cellule est soumise à un stress lorsqu'elle est soumise à différents types de contraintes imposées par son environnement. L'état de stress peut gêner voire empêcher la cellule de maintenir son homéostasie et de remplir sa fonction au sein d'un organisme en perturbant son métabolisme. Ces perturbations proviennent le plus souvent de dommages aux constituants cellulaires essentiels que sont les protéines, les acides nucléiques et les lipides. Toutefois, les cellules vivantes sont capables de mettre en place des mécanismes de défense afin de limiter ou de neutraliser l'impact de ces contraintes environnementales et maintenir leur intégrité. L'incapacité de la cellule à répondre efficacement à ces contraintes peut perturber définitivement le métabolisme cellulaire, ou conduire à l'apoptose, voire la nécrose dans le cas d'un changement d'environnement trop rapide ou trop important qui submerge complètement les défenses de la cellule.

Cependant, dans la pratique, pas ou très peu d'organismes vivent dans des conditions qui leurs sont adaptées. Au cours de l'évolution, les cellules et organismes ont développé des mécanismes leur permettant de faire face à un ensemble plus ou moins large de conditions environnementales. Ainsi, il n'est pas toujours facile de savoir où s'arrête la réponse aux conditions environnementales sous-optimales et où commence véritablement le stress. Une partie de la réponse est donnée par le fait que les cellules peuvent répondre à des conditions sous-optimales par des activités spécifiques et dédiées, alors que, d'une manière générale, pour répondre à un stress important, la cellule dispose plutôt de systèmes de défense moins spécifiques, rapidement inductibles, et qui servent avant tout à limiter et à traiter les dommages aux constituants cellulaires et les déséquilibres métaboliques engendrés. Cela permet de répondre efficacement à un ensemble large de variations environnementales, et confère une grande souplesse aux défenses de la cellule.

D'autre part, la capacité d'une cellule à répondre à un stress à un temps donné dépend de son état (densité de population, intégrité et activité de ses constituants), et varie donc selon le type cellulaire considéré et même des antécédents d'exposition à un stress de cette cellule. Il est donc nécessaire de considérer un stress non seulement selon sa force, sa nature, mais aussi selon les cellules impliquées : leur type, leur état et leurs antécédents.

En ce qui concerne le stress hyperthermique appliqué aux cellules humaines, les conditions d'étude dans la littérature impliquent des températures de 42-45°C rarement atteintes *in vivo*. Les dommages causés aux cellules viennent de la dénaturation des protéines par le surplus énergétique dû à la chaleur. Dans ce cas, le stress thermique sert d'avantage de modèle d'étude, car il provoque l'expression d'une grande partie des protéines de la réponse générale à un stress, notamment les protéines

HSP ("Heat Shock Proteins"), découvertes grâce au choc thermique, mais surexprimées dans beaucoup de types de stress différents. Nous nous concentrerons en grande partie sur ce modèle parce qu'il permet d'étudier efficacement et de façon générique comment l'épissage alternatif peut être régulé en condition de stress, mais aussi parce qu'il provoque l'expression d'ARN particuliers appelés ARN satellites III.

Dans un premier temps dans ce chapitre seront abordés succinctement les dommages causés par le stress thermique sur les différents constituants cellulaires, afin de mieux comprendre comment les cellules y répondent. Dans un deuxième temps, nous nous intéresserons aux protéines impliquées dans la réponse au stress, notamment les HSP. Enfin nous verrons plus globalement l'impact du stress sur les différents mécanismes cellulaires, et spécialement sur l'épissage constitutif et alternatif.

I.2. Dommages causés par le stress thermique

Le principal effet néfaste du stress thermique sur la cellule résulte de la perturbation de l'équilibre thermodynamique des protéines. En effet l'énergie nécessaire pour le repliement spontané ou le maintien de la forme native d'une protéine est directement reliée et dépend principalement de la température du milieu, y compris à quelques degrés près. On parle généralement de stress thermique *in cellulo* à partir d'une différence de 1 à 5°C mesurés par rapport à la température normale de croissance des cellules humaines (au-delà, la mort cellulaire devient trop importante pour la plupart des types cellulaires) (Carmo-Fonseca et al., 1993; Corton et al., 1994; Speit and Schütz, 2013). Une température trop faible ou trop élevée peut entraîner le mauvais repliement, la dénaturation, voire l'agrégation non spécifique des protéines dans les cellules du fait de l'apparition de régions hydrophobes en surface ou d'interactions inadéquates. L'accumulation de protéines dénaturées ou mal repliées est délétère voire létale pour la cellule.

Ainsi l'hyperthermie affecte de façon drastique la morphologie de la cellule et son fonctionnement global. Chez les eucaryotes, un des principaux dommages observés est la perturbation de l'organisation du cytosquelette. Une dizaine de minutes de choc thermique modéré (42°C à 43°C) suffisent à causer la désorganisation des réseaux protéiques du cytosquelette et l'agrégation autour du noyau de certaines protéines de la famille des filaments intermédiaire cytoplasmiques, constituant une partie du cytosquelette, comme la vimentine (pour revue, Toivola et al., 2010; Welch and Suhan, 1985). Une perturbation des réseaux d'actine ou de tubuline, la perte de localisation des organelles, la fragmentation de l'appareil de Golgi et du réticulum endoplasmique, et la diminution du nombre de lysosomes et de mitochondries ont également été observés (Welch and Suhan, 1986). Par ailleurs, plusieurs exemples montrent que chez les eucaryotes le stress thermique est aussi à l'origine d'une déplétion en ATP et d'une augmentation de l'AMPc intracellulaires chez les eucaryotes, à un faible taux à 42°C mais de manière plus marquée à 45°C (le ratio AMP/ATP est multiplié par un facteur d'environ 5 après 1 h à 45°C dans les hépatocytes de rat, et d'un facteur d'environ 1,5x après 1 h à 42°C) (Corton et

al., 1994; Patriarca and Maresca, 1990; Yasuda et al., 2006a). Cette diminution provient de la perte ou de la désorganisation des mitochondries provoquant un découplage de l'oxydation phosphorylante à l'origine de la synthèse d'ATP. Elle a un impact considérable sur le métabolisme énergétique de la cellule, mais aussi sur d'autres processus comme le transport nucléo-cytoplasmique, en particulier par les pores nucléaires et impliquant les protéines Ran, l'importine et l'exportine. En effet, ce transport nécessite le plus souvent l'hydrolyse d'ATP ou de GTP, dont le taux est diminué dans les cellules ayant subi un choc thermique (Miyamoto et al., 2004; Schwoebel et al., 2002).

Le choc thermique a aussi un effet sur les composants nucléaires : la taille des nucléoles augmente, et la synthèse des sous-unités des ribosomes peut être perturbée à tel point que des agrégats de protéines ribosomales deviennent visibles au microscope électronique (Boulon et al., 2010a; Welch and Suhan, 1985). Par ailleurs, le choc thermique a également des effets drastiques sur le déroulement du cycle cellulaire, de l'épissage et de la traduction, et provoque la formation de granules de stress cytoplasmiques et nucléaires (Nakahata et al., 2002; Yost and Lindquist, 1986). Nous reviendrons plus particulièrement sur ces effets lors du chapitre sur la réponse des cellules au stress.

De manière indirecte, en affectant les protéines, le choc thermique peut également affecter les autres constituants cellulaires. Une étude récente démontre par exemple une fragmentation indirecte de l'ADN à cause de la chaleur : la présence dans plusieurs lignées cellulaires de mammifères de foci formés par les histones H2AX phosphorylées (γ H2AX) par la protéine de réparation de l'ADN ATM indique la présence de cassures double brin de l'ADN (Takahashi et al., 2010a, 2010b). L'hyperthermie a aussi un effet néfaste sur certaines protéines de réparation de l'ADN qui sont sensibles à la chaleur, comme les sous-unités Ku70 et Ku80 du complexe DNA-PK, et peut mener à la persistance de cassures simple brin ou de régions simple brin dans l'ADN en réplication (Burgman et al., 1997; pour revue, Tomita, 2010). De plus, la chaleur perturbe le recrutement au niveau des cassures double brin de complexes de reconnaissance et de réparation de l'ADN, comme le complexe Mre11/Rad50/Nbs1 (M/R/N). Ce complexe est impliqué dans un type de réparation de l'ADN nommé NHEJ (pour "Non Homologous End Joining") mais surtout par recombinaison homologue, c'est-à-dire deux mécanismes essentiels de réparation des DSB chez les mammifères. Dans les cellules soumises à un choc thermique, il est exporté dans le cytoplasme grâce au transporteur nucléo-cytoplasmique exportine1 (CRM1) et ne peut plus être recruté sur les DSB (Seno and Dynlacht, 2004). Cependant on ignore les conséquences de cet export. En parallèle, il a été observé que le choc thermique peut provoquer une perturbation plus globale de l'organisation du génome, en particulier au niveau des centrosomes, qui sont désorganisés après un stress thermique (Nakahata et al., 2002; Pawlik et al., 2013).

Des perturbations de la structure secondaire des ARN lors d'un stress thermique ont aussi été décrites, notamment chez les bactéries. En effet, de part les faibles différences de stabilité thermodynamique entre les différentes conformations des ARN, des variations même minimales d'énergie par changement de la température peuvent induire des transitions structurales des ARN. Or, certaines

structures des ARN jouent des rôles importants dans l'expression des gènes. Ainsi les régions 5'UTR des ARNm issus des gènes *E. coli* rpoH en hyperthermie et cspA en hypothermie sont régulées par des structures d'ARN thermosensibles (Giuliodori et al., 2010; Morita et al., 1999). Ces régions fonctionnent comme des pièges du RBS ("ribosome binding site"). Elles changent de conformation sous l'effet d'une variation de température, ce qui permet la traduction de la protéine (Narberhaus, 2010). Etant donné les nombreuses régulations impliquant la structure secondaire des ARN chez les eucaryotes, les implications d'une augmentation de la température sur la traduction mais aussi l'épissage alternatif pourraient être nombreuses, mais ont été peu étudiées pour le moment.

D'autre part, une augmentation de la température perturbe la fluidité des lipides membranaires au niveau de microdomaines de la membrane plasmique, ce qui s'accompagne de changements dans la morphologie et dans le ratio protéines/lipides de la membrane, et qui résulte en une altération de sa perméabilité (Vigh et al., 2007). Ces modifications influencent l'homéostasie des ions et ont pour conséquence une acidification du pH cytosolique (Coote et al., 1991). Ces perturbations peuvent à leur tour affecter l'organisation de récepteurs associés à la membrane, comme la protéine TRPV (Bromberg et al., 2013).

L'ensemble de ces conséquences provoque l'arrêt du cycle cellulaire et une stagnation de la croissance et de la prolifération des cellules (Lindquist, 1980; Nakahata et al., 2002; Pawlik et al., 2013; Yost and Lindquist, 1986). Selon sa sévérité (température et durée), le stress thermique peut engendrer la mort cellulaire, mais un stress non létal peut mener à une tolérance accrue des cellules à des stress futurs, qu'ils soient d'ailleurs des stress thermiques ou d'une autre nature. Cette "cross-tolérance" est due à la synthèse ou à l'activation en cours de stress de protéines ayant un effet cytoprotecteur, qui font partie de la réponse au stress de la cellule. En effet, de nombreuses protéines chaperonnes ou régulatrices sont impliquées de manière récurrente dans la réponse à plusieurs types de stress (Kültz, 2003, 2005).

I.3. Les principaux mécanismes de réponse au stress thermique

I.3.1. Protéines dont l'expression est induite par le stress thermique

Le choc thermique induit l'expression de protéines qui peuvent être classées en plusieurs grandes catégories, en lien direct avec les types de dommages observés décrits dans le paragraphe précédent. Ces catégories ont été établies à partir des données de plusieurs études transcriptomiques réalisées pour étudier l'expression des gènes de réponse au choc thermique chez plusieurs organismes modèles : des bactéries, la levure *S. cerevisiae*, des archae, *C. elegans*, les plantes (Gasch et al., 2000; GuhaThakurta et al., 2002; Larkindale and Vierling, 2008; Matsuura et al., 2010; Richmond et al., 1999; Rohlin et al., 2005). Certaines de ces études ont également été réalisées sur des lignées cellulaires humaines (Furusawa et al., 2009; Kim et al., 2011; Murray et al., 2004; Tabuchi et al., 2008).

Il ressort de ces études que la catégorie largement prédominante des gènes dont l'expression est modifiée en réponse au stress thermique sont ceux des protéines HSP, qui sont des chaperonnes moléculaires, ainsi que les gènes de leurs cofacteurs. Ces chaperonnes empêchent la dénaturation des protéines. Chez l'homme, la seconde catégorie de gènes fréquemment activés comporte des gènes dont les produits contrecarrent la protéotoxicité des protéines dénaturées ou agrégées. Il s'agit de gènes codant des protéines impliquées dans les systèmes de dégradation des protéines tels que le protéasome. Une autre catégorie importante inclut des protéines régulatrices comme les kinases et les facteurs de transcription, qui sont nécessaires pour induire les voies de réponse au stress thermique et les cascades de signalisation cellulaire, ainsi que de nombreuses enzymes du métabolisme énergétique. Les changements dans l'expression de ces gènes doivent permettre de réajuster et de stabiliser la dépense énergétique et de participer à la détermination du destin de la cellule (survie ou apoptose) (Malmendal et al., 2006). Il a aussi été montré dans certaines lignées cellulaires que l'expression des protéines de la chromatine et la conformation de l'ADN sont également affectées par le choc thermique (Kim et al., 2011). En particulier, l'expression des histones de la famille H2A est réprimée, ce qui est probablement à mettre en relation avec l'apparition de foci H2AX résultant des cassures double brin de l'ADN. Certaines autres familles de gènes d'histones ont au contraire une expression augmentée. C'est en particulier le cas pour les histones de la famille H4, mais leur rôle dans la réponse au choc thermique reste inconnu. Enfin, l'expression des gènes codant des protéines de transport est aussi modifiée, ce qui est probablement en lien direct avec les modifications des propriétés des membranes entraînées par le choc thermique.

Toutefois, l'identité des gènes dont l'expression varie après un choc thermique peut beaucoup varier d'un type cellulaire à un autre chez l'homme (Murray et al., 2004). Actuellement, ces analyses globales sont encore incomplètes au vu de la complexité des régulations entraînées par le stress thermique. Elles ont néanmoins permis de mettre en évidence un certain nombre de mécanismes cellulaires activés en réponse au choc thermique. Les paragraphes suivants dressent un profil général de cette réponse, telle qu'elle est décrite aujourd'hui dans la littérature, afin de mieux comprendre par la suite quel rôle peuvent y jouer les régulations de l'épissage alternatif. L'élément central de cette réponse, qui est conservé au sein de tous les domaines du vivant, est l'induction de la synthèse des protéines chaperonnes HSP.

I.3.2. Les protéines du choc thermique HSP

Les protéines HSP sont présentes dans le cytosol, les mitochondries, le réticulum endoplasmique et le noyau (y compris le nucléole), et ont une durée de vie relativement longue (48 h environ dans les cellules de l'épiderme). Elles constituent une grande famille de protéines nommées et regroupées en sous-familles en fonction de leur poids moléculaires : HSP de faible poids moléculaire (sHSP ou HSPB), puis hsp40, 60, 70, 90 et 100kDa. La fonction des HSP est majoritairement d'aider ou de permettre le

repliement des protéines. Leur mode d'action ne repose pas sur un guidage stérique spécifique du repliement, mais plutôt sur une optimisation générale de l'efficacité de repliement, en jouant sur les paramètres thermodynamiques, parfois grâce à l'hydrolyse d'ATP. Pour cela, elles limitent le nombre d'intermédiaires de repliement, et couvrent les zones exposées ou non structurées des protéines, afin d'éliminer les interactions non productives et faciliter le repliement spontané de la protéine vers une forme native.

La propension des protéines à adopter un mauvais repliement, voire à générer des agrégats, est largement favorisée en cas de stress. C'est pourquoi l'expression des HSP de manière générale est massivement induite lorsque les cellules sont soumises à un stress, et leur premier rôle est alors de maintenir les protéines dans un état intermédiaire de repliement afin d'empêcher leur agrégation ou la formation de plaques amyloïdes, irréversibles et potentiellement cytotoxiques. Les HSP de haut poids moléculaire comme hsp70 et hsp90 provoquent même activement un retour de la protéine à l'état natif à l'aide de l'hydrolyse d'ATP, et lorsque ce n'est plus possible, elles facilitent l'élimination par le protéasome des protéines irrémédiablement dénaturées, avec la participation d'autres HSP et de co-chaperonnes (comme Hip48 et l'E3 ubiquitine ligase CHIP). L'expression des HSP est critique pour la résistance des cellules au stress thermique, mais aussi à divers stress, tels que l'ischémie et l'hypoxie, la déplétion en ATP, l'exposition à des espèces réactives de l'oxygène et à d'autres molécules toxiques comme l'éthanol, les métaux lourds, les analogues d'acides aminés, l'infection par certains virus, et de nombreuses autres conditions pathologiques (Collins and Hightower, 1982; Kabakov et al., 2002; Kiang and Tsokos, 1998; Kukreja et al., 1994; Levinson et al., 1980; Ostberg et al., 2002; Patel et al., 1995; Plesset et al., 1982; Richard et al., 1996).

De part leur fonction dans le repliement des protéines, les HSP jouent un rôle crucial dans le repliement des protéines dans la traduction, la modulation des interactions protéine-protéine, le transport des protéines et leur prise en charge par le protéasome, et ont également un rôle dans la régulation des processus apoptotiques et peuvent même être secrétées, probablement afin d'assurer un rôle dans la communication intercellulaire et l'inflammation (pour revue, Lanneau et al., 2007; De Maio, 2011; Schmitt et al., 2007). Ainsi, certaines HSP sont exprimées de manière constitutive en absence de stress, et d'autres HSP voient leur taux d'expression augmenter à certains stades du développement ou du cycle cellulaire sans l'intervention d'un stress. Par exemple l'expression d'hsp70 est plus particulièrement augmentée lors de la phase G1/S du cycle cellulaire, et en réponse à certaines hormones, au sérum, à des mitogènes ou des facteurs de croissance (Ferris et al., 1988; Milarski and Morimoto, 1986; Wu and Morimoto, 1985). En outre, en relation directe avec le rôle qu'elles ont lors d'un stress lors de la survie cellulaire, certaines HSP ont, comme nous le verrons, un rôle régulateur, et peuvent par exemple inhiber les protéines de signalisation de l'apoptose.

I.3.3. Principaux gènes et isoformes des protéines HSP et leurs fonctions associées

Chacune des sous-familles de HSP contient de nombreux gènes et isoformes différents, qui fonctionnent souvent en coopération dans le repliement des protéines. Les protéines de la famille sHSP, hsp40 et hsp60, par exemple, coopèrent activement avec les protéines de la famille hsp70 dans le repliement des protéines (pour revue, Kampinga et al., 2009). Seules les principales protéines de ces familles et leurs fonctions génériques seront brièvement présentées ici.

Chez les mammifères, la forme majeure induite par le stress est hsp70.1, issue du gène *HSPA1A*, mais d'autres gènes comme *HPSA1B* (hsp70.2), *HSPA6* et de façon moindre *HSPA7*, qui ont des fonctions similaires, sont également induits (Brocchieri et al., 2008; pour revue, Kampinga et al., 2009). En effet, les protéines hsp70 jouent un rôle central dans la prise en charge des protéines dénaturées en cas de stress. D'autre part, l'expression des protéines hsc70 (*HSPA8*) et hsp70-like 1 (*HSPAIL*) a lieu même en absence de stress. Elles favorisent le repliement des protéines naissantes lors de la traduction.

L'isoforme α de Hsp105 peut interagir directement avec le substrat de la protéine hsp70 avec laquelle elle coopère activement dans le processus de repliement. La forme α est constitutivement exprimée, mais aussi induite dans de nombreux stress et est majoritairement cytoplasmique. La forme β est spécifiquement produite lors d'un choc thermique moyen (environ 40-42°C) et est localisée dans le noyau, où elle provoque l'induction de l'expression de hsp70 (Saito et al., 2009a; Yamagishi et al., 2009a). Les deux isoformes sont nécessaires pour la survie face à des stress plus intenses.

Hsp90, chez les mammifères et la levure, est produite sous deux formes paralogues, α et β , cette dernière étant d'ailleurs indispensable à la viabilité des embryons de souris (Voss et al., 2000). En effet, hsp90 est constitutivement exprimée, de manière ubiquitaire et en très grande quantité, et elle intervient en continu dans le repliement d'une population spécifique du protéome incluant surtout des kinases, des récepteurs nucléaires aux hormones, des récepteurs membranaires, et des facteurs de transcription (Dezwaan and Freeman, 2008). Elle est souvent indispensable à son substrat pour l'acquisition d'une activité, et au laboratoire il a été montré avec l'équipe de E. Bertrand qu'elle intervient aussi dans la biogenèse de complexes ribonucléoprotéiques comme les snoRNP ou la snRNP U4, en interagissant avec un complexe protéique nommé R2TP (Boulon et al., 2008, 2012). La structure 3D de l'interaction entre la protéine Tah1 de ce complexe et hsp90 chez la levure vient d'être établie au laboratoire (Back et al., 2013). Le complexe hsp90-R2TP interviendrait également dans l'assemblage de nombreux complexes comme la biogenèse des ARN polymérase ou encore le chargement des miRNA et du complexe RISC sur leur substrat (Boulon et al., 2010b, 2012). Lors d'un stress thermique, l'induction de son expression suite à un stress est limitée comparée à celle d'hsp70, mais le manque de hsp90 active rend les cellules hyper-sensibles au stress thermique (Borkovich et al., 1989; Nakai and Ishikawa, 2001).

I.3.4. Régulation de l'expression des protéines de réponse au choc thermique par les facteurs HSF lors d'un stress

La transcription des HSP et de beaucoup de gènes répondant au choc thermique est spécifiquement activée par une famille particulière de facteurs transcriptionnels, les HSF ("Heat Shock Factors"). Il en existe 4 chez les mammifères, nommés HSF1 à HSF4. Les facteurs HSF sont des régulateurs transcriptionnels, constitués d'un domaine de liaison à l'ADN en N-terminal suivi par un domaine de trimérisation puis par un domaine de transactivation (Peteranderl and Nelson, 1992) (Figure I-1, A.). HSF1 est le facteur transcriptionnel majeur des régulations ayant lieu en réponse au stress. Il permet, entre autres, l'expression des formes inductibles de hsp70 et de hsp90 (Abravaya et al., 1991). HSF1 est produit à des niveaux constants, dans la plupart des lignées cellulaires, ce qui permet une induction bien plus rapide, de l'ordre de la minute, en cas de stress. Dans les cellules non stressées, HSF1 est localisée majoritairement dans le cytoplasme où elle adopte une forme monomérique inactive. Elle est liée à et réprimée par hsp90 dont la fixation est favorisée par le complexe hsc70-hsp40, et sa rétention dans le cytoplasme est favorisée par la protéine 14-3-3 (Wang et al., 2003; Zou et al., 1998). La quantité croissante de protéines dénaturées dans les cellules soumises à un stress provoque la titration des machineries HSP et leur réorganisation. Cela provoque la libération de HSF1 qui adopte alors une forme trimérique active, grâce à la formation d'interactions intermoléculaires de type "coiled-coil" entre les domaines d'oligomérisation (Peteranderl and Nelson, 1992). Cette forme est alors capable de pénétrer à l'intérieur du noyau et de se fixer sur l'ADN génomique au niveau de séquences spécifiques de régulation de la transcription, les éléments HSE (pour "Heat Shock Elements") (Figure I-1, B.). Ces séquences sont situées au niveau des promoteurs proximaux des gènes régulés. Ces HSE sont des séquences pentamériques 5'-nGAAn-3', arrangées en répétitions d'orientation inverse et adjacentes (Amin et al., 1988). Trois pentamères sont nécessaires pour obtenir un élément HSE fonctionnel (Xiao et al., 1991). L'activité de HSF1 est aussi régulée par des modifications post-traductionnelles complexes (voir paragraphe suivant). La forme monomérique inactive de HSF1 est phosphorylée au niveau des résidus Ser303 et 307, ce qui permet l'inhibition de l'activité transcriptionnelle par formation d'interactions intramoléculaires et qui permet aussi la rétention de HSF1 dans le cytoplasme par la protéine 14-3-3 ϵ (Chu et al., 1996; Kline and Morimoto, 1997; Wang et al., 2003). Comme nous le verrons, la forme active de HSF1 est hyperphosphorylée sur d'autres serines par plusieurs kinases, notamment la MAP kinase JNK, ce qui confère au trimère HSF1 une activité transcriptionnelle. Il a été observé toutefois que dans le cas d'un stress causé par des analogues d'acides aminés par exemple, l'hyperphosphorylation de HSF1 n'est pas nécessaire à son activation (Holmberg et al., 2001; Lee et al., 1995; Sarge et al., 1993). Une fois hyperphosphorylée, HSF1 peut également être sumoylée, processus qui implique hsp27, ou encore acétylée, ce qui inhibe son activité (Brunet Simioni et al., 2009; Hietakangas et al., 2003; Westerheide et al., 2009). Lors du retour des cellules à l'état non stressé, lorsque la concentration en protéines dénaturées diminue, hsp70-hsp40 et hsp90 ne sont plus titrées par les protéines dénaturées et

peuvent inhiber à nouveau HSF1 qui reprend sa forme monomérique inactive dans le cytoplasme. La protéine hsp27 (hspB1), qui se fixe au niveau de la région de trimérisation de HSF1, favorise ce retour à la forme monomérique par compétition avec HSF1 elle même (pour revue, Morimoto et al., 1992; Santoro, 2000) (figure I-1, B.). Il est à noter qu'une controverse existe quant à la localisation de HSF1 en absence de stress et que certains auteurs observent une localisation majoritairement nucléaire et non cytoplasmique de HSF1 et proposent que l'activation de HSF1 soit un processus nucléaire et non cytoplasmique (Mercier et al., 1999).

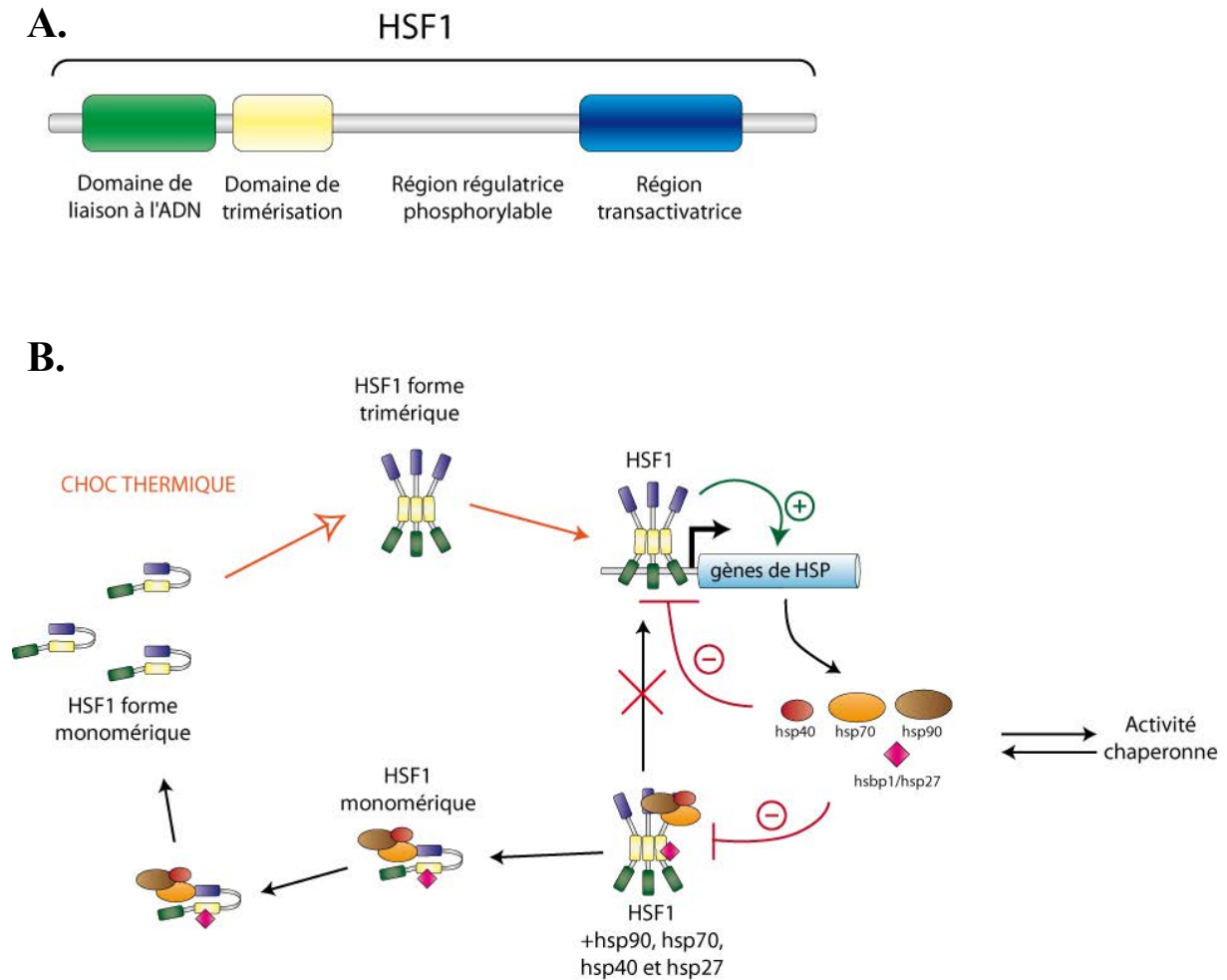


Figure I-1 : Régulation de la transcription des gènes de réponse au stress par HSF1. A. Les différents domaines de régulation de la protéine HSF1. B. Boucle d'autorégulation de la protéine HSF1 dans le noyau. Initialement, HSF1 est présente sous forme monomérique inactive. Sa trimérisation provoque la transcription de gènes de réponse au stress thermique, en particulier les HSP. La synthèse des HSP en grande quantité permet la rétro-inhibition de HSF1 à la fin du choc thermique, lorsque les protéines HSP ne sont plus titrées par les protéines dénaturées.

Même si HSF1 est au centre de ces régulations, et se fixe sur la quasi totalité des gènes de réponse au stress thermique, HSF1 et HSF3 sont toutes les deux requises pour une réponse efficace au choc thermique (Nakai and Morimoto, 1993). Il semble en effet exister une interdépendance complexe entre HSF1 et HSF3, puisque la présence de HSF3 est requise pour l'induction des gènes cibles de la

protéine HSF1, même si HSF1 est capable d'adopter une forme trimérique active dans les cellules dépourvues de HSF3 (Tanabe et al., 1998). HSF2 est principalement exprimée au début du développement embryonnaire chez la souris et le rat, et sous certaines conditions lors de la différenciation cellulaire, ou dans certains types cellulaires et tissus (Rallu et al., 1997; Sistonen et al., 1992). La mutation HSF2 *-/-* chez des souris cause des déficiences neuronales et de la reproduction. De plus, chez ces souris, l'activation des gènes de HSP par le stress est altérée : on note en effet une augmentation de l'expression des gènes *hsp70* et *hsp25* (équivalent de *hsp27* de l'homme), et une diminution de l'expression des gènes *hsp40* et *hsp105/110*. En effet, si l'activité de HSF2 n'est pas induite par les *stimuli* activant HSF1, il a cependant été montré que HSF2 est capable de s'associer avec HSF1 en un hétérotrimère, capable de se fixer à l'ADN et de moduler la transcription HSF1-dépendante, par exemple au niveau du gène *hsp70.1 (HSPA1A)* (Östling et al., 2007). La protéine HSF4 possède un domaine leucine-zipper supplémentaire en C-terminal, qui interférerait avec la trimérisation (Rabindran et al., 1993). De plus, elle se lie constitutivement à l'ADN, bien que son domaine transactivateur ne soit pas fonctionnel, ce qui indique qu'elle serait impliquée dans l'inhibition compétitive de l'activité des autres HSF (Nakai et al., 1997). Elle est notamment capable d'inhiber l'expression des gènes de *hsp27*, *hsp70* et *hsp90*.

Il semble donc exister une coopération (parfois inhibitrice) entre les différents facteurs HSF pour une régulation fine et complexe de la réponse au stress. Même si les modalités de cette coopération ne sont pas connues en détail, cela permet probablement une plus grande flexibilité de la réponse pour faire face à divers stress.

I.3.5. Importance des protéines de transduction du signal dans la réponse au stress thermique

Comme nous l'avons vu, l'activité des facteurs HSF est en effet également régulée par la présence de modifications post-traductionnelles, en particulier de phosphorylations. C'est aussi le cas pour les protéines HSP. Ainsi, les protéines des voies de signalisation "mitogen-activated protein kinases" (MAPK) phosphorylent de nombreux résidus Ser sur les sHSP, sur HSF1 et sur *hsp90* (pour revue, Garrido et al., 2012; Nadeau and Landry, 2007). Ces MAPK sont activées de façon très rapide après le choc thermique. Suite à ce choc les 3 MAPK ERK, JNK (c-Jun N-ter Kinase) et p38 sont activées suite à un choc thermique par phosphorylation selon 3 cascades de signalisation distinctes (pour revue, Chang and Karin, 2001; Takekawa et al., 2011) (figure I-2). ERK comme JNK peuvent activer directement HSF1 par phosphorylation (pour revue, Nadeau and Landry, 2007). On nomme parfois ces protéines les SAPK ("Stress-activated protein kinases", kinases activées par le stress). En effet, JNK et p38 peuvent également être activées par d'autres stress : le stress oxydant, des agents chimiques causants des dommages à l'ADN, l'exposition aux UV (Aikawa et al., 1997; Rosette and Karin, 1996;

Sachsenmaier et al., 1994). Les SAPK sont impliquées à la fois dans la progression du cycle cellulaire mais aussi dans l'apoptose. ERK est considérée comme un facteur anti-apoptotique, tandis que JNK et p38 sont des facteurs pro-apoptotiques (Desbiens et al., 2003; Tang et al., 2002; Tournier et al., 2000).

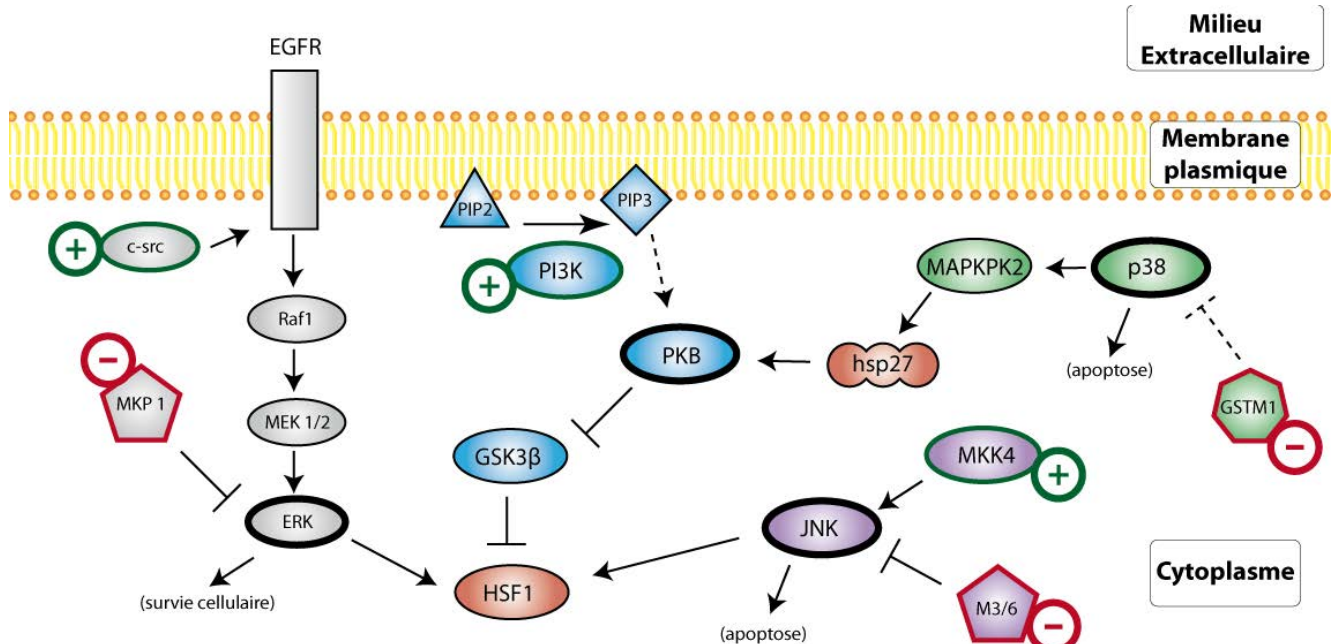
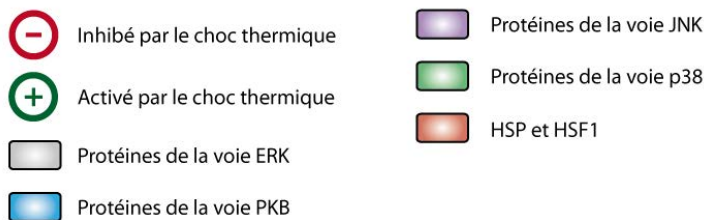


Figure I-2 : Activation des voies de signalisation des SAPK pendant le choc thermique. Les détails des régulations sont décrits dans le texte. L'activation d'une protéine par une autre est indiquée par une flèche, et l'inhibition par une barre. Les couleurs utilisées décrivent les différentes voies et les protéines activées ou inhibées suite au choc thermique de la façon suivante :



Lors d'un stress thermique, l'activation par phosphorylation des MAPK provient de l'inactivation thermique par dénaturation et agrégation de phosphatases impliquées dans la régulation des voies des MAPK, telles que la phosphatase de JNK, M3/6, ou les phosphatases de ERK, MKP3 et MKP1 (Palacios et al., 2001; Yaglom et al., 2003) (figure I-2). D'autre part, en particulier lors d'un stress thermique plus sévère, les MAPK peuvent être activées plus en amont par activation de kinase initiatrices de la cascade, telle que la kinase c-src (CSK) qui active EGFR, ce qui résulte à la phosphorylation de ERK (Lin et al., 1997). Ce type de régulation est également retrouvé lors d'autres stress tels que le stress lié aux UV ou le stress oxydant (Aikawa et al., 1997; Rosette and Karin, 1996; Sachsenmaier et al., 1994). De façon intéressante, hsp70 peut agir rétroactivement sur cette voie en stabilisant les phosphatases de ERK et

JNK et en bloquant les MAP2K (MAPK kinases) en amont (Park et al., 2001; Yaglom et al., 2003). Le mécanisme de l'activation de p38 par le choc thermique n'est pas totalement élucidé, mis à part le fait qu'il requiert la dissociation d'un inhibiteur de p38, la glutathione-S-transférase M1 (GSTM1), après le stress thermique (Dorion et al., 2002). La protéine hsp27/HspB1 est phosphorylée par MAPKAPK2, une cible directe de p38, ce qui active sa dimérisation et son activité chaperonne.

La voie impliquant l'activation de la kinase Phosphatidyl Inositol 3-Kinase (PI3K) et la protéine kinase B (PKB/Akt) est une autre voie de signalisation activée par le choc thermique, et qui favorise la survie cellulaire (figure I-2). L'activation de PI3K par le choc thermique génère des intermédiaires phosphatidylinositols PIP3, qui provoquent la translocation membranaire de PKB *via* l'activation de kinases PIP-dépendantes (PDK1 et 2) (Lin et al., 1997; Vanhaesebroeck and Alessi, 2000). L'activation de PKB par ce biais permet de limiter l'entrée en apoptose des cellules, en modulant l'activation de la voie JNK, la stabilité de p53, et l'inactivation de la protéine pro-apoptotique Bad. En outre, l'inhibition de PKB provoque une augmentation drastique de l'apoptose dans les cellules, même dans des conditions de choc thermique modéré (Bang et al., 2000). De plus, les dommages à l'ADN causés par le choc thermique provoquent l'induction des protéines H2AX et ATM, cette dernière ayant un lien direct avec les MAPK et l'activation de PKB (Tang et al., 2002).

L'inhibition de la glycogène synthase kinase 3β (GSK3 β) grâce à la phosphorylation par PKB peut être un autre phénomène important de la réponse au choc thermique, puisque la surexpression de GSK3 β empêche l'activation de HSF1 par phosphorylation au niveau du résidu Ser 303 et bloque donc l'activation de l'expression de HSP lors du choc thermique (figure I-2) (Bijur and Jope, 2000). De la même façon que pour les MAPK, les HSP peuvent agir rétroactivement sur cette voie. Hsp90 influence directement le repliement de PDK1 et PKB qui sont rapidement dégradées lorsque hsp90 n'est pas active (Fujita et al., 2002; Sato et al., 2000). Hsp27 interagit également avec PKB lors d'un stress et améliore son activité (Konishi et al., 1997).

Enfin, la protéine kinase C (PKC) est également activée en cas de stress thermique. Cette protéine est au centre de nombreux types de régulations métaboliques, notamment le métabolisme du Ca^{2+} auquel elle est sensible, et c'est un récepteur des esters de phorbols qui sont des promoteurs de tumeurs. La PKC est en effet impliquée dans de très nombreux processus métaboliques, notamment dans la prolifération des cellules endothéliales, la survie et l'apoptose (Lutzny et al., 2013)(pour revue, Breitkreutz et al., 2007). PKC est également à l'origine de l'hyperphosphorylation de la protéine HSF1. Cela explique pourquoi les agents activateurs de PKC, agissant sur le métabolisme du calcium (dont dépend PKC) peuvent également induire une réponse cellulaire similaire à celle du choc thermique (Ding et al., 1996; Kroeger et al., 1993). De plus, la surexpression de hsp70 provoque l'inhibition de l'activité de la PKC et donc diminue l'activité de HSF1 (Ding et al., 1998).

I.3.6. Arrêt du cycle cellulaire et apoptose induits par le choc thermique

Au travers de leur rôle de chaperonnes, les HSP sont aussi impliquées dans la régulation du cycle cellulaire et de la différenciation, qui sont des fonctions intervenant également dans la réponse au stress. Ainsi hsp27 limite l'activation de plusieurs procaspases telles que les procaspases 9 et 3, en captant et séquestrant le cytochrome c dès sa libération de la mitochondrie lors de l'enclenchement du processus apoptotique (Garrido et al., 1999; Pandey et al., 2000). Les protéines hsp70 et hsp90 sont aussi de puissants agents anti-apoptotiques, car elles interagissent avec de nombreuses protéines régulatrices des voies d'apoptose. L'accumulation des dommages cellulaires causés par la chaleur peut provoquer l'activation de réponses cellulaires conduisant à l'arrêt du cycle cellulaire puis à l'apoptose. Lorsqu'ils s'accumulent excessivement, ces dommages contrecarrent l'effet anti-apoptotique cytoprotecteur de certaines HSP. De plus, le choc thermique provoque la désorganisation des centromères, et dans les cellules tumorales, on observe également la reformation anarchique de ces centromères voire l'entrée en mitose non contrôlée des cellules (un phénomène nommé "mitotic catastrophe") pour les traitements les plus intenses, qui participerait à la mort cellulaire en induisant l'apoptose (Nakahata et al., 2002; Pawlik et al., 2013).

Néanmoins, les dommages à l'ADN ne semblent pas jouer un rôle dans le déclenchement de l'apoptose après un choc thermique, sans doute parce que le nombre de cassures double-brin de l'ADN reste limité en comparaison à celui induit par le traitement aux UV (Wong et al., 1995). Cela explique peut-être pourquoi l'apoptose n'est induite que par un stress thermique à des températures très élevées (fréquemment au delà de 46-48°C) et une durée d'exposition longue. Des expériences ont montré que l'exposition 60 à 120 min à 45°C de cellules humaines de poumon dérivées d'une lignée cancéreuse (A542) induit un nombre considérable de cassures double-brin de l'ADN, mais n'induit pas significativement la mort cellulaire. Par contre la mort cellulaire devient importante si le choc thermique est réalisé à 48°C (Speit and Schütz, 2013). Au contraire, il a été observé que les cellules exposées au choc thermique développent de façon surprenante une hypersensibilité aux radiations UV, contrairement à ce qui se passe avec la plupart des autres stress où le choc thermique a un effet cytoprotecteur. Cela peut s'expliquer par les déficiences de mécanismes de réparation de l'ADN causés par le choc thermique, évoquées au début de cette partie (pour revue, Tomita, 2010).

Par ailleurs, les HSP interviennent également dans la régulation de la progression à travers le cycle cellulaire. Initialement, il a été montré que le stress thermique intense s'accompagne de l'arrêt durable du cycle cellulaire en phase G1 ou G2, probablement à cause de la désorganisation des centromères (Nakahata et al., 2002). Cet arrêt du cycle n'est pas observé lors d'un stress thermique modéré, sauf lors de la déplétion des facteurs HSF1 et HSF3. Cela cause la réduction du taux de hsp90 dans les cellules, or hsp90 interagit directement et est nécessaire à l'activité de facteurs de progression dans le cycle cellulaire, comme les kinases cycline-dépendantes Cdc2 et Cdk4 (Nakai and Ishikawa, 2001; Stepanova et al., 1996). La réduction de hsp90 provoque une dénaturation thermique accrue de

Cdc2, ce qui résulte en l'arrêt du cycle cellulaire, principalement en phase G2/M mais aussi en phase G1 (Nakai and Ishikawa, 2001).

I.4. Valorisation de l'énergie anabolique après un stress thermique

Les nombreux dommages du stress thermique, essentiellement liés à la protéotoxicité, affectent indirectement une très grande partie des constituants et des fonctions cellulaires. Aussi, la réponse à ce stress nécessite la mobilisation de ressources conséquentes, ne serait-ce que pour l'expression massive des gènes de HSP. Il semblerait que l'énergie requise est obtenue dans la cellule par l'arrêt de certains processus moins indispensables, et par l'activation sélective et hiérarchisée des besoins essentiels à la survie de la cellule face au stress (pour revue, Niwa and Walter, 2000). Chez la levure par exemple, des études ont montré que l'activité transcriptionnelle des gènes codant des protéines dites de ménage ("housekeeping") est immédiatement diminuée pour laisser place à la transcription massive des gènes de réponse au choc thermique (Uffenbeck and Krebs, 2006).

I.4.1. Granules de stress cytoplasmiques et blocage transitoire de la traduction de certains ARNm

I.4.1.a. Formation des granules de stress cytoplasmiques

Lors d'un choc thermique, mais également lors d'autres stress comme le stress oxydant, l'infection virale, le stress généré par les UV et l'hypoxie, des granules de stress cytoplasmiques (cSG) sont formés dans le cytosol. Ils contiennent des ARNm dont la traduction coiffe-dépendante est inactivée, ainsi que des protéines, dont des HSP de faible poids moléculaire (Nover et al., 1983) (pour revue, Anderson and Kedersha, 2006). En effet, comme nous allons le voir, ces granules constituent un moyen d'inhiber sélectivement la traduction de certains ARNm (Nover et al., 1989). Ces ARNm sont pour la plupart issus des gènes codant des protéines de ménage normalement exprimés de manière constitutive.

La formation des cSG est enclenchée spécifiquement grâce à un processus d'avortement de la traduction des ARNm. Ce processus est initialisé par la phosphorylation de la sous-unité α du facteur d'initiation de la traduction eIF2 (eIF2 α) (figure I-3). La phosphorylation de ce facteur peut être catalysée par plusieurs kinases.

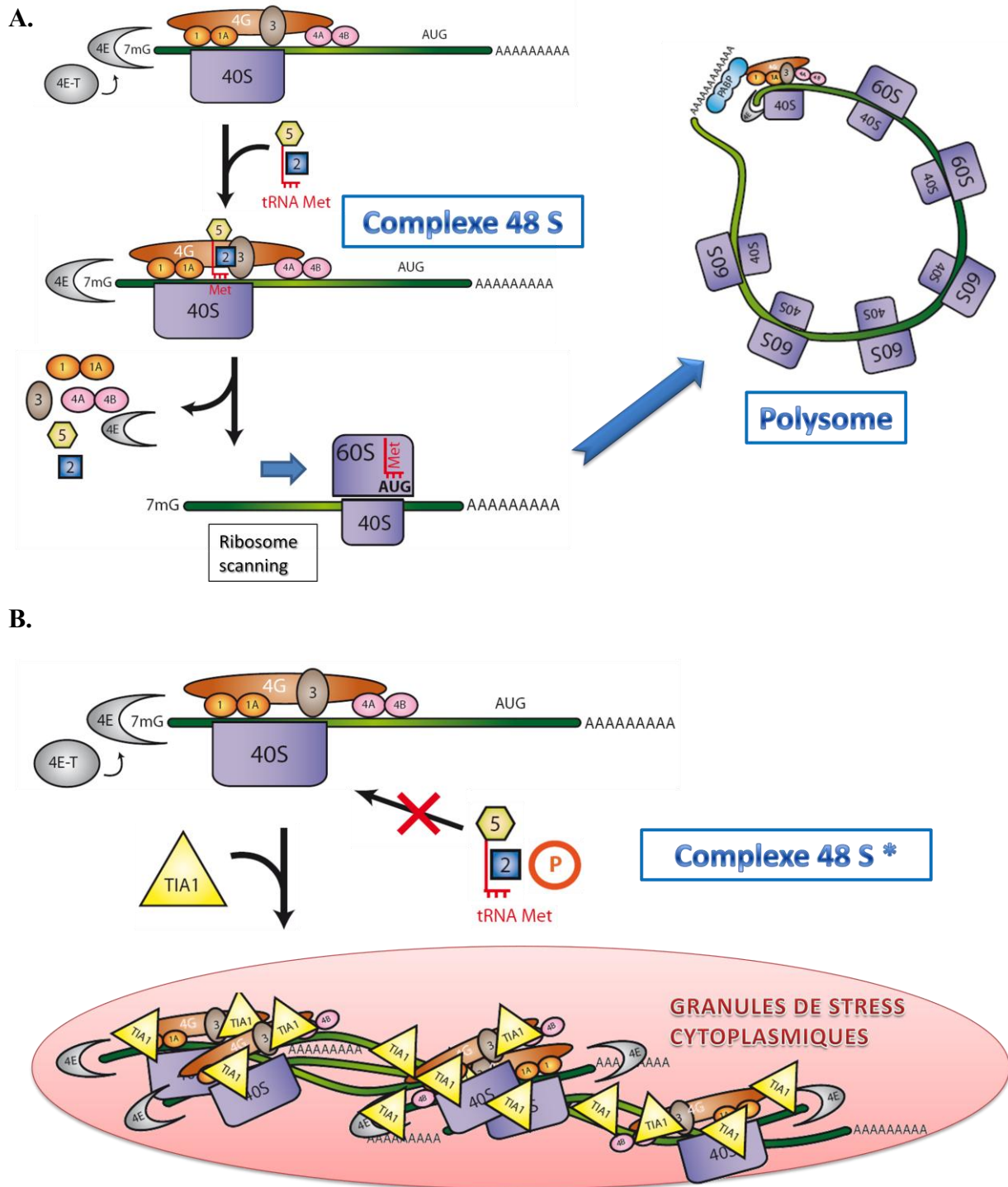


Figure I-3 : Formation des granules de stress cytoplasmiques. A. Déroulement normal de l'initiation de la traduction des ARNm. L'assemblage d'un complexe de traduction actif débute par le recrutement de facteurs d'initiation de la traduction eIF (notés ici 1 à 4G) et de la sous unité 40S du ribosome. Le chargement des facteurs eIF2 et eIF5 avec l'ARNt initiateur permettent la formation d'un complexe 48S actif, puis l'élimination des facteurs initiateurs, suivi par l'assemblage de la sous unité 60S du ribosome et le "ribosome scanning", jusqu'à la détection du codon initiateur AUG et le début de la traduction. **B.** Inhibition de la traduction et formation des granules de stress cytoplasmiques après un stress thermique. La phosphorylation (P) de eIF2 par la protéine PKR lors d'un stress thermique stoppe l'assemblage des complexes d'initiation de la traduction ce qui mène à la formation d'un complexe 48S inactif (48S*) et au recrutement de la protéine TIA-1. Les ARNm inhibés sont alors relocalisés, avec d'autres facteurs, pour former les granules de stress cytoplasmiques.

Dans le cas du stress thermique, elle est catalysée par la kinase PKR ("protéine kinase R" ou encore "Eukaryotic translation initiation factor 2-alpha kinase 2", EIF2AK2) (pour revue, Anderson and Kedersha, 2002). Cette kinase peut également être activée par la présence de grandes quantités d'ARN double brin en cas d'infection virale ou lors de traitements par les UV. Les kinases PERK, HRI et GCR2 peuvent assumer ce rôle dans le cas d'autres stress (Srivastava et al., 1998). La phosphorylation d'eIF2 α empêche le chargement de l'ARNt^{Met} initiateur dans les complexes d'initiation de la traduction devenus inactifs (figure I-3). Les ribosomes situés dans les polysomes lors du choc thermique finissent la synthèse de la protéine en cours puis sont libérés des ARNm, ce qui provoque le désassemblage des polysomes. Ces transcrits dont la traduction est inactive sont ensuite activement relocalisés dans les cSG à l'aide de protéines de liaison à l'ARN. Cela provoque notamment l'inclusion de la protéine TIA-1 au niveau des complexes d'initiation inactivés, qui participe à leur relocalisation avec d'autres facteurs protéiques spécifiques dans les cSG (figure I-3). D'autre part, ce mécanisme d'inhibition est aidé par la fixation spécifique de hsp27 sur eIF4G qui perturbe la formation des complexes d'initiation de la traduction, notamment en empêchant son interaction avec eIF4E et PABP1 (Cuesta et al., 2000; Ma et al., 2009). Les constituants protéiques principaux des cSG sont des complexes ribosomiques bloqués en phase de pré-initiation de la traduction, qui ne contiennent pas la sous-unité 60S du ribosome, mais qui contiennent les facteurs d'initiation eIF4E, eIF3, eIF4A, eIFG, et PABP ainsi que des protéines impliquées dans le métabolisme et la biogenèse des ARN, notamment la transcription et la dégradation des ARNm.

De plus, un certain nombre des protéines des cSG sont retrouvées également dans les P-bodies, qui sont des granules cytoplasmiques impliqués dans la dégradation des ARNm (Kedersha et al., 2002, 2005). Il existe un lien entre la composition, la répartition et la fonction des cSG et de celle des P-bodies, indiquant que les processus d'inhibition de la traduction pendant un stress et la dégradation des ARNm sont connectés. L'observation de granules hybrides entre cSG et P-bodies a d'abord conduit à l'hypothèse de l'existence d'un cycle des ARNm spécifique au stress, grâce à un échange entre P-bodies et cSG, permettant de passer de l'inhibition de la traduction des ARNm à leur dégradation par transition entre les granules (Buchan and Parker, 2009). Il s'avère cependant que les contacts entre P-bodies et cSG ont lieu uniquement pendant l'assemblage des cSG à proximité des P-bodies, et pas par rapprochement des deux structures préalablement assemblées. En plus, des cSG indépendants peuvent aussi se former ailleurs dans le cytoplasme (Mollet et al., 2008). Un point important mettant en avant l'aspect régulateur des cSG est que les transcrits non traduits ne sont pas à proprement parler séquestrés, puisque des échanges rapides de mRNP ont lieu entre les cSG et le cytosol, ainsi qu'avec les P-bodies lorsqu'ils sont adjacents aux cSG (Parker and Sheth, 2007a). C'est pourquoi l'inhibition de la traduction n'est que transitoire : l'activité traductionnelle des ARNm présents dans les cSG peut être réactivée.

1.4.1.b. Traduction indépendante de la coiffe et stress thermique

A cause de la complexité de leur assemblage et de leur composition, les détails du rôle fonctionnel des cSG ne sont pas encore entièrement compris. Il est cependant évident que les cSG participent au contrôle de l'expression des ARNm, probablement afin que la cellule concentre l'énergie disponible dans les mécanismes de survie. Par ailleurs, le choc thermique comme d'autres stress provoque une augmentation générale de l'utilisation d'IRES ("Internal Ribosome Entry Sites") pour la traduction de plusieurs ARNm de protéines impliquées dans la réponse au stress. Les IRES sont des régions très structurées présentes dans la partie 5'UTR de certains ARNm. En recrutant les ribosomes et certains des facteurs d'initiation de la traduction, elles permettent une initiation interne de la traduction, indépendante de la coiffe 5' terminale des ARNm. En plus de l'utilisation de certains facteurs d'initiation de la traduction, ces éléments utilisent d'autres facteurs nommés ITAF ("IRES Trans-acting factors"). La première mise en évidence de l'utilisation d'IRES lors d'un stress thermique a concerné l'ARNm de la HSP BiP/grp78 dans les cellules HeLa soumises à un choc thermique (Kim and Jang, 2002). Il a depuis été estimé par étude de fractions polysomales après choc thermique qu'au total environ 10% à 15% des ARNm traduits pourraient contenir un IRES (Spriggs et al., 2008; Stoneley and Willis, 2004). Il s'agit d'ARNm impliqués dans la régulation du cycle cellulaire, la mort cellulaire et de la réponse au stress. La traduction *via* les IRES expliquerait le maintien ou l'augmentation de la traduction des protéines essentielles à la réponse au choc thermique, dans des conditions où la traduction cap-dépendante est bloquée. Cela permettrait de reprogrammer la traduction pour produire en priorité les protéines indispensables à la réponse de la cellule, comme les HSP, afin de mieux utiliser l'énergie disponible. Ce concept de hiérarchisation de la production des protéines s'appliquerait aussi aux étapes précédentes de l'expression des gènes tels que l'épissage des ARNm et leur transport.

1.4.2. Blocage de l'épissage de certains ARNm

Un des phénomènes observés au début d'un choc thermique est l'arrêt de l'épissage de certains pré-ARNm. Comme l'épissage d'un pré-ARNm est indispensable pour son export dans le cytoplasme, le processus de synthèse des protéines correspondantes est également bloqué. Ce blocage au niveau de l'épissage est sélectif. Ainsi, on savait que les introns des transcrits des protéines hsp90 α , β et hsp27, sont correctement épissés dans des cellules soumises à un choc thermique. Par ailleurs, l'ARNm de hsp70.1 (*HSPA1A*) ne possède pas du tout d'intron (Jolly et al., 1999a), de même que les transcrits des gènes *HSPA1B*, *HSPA6*, et des sHSP *HSPB3* et *HSPB9* (source : IGD, "Intronless Genome Database" (Louhichi et al., 2011), <http://igd.bioinfo-cbs.org/>). Néanmoins, comme aucune étude à grande échelle n'avait été réalisée sur l'épissage après un choc thermique, la portion des transcrits bloqués à l'étape d'épissage n'était pas connue. De plus, ce blocage n'est pas observé systématiquement pour tous les types de stress.

Il avait été montré *in vitro* sur des extraits nucléaires qu'un stress thermique intense (43-46°C pendant 2h) provoque la formation de complexes spliceosomaux aberrants (Utans et al., 1992). Cette désorganisation est probablement due à l'inactivation d'un ou plusieurs facteurs d'épissage encore non définis mais désignés sous le nom de "Heat Shock Labile Splicing Factor" (HSLF). Ce ou ces facteurs, inactivés lors d'un choc thermique, sont de nature protéique puisque insensibles à la micrococcale nucléase, et il a été montré qu'ils sont impliqués dans l'assemblage de la tri-snRNP U4/U6/U5 (Shukla et al., 1990).

D'autre part, lors d'un stress hyperthermique sévère un mécanisme spécifique de régulation est mis en place. Il implique la protéine SR SRSF10 (SRp38), qui favorise la formation des complexes spliceosomaux, en renforçant l'interaction entre les snRNP U1 et U2 et le pré-ARNm (Feng et al., 2008). La déplétion en protéine SRSF10 provoque une inhibition globale de l'épissage (Shin et al., 2004). Le facteur SRSF10 est constitutivement phosphorylé, mais peut être déphosphorylé en réponse au stress thermique. La protéine SRSF10 déphosphorylée (dSRSF10) agit comme un répresseur de l'épissage, en interagissant avec la protéine U1-70K, composant de la snRNP U1 (figure I-4). Cela empêche l'interaction de U1-70K avec d'autres protéines SR à domaine RS, ce qui interfère avec l'assemblage des complexes spliceosomaux (Shin et al., 2004). Les enzymes qui déphosphorylent SRSF10 sont deux isoformes nucléaires d'une sérine-thréonine phosphatase des protéines SR, les phosphatases PP1 β et PP1 γ . Lors d'un stress thermique, la dissociation de l'inhibiteur constitutif de ces deux isoformes, NIPPI1 (Nuclear Inhibitor of PP1), permet leur activation (Shi and Manley, 2007).

Par ailleurs, en absence de stress, la protéine SRSF10 phosphorylée forme un complexe avec ses protéines de la famille 14-3-3, qui interagissent avec des résidus Ser/Thr phosphorylés de nombreuses cibles protéiques pour les protéger de la déphosphorylation. Le choc thermique provoque la dissociation de ce complexe ce qui libère la protéine SRSF10. Lorsque les cellules sont soumises à un stress thermique, c'est donc la conjonction de la dissociation de 14-3-3 consécutivement à l'activation de PP1 qui permet la déphosphorylation de SRSF10 (figure I-4).

PPR1 a de nombreux substrats parmi les protéines SR, mais seule SRSF10 est préférentiellement déphosphorylée lors d'un stress thermique. La kinase CLK/STY est responsable de la re-phosphorylation de dSRSF10 après un stress thermique (Ninomiya et al., 2011). Il a également été montré que hsp27 facilite la re-phosphorylation de SRSF10 et la reprise de l'épissage après un choc thermique sévère (1h à 45°C) (Marin-Vinader et al., 2006). En outre, l'épissage est moins affecté dans les cellules ayant subi un choc thermique sévère, lorsqu'elles sont soumises au préalable à un choc thermique moins violent et qu'elles produisent de grandes quantités de HSP (Yost and Lindquist, 1986). Tout ceci indique l'implication des HSP dans la récupération des capacités d'épissage après un choc thermique. Chez la levure, cette récupération est en effet favorisée par la protéine hsp70p, qui s'associe avec la tri-snRNP U4/U6/U5 dans les levures thermo-tolérantes, et par la protéine hsp105p (Bracken and Bond, 1999).

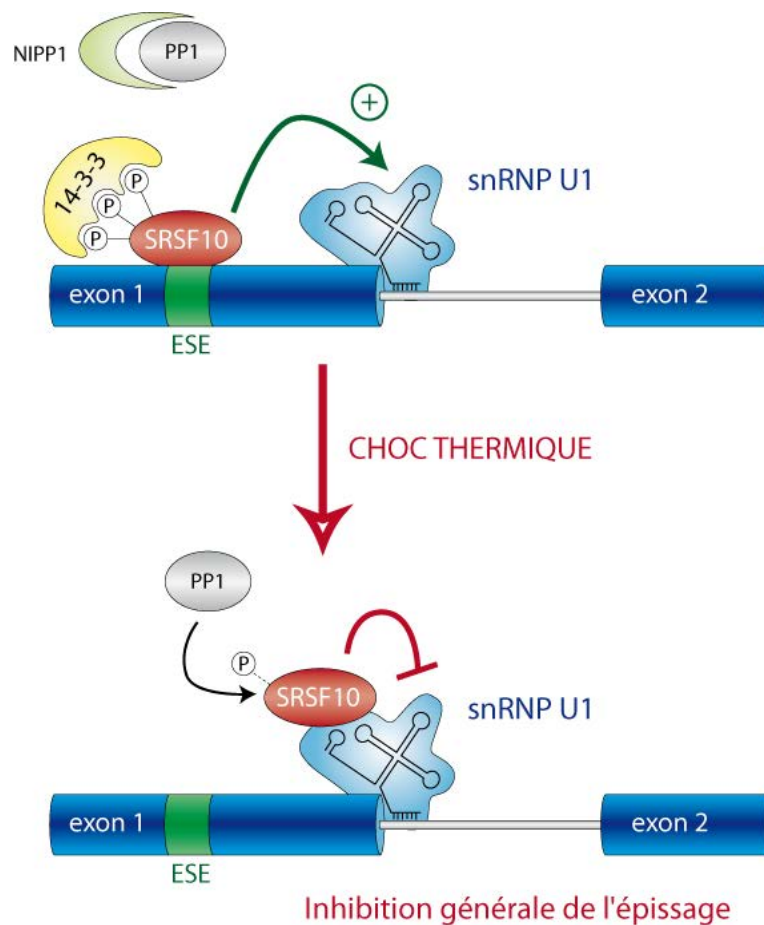


Figure I-4 : Mécanisme de répression de l'épissage dépendant de l'état de phosphorylation de SRSF10 (Srp38) (adapté de (Biamonti and Caceres, 2009)). Dans les cellules non stressées (partie supérieure), la protéine SRSF10 (Srp38) phosphorylée interagit avec des séquences de type ESE pour favoriser le recrutement de snRNP U1 et favoriser l'épissage à certains sites. Elle est protégée de la déphosphorylation par interaction avec la protéine 14-3-3. La phosphatase PP1 qui cible SRSF10 est inhibée par interaction avec NIPP1. Le choc thermique va provoquer la libération de NIPP1 et 14-3-3, ce qui conduit à la déphosphorylation de SRSF10. Dans cet état, SRSF10 agit comme un répresseur général de l'épissage en interagissant de manière stable avec la protéine U1-70K de la snRNP U1, empêchant ainsi son interaction avec ses partenaires.

Cette inhibition induite de l'épissage est souvent interprétée comme un moyen de la cellule pour mieux contrôler son métabolisme lors de la réponse au stress. Cette inhibition est maximum lors de stress sévères (45°C) et interviendrait de façon précoce dans le choc thermique. Cela permettrait notamment d'éviter les phénomènes d'épissage anormaux qui pourraient être délétères pour la cellule. En cas de stress sévère, l'inhibition de l'épissage permettrait de ralentir le métabolisme de la cellule, puis de réactiver successivement, de façon optimale et dans un ordre prédéfini, les mécanismes les plus essentiels. Certaines études vont dans le sens que ce phénomène pourrait être aidé dans divers stress dont le stress thermique par un contrôle de l'épissage alternatif.

I.4.3. Les régulations de l'épissage et de l'épissage alternatif suite à divers stress

Plusieurs types de stress provoquent des modifications de l'épissage et de l'épissage alternatif. Chez la levure, il a été observé grâce à des approches haut débit que l'activité d'épissage de sous-ensembles de pré-ARNm, comme ceux des protéines ribosomiques qui contiennent des introns, peut être augmentée ou diminuée en fonction du stress appliqué (déplétion en acides aminés du milieu, ou présence d'éthanol) (Pleiss et al., 2007). Par ailleurs, chez les plantes, l'épissage alternatif occupe une place prépondérante, et des travaux de ces dernières années indiquent qu'il constitue un élément important dans l'adaptation des plantes à un stress (pour revue, Reddy and Shad Ali, 2011). D'autres observations indiquent que l'épissage alternatif constitue également un élément important dans la réponse au stress chez les mammifères. Il a été observé qu'un stress génotoxique induit de différentes manières (inhibiteurs d'ADN topoisomérases, UV, molécules provoquant des dommages à l'ADN). Parmi ces événements on peut noter l'épissage alternatif de pré-ARNm de protéines intervenant dans la régulation de l'apoptose. La protéine Bcl-X par exemple dispose de deux isoformes, une courte pro-apoptotique Bcl-X_S, et une longue anti-apoptotique Bcl-X_L. Plusieurs molécules causant des dommages à l'ADN peuvent favoriser la production de la forme courte, via la régulation de l'utilisation de sites 5' d'épissage alternatif au niveau de l'exon 2 (Munoz et al., 2009). Des événements d'épissage alternatif sont également observés dans les pré-ARNm de plusieurs protéines de la famille des caspases (caspase 2, 8 et 9) menant à la production de protéines actives pro-apoptotiques (pour revue, Dutertre et al., 2011). Un autre des événements les plus notables concerne l'ARNm de la protéine MDM2, une ubiquitine ligase jouant le rôle de régulateur négatif de la protéine p53, induit par une panoplie d'agents génotoxiques (pour revue, Dutertre et al., 2011). Ces agents peuvent provoquer la perte de 8 exons de ce transcrite (sur 12), menant ainsi à la synthèse d'une forme inactive de MDM2 et permettant l'accumulation de la protéine p53 pour la réponse au stress. Un autre événement d'épissage avec un effet similaire a également été observé pour un autre inhibiteur de la protéine p53, MDM4 (Chandler et al., 2006).

L'épissage alternatif peut ainsi servir à la production de nouveaux isoformes protéiques avec une activité différente, mais aussi à la régulation du taux d'ARNm. C'est le cas par exemple pour le transcrite issu de l'oncogène *H-RAS*. La camptothécine (agent inhibiteur d'ADN topoisomérase I) peut induire l'inclusion de l'exon 5 dans l'ARNm H-ras, qui contient un codon stop prématuré et provoque la dégradation du transcrite grâce au système NMD. Cela empêche la production de la protéine p21^{H-ras}, impliquée dans la transduction par cascade de phosphorylation de signaux mitogènes extracellulaires, et limite donc la prolifération des cellules (Barbier et al., 2007).

Cependant, en ce qui concerne l'effet du choc thermique chez les mammifères, un seul événement d'épissage alternatif a été observé. Il s'agit de la sélection d'un site 5' d'épissage alternatif dans la région 5'UTR de l'ARNm de la protéine Hsp47 chez la souris (Takechi et al., 1994). Cet événement permet une traduction plus efficace de cette HSP. Les études sur ce type d'événements suivant un choc thermique semblent encore très rares chez l'homme, mais un événement en particulier indique que l'épissage alternatif pourrait assurer un rôle prépondérant dans la réponse au stress thermique : il s'agit de la formation de granules de stress nucléaires, suite à un stress thermique mais également divers autres stress.

I.5. Granules de stress nucléaires et contrôle de l'épissage alternatif

Dans le noyau environ une heure après le début d'un stress thermique, des structures appelées granules de stress nucléaires (nSG pour "nuclear Stress Granules") ou corpuscules de stress nucléaires (nSB pour "nuclear Stress Bodies") sont formées. Ces structures sont constituées notamment par des protéines intervenant pour la plupart dans l'épissage des ARNm. Comme nous allons le voir, une hypothèse fréquemment avancée dans la littérature est que ces protéines sont transitoirement piégées afin de moduler le profil d'épissage alternatif de la cellule, pour adapter la réponse au stress de la cellule et favoriser le retour progressif à un état métabolique non stressé. Ces protéines sont recrutées par des ARN spécifiques, les ARN satellite III, transcrits à partir de régions d'ADN satellite péri-centromériques que l'on a longtemps cru transcriptionnellement inertes.

I.5.1. ADN et ARN satellites péri-centromériques

Les régions centromériques et péri-centromériques des chromosomes contiennent des séquences d'ADN constituées de répétitions de tailles variables sur plusieurs milliers de nucléotides, appelées ADN satellites. Il existe différents types d'ADN satellites (α , β , γ , I, II, III) regroupés surtout en fonction de la longueur de leurs répétitions. Le caractère répétitif de ces séquences satellites rend difficile leur séquençage et leur positionnement dans le génome, en dépit des nombreuses initiatives de séquençage du génome humain. De ce fait, elles sont mal caractérisées. Les ADN satellite α et β sont retrouvés principalement dans les régions centromériques des chromosomes, tandis que les ADN satellite I, II, et III sont d'avantage retrouvés dans les régions péri-centromériques. Toutefois leur organisation est complexe et il est fréquent de retrouver plusieurs types d'ADN satellites amalgamés au sein d'une même région chromosomique (Finelli et al., 1996).

Les régions d'ADN satellite partagent la caractéristique d'être constitutivement organisées sous forme d'hétérochromatine très compacte, et portent plusieurs marques épigénétiques de la répression de la transcription. La principale marque épigénétique caractéristique de la chromatine silencieuse des

régions péricentromériques est la méthylation sur la lysine 9 de l'histone 3 des nucléosomes (H3-K9me), ainsi que l'association avec la protéine de l'hétérochromatine CBX5/HP1 (pour "Heterochromatin Protein 1"). Les régions centromériques sont elles associées avec des histones spécifiques des centromères, comme les protéines CENP-A (Craig, 2005; Peters et al., 2003).

Pendant longtemps, ces séquences ont été considérées comme inaccessibles aux facteurs de transcription et à l'ARN polymérase II tout au long du cycle cellulaire. Leur seul rôle avéré était leur implication dans l'ancrage des chromosomes au cytosquelette, la cohésion entre deux chromatides sœurs et leur séparation lors de la division cellulaire. Cette fonction est assurée grâce à l'interaction des séquences satellites avec des protéines spécifiques, notamment des cohésines et des protéines des kinétochores (Amor et al., 2004). Vers la fin des années 1960 il a été montré que la transcription de certaines séquences d'ADN péricentromériques était possible chez la souris (Harel et al., 1968). Depuis, cette transcription a été confirmée dans de nombreuses espèces du règne animal. Même s'il n'existe pas toujours un lien de cause à effet établi, la transcription de l'ADN péricentromérique a été observée dans plusieurs contextes, en particulier lors du développement embryonnaire chez la souris et l'homme, dans des contextes pathologiques où les régulations épigénétiques sont affectées (comme dans les cellules tumorales ou encore dans les cellules de patients progéria), dans les testicules où les régions péricentromériques ne sont pas transcriptionnellement réprimées, et, de façon intéressante, en cas de stress cellulaire (Eymery et al., 2009a, 2009b; Shumaker et al., 2006) (pour revue, Enuakashvily and Ponomartsev, 2013). En effet, suite à un choc thermique, la plupart des régions péricentromériques contenant des ADN satellites sont transcrites, consécutivement à la perte des marques épigénétiques de l'hétérochromatine péricentromérique H3-K9me et à la redistribution de la protéine HP1 de l'hétérochromatine péricentromérique. Ceci conduit à un remodelage de la chromatine permettant la transcription, qui par ailleurs est dépendante du facteur de transcription du choc thermique HSF1 décrit précédemment (Eymery et al., 2009a; Rizzi et al., 2004).

I.5.2. Description des ARN et des ADN satellites III

La transcription par l'ARN polymérase II des ARN satellite III (sat III) est spécifiquement activée suite à un stress thermique et également suite à divers autres stress comme l'exposition aux métaux lourds (Cadmium), le stress osmotique, le stress dû aux UV, le stress oxydant, et des stress provoqués par des molécules chimiques comme des molécules anti-tumorales bloquant la réplication de l'ADN (Jolly et al., 1999b; Sengupta et al., 2009; Valgardsdottir et al., 2007a). Les ARN sat III sont issus de l'ADN sat III des régions péricentromériques des chromosomes, majoritairement de la région q12 du chromosome 9 (9q12), mais également des régions centromériques des chromosomes 12 et 15 (Denegri et al., 2002; Jolly et al., 2002). Ces régions correspondent à des sites de transcriptions actifs de l'ARN polymérase II en cas de stress, et les transcrits sat III sont tous polyadénylés (Rizzi et al., 2004;

Valgardsdottir et al., 2005a). Une autre étude a également montré l'existence de sites de transcriptions mineurs des ARN sat III, mais également de régions contenant des ARN sat II, au niveau des chromosomes 4, 5, 7, 9, 10, 13, 14, 15, 17, 21, 22 et Y (Eymery et al., 2010). Certains ARN sat III ont pu être isolés et séquencés, et il apparaît qu'ils représentent un ensemble de transcrits polyadénylés de tailles très variables, mais possédant une organisation similaire. Ces ARN sont en effet constitués de deux types de répétitions. Les premières sont des séquences de pentanucléotides répétés en tandem, dont la séquence consensus, GGAAU, est retrouvée dans 60% des répétitions. Les 40% restants sont des séquences peu divergentes, elles présentent une variation à une seule position par rapport à la séquence consensus. Les secondes répétitions sont des séquences décanucléotidiques plus variables, dont la séquence consensus CAACCCGAGU est retrouvée dans 25% des répétitions. Les variations des répétitions de décanucléotides par rapport au consensus peuvent concerner de 1 à 4 nt. Les pentanucléotides et les décanucléotides sont organisés de façon particulière : dans tous les clones d'ARN sat III séquencés une unité structurelle comporte de 1 à 39 pentanucléotides, suivie d'un seul décanucléotide. Cette unité est répétée un certain nombre de fois pour former un ARN dont la taille varie, selon le nombre d'unités répétées, de quelques dizaines de bases à plus de 5 kb (figure I-5) (Rizzi et al., 2004; Valgardsdottir et al., 2005a).

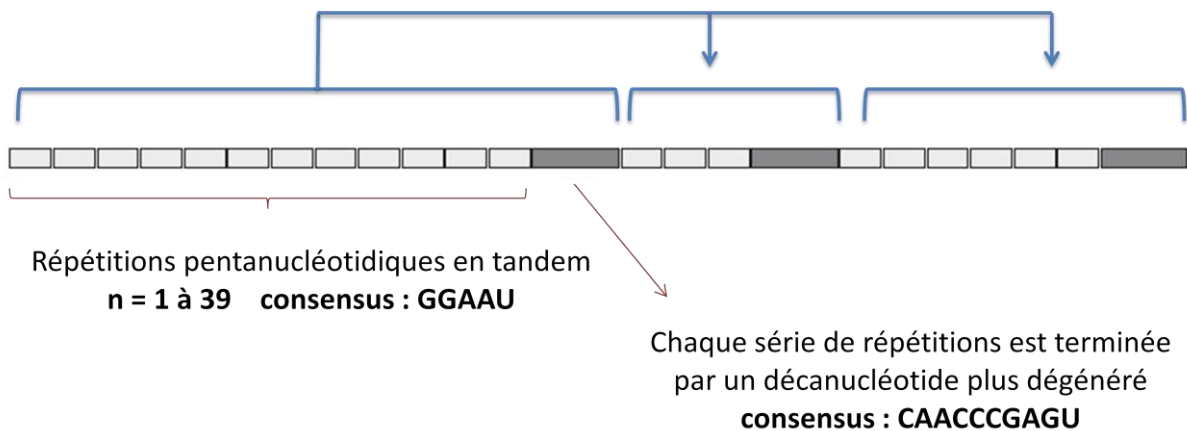


Figure I-5 : organisation des séquences d'ARN sat III (selon Valgardsdottir et al., 2005a). Chaque unité structurelle, constituée par des répétitions de pentanucléotides (rectangles clairs) suivies par un décanucléotide (rectangle foncé), est répétée au sein d'un ARN sat III pour former des ARN de tailles très variables (crochets et flèches bleues).

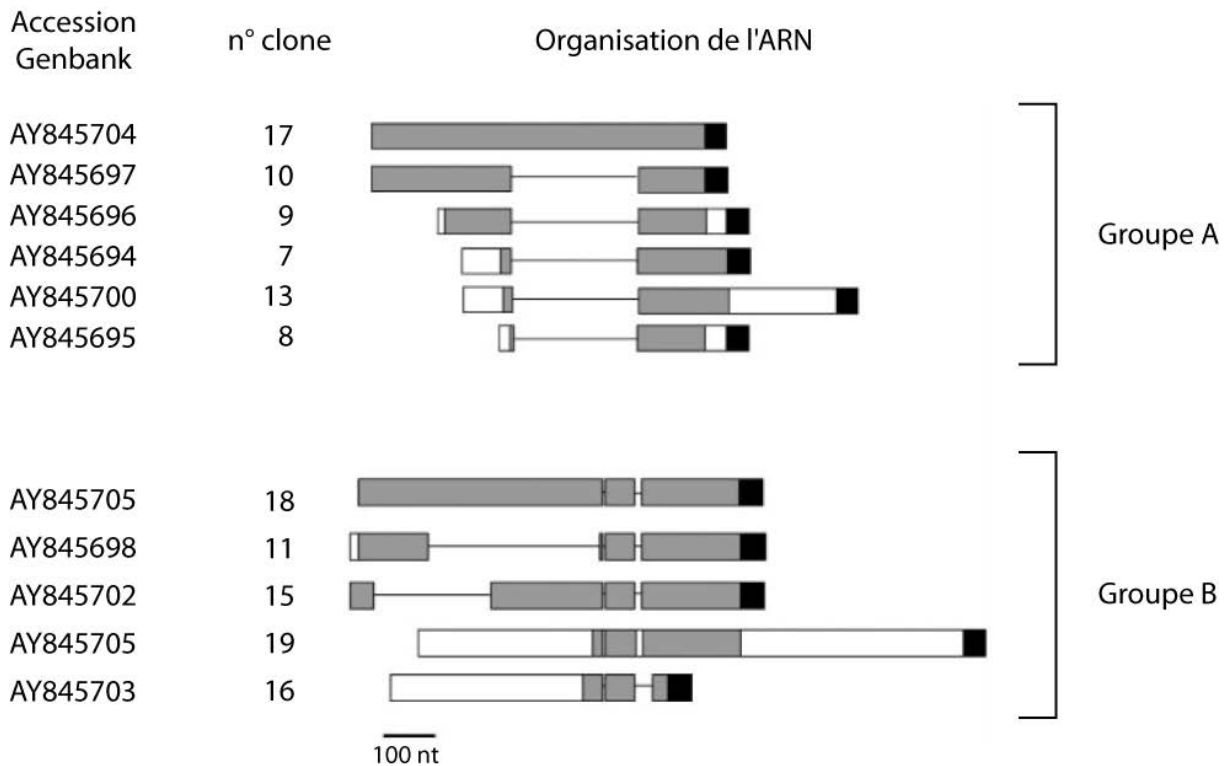


Figure I-6 : Homologies présentes dans les clones d'ARN sat III purifiés (tiré de Valgardsdottir et al., 2005a). En fonction de leur alignement, les séquences de certains des clones d'ARN sat III purifiés ont pu être classées en deux groupes (A et B). Les numéros des clones donnés par l'équipe à l'origine de ce travail sont indiqués à gauche, avec les références d'accèsion dans la base de donnée GenBank. La barre en dessous de la figure représente l'échelle pour 100 nt. Les symboles utilisés pour représenter les séquences sont les suivants :

- Queue polyA
- Séquences <95% d'homologie
- Séquences homologues
- Séquence absente ("Gap")

D'autre part, les différents ARN sat III présentent de grands blocs de séquences homologues ou similaires. L'alignement des séquences des clones connus a permis de les classer en deux groupes distincts, qui présentent des régions d'identité pouvant aller jusqu'à plusieurs centaines de nucléotides. De façon intéressante, la comparaison des séquences ARN d'un même groupe met en évidence des délétions de séquences rappelant l'épissage alternatif des introns (figure I-6). De ce fait la diversité des ARN sat III pourrait provenir à la fois de l'épissage alternatif d'un même transcrite et/ou de phénomènes de duplication de l'ADN sat III (Valgardsdottir et al., 2005a). L'hypothèse d'un épissage alternatif est en accord avec le fait que les UsnRNP peuvent être détectées au niveau des sites de transcription des ARN sat III (Metz et al., 2004). D'autre part, les séquences ADN péri-centromériques sont fréquemment sujettes à des phénomènes de duplication, ce qui rend la seconde hypothèse également envisageable (Humphray et al., 2004).

Enfin, il est intéressant de noter que dans certaines conditions, le brin anti-sens de l'ADN sat III peut aussi être transcrit (Valgardsdottir et al., 2007a). Le brin sens est par convention nommé "G-rich" (riche en G), car riche en séquences GGAAU, et c'est le brin principalement transcrit, tandis que le brin antisens est nommé brin "C-rich" (riche en C), car riche en séquences AUUCC. Cet ARN anti-sens est toujours co-localisé avec l'ARN sat III sens dans les cellules. Son induction est toutefois bien plus faible, car elle n'est que d'un facteur 10 tandis que l'ARN sat III G-rich atteint un facteur 1000 à 10 000x.

I.5.3. Transcription des ARN satellite III par HSF1 ou TonEBP/NFAT5

Comme il a été mentionné plus haut, lors de la plupart des stress, la transcription des ARN sat III est initiée grâce au facteur de transcription HSF1. Après sa translocation dans le noyau suite à un choc thermique ou à l'exposition au cadmium ou à l'analogue d'acide aminé azétidine, HSF1 forme des granules dans le nucléoplasme (Cotto et al., 1997; Jolly et al., 1997a; Metz et al., 2004; Sarge et al., 1993). La formation de granules HSF1 commence dès le début du choc thermique, et elle est optimale environ 30 min à 1 h après le début du stress. De plus, elle est réversible lorsque les cellules sont régénérées à 37°C (Jolly et al., 1999b; Metz et al., 2004). La co-localisation de ces granules HSF1 avec la région péricentromérique 9q11-9q12 du chromosome 9 contenant des séquences satellite III a permis de montrer que HSF1 est à l'origine de l'induction de la transcription des ARN sat III (Jolly et al., 2002; Metz et al., 2004; Rizzi et al., 2004).

Après une heure de choc, en général plus de 90%-95% des cellules présentent des granules, mais leur nombre et leur morphologie variable selon les noyaux, le type cellulaire et la température du choc thermique. Par exemple, dans les cellules humaines primaires comme les fibroblastes et les cellules épithéliales, après un choc à 42°C d'une heure, la grande majorité des cellules présentent 2 gros granules ressemblant à des clusters d'une dizaine de petits foci (Cotto et al., 1997; Jolly et al., 1997a). En revanche dans les lignées testées de cellules transformées comme les cellules HeLa, HOS et A431, dans des conditions similaires de choc thermique, il est possible d'observer jusqu'à une dizaine de foci par noyau, dont deux sont des gros granules d'environ 2 µm de diamètre, les autres étant des foci ponctuels et brillants de taille plus réduite (environ 1µm) (Alastalo et al., 2003; Cotto et al., 1997). Ces foci peuvent correspondre à d'autres sites de nucléation sur des chromosomes 9 supplémentaires, à cause de la polyploïdie aberrante de certaines cellules immortalisées, ou bien à des sites secondaires de formation des granules correspondant à d'autres régions d'ADN sat III ou sat II, présentes sur les régions péricentromériques de plusieurs autres chromosomes (Cooke and Hindley, 1979; Denegri et al., 2002; Eymery et al., 2010; Jolly, 2004). La transcription des ARN sat III repose vraisemblablement sur les similitudes existant entre les répétitions pentanucléotidiques 5'-GGAAT-3' des ADN sat III et les Heat Shock Elements (HSE) des promoteurs des gènes de HSP. Ces éléments sont en effet composés de répétitions de séquences 5'-nGAAn-3'. *In vitro*, HSF1 est capable d'interagir directement avec des ADN

double brin issus d'un clone d'ADN sat III connu (pHur98) (Jolly et al., 2002). De plus, l'induction de la transcription des ARN satellite III requiert à la fois la présence du domaine de liaison à l'ADN de HSF1 et son domaine de trimérisation (Jolly, 2004; Jolly et al., 2002; Metz et al., 2004). Les foci nucléaires contenant HSF1 sont également enrichis en histones H2A, H2B, H3 et H4 acétylées, qui sont des modifications épigénétiques fréquemment associées à une activité transcriptionnelle. Ils sont aussi co-localisés avec des foci enrichis en ARN polymérase II, en histone acétyltransférase (HAT) (catalysant l'acétylation des histones) et CBP (CREB Binding Protein), indiquant une forte augmentation de la transcription dans ces zones (Jolly, 2004; Rizzi et al., 2004). Ces granules ne possèdent pas non plus d'autres marqueurs caractéristiques de l'hétérochromatine, telle que la méthylation des lysines 9 des histones 3 (H3-K9me). Toutes ces données corroborent le fait que le facteur de transcription HSF1 active de façon spécifique la transcription des ARN sat III après le stress, par l'intermédiaire du recrutement de co-activateurs nécessaires à la décondensation de la chromatine, comme CBP, permettant l'accès de l'ARN polymérase II aux séquences ADN sat III.

Plusieurs résultats suggèrent fortement que HSF2 est impliquée dans la transcription des ARN sat III. L'implication de HSF2 dans la formation des nSB après un stress a été étudiée, et il en ressort plusieurs résultats contradictoires. Une première expérience avait montré l'absence de co-localisation entre HSF1 et HSF2, et même avec HSF3 et HSF4 après un choc thermique (Cotto et al., 1997; Nakai et al., 1997; Sarge et al., 1993). Cependant plusieurs autres études ultérieures, réalisées avec des cellules HeLa également, montrent que HSF1 et HSF2 sont co-localisées au niveau des nSB, et ce au bout de 15 minutes de choc. Cette co-localisation résulterait de la formation d'hétérotrimères HSF1-HSF2, *via* leurs domaines de trimérisation homologues, capables de se fixer au niveau des ADN sat III pour initier la transcription (Alastalo et al., 2003; Sandqvist et al., 2009). Au vu des relations fonctionnelles qu'il existe entre HSF1 et HSF2, il est probable que HSF1 et HSF2 coopèrent pour la régulation de la transcription des ARN sat III. En effet, en absence de stress, la co-expression de HSF1 et HSF2 à partir de plasmides, ou bien la surexpression de HSF2 dans les cellules permet l'activation de HSF1, sa translocation dans le noyau et la transcription de ses gènes cibles (Sandqvist et al., 2009; Sarge et al., 1993).

Alors que la relocalisation de HSF1 dans des granules permettant la transcription des ARN sat III a lieu suite à une variété de stress autres que le stress thermique (métaux lourds, UV, stress oxydant). Dans le cas du stress osmotique, un autre facteur de transcription, TonEBP ("Tonicity Element Binding Protein") est utilisé (Valgardsdottir et al., 2007a). Ce facteur, aussi connu sous le nom de OREBP ("Osmotic Response Element Binding Protein") ou NFAT5, interagit avec des séquences ADN présentes sur les promoteurs des gènes cibles (tel que celui de hsp70.2) nommées TonE (pour "Tonicity Element") (Ferraris et al., 2002). Ces séquences, dont le consensus est TGGAAANN(C/T)N(C/T), sont retrouvées dans plusieurs clones d'ADN satellite III.

Le fait que plusieurs facteurs de transcription spécifiques de réponse au stress (HSF1, HSF2 et TonEBP) soient impliqués dans l'induction de la transcription des ARN sat III suggère que ces ARN et

les granules qu'ils forment au sein du noyau (nSB) jouent un rôle important dans le processus d'adaptation général de la cellule en réponse au stress.

I.5.4. Organisation et fonction des nSB

L'initiation de la transcription des ARN sat III dans les granules HSF1 provoque la relocalisation dans ces granules de plusieurs facteurs protéiques, impliqués dans la maturation et l'épissage des pré-ARNm. En effet, les ARN sat III ne sont pas exportés hors du noyau, et restent associés à leur site de transcription, par un mécanisme encore non élucidé (Valgardsdottir et al., 2005a). HSF1, les ARN sat III et ces facteurs forment des domaines subnucléaires appelés nuclear Stress Bodies (nSB), dont la formation débute par la relocalisation de HSF1 et perdure jusqu'à plusieurs heures après la fin du stress thermique selon la durée du choc (Cotto et al., 1997; Jolly et al., 1997a, 1999b; Metz et al., 2004). Il est à noter que plusieurs agents physico-chimiques (chaleur, UV, la mitomycine C qui endommage l'ADN) sont capables d'initialiser l'appariement ("pairing") des deux régions chromosomiques homologues 9q12-13 autour desquelles sont formés les nSB (Abdel-Halim et al., 2006). Le rôle de cet appariement est inconnu, mais un phénomène de propagation dans les cellules non stressées a été observé (Abdel-Halim et al., 2006). Il pourrait intervenir dans la formation des nSB, ou contribuer à la protection des locus à proximité desquels se forment les nSB, ou bien influencer leur transcription.

Les facteurs hnRNP HAP, SRSF9 (SRp30c), SRSF1 (ASF/SF2), hnRNP M, Sam68 et dans une moindre mesure SRSF7 (9G8) sont connus pour être concentrés dans les nSB (Chiodi et al., 2000; Denegri et al., 2001; Gattoni et al., 1996; Metz et al., 2004; Weighardt et al., 1999). SRSF7 n'est que partiellement relocalisée dans les nSB et est également retrouvée dans les speckles. Toutes les protéines SR ou les facteurs d'épissage ne sont pas associés aux nSB, puisque après un choc thermique, SRSF2 (SC35) reste majoritairement dans les speckles, et hnRNP A1, A2, et C1/C2 ne sont pas retrouvées non plus dans les nSB (Denegri et al., 2001; Weighardt et al., 1999).

L'utilisation de siRNA dirigés contre l'ARN sat III provoque la dissociation des nSB, confirmant que l'ARN sat III joue un rôle important dans la formation de ces granules. Ce même type d'expérience a également montré qu'il est indispensable à la relocalisation de la hnRNP HAP et de SRSF1 dans les nSB (Chiodi et al., 2000; Valgardsdottir et al., 2005a; Weighardt et al., 1999). Des expériences réalisées avec des inhibiteurs de la traduction ont aussi montré que cette relocalisation ne semble pas dépendre de la traduction cap-dépendante, qui est inhibée lors d'un stress thermique (Weighardt et al., 1999). D'autre part, les ARN sat III sont très stables, puisqu'ils peuvent perdurer dans les nSB jusqu'à plus d'une dizaine d'heures après la fin du choc, y compris lors de la mitose après la reprise du cycle cellulaire (Jolly, 2004). A ce stade HSF1 commence déjà à se dissocier des granules. En effet, le recrutement de HSF1, puis des autres protéines relocalisées dans les nSB comme HAP pendant la synthèse de l'ARN sat III, se fait de manière séquentielle (figure I-7) (Metz et al., 2004; Weighardt et al., 1999).

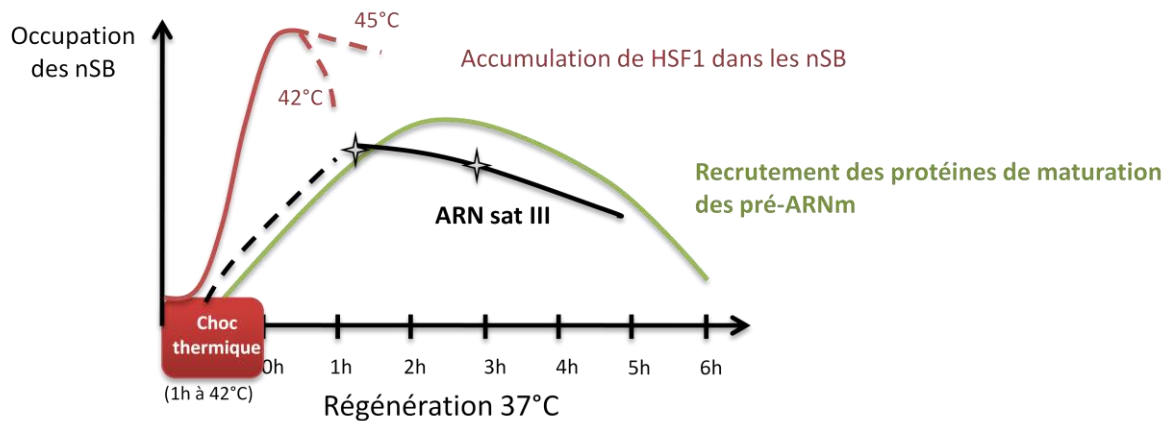


Figure I-7 : Diagramme schématisant la cinétique de formation des granules de stress dans le noyau d'après l'analyse des principaux composants.

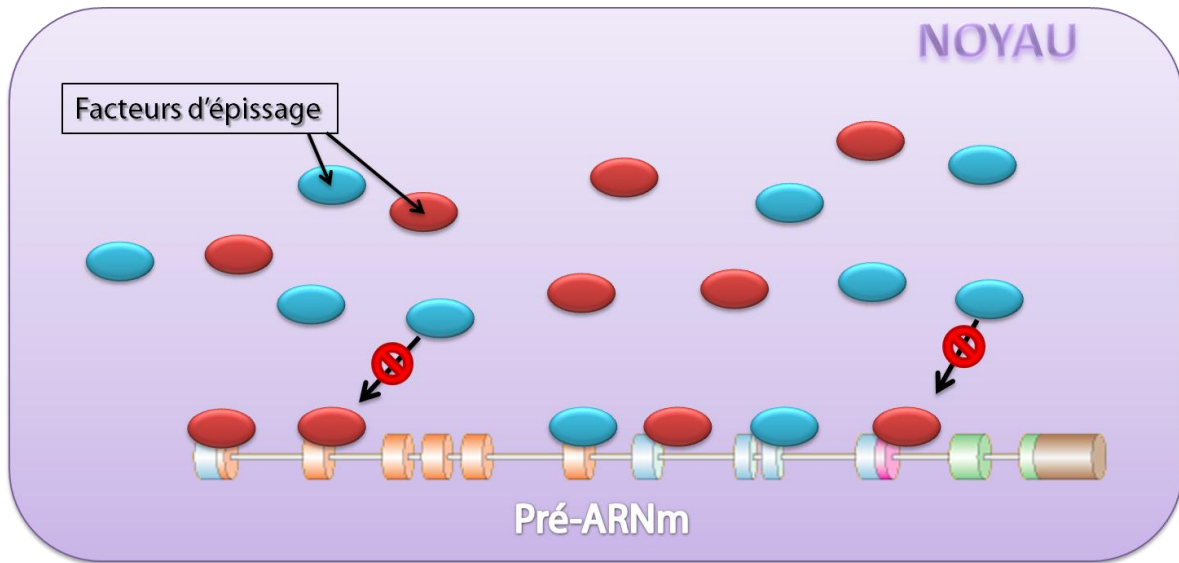
Bien que la transcription de l'ARN sat III et la formation des nSB soient observées pour différents types de stress, la cinétique, la morphologie et le nombre de nSB par noyau varient selon le type et la durée du stress. L'induction la plus importante et la plus précoce des nSB (c'est à dire de 1 à 2 h après la fin du stress) est provoquée par 1 à 2 h d'exposition à 42°C, bien qu'une exposition prolongée au cadmium (50 μ M pendant plus de 6 h) ait des effets similaires. Dans ce cas, l'ARN sat III est surexprimé environ 10000 x et plus de 90% des cellules HeLa présentent plusieurs granules de grande taille. D'autres stress ont un impact moins important, comme le stress liés aux UV-C (80J/m²), le stress oxydant (20 min d'exposition à H₂O₂) ou le stress hyperosmotique (0,8 M sorbitol), pour lesquels la formation des nSB est maximale environ une quinzaine d'heures après le stress. De plus, le nombre et la taille des granules changent également : 30 à 40% des cellules forment des nSB après le stress lié aux UV ou hyperosmotique, et environ 5% lors d'un stress oxydant. Ces granules sont également plus petits, très ponctuels, et moins nombreux, lors d'un stress oxydant ou de l'exposition à la drogue antitumorale étoposide. Enfin, certains stress comme l'hypoxie n'induisent pas du tout la formation de nSB. Un résumé des profils de formation des nSB obtenus suite à différents stress est donné dans le tableau I-1.

Facteur d'induction				
Temps d'induction	Induction faible	Induction moyenne	Induction forte	Pas d'induction
Précoce	MMS	Cadmium (1 h + régénération)	Choc thermique	Hypoxie
Moyenne	Etoposide Aphidicoline		Cadmium (6 h constant)	Cobalt Thymidine
Tardif	oxydation	UV-C Hyper-osmotique		Hypo-osmotique

Tableau I-1 : Effets de différents stress sur la production d'ARN sat III et la cinétique d'apparition des nSB. Adapté de (Valgardsdottir et al., 2007a). Les lignes correspondent aux temps de régénération suivants pour l'apparition des nSB : précoce : entre 2 et 4 h après la fin du stress, moyen : entre 6 et 8 h et tardif : plus de 10 h. Les colonnes correspondent aux facteurs d'induction de l'ARN sat III, au maximum de l'induction : faible : <100 x, moyen : de 100 à 1000 x, fort : > 1000 x. Les différents stress avaient été réalisés avec les expositions des cellules (pendant 1 h sauf indication) dans les conditions suivantes : UV-C : 80 J/m², stress hyperosmotique : 0,8 M Sorbitol, choc thermique : 42°C, Cadmium : 50 µM (pendant une heure suivi d'une période de régénération ou 6 h sans régénération), stress hypo-osmotique : milieu 30 % DMEM + 70 % eau, stress oxydant : 200 µM H₂O₂ pendant 20 min, hypoxie : 3 % O₂ ou 21% O₂ + 100 µM CoCl₂ (Chlorure de Cobalt), Méthyl Méthanosulfate (MMS) : 3,54 mM, aphidicoline : 2 µg/mL, étoposide : 100 µM, thymidine : 2 mM.

C'est entre autres la persistance des nSB après le choc thermique et différents autres stress qui a mené à proposer que les nSB agissent sur la régulation de l'épissage alternatif après un stress. En effet, la séquestration de plusieurs facteurs de l'épissage dans les nSB diminuerait leur concentration effective dans le nucléoplasme ce qui affecterait le profil d'épissage alternatif global dans la cellule après le stress (figure I-8). Cette hypothèse a été renforcée par l'observation que le profil d'épissage alternatif du pré-ARNm issu du gène rapporteur E1A de l'adénovirus est modifié suite à un stress thermique ou un traitement au sulfate de cadmium (Denegri et al., 2001). Il est intéressant de noter que le profil d'épissage alternatif de ce transcrit est modulé par l'antagonisme entre les protéines SRSF1 et SRSF7, présentes dans les nSB, et hnRNP A1 et SRSF2, qui n'y sont pas présentes (Caceres et al., 1994; Gattoni et al., 1991). D'un point de vue cinétique, cette régulation est optimale lors de la période de recouvrement à des temps compatibles avec la formation des nSB, que ce soit pour le stress thermique ou le traitement au cadmium. Ce type de régulation serait de plus en adéquation avec le blocage de l'épissage ayant lieu au début du choc thermique. Ce blocage cesse rapidement après le choc, et pendant la période de régénération, la reprise de l'épissage et l'épissage alternatif seraient alors contrôlés par la formation des nSB.

A.



B.

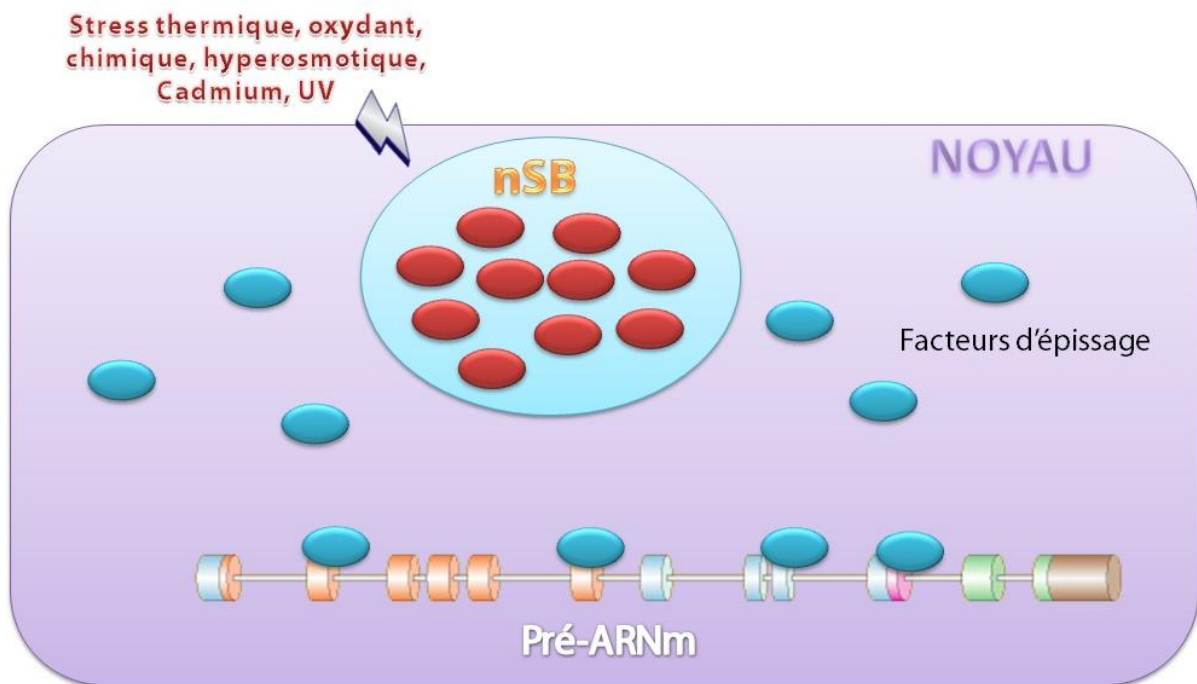


Figure I-8 : Modèle de la régulation du profil d'épissage alternatif global dans le noyau par formation des granules de stress nucléaires. A. Cellules non stressées. B. Cellules stressées.

Cela permettrait d'une part d'optimiser la production d'isoformes protéiques nécessaires à la régénération des cellules, et d'autre part d'éviter l'épissage aberrant provoqués par d'éventuels dommages causés par le stress. Ce modèle rejoint le principe du contrôle appliqué sur la traduction des ARNm grâce à la formation des granules de stress cytoplasmiques.

Des images de microscopie électronique ont montré que les nSB sont en partie constitués par des agglomérations de nano-structures nucléaires appelées granules périchromatiniens (Perichromatin Granules, ou PG) (Charlier et al., 2009; Chiodi et al., 2000). Les PG sont des structures denses et très ponctuelles (de 40 à 50 nm) observées en microscopie électronique, qui sont mal caractérisées car très peu nombreuses dans les cellules non stressées (de 500 à 2000 chez les mammifères) et de taille très réduite (Watson, 1962). Par ailleurs, il a été observé que le choc thermique, tout comme d'autres stress tels que l'inhibition de la traduction, de la transcription ou l'exposition aux ultraviolets provoquent une augmentation drastique du nombre de PG dans le noyau (Charlier et al., 2009; Daskal et al., 1975; Mähl et al., 1989; Miller et al., 2002). On parle même parfois en cas de stress thermique de hsiPG, pour "heat shock induced Perichromatin Granules". Les hypothèses sur les rôles des PG formés dans les cellules concernent le transport ou le stockage d'ARNm ou de pré-ARNm dont la maturation a été affectée suite à un stress (Iborra et al., 1998; Vázquez Nin et al., 1996, 1997). En effet, ces PG qui s'agglomèrent dans les cellules stressées hébergent des ARNm à longue durée de vie correspondant sûrement aux ARN sat III, et des facteurs d'épissage tels que protéines hnRNP et SR.

I.5.5. Protéines relocalisées dans les nSB

I.5.5.a. Protéines hnRNP et interactions avec la matrice nucléaire

Les premières observations de co-localisation de protéines avec HSF1 et les ARN sat III dans les nSB après un choc thermique concernaient les protéines hnRNP M et hnRNP HAP (pour hnRNP A1 Associated Protein). Ces deux protéines sont présentes à la fois dans les PG et dans les nSB après un choc thermique (Chiodi et al., 2000). HnRNP M est une protéine abondante, souvent associée aux complexes spliceosomaux en formation et impliquée dans la régulation de l'épissage alternatif de pré-ARNm spécifiques (Hovhannisyanyan and Carstens, 2007; Mähl et al., 1989). En temps normal, comme beaucoup d'autres protéines intervenant dans la maturation des ARN, hnRNP M s'associe aux transcrits naissants dans des particules ribonucléiques hétérogènes hnRNP. Elle est également présente dans des structures subnucléaires appelées paraspeckles (Marko et al., 2010). Les paraspeckles sont des structures de stockage de facteurs impliqués dans l'épissage, souvent associées spatialement et fonctionnellement aux speckles. Cependant, les partenaires de hnRNP M dans le noyau changent pendant et après un choc thermique : elle est alors retrouvée dans des fractions de coefficient de sédimentation plus importants, et son association avec la matrice nucléaire est renforcée (Gattoni et al., 1996; Mähl et al., 1989). Les mêmes observations ont été faites avec la protéine hnRNP HAP des nSB, et la protéine hnRNP 2H9/H3. Ce type de comportement lors d'un choc thermique distingue d'ailleurs ces protéines des autres hnRNP et constitue d'ailleurs une caractéristique qui les distingue des autres hnRNP (Mahé et al., 1997; Weighardt et al., 1999). A 37°C, hnRNP M forme des petits granules dans le nucléoplasme que l'on observe aussi durant le choc thermique. Toutefois, pendant la période de régénération qui suit, hnRNP M est également

relocalisée dans les agglomérats de PG observés après le choc thermique, et dans des granules de taille plus importante correspondant sans doute aux nSB (Chiodi et al., 2000; Gattoni et al., 1996; Mähl et al., 1989). Dans les cellules non stressées, hnRNP HAP est présente de façon diffuse légèrement granuleuse dans le nucléoplasme, hors du nucléole, et est relocalisée en grande majorité dans les nSB juste après un choc thermique (Denegri et al., 2001). Les expériences d'immunofluorescence ont montré que la co-localisation entre hnRNP HAP et hnRNP M n'est pas parfaite au sein des nSB (Chiodi et al., 2000).

La protéine hnRNP HAP fut initialement purifiée grâce à son interaction avec des éléments de séquence d'ADN de 300 à 3000 pb appelés SAR ou MAR ou encore S/MAR (pour "Scaffold" ou "Matrix Attachment Region", région d'attache à la matrice) (Renz and Fackelmayer, 1996). C'est pourquoi elle fut nommée alors SAF-B (pour "Scaffold Attachment Factor-B"). Les éléments S/MAR, par le biais de leur interaction avec la matrice protéique du noyau à laquelle appartiennent les protéines SAF, interviennent dans l'organisation topologique des domaines de la chromatine. L'interaction de HAP avec les éléments S/MAR participent à un effet isolateur des domaines chromosomiques par deux S/MAR en interaction. Ces domaines se présentent alors sous forme de boucles de chromatine décondensée où les gènes sont accessibles (pour revue, Cook, 1991). Cela permet la délimitation de zones de transcription active des gènes. Les éléments S/MAR sont peu conservés en séquence (pas de consensus), bien qu'ils aient la particularité d'être riche en nt A et T. Les modalités de fixation de SAF-B sur l'ADN au niveau de ces séquences ne sont pas connues. HAP/SAF-B est également connue sous le nom de HET (pour "Hsp27-ERE-TATA-binding protein"), car elle a été identifiée comme étant un facteur de transcription interagissant avec l'ADN au niveau du promoteur du gène hsp27 pour inhiber sa transcription (Oesterreich et al., 1997). Par ailleurs, des expériences de double hybride ont permis de montrer son interaction avec des protéines impliquées dans la transcription ou l'épissage des pré-ARNm. On retrouve parmi ses partenaires la protéine hnRNP A1, le domaine C-terminal de l'ARN polymérase II, les protéines SR et SR like SRSF1, SRSF9, SRSF7 et hTra2- β , la kinase des protéines SR CLK2, dont HAP est un substrat, et dans une moindre mesure les protéines U2AF et la protéine SR SRSF4 (SRp75) (Denegri et al., 2001; Nayler et al., 1998; Weighardt et al., 1999). En outre, hnRNP HAP est capable d'influencer le choix des sites d'épissage d'un gène rapporteur E1A de l'adénovirus *in cellulo* lorsqu'elle est surexprimée. Cela suggère qu'elle est capable d'influencer la fixation à l'ARN des protéines SR et hnRNP partenaires. Cette idée est en accord avec le fait que la délétion des domaines RS des protéines SRSF9 et SRSF1, normalement nécessaires pour l'interaction avec d'autres protéines SR, empêche leur interaction avec HAP dans des expériences de double hybride (Nayler et al., 1998).

1.5.5.b. Réseau d'interactions protéiques dans les nSB

Il est intéressant de constater que la protéine SAF-B/HAP, au vu de ses fonctions de protéine de la matrice nucléaire, pourrait être un facteur qui (1) relie les ARN sat III avec leur ADN matrice

péricentromérique afin de former les nSB, et (2) peut servir de plateforme d'assemblage des nSB. HAP semble bien posséder les propriétés de plateforme d'assemblage, d'après le grand nombre de partenaires avec lesquels elle interagit directement (figure I-9). La protéine hnRNP HAP est constituée en N-ter d'un domaine acide d'environ 300 aa, immédiatement suivi par un domaine RRM, pouvant refléter des capacités d'interaction directe avec l'ARN, en particulier avec les ARN sat III. Cette possibilité est renforcée par le fait que la présence d'ARN et d'une transcription active est requise pour l'association de SAF-B à la matrice nucléaire, ce qui suggère la formation de complexes tripartites matrice/SAF-B/ARN (Weighardt et al., 1999). Ceci pourrait expliquer pourquoi une transcription active et la présence des ARN sat III sont nécessaires à la relocalisation de HAP/SAF-B (mais également de SRSF1 et SRSF9) dans les nSB (Metz, 2004; Valgardsdottir et al., 2005a). Deuxièmement, HAP présente dans sa partie C-terminale une région protéique déterminante pour l'interaction avec de nombreux partenaires. Cette région est constituée de deux domaines, l'un riche en paires d'acides aminés S/K et l'autre en paires R/E, suivis à l'extrémité C-terminale d'un domaine riche en Gly (figure I-9) (Nayler et al., 1998; Weighardt et al., 1999). Le domaine riche en Gly permet l'interaction avec le domaine CTD de l'ARN polymérase II et SRSF1 tandis que le domaine R/E permet l'interaction avec SRSF9 et SRSF7. Cette région est également nécessaire pour le recrutement de la protéine HAP au niveau des nSB (Denegri et al., 2001). Certains des partenaires de la protéine hnRNP HAP *in vitro*, les protéines SRSF1, SRSF9, et dans une moindre mesure SRSF7, sont co-localisées avec HAP dans les granules après un stress (Denegri et al., 2001). Il est donc probable que ces 3 protéines et l'ARN sat III forment un complexe dans les nSB, et que l'ARN sat III soit le composant indispensable à la formation de ce complexe. De fait, des expériences de shift (EMSA) ont montré qu'une séquence d'une dizaine de nt présente dans l'ARN sat III, et qui est similaire à une séquence ESE, peut interagir directement avec la protéine SRSF1 *in vitro*. Cette interaction se fait par le biais de l'hélice α présente en N-ter du domaine RRM2 de la protéine SRSF1. De façon attendue, cette hélice est également nécessaire et suffisante pour la re-localisation de la protéine SRSF1 dans les nSB (Chiodi et al., 2004).

Dans les cellules non stressées, la protéine hnRNP HAP est également retrouvée partiellement co-localisée avec une protéine appelée Sam68 (pour "Src-activated during mitosis"), dans des granules de taille inférieure à celle des nSB, nommés SNB (Sam68 Nuclear Bodies) (Denegri et al., 2001). Dans les SNB on trouve également les protéines SLM-1 et SLM-2 (Sam68 Like Mammalian Protein 2), des homologues très proches de Sam68, et contenant comme elle un domaine STAR (pour "Signal Transduction and Activation of the RNA"), indispensable pour leur rôle de transduction du signal (Chen et al., 1997, 1999; Vernet and Artzt, 1997). En effet, Sam68 est une protéine exclusivement nucléaire qui intervient dans la progression du cycle cellulaire, grâce à la phosphorylation de résidus tyrosine C-terminaux par des kinases comme Src et Cdc2, qui sont des acteurs centraux dans le contrôle de la progression du cycle cellulaire (Resnick et al., 1997). Sam68 joue également un rôle dans l'export des ARNm et dans l'épissage alternatif de plusieurs pré-ARNm, comme celui de la protéine mTOR

impliquée dans la régulation du cycle cellulaire (Huot et al., 2012). Elle constituerait un lien important entre les phénomènes de transduction du signal par phosphorylation et le métabolisme de l'ARN (Barlat et al., 1997; Reddy et al., 2000; Vernet and Artzt, 1997).

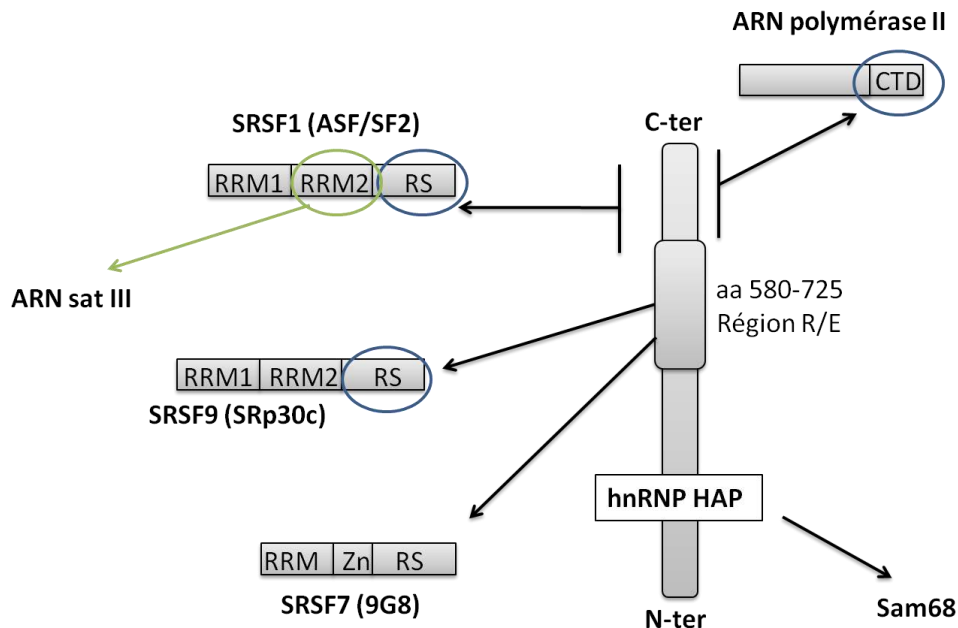


Figure I-9 : Représentation schématique du réseau d'interactions possibles entre les protéines des granules de stress nucléaires. La protéine hnRNP HAP est susceptible de servir de plateforme d'assemblage au vu des interactions qu'elle est capable de former avec les différentes protéines des nSB (voir texte). Les domaines d'interactions des partenaires avec hnRNP HAP sont entourés.

Des expériences de double hybride, de co-immunosélection et de pull-down chez le rat ont montré que les homologues de rat de Sam68 et de SLM-2 sont capables d'interagir *in cellulo* avec des protéines hnRNP et des protéines SR, et d'influencer l'épissage alternatif de différents mini-gènes de façon dose-dépendante (Stoss et al., 2001). Il est intéressant de noter que parmi ces partenaires, on retrouve les protéines HAP/SAF-B et SRSF9, présentes dans les nSB, ainsi qu'un partenaire commun avec HAP, Tra2 β , dont la présence dans les nSB n'est pas encore établie. En fait, après un choc thermique, Sam68 est relocalisée dans les nSB avec ses partenaires hnRNP HAP et SRSF9. C'est pourquoi les nSB ont également été qualifiés de "SNB induits par le stress" ("stress-induced SNB"). De plus, conformément à ce que suggère la présence d'un domaine de liaison à l'ARN KH ("hnRNP K homology"), Sam68 est capable d'interagir directement avec plusieurs espèces d'ARN issus d'ARN totaux marqués *in vitro* (Wong et al., 1992). Cette interaction avec l'ARN est directement liée à l'état de phosphorylation de Sam68 (Wang et al., 1995). Les SNB ne sont pas des sites de transcription active, mais en revanche, de la même manière que pour les nSB, une transcription active est nécessaire pour la formation et la maintenance des SNB, ce qui laisse supposer que l'intervention de certains ARN est nécessaire à leur formation (Chen et al., 1999).

I.6. Liens entre les mécanismes de défense au stress thermique et vieillissement

Les connaissances sur des mécanismes de défense face au stress sont utiles pour une meilleure compréhension du vieillissement. Par leurs propriétés cytoprotectrices, les HSP, outre la réponse au stress, protègent aussi les organismes contre les désordres métaboliques, neurodégénératifs et cardiaques (Morimoto and Santoro, 1998). En effet, les protéines mal repliées, agrégées ou qui forment des plaques amyloïdes, sont impliquées dans de nombreuses pathologies (pour revue, Morimoto and Cuervo, 2009). Ainsi, des mutations ou des défauts dans la production des HSP ont pu être reliés à des contextes pathologiques, notamment le cancer et les maladies neurodégénératives, mais sont également constatés lors de la sénescence des cellules ou lors du vieillissement.

Par ailleurs, un certain nombre de gènes codant des protéines impliquées dans la défense face aux stress protéotoxiques, comprenant de nombreux gènes de HSP, sont considérés aujourd'hui comme critiques dans le processus de vieillissement et parfois regroupés sous le terme "vitagènes" (Calabrese et al., 2011, 2012; Rattan, 1998). Or, il a été noté une diminution de l'expression de hsp70 qui est proportionnelle à l'âge de l'organisme dans les cellules hépatiques de rat (Heydari et al., 1993). Dans plusieurs tissus, l'induction des HSP suite à un stress thermique est fortement diminuée chez les rats âgés par rapport aux jeunes rats (Blake et al., 1991). Le même phénomène est observé dans les fibroblastes d'humains âgés, et a pu être relié à la diminution globale de l'occupation des éléments HSE par la protéine HSF1 (Choi et al., 1990; Liu et al., 1989). En parallèle, la sénescence induite par les dommages à l'ADN est à l'origine d'une diminution de l'expression de HSF1, ce qui renforce le phénotype sénescence (Kim et al., 2012).

D'autre part, la transcription des ARN sat III à partir du chromosome 1, mais pas du chromosome 9, a été observée dans des cellules tumorales et de façon plus importante dans des fibroblastes sénescents ainsi que dans des fibroblastes de patients atteints d'une maladie neurodégénérative, l'ataxia telangiectasia, dont l'une des caractéristiques est le vieillissement prématuré (Erukashvily et al., 2007). Cette activation de la transcription est consécutive à la déméthylation et la décondensation de la chromatine, et les régions ainsi décondensées s'associent avec plusieurs autres domaines chromosomiques du noyau, notamment les régions péri-centromériques du chromosome 9, avec lesquelles elles partagent des séquences conservées (Erukashvily et al., 2007). D'autre part, dans les cellules de patients atteints de la progéria, de nombreux changements au niveau des marqueurs épigénétiques sont observés, notamment des marqueurs épigénétiques de l'hétérochromatine de la région 9q12 contenant les ARN sat III (Shumaker et al., 2006). Il en résulterait la transcription des ARN sat III et la formation de granules de stress nucléaires, y compris en absence de stress. Le lien entre l'expression de l'ARN sat III et le phénotype observé dans ces maladies est encore inconnu.

Partie 2 - Objectifs

Comme nous venons de le voir dans le chapitre introduction de cette partie, de nombreux stress conduisent à l'expression des ARN satellite III qui forment des granules de stress nucléaires (nSB). Nous nous sommes au départ intéressés aux nSB du fait de leur détection dans les cellules de patients HGPS (Shumaker et al., 2006). Etant donnée la re-localisation de certaines protéines SR, et en particulier de SRSF1, que nous avons montrée impliquée dans la régulation de l'épissage des pré-ARNm LMNA, dans les nSB, on pouvait s'attendre à avoir un effet global de différents stress sur l'épissage et notamment sur l'épissage du gène LMNA. L'examen de la littérature révélait que peu de données étaient disponibles quant aux effets de la formation des nSB sur l'épissage alternatif. C'est pourquoi nous avons entrepris d'étudier l'impact de la formation des nSB sur l'épissage alternatif à l'échelle du génome humain. Le modèle de stress que nous avons retenu pour mener cette étude est le choc thermique. Le taux de formation des nSB est en effet très élevé dans les cellules humaines en culture soumises au choc thermique et ce stress constitue donc un bon modèle d'étude de la formation des nSB et des conséquences de leur formation sur l'expression des gènes. L'idée était qu'à terme ces informations puissent compléter les connaissances sur la progéria. Cette étude à haut débit a été réalisée en employant des micropuces spécifiques des jonctions exoniques.

Par ailleurs, les connaissances sur la structure secondaire des ARN satellite III et leur mode d'interaction avec les protéines des granules étaient incomplètes. Ayant des compétences en matière d'étude de la structure secondaire des ARN et d'étude des interactions ARN-protéine, nous avons aussi décidé d'entreprendre une étude de la structure secondaire des ARN satellite III et de leur interaction avec les protéines nucléaires.

L'ensemble de ces travaux concernant la seconde partie de ma thèse est présenté ci-après.

PARTIE 2

STRESS THERMIQUE ET EPISSAGE

(Résultats)

II. Etude des ARN satellites III et recherche de leurs partenaires protéiques

II.1. Analyse de la structure secondaire de l'ARN satellite III

II.1.1. Stratégie d'analyse et mise au point

Comme il a déjà été mentionné, les ARN sat III comportent des enchaînements de répétitions en tandem de pentanucléotides (GGAAU) entrecoupés de décanucléotides (CAACC/ACGAGU) (Valgardsdottir et al., 2005b). On pourrait se demander si ces modules de séquence constitutifs des ARN sat III peuvent adopter une structure secondaire particulière. Des travaux réalisés précédemment au laboratoire et ceux d'autres équipes avaient par exemple montré que les répétitions de triplets CUG présents dans les transcrits DMPK, chez les patients atteints de dystrophie myotonique de type I, sont capables d'adopter une structure en double hélice de type ADN, dont dépend le recrutement et la séquestration de la protéine MBNL1 dans des structures granulaires comparables aux nSB (Holt et al., 2007). Nous avons donc tenté d'établir la structure secondaire de modules d'ARN sat III. Le clone pHur98 que nous avons utilisé pour cela nous avait été fourni par l'équipe de G. Biamonti. Il est issu d'une banque de séquençage de la région 9qh (chromosome 9) (figure II-1) (Moyzis et al., 1987).

5'-AAUCAACCCGAGUGCAAUCGAAUGGAAUCGAAUGAAUGGAAUGCAAUGGAAUGGAAU
UCAACUUGAAUGGAAUGGAAAGAAUGGAAUCAACACGAGUGGAAUGGCAUGGAAUUGGA
AUGGAAUGGAAUGGAAUCAACCCGAGUACAGGAAUGGAAUGGAA-3'

Figure II-1 : Séquence de l'ARN satellite III issu du clone pHur98 (158 nt). Cet ARN renferme 4 décanucléotides caractéristiques des ARN sat III (soulignés), l'un d'entre eux ayant une séquence assez dégénérée par rapport au consensus établi (CAACC/ACGAGU). Des répétitions de 9, 4 et 7 pentanucléotides caractéristiques des ARN sat III (GGAAU), plus ou moins dégénérés, séparent respectivement les 4 décanucléotides. Dans ce schéma, les décanucléotides sont soulignés, les séquences en gras correspondent à des pentanucléotides incomplets (un à deux nucléotides délétés) qui sont naturellement présents dans le clone pHur98. La séquence d'hybridation de l'oligonucléotide amorce 7239 utilisé pour les études de cet ARN est encadrée.

Afin de déterminer la structure secondaire de cet ARN, nous avons utilisé la même stratégie que pour la région de l'intron 11 du pré-ARNm LMNA dans le chapitre précédent, à savoir l'utilisation de sondes enzymatiques spécifiques de régions d'ARN en simple brin ou en double brin. Dans un premier temps, j'ai essayé d'utiliser la transcriptase inverse pour identifier les positions de clivage mais comme nous allons le voir, cette approche n'a pas été satisfaisante. Nous avons essayé de choisir un

oligonucléotide (7239) qui puisse s'hybrider spécifiquement à l'extrémité 3' de l'ARN en dépit des nombreuses répétitions de séquence. L'avantage de cette extrémité est qu'elle comporte un triplet atypique ACA. Néanmoins l'hybridation ne s'est pas avérée assez spécifique pour utiliser cette approche pour l'identification des positions de clivage. L'ajout d'une séquence étiquette en 3' de l'ARN pHur98 aurait peut-être permis d'hybrider de façon plus spécifique une amorce mais cela risquait de modifier la structure de l'ARN. C'est pourquoi nous avons préféré réaliser les expériences en utilisant des molécules d'ARN pHur98 radiomarquées en 5' afin de nous affranchir des problèmes de lecture provenant de la rétrotranscription.

II.1.2. Analyse de la structure secondaire de l'ARN sat III pHur98

L'identification des sites d'hydrolyse par les sondes enzymatiques en utilisant un ARN directement radiomarké se fait par comparaison avec les produits d'une hydrolyse par la RNase T1 effectuée sur un ARN dénaturé (identification des résidus G), ainsi que les produits d'une hydrolyse alcaline ménagée, permettant le clivage à chaque position de l'ARN (figure II-2). Comme pour l'ARN LMNA 638 précédemment, les données issues des hydrolyses enzymatiques ont été utilisées comme contraintes dans le logiciel M-fold, afin de rechercher une structure secondaire optimale compatible avec les sites de clivage. L'utilisation de durées différentes d'électrophorèse nous a permis de couvrir l'ensemble de la molécule.

L'analyse expérimentale réalisée met en évidence l'existence d'une structuration de l'ARN pHur98 à la fois par la présence d'un certain nombre de clivages par la RNase V1, et par l'absence de clivages par les RNases dans une grande partie de la molécule (correspondant aux nt 85 à 125 environ). Dans l'état actuel des données, il est difficile de prédire précisément quelle est la structure adoptée par l'ARN. Plusieurs appariements alternatifs étaient possibles du fait de la redondance des séquences. Un point intéressant à noter est la localisation assez prédominante des clivages par la RNase V1 dans 2 des 4 décanucléotides, et l'absence de clivages au sein des deux autres décanucléotides, ce qui suggère que ces décanucléotides pourraient jouer un rôle important dans la structuration des ARN sat III. Certains d'entre eux sont en effet partiellement auto-complémentaires. Une des structures possibles de l'ARN, qui est le modèle le plus satisfaisant obtenu à partir du logiciel M-fold, est représentée dans la figure II-2. Ce modèle n'est pas optimal car il n'est pas totalement en accord avec les données expérimentales, notamment par la présence d'une boucle aux positions 101 à 105. Quoiqu'il en soit, contrairement à ce qui était espéré, l'ARN étudié ne semble pas adopter une structure régulière en lien avec la présence des répétitions pentanucléotidiques. Toutefois, nous ne pouvons pas exclure la possibilité que la structure du clone pHur98 ne soit pas représentative de la structure générale des ARN sat III. Elle met peut-être en œuvre des interactions à plus longue distance ou repose peut-être sur l'interaction avec des protéines.

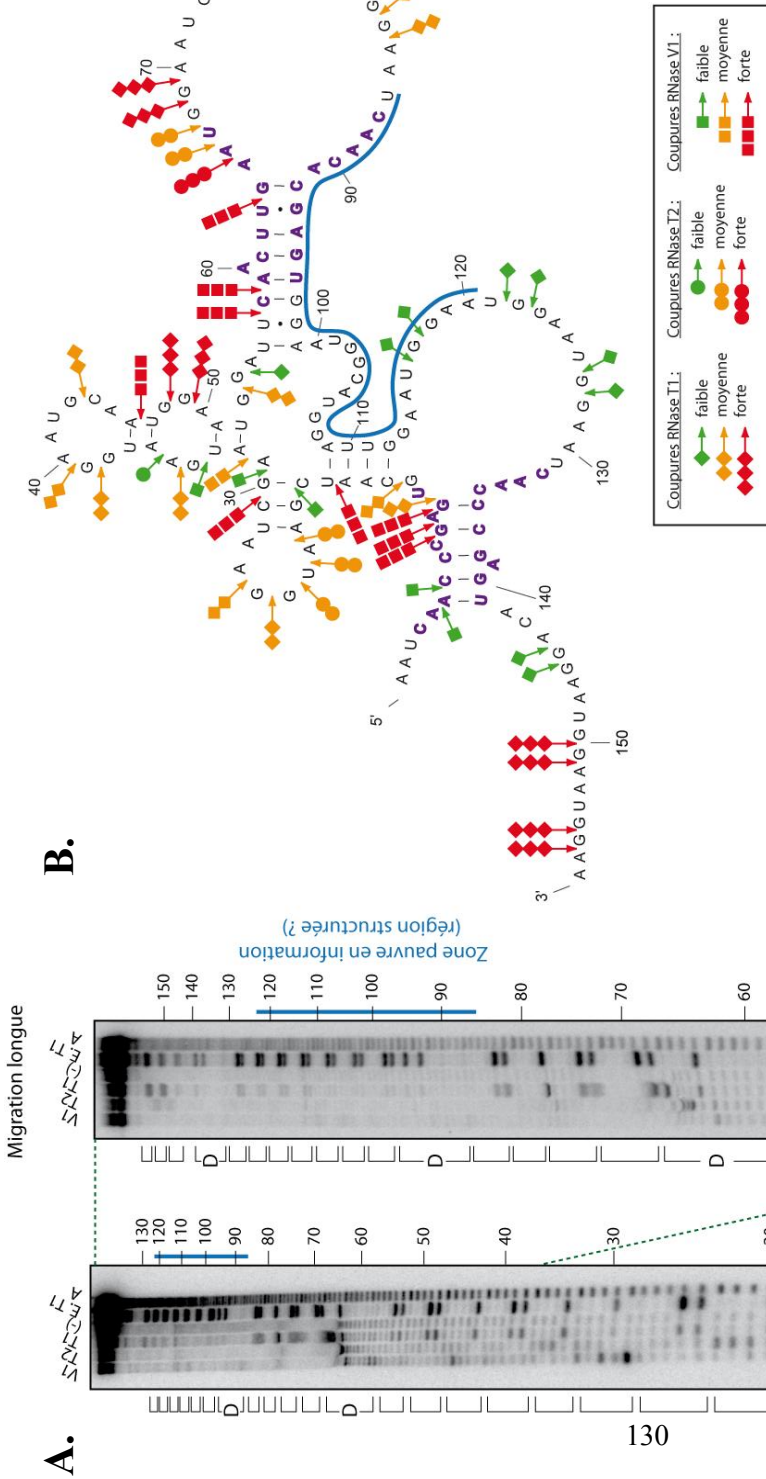


Figure II-2 : Etude de la structure secondaire adoptée par l'ARN satellite III pHur98 par emploi de sondes enzymatiques en solution. A. Après renaturation *in vitro*, l'ARN radiomarqué en 5' a été soumis à des hydrolyses ménagées par les RNases T1, T2 et V1, ou incubé sans enzymes (-) (voir matériel et méthodes). Les positions des coupures ont été identifiées grâce à une hydrolyse alcaline induisant des clivages à chaque position (A), et une hydrolyse par la RNase T1 sur un ARN dénaturé, permettant d'identifier les résidus G (E. T1). Les différents produits d'hydrolyse des ARN ont été fractionnés sur gels de polyacrylamide 10%. Les bandes d'ARN ont été révélées par autoradiographie. L'autoradiogramme de gauche correspond à un temps de migration de 2h et celui de droite de 4h. Les positions des nt dans la séquence sont indiquées à droite des autoradiogrammes. Les pentanucléotides et les décanucléotides (D) sont indiqués à gauche par des crochets. Un trait bleu à droite du gel indique un segment fortement protégé contre l'action des RNases. B. Modèle de structure secondaire possible de l'ARN pHur98 établi à partir des données d'hydrolyses enzymatiques et du logiciel Mfold. L'intensité et le type des coupures enzymatiques sont indiquées (voir encadré). Les décanucléotides sont indiqués en violet. La séquence correspondant à la zone du gel pauvre en clivages enzymatique est indiquée, ainsi que la position et l'intensité des coupures (cf encadré).

II.2. Recherche de partenaires protéiques des ARN sat III

La compréhension des mécanismes à l'origine de la formation des nSB et la caractérisation de l'impact de la formation de ces granules sur l'épissage alternatif nécessitent, entre autres, la caractérisation du contenu en protéines de ces granules et l'étude de l'interaction de ces protéines avec les ARN sat III. La purification des nSB suivie de la caractérisation de leur contenu aurait été la méthode la plus performante mais n'était pas applicable du fait des difficultés de la purification de ces granules. Nous avons donc utilisé deux approches *in vitro*, la première était basée sur la prédiction informatique de sites potentiels de fixation des protéines SR couplée à des tests par retard sur gel des capacités de fixation à l'ARN de protéines recombinantes. La seconde approche était basée sur la technique de chromatographie d'affinité MBP-MS2 déjà décrite dans la partie précédente.

II.2.1. Utilisation de programmes informatiques pour prédire la présence de sites de fixations des protéines SR aux ARN sat III

Au cours des dernières années de nombreux programmes informatiques permettant la prédiction de sites de fixation de protéines régulant l'épissage ont été développés. Parmi ces programmes, le logiciel ESEFinder (http://rulai.cshl.edu/cgi-bin/tools/ESE3/ese_finder.cgi) utilise des données expérimentales obtenues par l'approche SELEX pour certaines protéines SR, afin de prédire des sites d'interactions potentiels de ces protéines au sein de séquences ARN (Cartegni et al., 2003; Smith et al., 2006). L'utilisation de ce logiciel avait notamment permis à l'équipe de G. Biamonti de prédire une interaction directe entre la protéine SRSF1 (ASF/SF2) et un petit fragment de l'ARN pHur98, qui avait par la suite été confirmée expérimentalement (Chiodi et al., 2004). Ce logiciel permet aussi de prédire la présence de sites de fixation pour les protéines SR suivantes : SRSF1 (ASF/SF2), SRSF2 (SC35), SRSF5 (SRp40) et SRSF6 (SRp55). D'autres programmes tels que SpliceAid (<http://www.introni.it/splicing.html>) (Piva et al., 2009) ou RegRNA (<http://regrna2.mbc.nctu.edu.tw/statistics.html>) (Chang et al., 2013) utilisent des bases de données répertoriant des motifs ARN expérimentalement démontrés comme interagissant avec certains facteurs nucléaires impliqués dans des régulations d'épissage. Ces programmes couvrent un nombre beaucoup plus large de protéines que ESEFinder (71 et 25, respectivement). Nous avons utilisé ces 3 programmes afin de rechercher des sites potentiels de fixation de protéines SR et hnRNP au sein de la séquence de l'ARN pHur98. Les prédictions les plus fortes que nous avons obtenues concernaient les protéines SRSF1 (ASF/SF2), SRSF7 (9G8), SRSF5 (SRp40), SRSF10 (SRp38) et Tra2 β . Parmi ces protéines SRSF1 et SRSF7 sont déjà connues pour relocaliser au sein des granules de stress nucléaires en cas de choc thermique (Denegri et al., 2001). Une interaction directe entre SRSF1 et les ARN sat III avait d'ailleurs été démontrée alors qu'aucunes données concernant une interaction directe possible entre SRSF7 et les ARN sat III n'étaient disponibles (Chiodi et al., 2004). En ce qui concerne la protéine SRSF5, des données non publiées de l'équipe de G. Biamonti mentionnées dans l'article de Chiodi et al.,

publié en 2004, indiquaient sa localisation au sein des granules. Enfin, aucune donnée expérimentale ne faisait état d'une éventuelle localisation des protéines SRSF10 et Tra2 β dans les granules de stress nucléaires. Cependant SRSF10 est un partenaire protéique connu de la protéine hnRNP HAP qui est un constituant des granules.

Etant donné que nous disposons au laboratoire de protéines SRSF7 et SRSF5 recombinantes, produites en baculovirus par l'équipe de J. Stévenin, j'ai entrepris de tester si ces protéines pouvaient se fixer directement sur les ARN sat III. J'ai pour cela réalisé des expériences de retard sur gel en utilisant l'ARN pHur98 produit par transcription *in vitro* (figure II-3).

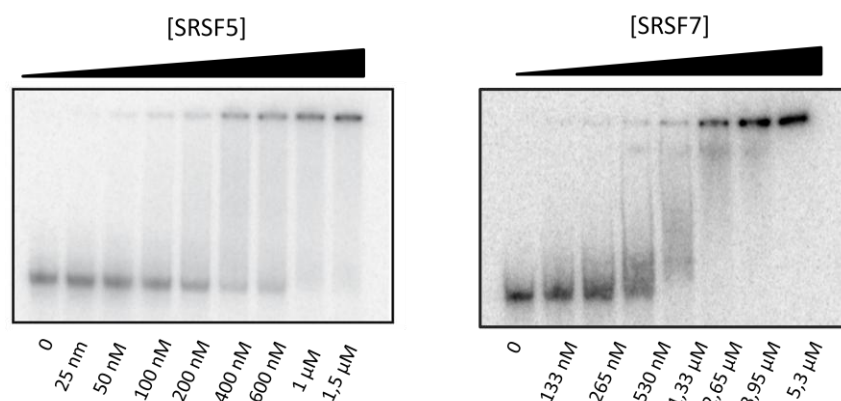


Figure II-3 : Expériences de retard sur gel réalisées avec l'ARN pHur98 et les protéines SRSF5 et SRSF7. L'ARN pHur98 non radiomarqué a été incubé, en présence d'une fraction d'ARN radiomarqué et d'ARNt en excès molaire 100x, avec des concentrations croissantes de protéine SRSF5 ou SRSF7 qui sont indiquées en dessous des pistes du gel.

Un retard de migration de l'ARN pHur98 radiomarqué apparaissait de façon concomitante à l'ajout de protéine SRSF5 ou SRSF7. Ce résultat montrait la capacité de ces deux protéines à se fixer sur l'ARN *in vitro*. La mise en évidence de ces interactions suggère que le recrutement des protéines SRSF5 et SRSF7 au sein des nSBs pourrait se faire *via* une interaction directe avec l'ARN sat III. Cela est en accord avec les observations antérieures de leur localisation dans les nSB. Pour SRSF10 et Tra2 β , aucune étude de localisation dans les nSB n'a été effectuée, et je n'ai pas moi même réalisé d'expérience de retard sur gel car nous ne disposons pas des clones permettant de les produire sous forme recombinante.

II.2.2. Purification et identification des partenaires d'un ARN sat III *in vitro*

Pour réaliser une étude plus exhaustive des protéines partenaires des ARN sat III j'ai utilisé la technique de la chromatographie d'affinité MBP-MS2 couplée à une analyse par spectrométrie de masse

des protéines isolées. Le nombre de protéines identifiées par cette approche étant directement proportionnel à la taille de l'ARN utilisé, j'ai produit un ARN sat III de petite taille afin de limiter le risque d'isoler des protéines associées de manière aspécifique. L'ARN utilisé (nommé ARN MiniHur) correspondait aux positions 97 à 141 de l'ARN pHur98 et comprenait 7 répétitions pentanucléotidiques et un décanucléotide, soit un total de 44 nt. La stratégie expérimentale utilisée pour produire cet ARN est illustrée dans la figure II-4. Brièvement, un oligonucléotide matrice (7741) contenant la séquence 97-141 encadrée de sites de restriction *Bam*HI (en 5') et *Kpn*I (en 3'), a été amplifié par PCR et le produit d'amplification obtenu a été cloné dans le plasmide pUC18::T7-MS2 (figure II-4). Le plasmide pUC18::T7-MiniHur-MS2 obtenu permettait la production, par l'ARN polymérase du phage T7, de l'ARN MiniHur fusionné à 3 étiquettes MS2 (MiniHur-MS2). Cet ARN a été fixé sur une résine d'amylose grâce à la protéine MBP-MS2, puis incubé dans un extrait nucléaire de cellules HeLa. L'ARN MS2 a été traité de la même façon et a été utilisé en tant que témoin négatif. Les protéines isolées lors de la purification des complexes ont ensuite été fractionnées sur gel SDS-PAGE et identifiées (découpage systématique) par spectrométrie de masse par l'équipe d'A. Van Doerssealer à Strasbourg (figure II-5). La protéine MBP-MS2 présente en grande quantité correspond à la bande sur le gel aux environs de 55kDa. Cette technique a permis d'identifier des peptides correspondant à 169 protéines, dont 113 n'étaient pas retrouvées dans la fraction contrôle correspondant à l'ARN MS2.

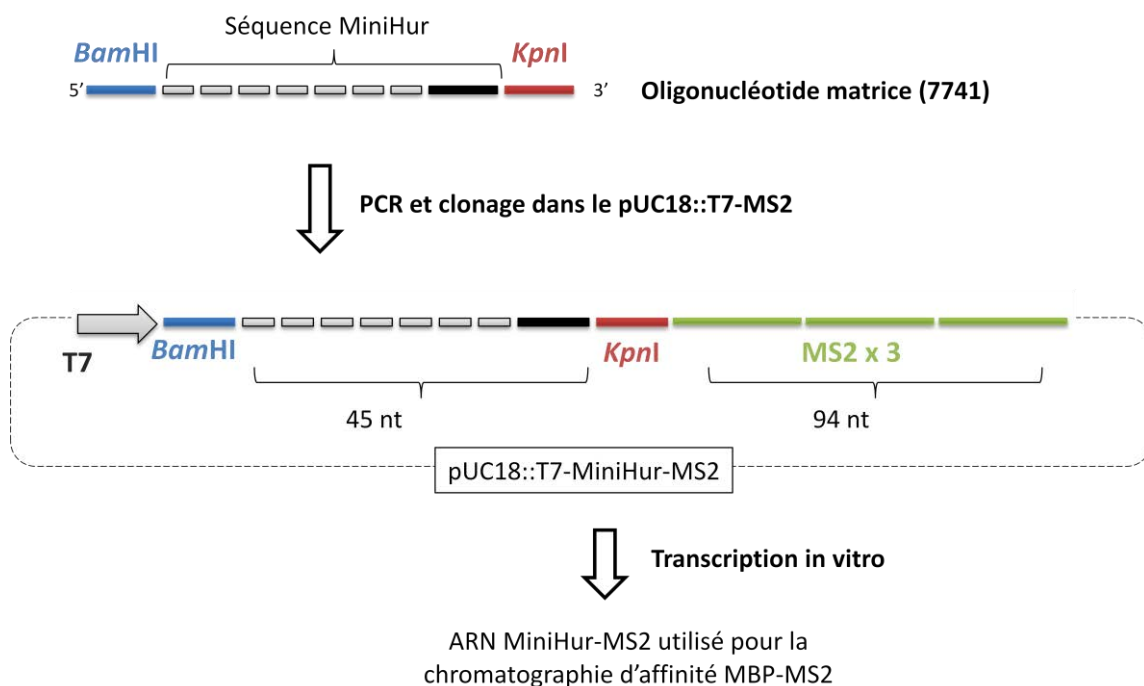


Figure II-4 : Stratégie de clonage de l'ARN MiniHur-MS2. La présence de séquences générant des sites de restriction *Bam*HI et *Kpn*I dans l'oligonucléotide matrice nous a permis de cloner le fragment PCR dans le plasmide pUC18::T7-MS2. La transcription *in vitro* de l'ARN MiniHur-MS2 peut être réalisée à partir du plasmide recombinant obtenu, après synthèse d'une matrice par PCR (voir Matériel et méthodes).

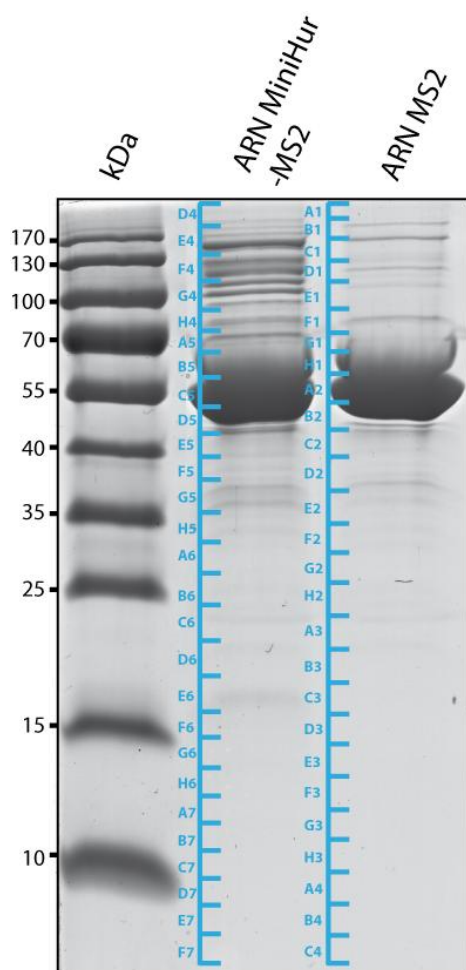


Figure II-5 : Séparation par électrophorèse SDS-PAGE des protéines associées aux ARN MS2 et MiniHur-MS2 issues de la chromatographie d'affinité MBP-MS2. Les fractions protéiques issues des chromatographies d'affinité ont été déposées fractionnées sur gel SDS-PAGE, en parallèle avec un marqueur de taille (piste 1 - les poids moléculaires sont indiqués à gauche du gel en kDa). Des bandes ont été découpées dans chaque piste tous les 5 mm (échelle bleue). Chaque bande a été soumise à une digestion trypsique afin de permettre la détection des peptides obtenus par spectrométrie de masse (collaboration avec l'équipe d'A. Van Doerssealer).

Parmi ces 169 protéines nous n'avons pris en considération que les protéines possédant un nombre de peptides dans la fraction MiniHur-MS2 supérieur au nombre de peptides détectés dans la fraction MS2, ce qui représente 149 protéines (19 protéines non prises en compte) (tableau II-1). L'approche de spectrométrie de masse utilisée ne fournissait pas des données quantitatives. En effet, le nombre de peptides identifiés par cette approche dépend d'un grand nombre de paramètres : la quantité de cette protéine présente dans le mélange mais aussi sa taille, le nombre de sites de clivage par la trypsine, les capacités d'ionisation des peptides libérés... Nous ne pouvons donc pas considérer que le nombre de peptides identifiés pour chacune des protéines soit corrélé directement à la quantité de ces protéines dans les complexes. Néanmoins, la détection d'un faible nombre de peptides pour une protéine donnée suggèrerait sa faible abondance dans les complexes purifiés. Sur les 149 protéines détectées dans l'essai pour lesquelles le nombre de peptides détectés est inférieur dans la fraction contrôle, seules 56 protéines avaient été détectées par identification d'au moins 3 peptides dans la fraction MiniHur-MS2. Ce sont donc ces protéines que nous avons prises en considération. Beaucoup d'entre elles (soit 20 protéines) avaient un lien avec l'épissage, selon la base de données Gene Ontology (<http://amigo.geneontology.org>). Parmi elles, on comptait beaucoup de protéines hnRNP et d'ARN hélicases, ainsi que d'autres facteurs de régulation de l'épissage (tableau II-1).

Nom de la protéine	PM (kDa)	Symbole	peptides détectés		% de couverture	
			MS2	MiniHur	MS2	MiniHur
116 kDa U5 snRNP component	109,4	EFTUD2		2		2,88
14-3-3 protein epsilon	29,2	YWHAE	1	3	3,92	15,7
40S ribosomal protein S13	17,2	RPS13		1		7,95
Actin, cytoplasmic 1	41,7	ACTB		5 (4)		20
Alpha-actinin-4	104,8	ACTN4	8	18	11,2	25,7
ATP-dependent RNA helicase A	140,9	DHX9	32	41	32,4	42,4
ATP-dependent RNA helicase DDX50	82,5	DDX50		4 (2)		8,14
Cellular nucleic acid-binding protein	19,4	CNBP		5		33,9
Cold-inducible RNA-binding protein	18,6	CIRBP		3		26,7
DNA topoisomerase 1	90,7	TOP1		5		6,8
DNA topoisomerase 2-alpha	174,4	TOP2A		8		7,38
DNA-binding protein A	40,1	CSDA		3 (2)		12,6
DNA-directed RNA polymerase I subunit RPA1	194,8	POLR1A		10		10,3
Double-stranded RNA-specific adenosine deaminase	136,0	ADAR		4		4,49
Double-stranded RNA-specific editase 1	80,7	ADARB1		3		7,83
Drebrin	71,4	DBN1	1	4	1,85	10
FACT complex subunit SSRP1	81,1	SSRP1		3		5,22
Glycogen phosphorylase, brain form	96,7	PYGB		3 (2)		5,34
Glycogen phosphorylase, liver form	97,1	PYGL	1	5 (2)	1,65	8,26
Heterogeneous nuclear ribonucleoprotein A1	38,8	HNRNPA1	4	6	15,9	20,4
Heterogeneous nuclear ribonucleoprotein A3	39,6	HNRNPA3	3	4	11,1	16,9
Heterogeneous nuclear ribonucleoprotein H	49,2	HNRNPH1		5 (2)		17,4
Heterogeneous nuclear ribonucleoprotein H2	49,2	HNRNPH2		3 (2)		9,8
Heterogeneous nuclear ribonucleoprotein M	77,5	HNRNPM		5		11,2
Heterogeneous nuclear ribonucleoprotein Q	69,6	SYNCRIP	4 (1)	20 (2)	9,31	32,1
Heterogeneous nuclear ribonucleoprotein R	70,9	HNRNPR	3 (1)	14 (2)	6,64	29,1
Heterogeneous nuclear ribonucleoprotein U	90,6	HNRNPU	5	11	10,7	20
Heterogeneous nuclear ribonucleoprotein U-like 2	85,1	HNRNPUL2		3		4,28
Heterogeneous nuclear ribonucleoprotein A2/B1	37,4	HNRNPA2B1	1	3	4,53	11,6
Insulin-like growth factor 2 mRNA-BP 1	63,4	IGF2BP1		5 (1)		9,88
Interferon-induced, ds RNA-activated kinase	62,1	EIF2AK2	5	11	12,5	27,6
Interleukin enhancer-binding factor 2	43,0	ILF2	6	13	22,3	60,5
Interleukin enhancer-binding factor 3	95,3	ILF3	8	21 (4)	12,4	27,4
Myb-binding protein 1A	148,8	MYBBP1A	2	17	1,96	16,4

Nom de la protéine	PM (kDa)	Symbole	peptides détectés		% de couverture	
			MS2	MiniHur	MS2	MiniHur
Nuclear cap-binding protein subunit 1	91,8	NCBP1	2	5	3,29	7,72
Nuclease-sensitive element-binding protein 1	35,9	YBX1		5 (2)		26,2
Nucleolar RNA helicase 2	87,3	DDX21	2	27 (2)	2,81	40,6
Nucleolin	76,6	NCL	4	15	7,32	23,4
Nucleophosmin	32,6	NPM1	1	3	7,14	16,7
Probable ATP-dependent RNA helicase DDX5	69,1	DDX5	5	8 (1)	10,3	15,5
Probable ATP-dependent RNA helicase DHX36	114,8	DHX36	4	19	5,06	26,4
Probable ATP-dependent RNA helicase YTHDC2	160,2	YTHDC2		11		12,1
Pumilio domain-containing protein C14orf21	69,4	C14orf21		6		16,4
Putative ribosomal RNA methyltransferase NOP2	89,3	NOP2		4		6,53
Ras GTPase-activating-like protein IQGAP1	189,2	IQGAP1	2	4	1,57	4,1
Ribosomal L1 domain-containing protein 1	55,0	RSL1D1	1	8	2,65	23,5
RNA-binding protein FUS	53,4	FUS	4	5	9,89	15,8
S1 RNA-binding domain-containing protein 1	111,8	SRBD1	1	9	1,01	10,9
Serine/arginine-rich splicing factor 1	27,7	SRSF1		6		28,2
Spermatid perinuclear RNA-binding protein	73,6	STRBP		7 (4)		9,82
Superkiller viralicidic activity 2-like 2	117,8	SKIV2L2		4		7,2
Transcriptional activator protein Pur-beta	33,2	PURB		4		20,8
Tropomyosin alpha-3 chain	32,8	TPM3	2	3	7,39	12
Zinc finger RNA-binding protein	117,0	ZFR		6		9,4

Tableau II-1 : Protéines les plus représentées dans les fractions d'éluion issues de la chromatographie MBP-MS2. Ces expériences ont été effectuées avec les ARN MS2 (témoin) ou MiniHur-MS2. Le nombre de peptides retrouvés dans les deux fractions pour chaque protéine est indiqué, et le chiffre entre parenthèses indique le nombre de peptides détectés qui ne sont pas strictement spécifiques d'une protéine donnée et qui pourraient être communs avec d'autres protéines détectées. Les protéines prises en considération sont celles pour lesquelles plus de 3 peptides sont retrouvés dans la fraction MiniHur-MS2. Le pourcentage de couverture de la séquence de la protéine par les peptides détectés dans les deux fractions est également indiqué pour chaque protéine. Les 20 protéines impliquées dans l'épissage (selon la base de données Gene Ontology) sont indiquées en gras.

A noter qu'en accord avec le recrutement de la protéine SRSF1 dans les nSB après un choc thermique et son interaction directe avec les ARN sat III *in vitro*, via son domaine RRM2 (Chiodi et al., 2004; Denegri et al., 2001; Metz et al., 2004), SRSF1 était détectée par spectrométrie de masse en association avec l'ARN MiniHur-MS2 (6 peptides) mais pas avec l'ARN MS2. Ce qui validait la méthodologie employée, d'autant que SRSF1 (ASF/SF2) fait habituellement partie, avec les autres protéines SR, des protéines difficiles à détecter en spectrométrie de masse du fait d'un faible nombre de sites d'hydrolyse par la trypsine et d'une mauvaise ionisation des peptides.

Nous avons pu rassembler plusieurs anticorps spécifiques dirigés contre certaines des protéines considérées dans le tableau II-1 (principalement des hnRNP) ce qui nous a permis de vérifier leur

présence par western blot dans les complexes purifiés. Pour chaque protéine un test a été réalisé sur les complexes purifiés avec l'ARN MiniHur-MS2 et sur ceux purifiés dans l'expérience contrôle réalisée avec l'ARN MS2.

II.3. Confirmation par western blot de la présence des protéines identifiées dans les fractions de purification

Etant donné la masse hétérogène de données fournies par l'approche de spectrométrie de masse que nous avons utilisée, il était nécessaire de réaliser des contrôles par l'approche western blot. En plus, une fois la présence d'une protéine confirmée dans les complexes, cette approche peut permettre de tester rapidement et quantitativement l'effet des mutations dans l'ARN sur la fixation de cette protéine.

Nous avons testé un certain nombre d'anticorps commerciaux dirigés contre des protéines hnRNP ou des protéines SR que nous avons détectées par spectrométrie de masse (hnRNP U, H, R, Q et SRSF1), ou que l'on savait se localiser totalement ou partiellement dans les nSB formés lors d'un choc thermique (Sam68, hnRNP HAP, SRSF7). En plus, nous avons introduit dans les essais SRSF2 (SC35) en tant que contrôle négatif car on sait qu'elle n'est pas localisée dans les nSB après un choc thermique (Denegri et al., 2001).

Comme le montre la figure II-6, les expériences de western blot confirmaient bien la présence dans les complexes formés avec l'ARN MiniHur-MS2 des protéines hnRNP U, H, R et Q qui avaient été identifiées avec un nombre de peptide supérieur à 3 par spectrométrie de masse (tableau II-1). Ces associations étaient bien spécifiques à la partie MiniHur de l'ARN hybride puisqu'elles n'étaient pas ou beaucoup moins détectées dans les complexes formés avec l'ARN contrôle. Comme attendu, il en était de même pour la protéine SRSF1. Comme attendu également, la protéine SRSF3 (SRp20) pour laquelle un nombre faible de peptide (2 peptides) avait été détecté dans les complexes formés avec l'ARN MiniHur, n'était pas détectée dans les complexes par Western blot (figure II-6).

La très faible détection par western blot de SRSF2 dans les complexes formés avec l'ARN MiniHur (seulement après une forte exposition) confirmait bien que les ARN sat III ne doivent pas avoir une forte affinité pour cette protéine contrairement à SRSF1. Elle était vraisemblablement présente dans les complexes à faible taux du fait de son interaction possible avec des protéines liées à l'ARN. Les données que nous avons obtenues pour SRSF7, que nous avons montrée avoir une affinité moyenne pour l'ARN pHur98 *in vitro*, ne sont pas claires du fait de la mauvaise qualité de l'anticorps. Nous observons néanmoins la présence à un faible taux de cette protéine dans les complexes formés avec l'ARN MiniHur-MS2 mais pas dans ceux formés avec l'ARN contrôle. Cette protéine est connue pour être partiellement localisée dans les nSB après un choc thermique. Cette localisation pourrait donc s'expliquer par son affinité pour les ARN sat III et sans doute du fait des capacités d'interaction avec

d'autres protéines des nSB. L'affinité de SRSF7 pour les ARN sat III ne semble pas très forte en comparaison de celle de SRSF1. Néanmoins, nous ne pouvons pas exclure que des modifications post-traductionnelles portées par SRSF7 en cas de choc thermique augmente son affinité pour les ARN sat III. Rappelons que l'extrait nucléaire employé pour ces expériences n'avait pas été soumis à un choc thermique.

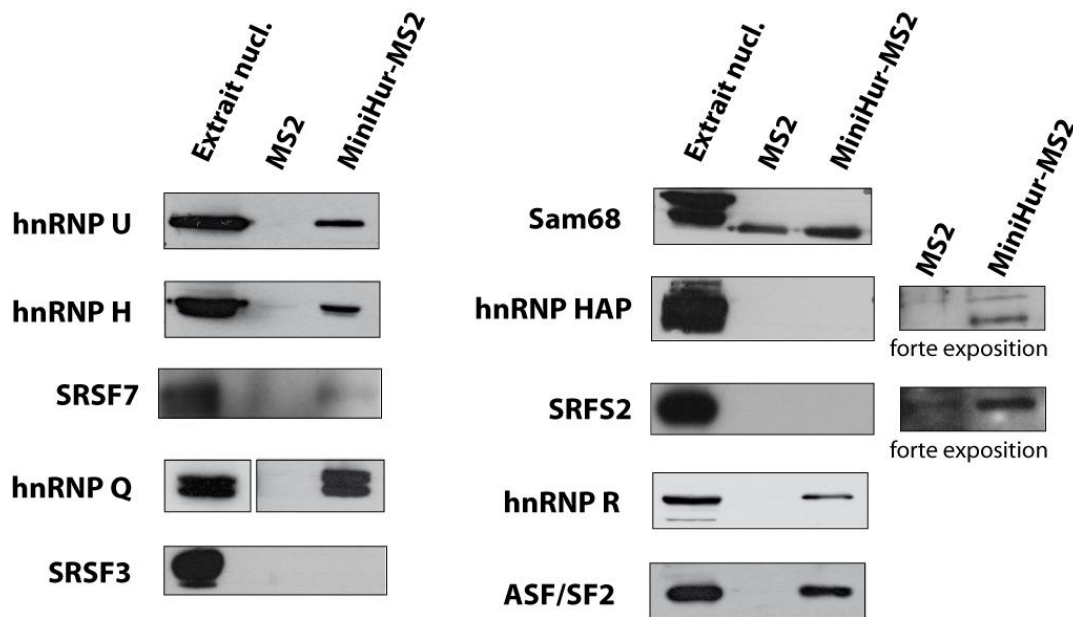


Figure II-6 : Expériences de western blot réalisées sur les complexes formés avec les ARN purifiés par chromatographie d'affinité MBP-MS2. Les protéines contenues dans les complexes issus des chromatographies d'affinité ont été éluées puis déposées sur gel SDS-PAGE, transférées sur membrane et la présence de différentes protéines a été testée par les anticorps indiqués.

La protéine hnRNP HAP est connue pour être présente en grande quantité dans les nSB formés en cas de choc thermique. Alors que cette protéine est présente en quantité importante dans les extraits nucléaires de cellules HeLa que nous avons utilisés, seule une très forte exposition permet d'en détecter des traces dans les complexes formés avec l'ARN MiniHur-MS2. Au moins 2 formes sont visibles traduisant sans doute la présence de modifications post-traductionnelles. Une seule de ces deux formes est renforcée dans l'essai par rapport au contrôle. Mais clairement, en accord avec les données de spectrométrie de masse, la quantité de cette protéine dans les complexes formés avec l'ARN MiniHur est très faible. Il est possible que hnRNP HAP soit présente dans les nSB du fait de son interaction avec des protéines partenaires des ARN sat III, comme mentionné dans l'introduction de cette partie, plutôt qu'avec les ARN sat III directement, ce qui pourrait expliquer sa présence limitée dans les complexes purifiés avec l'ARN MiniHur-MS2.

Enfin Sam 68 que l'on sait aussi être localisée dans les nSB, semble avoir une affinité assez générale pour les ARN ou des capacités d'interaction avec des protéines présentes dans les complexes

ayant une affinité générale pour les ARN. En effet, l'une de ses 2 formes détectées avec l'anticorps que nous avons utilisé est détectée à la fois dans les complexes MS2 et le complexe MiniHur-MS2.

En conclusion, l'ensemble de nos données suggérait qu'en plus de SRSF1 plusieurs protéines hnRNP avaient de fortes capacités d'association avec les ARN sat III : hnRNP U, H, R et Q. En particulier, pour Q, R et U, le nombre de peptides détectés par spectrométrie de masse (20, 14 et 11 respectivement) était très élevé en comparaison de l'expérience contrôle.

Ayant mis en évidence des capacités d'association des protéines hnRNP U, H, Q avec l'ARN MiniHur *in vitro*, il était important d'essayer d'avoir une confirmation de leur présence *in cellulo* dans les nSB formés après choc thermique, ceci par des approches d'imagerie cellulaire. Pour cela, il nous fallait mettre au point des conditions satisfaisantes de formation des nSB après choc thermique dans des cellules HeLa, dont nous avons utilisé des extraits dans les expériences *in vitro*.

II.4. Etude de la localisation des protéines candidates dans des cellules HeLa soumises à un choc thermique

II.4.1. Mise au point des conditions de choc thermique

Nous avons utilisé des cellules HeLa (S3), qui, avec des lignées de fibroblastes, sont les plus fréquemment utilisées dans la littérature comme modèle pour la formation des granules de stress nucléaires. On savait d'après la littérature que le choc thermique à 42°C est le stress le plus adapté à la formation des granules, puisqu'il permet d'obtenir dans les cellules HeLa les granules de la plus grande taille, et ceci dans une grande proportion de cellules et avec une grande quantité d'ARN sat III produits (Valgardsdottir et al., 2007b). Dans un premier temps, nous avons essayé de reproduire ces résultats avec notre lignée cellulaire. Pour cela, j'ai évalué le taux induction de l'ARN sat III et de l'ARNm de la protéine hsp70 par RT-qPCR. La normalisation des expériences de RT-qPCR a été faite grâce à l'amplification de l'ARNm de la protéine ribosomique rplp0, un gène de référence stable (Fu et al., 2010). Les ARNm hsp70 et rplp0 ont été rétrotranscrits à l'aide d'un oligonucléotide constitué exclusivement de désoxythymidine (oligonucléotide dT16) permettant la rétrotranscription des ARN polyadénylés, principalement les ARNm. Les oligonucléotides que nous avons utilisés pour l'amplification par qPCR de l'ADNc hsp70 sont les mêmes que ceux utilisés par l'équipe de G. Biamonti (7828 et 7829 dans le matériel et méthodes) (Valgardsdottir et al., 2007b). En théorie, l'oligonucléotide dT16 devrait permettre la détection des ARN sat III par RT-qPCR puisque ces transcrits sont polyadénylés, mais l'évaluation du taux de ces ARN par RT-qPCR nécessite l'emploi d'un protocole plus spécifique, du fait que les ARN sat III correspondent à une population hétérogène d'ARN et que la quantification par RT-qPCR nécessite l'obtention d'un produit final unique. Nous avons donc utilisé la stratégie de détection d'ARN sat III par RT-qPCR décrite par l'équipe de G. Biamonti (Valgardsdottir et al., 2005b). Il s'agit d'une "hemi-nested"

RT-qPCR, utilisant pour la rétrotranscription un oligonucléotide (7796, aussi appelé RSM13 dans la publication) s'hybridant en 3' sur une séquence unique du clone pHur98 d'ARN sat III (figure II-7). Cet oligonucléotide possède en 5' une séquence dérivée du phage M13, ne s'hybridant pas avec l'ARN sat III ciblé, mais permettant d'ajouter cette séquence au produit de rétrotranscription. Cela permet d'effectuer à l'étape suivante une PCR ou qPCR de manière plus ciblée, en utilisant un oligonucléotide correspondant à la séquence M13 (8539) conjointement avec un autre oligonucléotide (7795, aussi appelé Hu98R dans la publication) s'hybridant spécifiquement en 5' du produit de rétrotranscription obtenu (figure II-7).

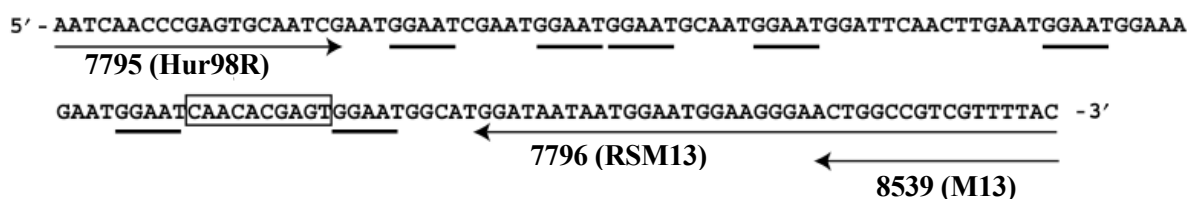


Figure II-7 : Stratégie de détection des ARN sat III par RT-qPCR (tiré de (Valgardsdottir et al., 2007b)). La première étape est une rétrotranscription effectuée à l'aide de l'oligonucléotide 7796 (ou RSM13), dont seule la moitié 3' s'hybride sur sa cible, et qui permet grâce à la partie restante d'ajouter en 3' du produit ADNc obtenu la séquence M13. Cette séquence va aider à l'amplification et la quantification plus spécifique du produit par qPCR, grâce au couple d'oligonucléotides 8539 (ou M13) et 7795 (ou Hur98R). Dans la séquence ARN sat III cible, les répétitions pentanucléotidiques identiques au consensus des ARN sat III sont soulignées, et l'unique décanucléotide présent est encadré.

Sur la base des données de la littérature, nous avons choisi de comparer les taux d'ARNm hsp70 et d'ARN sat III dans des cellules HeLa ayant subi une incubation soit d'1 h soit de 2 h à 42°C, suivie de différentes périodes de régénération à 37°C (1 h à 4 h) (Rizzi et al., 2004; Valgardsdottir et al., 2007b). Par ailleurs j'ai testé plusieurs méthodologies d'incubation des cellules à 42°C. Les deux premières solutions envisagées ont été l'incubation soit en atmosphère contrôlée à 42°C en étuve à CO₂, soit l'incubation dans un four à hybridation en atmosphère non contrôlée. Dans les deux cas, les résultats obtenus ne se sont pas avérés reproductibles, et les chocs thermiques ainsi réalisés n'avaient souvent peu ou pas d'effet sur l'expression des gènes étudiés, probablement du fait que la transition de température (de l'air au milieu de culture) était trop lente, et la déperdition en chaleur lors de l'ouverture de la porte de l'étuve trop élevée. C'est pourquoi nous avons opté pour une autre solution, proposée par A. Metz qui a utilisé cette technique dans le laboratoire de C. Vourc'h, à savoir une incubation au bain-marie à 42°C, après avoir enveloppé les boîtes de culture de cellules dans du parafilm. Ceci permet un transfert de chaleur de liquide à liquide bien plus rapide. Les chocs thermiques ont été beaucoup plus efficaces avec cette technique. Afin de nous assurer de l'absence d'effet de l'immersion des boîtes dans un bain marie par rapport à la culture en étuve, nous avons comparé les taux d'ARN sat III et d'ARNm hsp70 de cellules immergées dans un bain marie à 37°C par rapport à ceux de cellules en culture en étuve à 37°C. Par ailleurs, la mesure du pH après incubation nous a permis d'établir qu'aucun changement de pH du

milieu n'apparaissait que ce soit après 1 h ou 2 h d'immersion au bain marie à 37°C ou à 42°C. Nous avons donc retenu le protocole d'immersion des boîtes dans un bain marie pour la suite des travaux.

L'analyse par RT-qPCR du niveau d'expression des ARN hsp70 et des ARN sat III a permis de déterminer les temps de choc thermique et de régénération conduisant à une induction optimale de l'expression des ARN sat III (figure II-8). On note que, de façon systématique, deux heures d'immersion au bain marie à 42°C résultent en une induction de l'ARNm hsp70 et de l'ARN sat III plus forte qu'avec seule heure d'immersion suivie des mêmes temps de régénération (données non présentées).

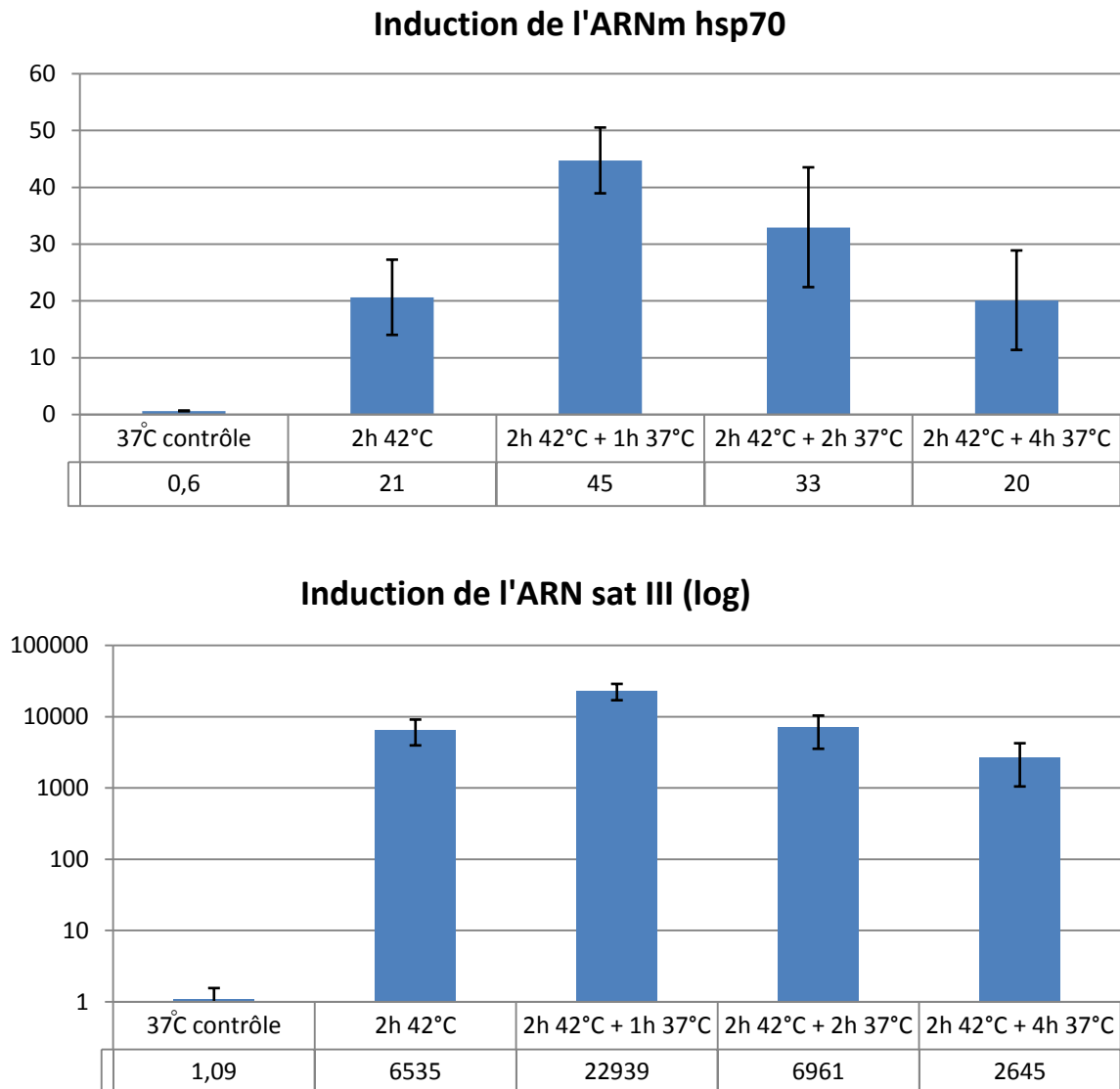


Figure II-8 : Détermination par RT-qPCR du taux d'induction de l'ARNm de hsp70 et de l'ARN sat III après un choc thermique de 2 h à 42°C suivi de différents temps de régénération à 37°C. Les chocs thermiques ont été effectués au bain marie, et les temps de régénération sont indiqués à côté du temps de choc thermique (+0 h, +1 h, +2 h ou +4 h de régénération à 37°C). Les taux d'induction calculés sont indiqués en dessous des graphiques. Le contrôle à 37°C a été obtenu par immersion des cellules pendant 2 h dans un bain marie à 37°C. Les valeurs indiquées correspondent au taux d'induction de la quantité d'ARN sat III après le traitement par rapport à l'expérience sans traitement (culture à 37°C), qui a été fixé à 1. Ces données sont issues des moyennes calculées à partir des résultats de 3 expériences distinctes, et l'erreur type a été calculée.

Par ailleurs 1 h de régénération à 37°C après la fin du choc thermique conduit à une induction maximale des ARN sat III et de l'ARNm hsp70. Dans ces conditions les taux d'induction calculés étaient respectivement de plus de 20 000 fois pour les ARN sat III et de 45 fois pour les ARNm hsp70 ce qui est en accord avec les données de la littérature pour ce type d'expériences (Valgardsdottir et al., 2007b).

II.5.2. Test de la présence des protéines purifiées avec l'ARN sat III *in vitro* dans les granules de stress nucléaires

Ayant déterminé les conditions expérimentales optimales pour la formation des nSB dans nos cellules, j'ai alors entrepris d'analyser par immunofluorescence la localisation, suite à un choc thermique, des protéines identifiées lors des expériences de chromatographie d'affinité *in vitro*. La relocalisation de ces protéines au sein des nSB était vérifiée par co-localisation des signaux obtenus pour ces protéines avec ceux des ARN sat III détectés par RNA-FISH (Fluorescence *In Situ* Hybridization) ou ceux des protéines des nSB SRSF1 ou hnRNP HAP, détectées par immunofluorescence.

II.5.2.a. Mise au point de la détection des granules de stress nucléaires par RNA-FISH

La technique de RNA-FISH consiste à localiser par microscopie de fluorescence un ARN dans les cellules, grâce à l'emploi de sondes oligonucléotides modifiées pouvant être détectées par fluorescence. On fait pénétrer ces sondes dans les cellules après perméabilisation. Elles s'hybrident spécifiquement à l'ARN cible dont la position est révélée par le fluorophore. Cette technique, tout comme l'immunofluorescence des protéines, nécessite l'emploi d'un agent fixateur avant perméabilisation, qui stoppe le métabolisme des cellules et les fige en l'état, afin de faciliter leur observation par imagerie. Nous avons utilisé du formaldéhyde ou du paraformaldéhyde qui permettent la fixation covalente de façon massive et non spécifique des protéines (par "crosslink") dans les cellules. Les oligonucléotides sondes utilisés étaient biotinylés. Les résidus biotinylés pouvant être reconnus par l'avidine ou la streptavidine, l'incorporation dans les cellules perméabilisées et fixées d'avidine D couplée au fluorophore FITC (Fluorescéine IsoThioCyanate) ou de streptavidine couplée au fluorophore Alexa488 (LifeTechnologies) permet de détecter l'ARN cible au sein des cellules par détection de la fluorophore (Jolly et al., 1997a; Metz et al., 2004; Rizzi et al., 2004). Néanmoins, comme le fluorophore FITC est peu stable, dans ce cas une étape supplémentaire d'incubation des cellules avec des anticorps anti-avidine D biotinylés est nécessaire pour amplifier le signal fluorescent. La détection de la sonde en interaction avec le fluorophore Alexa488 peut être réalisée en une seule étape. Cette technique s'est révélée être la plus efficace et la moins contraignante et c'est celle que nous avons retenue après les premiers tests.

En ce qui concerne la nature des sondes oligonucléotides employées, sur la base des données de la littérature, nous avons dans un premier temps utilisé des sondes synthétiques d'une vingtaine de nucléotides porteuses d'un nucléotide biotinylé à l'extrémité 5'. Ce type de sondes avait été utilisé par l'équipe de G. Biamonti (Valgardsdottir et al., 2005b). Ces sondes se sont avérées néanmoins peu efficaces dans nos conditions pour la détection des ARN sat III, vraisemblablement du fait de leur courte taille peu adaptée pour l'approche RNA-FISH, car la spécificité d'hybridation *in cellulo* était faible. Nous avons par la suite utilisé le même type de sonde que l'équipe de C. Vourc'h, qui correspond à la séquence pleine longueur du clone pHur98 (158 nt) mentionné précédemment (cf figure II-1) (Jolly et al., 1997a, 1997b). Cette sonde a été biotinylée par l'emploi de la technique dite du "random priming". Cette méthode est basée sur la réplication de l'ADN matrice par le fragment de Klenow de l'ARN polymérase d'*E. coli* qui permet l'incorporation de nucléotides biotinylés lors d'une extension d'oligonucléotides amorces de séquences aléatoires. Ces sondes ADN renfermant de la biotine ont été obtenues à partir du plasmide recombinant pUC18::T7-pHur98, dans lequel a été inséré l'ADN pHur98. Un témoin négatif a été créé par "random priming" en utilisant comme matrice pour la réplication le plasmide pUC18::T7 vide.

Les sondes obtenues ont été utilisées dans des expériences d'imagerie cellulaire (figure II-9). En parallèle, j'ai réalisé des expériences de RT-qPCR en vue d'évaluer le taux d'expression des ARN sat III dans les cellules. La combinaison de ces deux approches nous a permis de vérifier la présence de *foci* contenant les ARN sat III dans les cellules HeLa ayant subi un choc thermique de 2 h suivi de 1 h de régénération à 37°C.

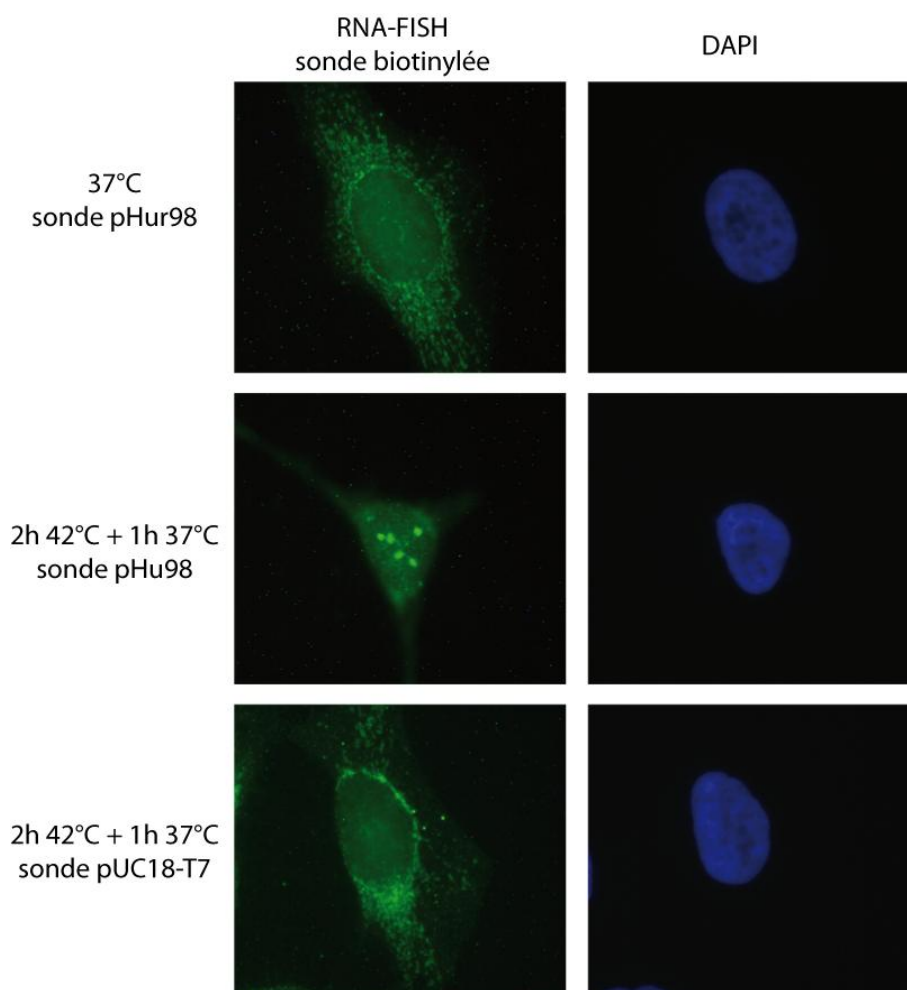


Figure II-9 : Visualisation par RNA-FISH des granules de stress nucléaires dans des cellules HeLa soumises à un choc thermique. Ces expériences ont été réalisées avec des sondes biotylées obtenues par "random priming" à partir des plasmides pUC18::T7-pHur98 et pUC18::T7 (témoin négatif). Les noyaux des cellules ont été observés par coloration au DAPI. Les cellules ont été cultivées à 37°C ou bien soumises à un choc thermique pendant 2 h à 42°C, puis régénérées une heure à 37°C avant d'être fixées à l'aide de formaldéhyde.

Cependant, nous constatons la présence d'un fort bruit de fond à la fois dans le noyau et dans le cytoplasme des cellules. Nous avons alors tenté, sans succès, de remédier à ce problème en employant des conditions plus stringentes lors des étapes d'hybridation et de lavages des sondes, essentiellement en augmentant la concentration en formamide jusqu'à 65%, la température d'hybridation des sondes jusqu'à 55°C et la température des solutions de lavage jusqu'à 45°C. Les granules observés dans toutes les conditions testées étaient souvent de petite taille et difficilement visibles, et le bruit de fond restait important. Nous notons également un manque de reproductibilité : pour une même condition donnée les granules d'ARN sat III pouvaient être clairement visibles ou peu voire pas observables après le choc thermique, alors que les expériences de qRT-PCR réalisées sur des échantillons de cellules ayant servi à faire l'expérience de RNA-FISH montraient une forte induction de l'expression des ARN sat III.

Nous avons supposé que les problèmes de spécificité observés étaient probablement dus au fait que la sonde pHur98 employée était produite par "random priming" à partir du plasmide pUC18::T7-pHur98 et qu'en réalité une multitude de sondes étaient produites à partir de ce plasmide, certaines d'entre elles correspondant à des régions du plasmide n'ayant rien à voir avec les ARN sat III. Ces sondes « parasites », bien que supposées ne rien cibler au sein des cellules HeLa, pouvaient augmenter le bruit de fond. Afin d'essayer de résoudre ce problème, j'ai alors tenté d'utiliser un oligonucléotide (n°967) spécifique de la séquence du promoteur T7, qui s'hybride à proximité de la séquence pHur98 sur le plasmide, à la place de l'oligonucléotide dégénéré utilisé dans les expériences précédentes, mais cela n'a pas réduit le bruit de fond ni amélioré la détection des ARN sat III.

Ne disposant pas de clones exprimant des ARN sat III de plus grande taille, nous avons alors décidé d'utiliser des sondes oligonucléotides de type LNA ("Locked Nucleic Acids", ou "acides nucléiques verrouillés"). En temps normal, même au sein d'une double hélice, les résidus riboses des nucléotides passent alternativement d'une conformation S (2'-endo) à une conformation N (3'-endo), ce qui déstabilise l'interaction entre les deux brins. La technologie LNA consiste à ajouter un pont méthylène (-CH₂-O-) entre le groupement OH porté par le carbone 2' du ribose et son carbone 4'. Cela permet de figer le ribose dans la conformation la plus favorable pour la formation de liaisons Watson-Crick, ce qui stabilise fortement les interactions établies entre les oligonucléotides LNA et leurs cibles (la température de fusion d'un duplex d'acides nucléiques (T_m) augmente de 2°C à 8°C par l'ajout d'un seul nucléotide LNA) et augmente fortement la spécificité de cette interaction. Nous avons utilisé une sonde LNA complémentaire d'une séquence correspondant à une succession de 4 répétitions "GGAAU" canoniques des ARN sat III (20 nt), le T_m de l'interaction étant de 84°C. Cet enchaînement de 4 répétitions est présent, au moins une fois, dans chaque clone référencé d'ARN sat III de taille suffisante (supérieure à 100 nt), ce qui représente 17 clones sur les 20 connus (Valgardsdottir et al., 2005b). Les sondes LNA utilisées étaient biotinylées en 5' et en 3' afin de permettre leur détection grâce à la streptavidine couplée au fluorophore Alexa 488. La révélation des ARN sat III grâce aux sondes LNA s'est avérée extrêmement efficace après un choc thermique réalisé dans les mêmes conditions que précédemment (figure II-10).

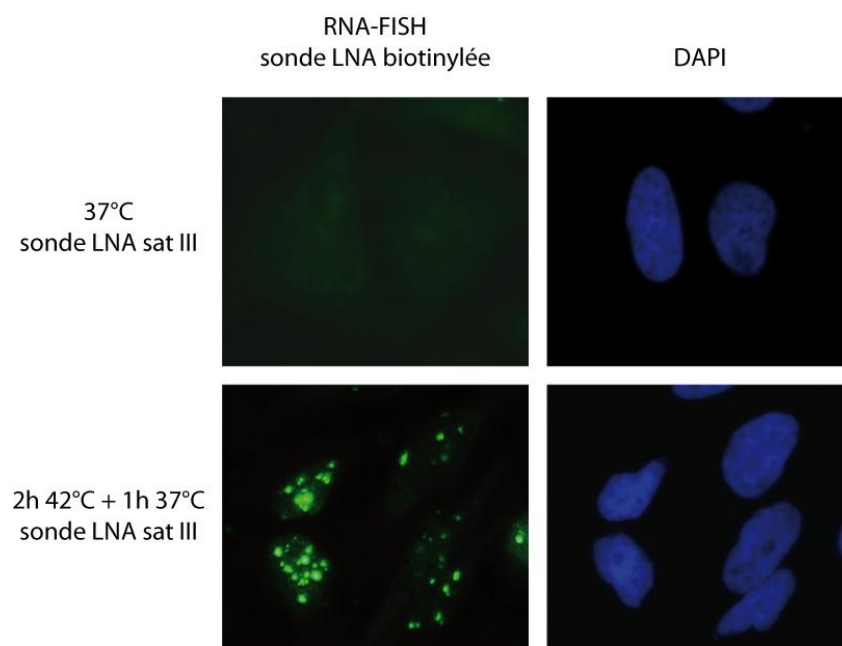


Figure II-10 : Amélioration de la visualisation des granules de stress nucléaires après un choc thermique par utilisation de sondes LNA pour les expériences de RNA-FISH. Ces expériences ont été réalisées dans les mêmes conditions que celles présentées dans la figure II-9. Les cellules ont été incubées à 37°C ou bien soumises à un choc thermique de 2 h à 42°C, puis régénérées une heure à 37°C. Les ARN sat III ont été révélés avec une sonde LNA synthétique (Exiqon) et les noyaux révélés par du DAPI.

Les granules de stress observés sont de grande taille, fortement fluorescents, et plus nombreux qu'avec la plupart des protocoles décrits dans la littérature, où généralement 2 granules principaux sont observés, parfois accompagné d'un ou plusieurs autres granules de stress minoritaires dans le cas de cellules transformées. Aucun signal n'était par contre détecté dans les cellules non exposées à la chaleur, et très peu de bruit de fond était observé. De plus, les données étaient très reproductibles. Avec l'emploi de ces sondes nous disposons donc de conditions optimales d'observation des granules permettant d'évaluer la co-localisation possible des protéines partenaires potentiels des ARN sat III identifiées dans les expériences *in vitro* décrites ci-dessus.

II.5.2.b. Test de la présence des partenaires protéiques potentiels de l'ARN sat III dans les granules de stress nucléaires par immunofluorescence

Pour chacune des protéines testées, j'ai dans un premier temps établi les conditions expérimentales permettant leur détection par immunofluorescence. J'ai ensuite comparé au microscope à épifluorescence la localisation de ces protéines dans des cellules HeLa soumises ou non à un choc thermique. Puis, pour les protéines présentant une localisation différente après un choc thermique, j'ai testé leur présence dans nSB en évaluant leur co-localisation soit avec les ARN sat III détectés par RNA-

FISH, soit avec les protéines hnRNP HAP (SAF-B) ou SRSF1 (ASF/SF2) utilisées comme marqueurs protéiques des nSB.

Dans un premier temps j'ai dû établir les conditions d'immunofluorescence permettant la détection de la localisation des protéines contrôles HAP et SRSF1. J'ai ainsi pu observer au microscope à épifluorescence la formation des granules de stress nucléaires via la relocalisation de ces protéines dans des *foci* bien délimités dans plus de 90% des cellules (figure II-11). La protéine SRSF2 (SC35), présente dans les speckles et absente des granules de stress nucléaires, a été utilisée comme témoin négatif. Comme attendu, le profil observé pour la protéine SRSF2 ne change pas avant ou après le choc thermique.

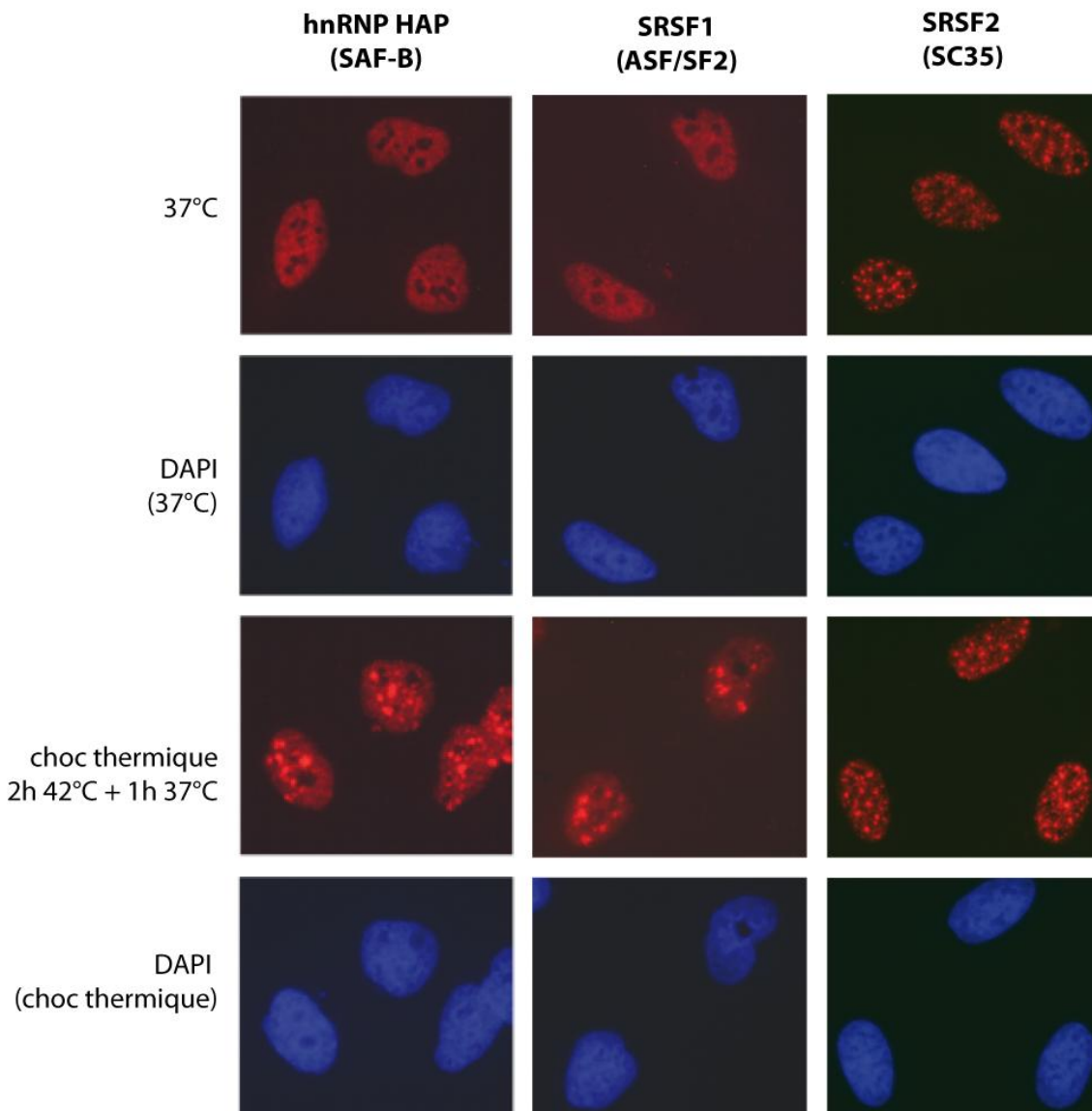


Figure II-11 : Visualisation de la formation de granules de stress nucléaires par analyse de la localisation des protéines SRSF1 et SAF-B. Les cellules HeLa S3 ont été soumises ou non à un choc thermique de 2 h à 42°C suivi d'une heure de régénération à 37°C, puis fixées à l'aide de formaldéhyde. Les protéines SRSF1 et hnRNP HAP ont été visualisées au microscope à épifluorescence par immunofluorescence. Les noyaux ont été colorés grâce au

DAPI. Un anticorps dirigé contre la protéine SRSF2 (SC35) phosphorylée, présente au sein des speckles servait de contrôle négatif dans cette expérience.

Des expériences de RNA-FISH suivies de tests d'immunofluorescence ont ensuite permis de vérifier au microscope confocal la co-localisation des protéines SRSF1 et hnRNP HAP avec les ARN sat III suite à un choc thermique (figure II-12). On note que la co-localisation de l'ARN sat III et de la protéine SRSF1 est quasiment parfaite, bien que SRSF1 forme un léger bruit de fond hors des granules dans le nucléoplasme. De la même manière, après le choc thermique la protéine HSF1 est co-localisée avec les ARN sat III au sein des nSB. La protéine hnRNP HAP est également en grande majorité contenue dans les granules de stress sauf à certains endroits où elle forme de minuscules *foci*. Il s'agit potentiellement des mêmes *foci* que ceux formés par hnRNP HAP à 37°C, à peine visibles au microscope à épifluorescence à cause du plus fort signal nucléoplasmique (figure II-11). Il est en effet connu qu'à 37°C, la protéine hnRNP HAP peut également être comprise dans des *foci* nucléaires, qui contiennent également la protéine Sam68, et qui sont par conséquent nommés "Sam68 nuclear bodies" (SNB) (Denegri et al., 2001). Je n'ai pas pu confirmer qu'il s'agissait bien de SNB du fait des difficultés rencontrées avec l'utilisation de l'anticorps dirigé contre Sam68.

Ayant validé le système d'étude, nous avons alors analysé la localisation après un choc thermique des protéines hnRNP R, U, H et Q capables de s'associer *in vitro* avec les ARN sat III dans des extraits de cellules HeLa (figures II-13, II-14 et II-15). Les protéines SR SRSF3 (SRp20) et SRSF4 (SRp75) ont aussi été incluses dans cette liste (données non présentées). L'intensité des signaux obtenus pour les protéines hnRNP Q (figure II-13), SRSF3 et SRSF4 était trop faible pour que nous en tirions des conclusions. Le signal obtenu avec les anticorps dirigés contre la protéine hnRNP R était fort et réparti de façon uniforme dans tout le nucléoplasme (figure II-13). Ce profil était identique avant ou après un choc thermique, indiquant que hnRNP R n'était pas relocalisée dans les nSB *in cellulo* lors de ce stress. La protéine hnRNP U adoptait également un profil identique dans les cellules non stressées et les cellules soumises au choc thermique. Elle était localisée autour des nucléoles et adoptait dans la plupart des cellules une répartition préférentielle sur les bords des noyaux. Cependant rien n'indiquait que cette protéine puisse être relocalisée dans les nSB après le stress (figure II-13). De façon intéressante, les signaux obtenus pour la protéine hnRNP H révélait la formation de *foci* de nature indéterminée, formés à proximité des nucléoles (figure II-14). Toutefois, l'absence de co-localisation avec la protéine SRSF1 ou avec hnRNP HAP indiquait que ces *foci* n'étaient probablement pas des nSB (figure II-14).

Choc thermique 2 h à 42°C + 1 h à 37°C

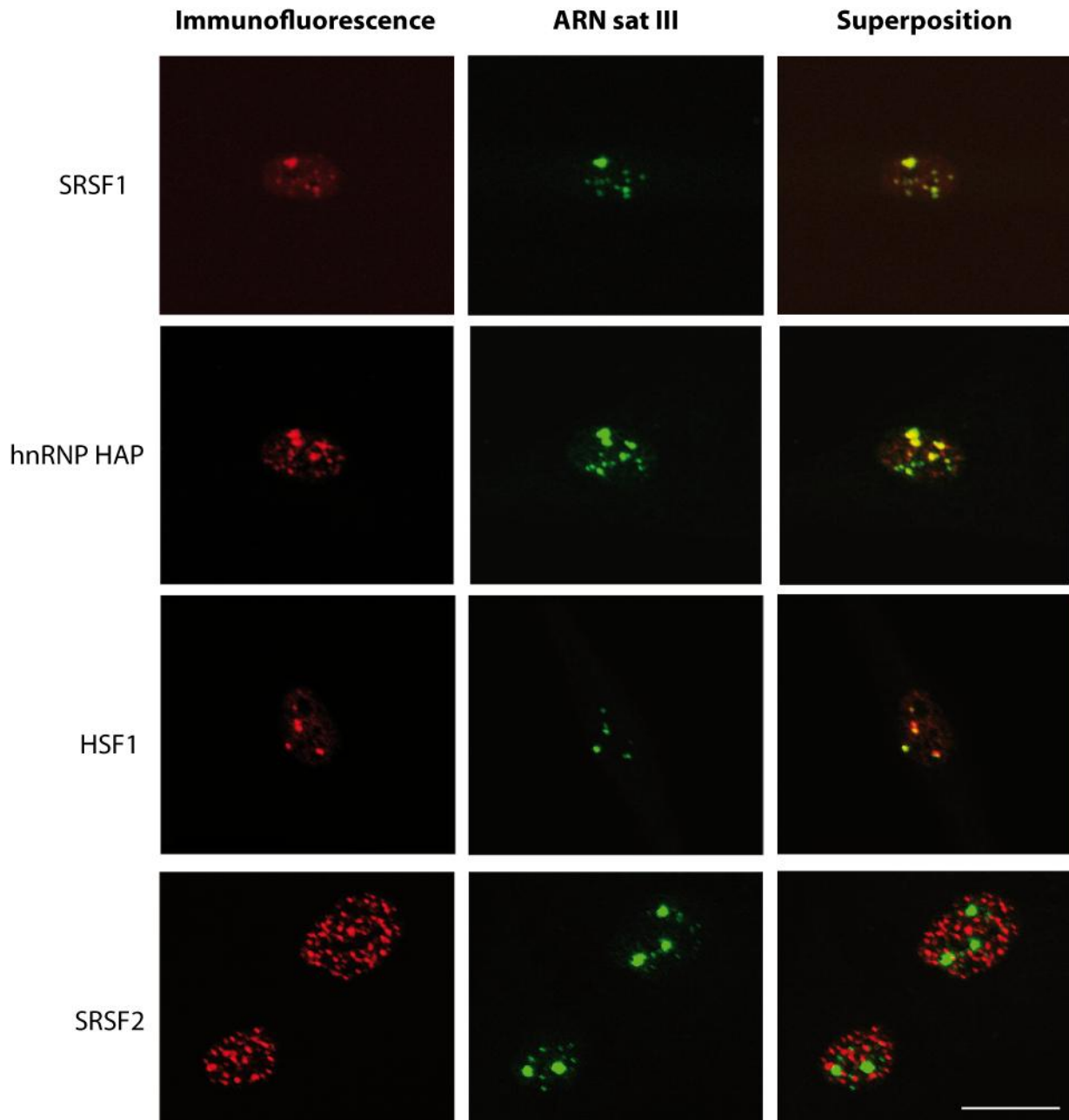


Figure II-12 : Visualisation par microscopie confocale de la co-localisation entre les ARN sat III et les protéines présentes dans les nSB après choc thermique (SRSF1, hnRNP HAP, HSF1). Les cellules HeLa S3 ont été soumises ou non à un stress thermique de 2 h à 42°C suivi d'une heure de régénération à 37°C, puis fixées à l'aide de formaldéhyde. Les ARN sat III ont été visualisés par RNA-FISH grâce à l'oligonucléotide LNA, dans les mêmes conditions que pour la figure II-12. La protéine SRSF2 localisée dans les speckles a été utilisée comme témoin négatif. Barre : 10µm.

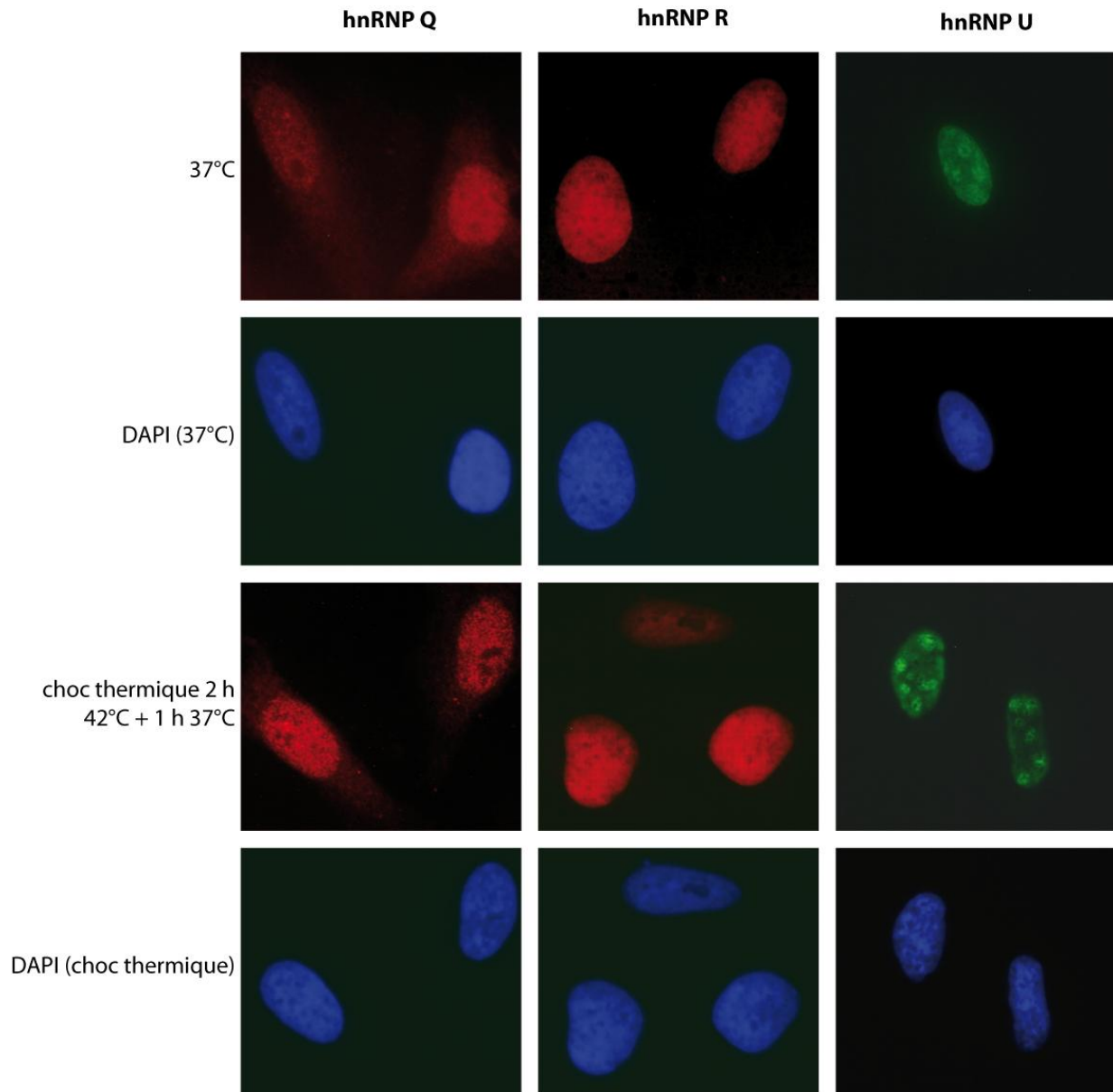


Figure II-13 : La localisation des protéines hnRNP Q, R et U est identique avant ou après un choc thermique. Les cellules HeLa S3 ont été soumises ou non à un choc thermique d'une durée de 2 h à 42°C suivi d'une heure de régénération à 37°C, puis fixées à l'aide de formaldéhyde. Les protéines ont été visualisées par microscopie à épifluorescence grâce à la technique d'immunofluorescence et les noyaux par coloration au DAPI (en bleu).

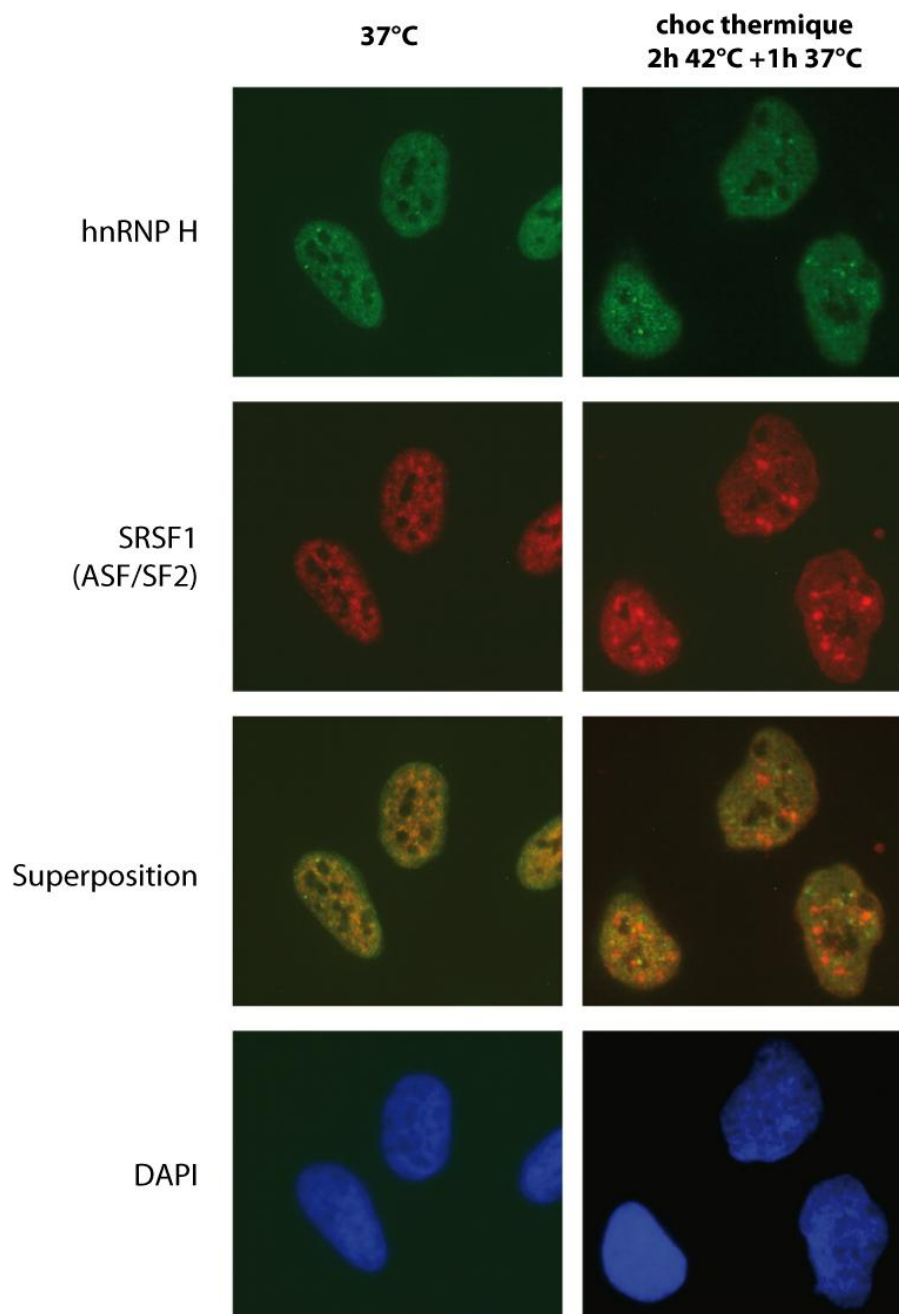


Figure II-14 : Différence de localisation entre les protéines hnRNP H et SRSF1 après un choc thermique. Les cellules HeLa S3 ont été soumises ou non à un choc thermique d'une durée de 2 h à 42°C suivi d'une heure de régénération à 37°C, puis fixées à l'aide de formaldéhyde. Les protéines ont été visualisées par immunofluorescence au microscope à épifluorescence et les noyaux par coloration au DAPI (en bleu).

Afin de confirmer ce point, j'ai répété les expériences d'immunofluorescence en modifiant les conditions de fixation des cellules, puisqu'il est établi que l'efficacité de révélation et l'aspect du signal des protéines par immunofluorescence dépendent de nombreux facteurs liés aux différentes étapes du déroulement de l'expérience, en plus de la spécificité intrinsèque de l'anticorps primaire *in cellulo*. Dans les expériences précédemment décrites, l'étape de fixation des cellules était réalisée par emploi de formaldéhyde ou de paraformaldéhyde, puis les cellules étaient perméabilisées à l'aide d'une solution de

détergent ou par une série d'incubations brèves dans l'azote liquide. Nous avons testé si la fixation par le méthanol pouvait être plus performante. Le méthanol permet de fixer les cellules en faisant précipiter les composants cellulaires sur la lame, tout en rendant les cellules suffisamment perméables pour permettre le passage des anticorps au travers des membranes. En utilisant ce procédé, nous avons pu observer une co-localisation partielle entre les protéines hnRNP H et hnRNP HAP. Cependant, l'aspect granulaire des deux protéines, identique avant et après le choc thermique, nous empêchait de distinguer les nSB. Un exemple de ce type de profil est donné dans la figure II-15. On observe par ailleurs des *foci* hnRNP H isolés qui ne sont pas co-localisés avec les *foci* SAF-B et vice-versa.

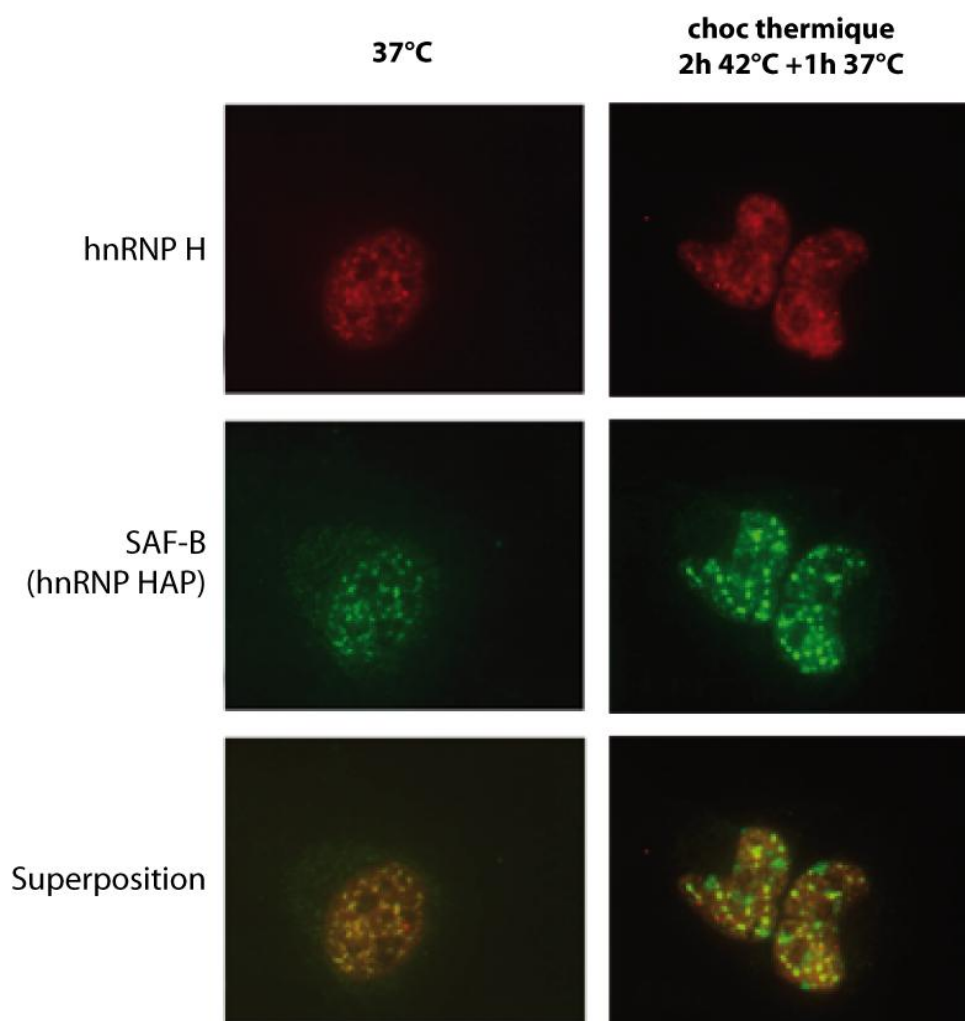


Figure II-15 : Test de la co-localisation des protéines hnRNP H et SAF-B (hnRNP HAP) après fixation des cellules au méthanol. Les cellules HeLa S3 ont été soumises ou non à un choc thermique d'une durée de 2 h à 42°C suivi d'une heure de régénération à 37°C, puis fixées et perméabilisées simultanément dans une solution de méthanol froid. Les protéines ont été visualisées au microscope à épifluorescence grâce à la technique d'immunofluorescence et les noyaux par coloration au DAPI (en bleu).

Même s'il n'est pas exclu que la protéine hnRNP H soit aussi présente dans les nSB, la formation de granules à 37°C indique qu'il s'agit probablement d'un biais lié à la précipitation des composants sur la

lame inhérent au procédé de fixation utilisé. Ainsi, une co-localisation des protéines candidates avec les nSB n'a pas pu être observée, y compris en utilisant d'autres solutions pour l'incubation avec les anticorps, comme le tampon PHEM, ou en augmentant la concentration de BSA ou en sérum, afin de diminuer le bruit de fond potentiel grâce à la saturation des sites de fixation aspécifiques des anticorps (données non présentées).

II.6. Test préliminaires de l'interaction des protéines SRSF1, hnRNP Q, hnRNP H et hnRNP U avec les ARN sat III *in cellulo*

En parallèle à ces expériences d'immunolocalisation, j'ai débuté une analyse de l'interaction avec les ARN sat III *in cellulo* de certaines des protéines s'associant avec l'ARN MiniHur-MS2 *in vitro*. Les protéines étudiées ont été immunosélectionnées à partir d'extraits totaux de cellules HeLa soumises au non à un choc thermique et la présence des ARN sat III au sein de la fraction immunosélectionnée a été testée par RT-qPCR. En effet, si ces protéines étaient associés *in cellulo* avec l'ARN sat III, lors de leur immunosélection nous pouvions espérer co-purifier des ARN sat III. Dans ces expériences, nous nous sommes plus particulièrement intéressés aux protéines hnRNP H, hnRNP Q et hnRNP U, pour lesquelles nous disposions d'anticorps supposés fonctionner en immunosélection. Les immunosélections ont été réalisées en immobilisant les anticorps sur des billes de G-sépharose. Les immunosélections des protéines SRSF1 (ASF/SF2) et SRSF2 (SC35) ont été respectivement utilisées comme témoins positif et négatif d'une interaction avec les ARN sat III. Une réaction contrôle réalisée en absence d'anticorps (billes de G-sépharose seules) a également été incluse.

Après incubation avec les extraits totaux de cellules HeLa, les billes ont été lavées, puis chauffées pour détacher les complexes immunosélectionnés. Un aliquot a été traité par la protéinase K, afin d'hydrolyser les protéines, et les ARN retenus ont été extraits au phénol avant d'être soumis à des expériences de RT-qPCR, tandis qu'un autre aliquot a servi à confirmer la présence des protéines immunosélectionnées par western blot. Les protéines hnRNP H, hnRNP U, SRSF2 et SRSF1 ont été efficacement immunosélectionnées, mais pas la protéine hnRNP Q (données non présentées).

La quantification des ARN sat III co-immunosélectionnés a été réalisée en utilisant le même protocole de RT-qPCR utilisé pour la mise au point des conditions de choc thermique. Cependant dans ce cas le signal obtenu pour les ARN sat III n'a pu être normalisé en utilisant l'ARNm de la protéine rplp0, qui n'était pas immunosélectionné avec les protéines candidates. Les résultats obtenus sont illustrés dans la figure II-16. Ils sont exprimés en facteur d'enrichissement du signal des ARN sat III entre la condition de choc thermique et la condition contrôle des cellules en culture à 37°C. Ce calcul était possible du fait de l'observation d'efficacités d'immunoselection similaires dans ces deux conditions. Comme attendu, nous observons un enrichissement important en ARN sat III immunosélectionné avec la protéine SRSF1 après un choc thermique (environ 200 fois par rapport aux

cellules non stressées), confirmant l'association des ARN sat III avec la protéine SRSF1 *in cellulo* et validant notre approche expérimentale. En accord avec ce qui était attendu, nous mesurons un faible facteur d'enrichissement lors de l'immunosélection de la protéine SRSF2 et lorsque l'immunosélection est réalisée en absence d'anticorps (Ctrl) (la figure II-16). De façon intéressante un enrichissement d'un facteur 1500, soit 7,5x plus importante qu'avec SRSF1, est observé avec la protéine hnRNP H. De même, nous constatons un enrichissement important avec la protéine hnRNP U, qui est d'un ordre juste deux fois inférieur à celui de SRSF1. Ces résultats, bien que préliminaires, suggèrent donc que parmi les protéines candidates testées, les protéines hnRNP H et hnRNP U peuvent être associées aux ARN sat III *in cellulo*.

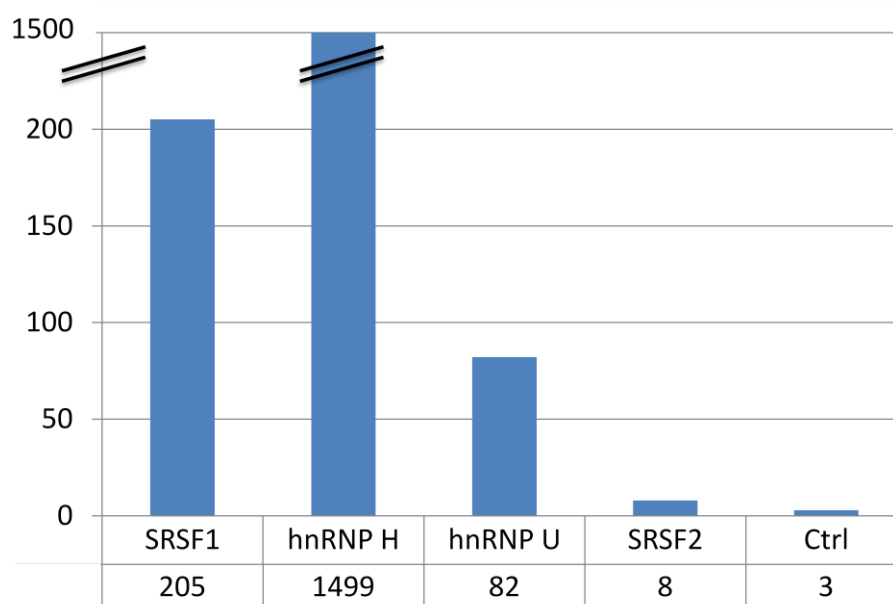


Figure II-16 : Les protéines hnRNP H et hnRNP U interagissent avec l'ARN sat III *in cellulo*. Les expériences réalisées ici sont des RT-qPCR réalisées sur des fractions récupérées après des immunosélections grâce à des anticorps dirigés contre des protéines que nous avons testées à cause de leur interaction avec l'ARN sat III *in vitro*. Le facteur d'induction calculé indique l'enrichissement en ARN sat III immunosélectionné dans les fractions après un choc thermique par rapport aux fractions réalisées sur des cellules en culture à 37°C. "Ctrl" représente les fractions d'immunosélection réalisées en absence d'anticorps.

II.7 Bilan et perspectives

Bien que l'on sache que les ARN satellite III jouent un rôle essentiel dans la formation des nSB, puisqu'en absence de leur transcription aucun granule n'est formé, peu d'informations étaient disponibles sur le rôle possible de ces ARN dans la formation des nSB.

Un ARN peut avoir un rôle dans l'assemblage de macrocomplexes, tels que les nSB, à la fois au travers de sa séquence et de sa structuration qui définissent ensemble ses capacités d'interaction avec

d'autres composants cellulaires. C'est pourquoi nous avons essayé à la fois d'obtenir des informations sur les capacités de structuration et d'interaction avec des protéines nucléaires des ARN sat III.

La nécessité d'adopter une structuration bien définie ou de s'associer à des partenaires définis génère des contraintes évolutives au niveau d'un ARN. Dans ce cadre, il est important de noter que dans l'état actuel des connaissances on sait que les séquences d'ADN codant les ARN sat III sont au moins soumises à une double contrainte vue la dualité de leur rôle. En effet, en plus de coder les ARN sat III, ces séquences d'ADN, comme d'autres séquences satellites du génome humain, sont en particulier impliquées dans l'ancrage des chromosomes avec des protéines du kinétochore lors de la mitose. De ce fait, il n'est pas facile par simple examen des séquences des ARN sat III de définir quels motifs sont impliqués dans la fonction de l'ARN en lui-même et quels motifs sont impliqués dans les fonctions de l'ADN correspondant. Cette problématique est commune à beaucoup de grands ARN non codants, comme c'est le cas pour l'ARN Xist, car à la fois eux même et leur matrice ADN ont des fonctions cellulaires.

Contrairement aux séquences ADN répétées de type Alu du génome humain qui sont apparues tardivement au cours de l'évolution, les séquences ADN répétées de type sat II et sat III, majoritairement localisées dans les régions centromériques des chromosomes, sont fortement conservées chez les eucaryotes. Les enchainements (GCAAT)_n des séquences des ARN sat II et sat III sont similaires aux séquences des éléments (CDE)_{III} des centromères de *S. cerevisiae* (Grady et al., 1992). Ceci souligne leur importance fonctionnelle possible dans des mécanismes cellulaires essentiels chez ces espèces en particulier dans la formation des centromères. Ce rôle pourrait en particulier reposer sur leurs propriétés thermodynamiques. En effet une étude thermodynamique réalisée sur des segments d'ADN (GGAAT)_n synthétiques a montré qu'ils ont une stabilité thermique inhabituelle. Le brin riche en G (G-rich) des ADN sat III a une stabilité thermodynamique similaire à celle du duplex d'ADN, ce qui reflète des capacités de structuration de ce brin sur lui-même, qui sont sans doute importantes lors de la séparation des paires de chromosomes lors de la mitose.

Sachant que l'ADN sat III doit posséder ces propriétés et en plus être capable de reconnaître spécifiquement certaines protéines, la question qui se pose est de savoir si la séquence et la structuration des ARN sat III est simplement la résultante des contraintes fonctionnelles de l'ADN sat III, ou si un second type de pression de sélection s'exerce sur au moins certains des ARN sat III afin qu'eux aussi puissent assumer leur fonction, en particulier dans les mécanismes de résistance au stress. La fonction supposée des séquences ADN sat II ou sat III suggère que la pression exercée au niveau de la séquence ADN est surtout de maintenir la régularité des motifs pentamériques. On s'attend à ce qu'une pression évolutive au niveau des transcrits sat III concerne à la fois les capacités de formation d'une structure secondaire ou tridimensionnelle et d'association avec des protéines partenaires.

Actuellement peu d'ARN sat III ont été séquencés mais lorsque c'est le cas, on note l'existence d'un degré non négligeable de variations au niveau des motifs GGAAU et encore plus au niveau des

décamères. Les variations ont pu apparaître soit de manière aléatoire, soit avoir été sélectionnées pour un avantage fonctionnel qu'elles confèrent.

II.7.1 Structuration des ARN sat III

Un premier point très clair qui ressort de nos données est que l'ARN sat III que nous avons étudié, présente un certain degré de structuration comme en attestent les nombreux clivages par la RNase V1, coupant spécifiquement les régions appariées ou contenant des empilements de bases. Un autre argument fort en faveur d'une structuration est la présence de très faible taux de clivage par les RNases aussi bien V1 que T1 et T2 dans d'autres segments. De manière intéressante les 4 décimères présents dans l'ARN étudié sont soit fortement clivés par la RNase V1, soit non hydrolysés, ce qui suggère un rôle de ces décimères, dont la séquence est moins conservée, dans l'organisation structurale des ARN sat III.

Bien que sur la base de nos données, plusieurs modèles de structures secondaires puissent être proposées pour l'ARN étudié, la structure tige boucle que nous proposons pour le segment s'étendant des positions 56 à 99 nous paraît tout à fait satisfaisante par rapport aux données. Bien que tous les résidus G de la boucle terminale soient clivés par les RNases T1, il n'est pas exclu que des résidus de cette boucle établissent entre eux des interactions non canoniques de plus faible stabilité.

Les deux autres décanucléotides aux extrémités de la molécule peuvent former un appariement mais il est plus irrégulier et moins stable. Cela s'explique peut être par le fait que la séquence clonée pour l'ARN pHur98 n'est pas complète et que la séquence partenaire du décanucléotide en 5' soit située en amont et non contenue dans le clone, et de la même manière que le partenaire du décanucléotide en 3' soit un décanucléotide situé en aval non contenu dans le clone. On pourrait en effet imaginer que les décanucléotides s'apparient par paires successives formant de grandes boucles, ou des régions de structures irrégulières, contenant les pentamères. Ainsi on constate que le segment allant des positions 14 à 52 peut former un certain nombre de petites structures tige-boucles et d'appariements à longue distance au sein de la molécule, qui témoignent de capacités claires de structuration impliquant des appariements de type Watson-Crick.

Notre étude suggère donc que les ARN sat III n'ont sans doute pas tous la même architecture. Elle doit beaucoup dépendre du nombre et du degré de divergence des enchaînements de pentamères et de la complémentarité existante entre leurs décimères. Si bien que différents types de motifs structuraux coexistent sans doute au sein d'un même ARN sat III. En plus l'interaction de ces ARN avec des partenaires protéiques peut favoriser ou défavoriser la formation de certaines structures.

Pour compléter nos données, il serait important d'une part d'étudier des sous fragments de l'ARN pHur98 en particulier un ARN allant des positions 53 à 109 afin de vérifier la formation de la structure tige-boucle identifiée. Par ailleurs, une étude de la structure secondaire d'autres ARN sat III

nous permettra de vérifier si nos conclusions peuvent être généralisées à l'ensemble de ces ARN. Enfin, il serait intéressant de tester l'impact de la structure secondaire sur l'interaction de l'ARN pHur98 avec les partenaires que nous avons identifiés.

II.7.2 Interaction des ARN sat III avec des protéines

Comme la purification des nSB n'avait jamais été réalisée et semblait difficile à mettre en œuvre, nous avons développé différentes approches en vue de définir les capacités d'interaction des ARN sat III avec des protéines nucléaires.

Les approches *in vitro* que nous avons d'abord réalisées visaient à identifier des protéines qui peuvent se lier directement aux ARN sat III, ou s'associer à des protéines se liant directement à ces ARN. Nous avons confirmé tout au long de cette étude une interaction entre l'ARN MiniHur-MS2 aussi bien par spectrométrie de masse ou par western blot. On savait qu'un court segment de l'ARN pHur98 se liant directement mais avec une affinité assez faible à SRSF1 avait été identifié (Chiodi et al., 2004). De manière intéressante, il correspond au troisième décanucléotide de l'ARN pHur98. Dans notre modèle de structure secondaire ce décanucléotide est partiellement apparié. Néanmoins, on sait aussi que SRSF1 a de fortes capacités d'interaction avec des séquences riches en purines. Or les ARN sat III comportent justement un très grand nombre de motifs GGAAU qui doivent donc participer à la fixation de cette protéine.

Notre recherche informatique de sites de fixation d'autres protéines SR dans l'ARN pHur98 suggérait que SRSF5 et SRSF7 pouvaient avoir des sites de liaison dans cet ARN. SRSF7/9G8 est connue pour se relocaliser partiellement dans les nSB (Denegri et al., 2001). La relocalisation partielle de SRSF5/SRp40 dans les nSB a aussi été mentionnée dans un article sans démonstration présente (Chiodi et al., 2004). Plusieurs sites de fixation de SRSF5 sont prédits dans l'ARN pHur98 par ESEFinder, 3 d'entre eux chevauchent des séquences de décarnères dont encore le troisième décarnère de cet ARN. La fixation simultanée de plusieurs molécules de SRSF5 sur l'ARN pHur98 explique peut être la rétention du complexe ARN-protéine dans les poches du gel que nous avons observé.

Des sites potentiels de fixation de SRSF7/9G8 sont aussi situés dans les décanucléotides (dont le consensus est GAACC/ACGAGU) qui renferment des séquences UCAACP_y ou UCAACA, ce qui correspond à la séquence de fixation de SRSF7 identifiée par pontage aux UV (tableau I-2 - 1ere partie) (pour revue, Bourgeois et al., 2004; Long and Caceres, 2009). Ceci peut expliquer notre obtention d'un retard sur gel clair lorsque nous avons incubé l'ARN pHur98 avec des concentrations croissantes en protéines SRSF7. Donc comme SRSF1, SRSF7 pourrait être relocalisée dans les nSB du fait de ses capacités d'interaction directe avec les ARN sat III. Les protéines SRSF1 et SRSF7 agissent souvent de concert en se liant sur des sites voisins dans les pré-ARN^m pour réguler l'épissage alternatif. On pourrait imaginer leur fixation à proximité l'une de l'autre sur les ARN sat III, l'une reconnaissant les séquences

UCAACA/Py des décanucléotides, l'autre le reste de la séquence des décanucléotides et les éléments riches en purines des pentamères.

Constatant que les 3 protéines SR connues pour être localisées dans les nSB possèdent un domaine RRM2 ou RRMH et que ce domaine est nécessaire à la fixation de SRSF1, l'équipe de G. Biamonti avait proposé que la présence de ces protéines dans les nSB soit due à leur liaison directe avec l'ARN et ceci par leur domaine RRM2 (Chiodi et al., 2004). Cela pourrait être le cas pour SRSF5 d'après nos données. Une recherche par blast sur le site internet de NCBI identifie trois protéines possédant des domaines RRM2 très similaires à celui de SRSF1 : les protéines SRSF4 (52% d'identité), SRSF6 (49% d'identité) et hnRNP M (34% d'identité). hnRNP M a bien été identifiée dans les nSB, et elle était détectée par spectrométrie de masse dans les complexes formés avec l'ARN MiniHur-MS2.

Par contre, nous n'avons pas détecté SRSF4 et SRSF6 par spectrométrie de masse. La présence d'un domaine RRM2 homologue à celui de SRSF1 n'est donc peut-être pas une condition suffisante pour une interaction directe avec l'ARN sat III. Cette absence peut également être due à la complexité et la diversité des modes de reconnaissance des ARN par les domaines RRM.

Pour compléter cette partie de notre étude, il sera intéressant de définir les sites respectifs des protéines SRSF1, SRSF7 et aussi SRSF5 sur l'ARN pHur98 par expérience d'empreintes avec des RNases, ceci en utilisant un ARN sauvage ou muté. Etant donné que le taux de SRSF1 détecté par western blot dans nos expériences de purification des complexes formés par l'ARN MiniHur-MS2 dans un extrait nucléaire de cellules HeLa semblait révéler une plus forte fixation de la protéine SRSF1 par rapport à SRSF7, il sera intéressant de tester si SRSF1 et SRSF7 entrent en compétition lors de leur fixation à l'ARN pHur98, ou si au contraire leur fixation se fait en synergie. Pour cela, des expériences de pontage aux UV réalisées avec un ARN uniforme marqué suivies de l'hydrolyse de l'ARN ponté, la purification des protéines portant des nucléotides marqués par immunosélection et la détection des protéines marquées par électrophorèse pourront être effectuées avec des mélanges de protéines SR dont les concentrations relatives varient.

Notre recherche des partenaires de l'ARN MiniHur-MS2 au sein d'un extrait nucléaire de cellules HeLa par chromatographie d'affinité, suivie d'une analyse par spectrométrie de masse et par western blot, nous a permis de mettre en évidence des capacités d'association *in vitro* des protéines hnRNP Q, R, U et H avec cet ARN (figure II-6). Les séquences ARN préférentiellement reconnues par hnRNP Q et R sont encore mal définies. Ces protéines contiennent plusieurs domaines RRM et un domaine riche en gly (tableau I-3 - 1ère partie). La protéine hnRNP H reconnaît des séquences de type poly(G) et GGGA/C. L'ARN pHur98 ne contient pas de longs enchainements de résidus G mais beaucoup de motifs GGAAU du fait de la présence des répétitions de pentamères. Vu notre forte détection d'ARN sat III dans les fractions obtenues par immunosélection de hnRNP H dans des extraits de cellules HeLa soumises à un choc thermique, il est possible que cette protéine s'associe directement aux ARN sat III, ou du moins à une protéine qui est fortement associée à ces ARN.

Peu de données sont disponibles sur un rôle possible dans protéines hnRNP H, H' et F dans la réponse au choc thermique. Par contre, un lien existe entre une série de protéines appartenant à la même famille : les protéines 2H9 (tableau I-3 - 1ere partie) produites par épissage alternatif du transcrite du gène *HNRNP3*, contiennent seulement 1 ou 2 domaines RRMH, à la différence des protéines hnRNP H, H' et F qui en contiennent 3. Ces protéines hnRNP 2H9 reconnaissent aussi des séquences GGGGA. Un anticorps dirigé contre 2H9 inhibe l'épissage *in vitro*, suggérant que certaines de ces protéines sont nécessaires au mécanisme général d'épissage (Mahé et al., 1997). Dix minutes après un choc thermique à 45°C ces protéines sont relocalisées des complexes hnRNP vers la matrice nucléaire et elles sont aussi détectées dans les corps nucléaires, ce qui a conduit à proposer qu'elles participeraient au blocage précoce de l'épissage produit par le choc thermique (Mahé et al., 1997). Il sera donc intéressant d'étudier de manière plus approfondie avec des anticorps plus spécifiques quelles sont les protéines parmi hnRNP H, H', F et 2H9 qui ont le plus d'affinité pour les ARN sat III, lorsque nous répéterons l'expérience d'immunosélection à partir d'un extrait de cellules HeLa soumises à un choc thermique. Par ailleurs, nous disposons de clones au laboratoire permettant de produire soit le domaine terminal de hnRNP H (RRMH1 + RRMH2) soit la protéine entière, nous pourrions tester leur capacité d'interaction directe avec l'ARN pHur98 par expérience de retard sur gel et par expérience d'empreintes.

Par immunosélection, nous avons aussi observé une association très significative de la protéine hnRNP U/SAF-A avec ARN sat III après un choc thermique. Cette protéine a une spécificité large de reconnaissance des ARN. Elle se fixe à la fois sur des enchainements des résidus G et sur des éléments riches en A et U. Les séries de pentanucléotides renferment beaucoup de résidus A, U et aussi G, son interaction avec les ARN sat III pourrait être directe.

Cette observation d'une association des ARN sat III avec hnRNP U pourrait avoir une signification fonctionnelle et ceci sur la base de différentes données de la littérature. Tout d'abord, comme SAF-B présente en grande quantité dans les nSB, hnRNP U/SAF-A est une protéine nucléaire interagissant spécifiquement avec les sites d'attachement de l'ADN de la matrice nucléaire (MAR). Chez la souris il a été montré qu'elle se lie aux séquences d'ADN satellite majeures (Kukalev et al., 2009; Lobov et al., 2000). Plus précisément, elle serait comme la lamine B capable de distinguer les séquences péracentromériques par rapport aux séquences centromériques car ces séquences satellites présentent une courbure de l'ADN (Lobov et al., 2001). L'ARN Xist requis pour qu'il y ait inhibition d'un des chromosomes X chez les mammifères reste associé à la chromatine après transcription comme c'est le cas pour les ARN sat III (Clemson et al., 1996). HnRNP U/SAF-A est requise pour cette rétention et l'accumulation de l'ARN Xist sur le chromosome X à inhiber (Hasegawa et al., 2010). On pourrait donc imaginer que hnRNP U participe aussi à la rétention des ARN sat III sur leur site de transcription. Un autre point intéressant concernant hnRNP U en lien avec la réponse au stress, est son implication dans plusieurs processus de réparation de l'ADN comme le système d'excision réparation des bases oxydées (Hegde et al., 2012). En plus hnRNP U/SAF-A participe à un mécanisme général de régulation de la

transcription par l'ARN polymérase II (Vizlin-Hodzic et al., 2011). Elle est aussi impliquée dans un contrôle global de l'épissage en régulant la maturation des snRNP U2 dans les corps de Cajal (Xiao et al., 2012). Non seulement hnRNP U joue un rôle majeur dans l'organisation de la chromatine durant tout le cycle cellulaire et ceci en liaison avec l'hélicase p68 mais en plus, elle joue un rôle majeur dans l'attachement des microtubules aux kinétochores et contribue à la localisation de la kinase Aurora-A aux fuseaux mitotiques (Ma et al., 2011; Pahlich et al., 2009). hnRNP U intervient donc dans un grand nombre de mécanismes liés à l'ADN : organisation, transcription, réparation, division.

Enfin, hnRNP M présente dans les nSB a été retrouvée plusieurs fois en association avec hnRNP U (Pahlich et al., 2009). Il n'est donc pas illogique de proposer sur la base de nos résultats qu'elle puisse aussi intervenir dans la réponse au stress thermique en maintenant les ARN sat III à leur site de transcription, ceci en étant capable d'interagir à la fois avec la chromatine et avec ces ARN. Les expériences futures devront être désignées afin de vérifier ces deux hypothèses. Néanmoins, d'après nos résultats de co-localisation des protéines dans les nSB, si hnRNP U intervient dans l'activité des nSB, elle n'y est pas relocalisée de façon massive comme SRSF1, SAF-B ou HSF1.

En effet, la technique performante que nous avons mise en point pour étudier la co-localisation des ARN sat III avec des protéines ne nous a pas permis de mettre en évidence une forte co-localisation d'aucune protéines autre que SRSF1, SAF-B et HSF1 dans les nSB. Après un choc thermique, hnRNP U, en plus d'une localisation diffuse dans le noyau, était proche des nucléoles. Notre absence de détection de granules formés par hnRNP U et de co-localisation entre les protéines hnRNP HAP et hnRNP H et dans les nSB par immunofluorescence, alors que nous détectons leur interaction par immunosélection des protéines à partir d'un extrait de cellules ayant subi un choc thermique, pourrait s'expliquer par une trop faible concentration de ces protéines dans les nSB et un manque de sensibilité des anticorps utilisés. Pour remédier à ce problème, nous envisageons de répéter les expériences d'immunofluorescence avec des cellules HeLa exprimant des protéines portant une étiquette (His, Flag ou GFP).

Cette approche n'avait initialement pas été envisagée, car nous craignons que le stress généré par la transfection n'affecte la localisation des protéines étudiées voire même la formation des nSB. Cependant des tests préliminaires réalisés au laboratoire ont montré que nos conditions de transfection des cellules n'induisaient pas la formation des nSB et, après un choc thermique, ne semblaient pas affecter la formation des nSB ainsi que leur détection par FISH. Les transfections seront réalisées de manière à ne pas surexprimer à un taux trop élevé les protéines candidates et la présence d'étiquette permettra d'utiliser un anticorps dirigé contre l'étiquette efficace en immunofluorescence ou d'utiliser la fluorescence intrinsèque de la GFP en imagerie. Nous disposons d'ores et déjà de plasmides permettant d'exprimer des versions étiquetées des protéines hnRNP Q, SRSF3, SRSF5 et hnRNP H.

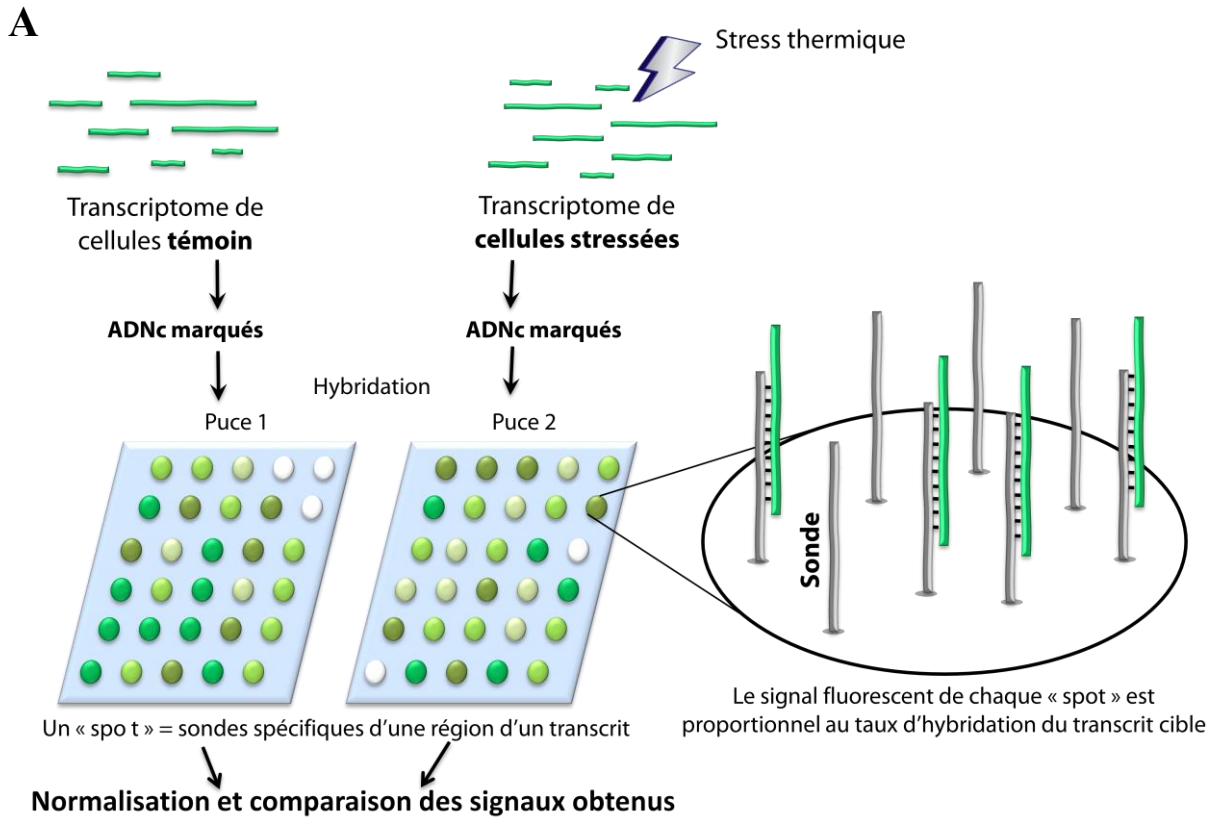
III. Etudes des conséquences du choc thermique sur l'épissage alternatif

Dans l'état actuel des connaissances on sait que au moins 3 protéines SR sont totalement (SRSF1, SRSF9) ou partiellement (SRSF7) relocalisées dans les nSB et que dans le cas de SRSF1 et aussi vraisemblablement des deux autres, cette localisation serait liée à leur capacité d'interaction avec les ARN sat III. Il était donc important de déterminer quel est l'impact de cette relocalisation sur les régulations de l'épissage alternatif. On ne savait pas en effet si cette relocalisation correspondait à une séquestration de ces protéines au sein des nSB, limitant leur utilisation pour des régulations de l'épissage, ou s'il s'agissait simplement d'une concentration forte de ces protéines au niveau des sites de transcription les plus actifs après un choc thermique. Ainsi dans le cas des *foci* formés autour des ARN renfermant des répétitions CUG dans les noyaux des cellules de patients atteints de dystrophie myotonique de type I, il a été montré par des expériences de FRAP que des échanges rapides ont lieu entre ces granules et le nucléoplasme, et que les protéines relocalisées dans ces granules ne sont pas séquestrées de façon irréversible. Ce type d'expérience n'a pas été réalisé dans le cas des nSB, ce qui ne permet donc pas de savoir si les protéines SR qui y sont localisées s'échangent rapidement avec celles du nucléoplasme. On sait néanmoins d'après des études que des modifications de régulation d'épissage alternatif *in cellulo* d'une construction test comportant le gène E1A de l'adénovirus apparaissent en parallèle de la formation des nSB (Denegri et al., 2001).

Afin d'avoir une vision plus globale de l'effet du choc thermique sur l'épissage alternatif des cellules, j'ai donc entrepris de réaliser des expériences à l'échelle du génome, par emploi de puces ADN spécifiques des jonctions exoniques. Je commencerai par préciser le principe de ces expériences.

III.1. Description des puces à ADN utilisées spécifiques des jonctions exoniques

Les puces à ADN spécifiques des jonctions exoniques nommées HJAY (Human Junction Array) que nous avons utilisées étaient des prototypes réalisés par Affymetrix que nous avons obtenus grâce à notre participation au réseau européen d'excellence EURASNET (EUROpean Alternative Splicing NETwork). Ces puces contiennent des sondes spécifiques de tous les exons et de toutes les jonctions exon-exon identifiées dans les ARNm humains, soit 315 137 exons et 260 488 jonctions. Le principe général des puces et des puces à jonctions exon-exon est schématisé dans la figure III-1.



B.

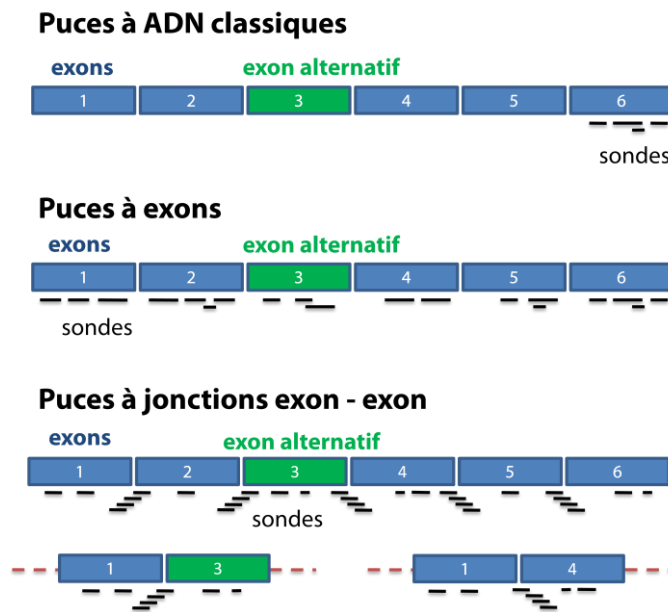


Figure III-1 : Principe des puces à jonctions exon-exon que nous avons utilisées. A. Stratégie générale pour la comparaison des transcriptomes de cellules stressées et non stressées par emploi de puces à ADN. Les ARN totaux à comparer sont rétrotranscrits et marqués grâce à un fluorophore, puis hybridés sur les micropuces. Chaque élément ("spot") de la puce comporte des sondes s'hybridant à une région définie d'un transcrit donné. Le signal fluorescent obtenu pour chaque élément est proportionnel au nombre de molécules d'ADNc cibles qui s'hybrident

sur cet élément. La comparaison des signaux fluorescents de chaque puce permet de quantifier la différence de taux d'expression des différents segments pour un transcrit donné. **B.** Types de sondes utilisés dans les puces à jonctions exon-exon. Contrairement aux puces ADN classiques ou aux puces à exon (en haut), les puces à jonctions exon-exon comportent des sondes s'hybridant au niveau des jonctions entre les exons, pour toutes les isoformes connues d'un même ARNm.

Chacun des exons et chacune des jonctions exon-exon est ciblé par un ensemble (probeset) de 8 sondes. Dans le procédé Affymetrix les sondes sont synthétisées directement sur la puce. Plusieurs de ces sondes identiques sont synthétisées en un point donné de la puce. Chaque point, après l'hybridation des ADNc marqués, renvoie un signal fluorescent proportionnel au nombre de fragments d'ADNc cibles qui s'y sont hybridés, ce qui reflète la concentration du segment d'ARNm correspondant présent dans l'échantillon. Ainsi, nous pouvons avoir accès de manière relativement précise aux informations concernant non seulement la fréquence d'utilisation des exons mais également à l'identité des isoformes produites, c'est à dire les combinaisons d'exons produites. Par ailleurs, l'utilisation de ces puces permettait d'accéder aux données de régulation de l'expression des gènes en plus des données relatives à l'épissage alternatif. L'emploi de ces puces, couplé à celui de puces spécifiques des ARN non codant (GeneChip® miRNA 3.0, Affymetrix), nous permettait donc de réaliser une analyse transcriptomique complète. Les manipulations de ces puces ont été réalisées en collaboration avec l'équipe de L. Vallar (CRP, Luxembourg). L'analyse statistique des données issues des puces à jonctions exon-exon a été faite par la société Genosplice, tandis que celle des puces microARN a été faite par l'équipe de L. Vallar.

III.2. Nature des échantillons analysés sur les puces à jonctions exon-exon

Des hybridations avec les puces à jonctions exon-exon que nous venons de décrire ont été réalisées avec des ARN provenant de 3 expériences distinctes incluant chacune l'isolement d'ARN à partir de cellules HeLa contrôles et de cellules HeLa soumises à un choc thermique. Le choc thermique a été réalisé par incubation des cellules 2 h à 42°C suivi d'1 h de régénération à 37°C, dans les conditions expérimentales établies dans le chapitre précédent. Pour chacune de ces expériences, nous avons validé l'expression de l'ARN sat III et de l'ARNm hsp70 par RT-qPCR. Les ARN totaux extraits des cellules étaient soumis à différents traitements permettant de les convertir en ADNc marqués utilisés pour l'hybridation (voir Matériel et méthodes). Une fois l'hybridation réalisée et les lavages effectués, les puces ont été scannées et le signal lumineux obtenu pour chacun des spots a été quantifié. Pour chaque "spot", la moyenne des signaux obtenus pour les 3 échantillons correspondants aux cellules non stressées a été comparée à la moyenne des signaux obtenus pour les 3 échantillons de cellules soumises au choc thermique. Une différence de signal indiquait une différence du taux d'expression du segment d'ARNm correspondant à la sonde contenue dans le "spot". Comme mentionné préalablement, l'analyse statistique

des puces à jonctions exon-exon a été confiée à la société Genosplice, qui nous a fourni les résultats sous forme d'un tableau récapitulant les variations d'expression des gènes et d'épissage alternatif qui étaient statistiquement significatives.

III.3. Données concernant les variations d'expression des gènes

III.3.1. Identification préliminaire des gènes induits et réprimés

Du point de vue de la variation globale de l'expression des transcrits, une différence d'expression significative d'un facteur d'au moins 1,5x a été mise en évidence pour 1159 gènes (avec une p value <0,05). Parmi ces gènes, 692 étaient régulés positivement et 467 étaient régulés négativement :

	Régulés positivement	Régulés négativement	Total
Nombre de gènes régulés	692	467	1159
Moyenne des facteurs de régulation	3,85	1,86	3,05
Médiane des facteurs de régulation	2,08	1,72	1,86
Facteur de régulation max	358,22	4,03	

Tableau III-1 : Données des puces à jonctions exon-exon concernant la détection des gènes dont l'expression est induite ou réprimée après un choc thermique de 2 h à 42°C suivi d'1 h de régénération à 37°C.

On remarque que l'activation de la transcription des gènes semble plus importante que leur inactivation lors d'un choc thermique, non seulement en termes de nombre de gènes dont l'expression varie, mais également en termes d'intensité de la variation, la moyenne et la médiane des facteurs d'induction étant plus élevées dans le cas des activations. Cette moyenne est en effet augmentée par la très forte induction de protéines de réponse au stress, en particulier les HSP. Les gènes les plus induits étaient *HSPA6* (358x) et *HSPA7* (113x), qui sont des gènes codant des protéines chaperonnes de la famille hsp70. Plusieurs autres gènes de protéines HSP font partie des gènes les plus induits (plus de 2x), notamment les gènes des protéines hsp70-1 et hsp70-2, toutes deux induites d'un facteur d'environ 3,5x, ainsi que les protéines hsp70-like 1 et 4 (respectivement 2,5 et 3,5x) et hsp70-4 (1,8x). On retrouvait également parmi les gènes induits les gènes codant les protéines hsp105/110 (~5x), la protéine hsp60 (~2x), et des gènes de "small HSP" de faibles masses moléculaires, comme des membres de la famille hsp40/DnaJ (les membres 1, 2, 4 et 6, induits d'un facteur entre 2 et 6x), les protéines hsp27/hspB1 et hsp27-2/hspB2 (respectivement 1,9x et 1,6x) et hsp22/hspB8 (2,3x). L'induction de toutes ces protéines HSP confirmait qu'il y avait bien eu la réponse des cellules à un stress. Cependant, il est à noter que,

contrairement à ce qui est décrit dans la littérature, dans nos cellules les protéines hsp70-6 et hsp70-7 (et non hsp70-1) étaient les formes les plus exprimées parmi les protéines hsp70. On notera également l'absence de variation d'expression des gènes codant les protéines de la famille hsp90. Ces différences pourraient relever d'un profil d'expression particulier des cellules après une période de régénération à 37°C relativement courte par rapport à ce qui avait été étudié dans la littérature. Cela explique peut-être aussi le fait que le gène *HSF2* soit régulé négativement d'après nos données (d'un facteur 1,9x environ). En effet, la protéine HSF2 interviendrait comme un régulateur de l'activité de HSF1, dont l'activité doit commencer à baisser peu après la fin du choc thermique. Par ailleurs, il est étonnant que l'augmentation du taux d'expression de hsp70-1 soit de 3,5x environ d'après l'analyse par puce alors qu'il est de 40x d'après les analyses par RT-qPCR, même si les données issues de puces sont moins précises que les données de RT-qPCR, et fournissent d'avantage des tendances générales.

De façon intéressante, on retrouve en position 3 à 11 du classement des gènes les plus induits des EST dont les gènes ne sont pas définis. La plupart ont disparu ou ont été déplacées dans les bases de données d'assemblage des séquences du génome humain la plus récente (nommée GRCh37/hg19, datant de septembre 2009) par rapport à celle utilisée pour référence dans l'analyse par la société Genosplice (NCBI36/hg18, datant de mars 2006). Cependant, ces EST sont des séquences identiques ou très proches des ARN sat III transcrits à partir du brin "G-rich" ou "C-rich", ce qui confirme également que le choc thermique réalisé avait été efficace.

En ce qui concerne les gènes les plus négativement régulés, parmi les 5 premiers de la liste en termes de facteur de répression, nous retrouvons les histones de la famille H1B, H1D, H3A et H3C, qui sont régulées négativement entre 3,6 et 4x environ. Plus loin dans cette liste, on retrouve également plusieurs autres gènes d'histones appartenant à plusieurs familles, ainsi que des histones déacétylase ou méthyltransférase, comme l'enzyme responsable de la méthylation de la lysine 9 de l'histone 3. L'ensemble de ces gènes a été regroupé dans le tableau III-2.

Symbole	Description	facteur de répression	P-value
H1	Histone 1, H1b (HIST1H1B) Histone H1.5 (Histone H1a)	3,94	9,32E-03
H1.3	Histone 1, H1d (HIST1H1D) Histone H1.3 (Histone H1c)	3,78	4,80E-03
H3/A	Histone 1, H3a (HIST1H3A) Histone H3.1	3,69	2,73E-02
H3.1	Histone 1, H3c (HIST1H3C) Histone H3.1	3,65	1,83E-02
H3.1	Histone 1, H3e (HIST1H3E) Histone H3.1	3,02	2,97E-02
H2A/m	Histone 1, H2ab (HIST1H2AB) Histone H2A type 1-E (H2A.2)	2,97	3,48E-02
HIST3H2BB	Histone 3, H2bb (HIST3H2BB) Histone H2B type 12	2,84	2,38E-02
HIST3H2A	Histone 3, H2a (HIST3H2A) Histone H2a	2,27	1,46E-02
H1.1	Histone 1, H1a (HIST1H1A) Histone H1.1	2,27	1,87E-02

Symbole	Description	facteur de répression	P-value
H3.f	Histone 1, H3i (HIST1H3I) Histone H3.1	2,19	3,64E-02
H1.2	Histone 1, H1c (HIST1H1C) Histone H1.2 (Histone H1d)	2,15	2,81E-02
HIST1H2BC	Histone 1, H2bg (HIST1H2BG) Histone H2B	2,15	4,17E-02
H2B/r	Histone 1, H2bj (HIST1H2BJ) Histone H2B.r (H2B/r) (H2B.1)	2,13	3,18E-02
SIN3A	Histone deacetylase complex subunit Sin3a	2,11	1,06E-03
H4/a	Histone 1, H4a (HIST1H4A) Histone H4	1,93	3,54E-02
H3/h	Histone 1, H3g (HIST1H3G) Histone H3.1	1,86	4,49E-02
ESET	Histone-lysine N-methyltransferase, H3 lysine-9 specific 4	1,70	4,19E-02
H2B.1	Histone 1, H2bb (HIST1H2BB) Histone H2B.f (H2B/f) (H2B.1)	1,69	3,35E-02

Tableau III-2 : Gènes d'histones, d'une histone déacétylase et d'une histone méthyltransférase régulés négativement lors d'un choc thermique selon l'analyse des puces à jonctions exon-exon.

Comme il a été mentionné dans le chapitre d'introduction, la variation de l'expression de certains gènes d'histones avait déjà pu être observée lors d'un choc thermique. Cependant, la signification biologique de cette régulation n'a pas encore été étudiée en détail, et pourrait résulter d'un grand nombre de processus cellulaires.

III.3.2. Analyse des voies de régulation et des groupes fonctionnels auxquels appartiennent les gènes dont l'expression est modifiée lors d'un choc thermique

III.3.2.a. Analyses des voies de régulations affectées par emploi des outils d'analyse des bases de données KEGG et REACTOME

Plusieurs outils permettent d'analyser les voies de régulations influencées par des changements de conditions de croissance des cellules humaines à partir de données d'expériences à haut débit. Dans notre cas la société GenosplICE a réalisé cette analyse en utilisant des outils informatiques permettant de traiter un grand nombre de données et de les comparer aux données des bases de données KEGG et REACTOME. Ces outils permettaient d'identifier les processus cellulaires (dont la liste est spécifique à chacune de ces bases de données) pour lesquelles les données de puces identifient un nombre significatif de gènes dont l'expression varie entre les deux conditions étudiées. L'analyse par l'outil de la base de données KEGG n'a pas permis de mettre en évidence de variations très significatives (bien que les valeurs maximales de p-values minimales obtenues restent inférieures à 0,05). Les résultats obtenus avec l'outil utilisant la base de donnée REACTOME étaient plus conséquents puisque 134 réseaux de

régulation comprenaient un nombre significatif de gènes dont l'expression variait dans nos données, avec des p-values estimées entre $4,6.10^{-2}$ et $2,82.10^{-7}$. Dans les deux cas, les voies répertoriées ne sont pas spécifiquement représentatives du choc thermique. Cela résulte probablement du type de processus répertoriés dans cette base de données, qui sont prédéterminés. Ces deux outils nous ont toutefois permis de mettre en évidence des voies de régulations en concordance avec les données de la littérature en ce qui concerne la réponse au choc thermique, comme l'activation des MAPK ("MAPK signaling pathway", p-value= $5,97.10^{-3}$ dans KEGG). Parmi les 21 gènes régulés appartenant à cette voie, on trouve plusieurs gènes de HSP, comme ceux de hsp70-1, hsp105/110, hsp70-like et hsp70-6, ainsi que une MAP2K, la ERK kinase 3, et le facteur de transcription Jun. On note également l'activation de la voie d'activation de p53 après choc thermique, ce qui est cohérent avec le fait que cette voie est activée lorsque des dommages sont subis par l'ADN. De façon intéressante, on retrouve parmi les gènes induits qui appartiennent cette voie, les gènes codant deux protéines GADD45 α et GADD45 β , qui sont des protéines impliquées dans l'activation de la MAP3K MTK1, elle-même impliquée dans l'activation par le stress des SAPK (pour revue, Takekawa et al., 2011). Il est connu que les trois protéines paralogues GADD45 α , β et γ sont induites par le stress oxydant, le stress généré par les UV ou par le MethylMéthaneSulfonate (MMS), une molécule causant des dommages à l'ADN. Le profil d'induction est néanmoins différent dans chaque cas.

III.3.2.b. Regroupement fonctionnel des termes ontologiques grâce au logiciel DAVID

Afin de caractériser de manière plus précise les grandes catégories de mécanismes cellulaires modifiés dans les cellules soumises à un choc thermique, nous avons utilisé une autre approche, utilisant simultanément de nombreuses bases de données de termes ontologiques associés aux gènes dont l'expression variait. Par exemple les domaines protéiques (bases de données Interpro et Pir superfamily), les termes "Gene Ontology" (GO), les voies de régulation (KEGG, REACTOME) (Huang et al., 2009). Les termes GO sont issus de la base de données <http://geneontology.org> (Ashburner et al., 2000). Ce sont des dénominations standardisées attribuées à chaque gène suivant les caractéristiques de leurs produits. Un ou plusieurs termes GO sont ainsi associés à un produit de gène, sur la base de 3 caractéristiques le concernant : (1) le processus biologique dans lequel il intervient (exemple, GO:0008283 : prolifération cellulaire), (2) sa fonction moléculaire (exemple, GO:0003677 : fixation à l'ADN), et (3) sa localisation cellulaire (exemple, GO:0005634 : noyau). Chacun de ces termes est classé au sein d'une arborescence complexe permettant de resituer la fonction par rapport à d'autres termes GO et de créer une filiation entre les gènes. Par exemple, "fixation à l'ADN" est un terme parent de "GO:0000182 : fixation à l'ADN ribosomique", qui est plus précis. Cette arborescence est non linéaire et redondante. C'est à dire que plusieurs termes GO d'un même produit de gène peuvent appartenir à une même catégorie parente, ou au contraire qu'un seul terme GO de ce produit peut appartenir à plusieurs catégories parentes. Au total, 331

termes GO ont été répertoriés par GenosplICE à partir de nos données expérimentales. Le traitement de ces termes a également été réalisé avec un autre outil, CateGORizer, et est disponible en annexe 1 à la fin de cette partie (après la p. 204).

Le logiciel DAVID permet ensuite de regrouper les termes ontologiques biologiquement reliés entre eux en groupes fonctionnels ("fonctionnal clustering"). La pertinence des variations observées au sein de chaque groupe est calculée de la même manière que pour les bases de données KEGG et REACTOME décrites plus haut, c'est à dire en déterminant un taux d'enrichissement pour chaque terme et en évaluant sa significativité grâce à un test statistique (test de Fisher). Un score est ensuite attribué à la totalité du groupe en fonction des p-values obtenues pour chaque terme. Plus ce score est élevé, plus l'enrichissement constaté pour les termes réunis dans ce groupe fonctionnel est significatif. En soumettant la liste des gènes dont l'expression variait d'après l'analyse de nos puces à jonctions exon-exon, une série de groupes fonctionnels ressort comme ayant le meilleur score statistique. Les deux premiers groupes, qui possèdent un score statistique DAVID très élevé, sont présentés dans le tableau III-3.

Groupe fonctionnel (régulation de la transcription)

Score d'enrichissement DAVID : 14,48

Catégorie	Terme	Nombre	%	Enrich.	PValue
Terme GO_fonction moléculaire	GO:0003677~DNA binding	193	21,5	1,8	1,02E-18
Mot-clé Pir Superfamily	Nucleus	297	33,1	1,6	1,87E-18
Mot-clé Pir Superfamily	Transcription	172	19,2	1,9	6,05E-17
Mot-clé Pir Superfamily	dna-binding	160	17,8	1,9	6,30E-17
Mot-clé Pir Superfamily	transcription regulation	169	18,8	1,9	1,92E-16
Terme GO_processus biologique	GO:0006350~transcription	174	19,4	1,8	2,00E-16
Terme GO_processus biologique	GO:0045449~regulation of transcription	200	22,3	1,7	9,03E-16
Terme GO_fonction moléculaire	GO:0030528~transcription regulator activity	127	14,2	1,8	4,52E-12
Terme GO_processus biologique	GO:0051252~regulation of RNA metabolic process	144	16,1	1,8	5,16E-12
Terme GO_processus biologique	GO:0006355~regulation of transcription, DNA-dependent	141	15,7	1,8	8,69E-12
Terme GO_fonction moléculaire	GO:0003700~transcription factor activity	93	10,4	2,1	8,86E-12

Groupe fonctionnel (réponse au stress et aux protéines dénaturées)**Score d'enrichissement DAVID : 7,14**

Catégorie	Terme	Nombre	%	Enrich.	PValue
Mot-clé Pir Superfamily	stress response	18	2,0	6,2	2,43E-09
Terme GO_processus biologique	GO:0051789~response to protein stimulus	21	2,3	4,3	5,63E-08
Terme GO_processus biologique	GO:0006986~response to unfolded protein	15	1,7	4,7	2,86E-06

Tableau III-3 : Groupes fonctionnels de gènes pour lesquels des variations significatives d'expression sont détectées avec les meilleurs scores statistiques par le logiciel DAVID. A chaque terme ontologique est associée la base de données dont il provient (Catégorie), le nombre de gènes concernés dans nos données (nombre), le pourcentage de gènes que cela représente (%) sur la totalité des données analysées. DAVID permet également de calculer le facteur d'enrichissement de ces termes par rapport à la proportion que ce terme représente dans la totalité du génome humain, et détermine grâce à une p-value à quel point cet enrichissement est significatif. A partir des différentes p-values obtenues, il attribue un score au groupe fonctionnel (indiqué dans la première ligne de chaque groupe fonctionnel).

Comme on peut le constater, les groupes fonctionnels concernent directement la régulation de la transcription (groupe fonctionnel 1) et, comme on pouvait s'y attendre, la réponse au stress, en particulier à la présence de protéines dénaturées.

Etant donné la prépondérance de la régulation de la transcription dans les termes GO (voir annexe 1) et des regroupements fonctionnels DAVID réalisés à partir de nos données, nous avons utilisé un dernier outil permettant de savoir quels facteurs de transcription ont un nombre important de cibles parmi les variations d'expression les plus marquantes après un choc thermique.

III.3.2.d. Analyse par le logiciel Ingenuity

Grâce à nos collaborateurs de l'équipe de L. Vallar au Luxembourg, nous avons eu accès au logiciel Ingenuity, utilisant une base de données reposant uniquement sur les résultats expérimentaux publiés et contrôlés. Une de ses nombreuses fonctions intéressantes est l'analyse dite en amont ("upstream analysis"). Cette fonction permet d'identifier les facteurs régulateurs susceptibles d'avoir provoqué l'induction ou l'inhibition d'ensembles de produits de gènes observés dans nos résultats expérimentaux, ceci à partir de données répertoriés dans la base Ingenuity. A chacun de ces facteurs régulateurs est attribué une p-value, selon les mêmes principes que précédemment. Cette valeur reflète la probabilité que l'enrichissement constaté dans nos données en expression des gènes cibles de l'élément régulateur soit dû au hasard ou soit significatif : plus elle est faible, moins l'enrichissement est supposé être dû au hasard. Les éléments régulateurs considérés peuvent être des facteurs de transcription, ce qui nous intéressait particulièrement ici, mais également des microARN, des kinases ou des enzymes.

Ingenuity nous a permis d'identifier 72 gènes de protéines impliquées dans la régulation de la transcription, pour lesquels l'expression de leurs cibles variait significativement dans nos données expérimentales (voir liste complète en Annexe 1, à la fin de cette partie). Le premier facteur identifié avec le meilleur score est Nupr1 (= p8 ou COM1) (p-value $1,95.10^{-16}$). Soixante-six gènes cibles répertoriées de cette protéine ont une expression modifiée après un choc thermique selon nos données expérimentales. Cinquante-quatre d'entre elles varient d'une manière qui correspond à une activation de Nupr1. En lien avec cette observation, une surexpression des ARNm Nupr1 a été observée dans le foie de cabillaud lors d'un choc thermique et l'inactivation du gène Nupr1 chez des souris modèles de cancer du pancréas augmente la sensibilité des cellules pancréatiques à un stress thermique (Cano et al., 2013; Hori et al., 2010). Aucune variation significative du taux de transcription de Nupr1 n'était détectée sur nos puces, mais le taux de traduction ou l'activité du facteur Nupr1 pouvait être augmentée lors de la période de régénération des cellules après un stress thermique. On savait déjà que Nupr1 est capable de se fixer sur la chromatine pour réguler l'expression de gènes impliqués dans la résistance des cellules au stress, notamment le stress génotoxique (Clark et al., 2008; Jin et al., 2009; Vincent et al., 2012). C'est pourquoi, tout comme certaines HSP, ce facteur est impliqué dans la migration, la croissance et la résistance de cellules tumorales issues de différents tissus (Du et al., 2013; Jiang et al., 2006; Li et al., 2012; Sandi et al., 2011). Par ailleurs, Nupr1 interagit directement avec le facteur p53 et peut être activé dans diverses conditions de stress par la protéine kinase ATF4. ATF4 et p53 sont également suggérés comme éléments régulateurs en amont de gènes variant dans nos expériences d'après le logiciel Ingenuity (p-values $1,24.10^{-4}$ et $8,15.10^{-3}$, respectivement) (Clark et al., 2008; Jin et al., 2009). Parmi les régulateurs de la transcription identifiés grâce à Ingenuity, deux en particulier ont été mis en évidence lors de l'analyse : HSF1, le facteur de régulation général des HSP (p-value $4,28.10^{-11}$) et p53, le facteur de régulation lié aux dommages à l'ADN.

Il est à noter qu'Ingenuity a aussi identifié d'autres éléments régulateurs en amont. Il peut s'agir de protéines régulatrices, de type kinase notamment, situées en amont dans les cascades de transduction du signal, voire même de microARN. Parmi ces autres éléments régulateurs, celui identifié avec la p-value la plus faible est le complexe PDGF B (p-value $1,69.10^{-15}$). Ce facteur de croissance est impliqué dans le contrôle de la prolifération et de la migration cellulaire, et sa fixation à son récepteur PDGFR est capable d'induire la signalisation intracellulaire par les voies de MAPK, PI3-K et PKC γ , 3 voies potentiellement impliquées dans la réponse au stress thermique (Fang et al., 2013)(pour revue, Heldin and Lennartsson, 2013). Il est donc possible que PDGF B soit un acteur utilisé en amont pour la stimulation de ces voies, ou bien que son identification par Ingenuity soit due à la régulation de plusieurs de ses cibles en aval, ne résultant pas nécessairement de la présence de PDGF B, mais plutôt d'autres mécanismes activés par le choc thermique. En effet, il a par exemple été montré que l'activation d'une autre kinase régulatrice, PKB, repose sur deux mécanismes bien distincts selon qu'elle ait lieu après stress thermique ou grâce à des facteurs de croissance de type PDGF (Matsuzaki et al., 2004). D'autre

part, on retrouve bien également avec des p-values parmi les plus faibles le groupe de kinases ERK (p-value $4,64.10^{-9}$), et de façon plus générale le groupe des MAP2K (p-value $6,03.10^{-6}$), ainsi que la kinase p38 ($3,36.10^{-5}$), et PI3K ($6,74.10^{-4}$), Jun (1.10^{-3}) et le groupe de kinases PKC ($1,95.10^{-3}$), appartenant toutes au système de régulation des Stress-induced Protein Kinase (SAPK). De nombreux autres éléments ont été identifiés, et il conviendra de les étudier plus en détails lors d'expérimentations ultérieures. Etant donné la nature de cette étude, une fois cet aperçu des processus de régulation de la transcription établi, nos expériences ont porté en premier lieu sur les événements d'épissage alternatifs détectés après le choc thermique.

III.4. Données des puces concernant l'épissage alternatif

III.4.1. Analyse globale des transcrits soumis à l'épissage alternatif

Les événements d'épissage modifiés après le choc thermique repérés grâce aux puces ont été classés en 3 catégories : les exons cassette, les exons mutuellement exclusifs, et les sites 5'/3' alternatifs (tableau III-4). Un indice de confiance a été attribué à la détection de ces événements, plutôt qu'un indice de probabilité, car la détection et la validation de ces événements est plus complexe que pour la variation d'expression des ARNm. Cet indice est dépendant : 1) de la probabilité que la modification d'intensité observée à chaque site d'épissage impliqué dans l'événement d'épissage observé soit significative ($p < 0,05$ pour tous les sites impliqués), 2) du taux d'expression de l'ARNm concerné par cet événement (un taux faible faisant diminuer l'indice de confiance), mais aussi 3) de la cohérence ou au contraire des contradictions des données obtenues avec les différentes sondes testant cet événement d'épissage (accord entre le signal obtenu pour les exons et les jonctions avoisinantes). En effet, en particulier lors de la détection de sites 5'/3' alternatifs, il arrive par exemple que les données sur l'ensemble des sondes testant la présence de la forme longue de l'exon ne soient pas disponibles, ce qui fait diminuer l'indice de confiance.

Indice de confiance	Exons cassette	Exons mutuellement exclusifs	Sites 5'/3' alternatifs
HAUT	96	0	5
MOYEN	165	7	33
BAS	84	15	2
TRES BAS	9	3	1
TOTAL	354	25	41

Tableau III-4 : Données globales obtenues avec les puces à jonctions exon-exon sur la détection d'événements d'épissage alternatifs nouveaux après un choc thermique de 2 h à 42°C suivi d'une heure de régénération à 37°C.

Près de 85% de l'ensemble des événements d'épissage alternatif nouveaux détectés après choc thermique et 96% des événements détectés avec un indice de confiance élevé, étaient de type exon cassette. C'est pourquoi nous avons dans un premier temps focalisé notre analyse sur cette catégorie d'exons, en ne prenant en compte que les événements possédant un indice de confiance moyen ou haut, ce qui concernait 249 gènes. Aucun terme GO n'ayant été fourni par la société Genosplice, nous avons utilisé à nouveau l'outil DAVID pour définir dans un premier temps à quels types de processus biologiques selon la nomenclature GO appartenaient les gènes dont les transcrits présentaient de nouveaux événements d'épissage après choc thermique. Nous avons ainsi pu définir quels groupes fonctionnels étaient les plus affectés par des modifications de l'épissage. Nous avons utilisé les termes GO parentaux les plus représentatifs de chaque groupe et reporté la p-value pour chacun de ces termes (tableau III-5). Le détail des protéines de chaque catégorie est fourni en annexe 3 à la fin de ce chapitre.

Terme GO (processus biologique)	Nombre	%	Enrich.	P-value	score DAVID
Cycle cellulaire	26	10%	2,4	8,00E-05	1,97
Transport des protéines	24	10%	2,2	4,00E-04	1,7
Maturation des ARN	19	8%	2,5	6,60E-04	1,33
Réponse cellulaire au stress	19	8%	2,4	9,80E-04	2,23
Processus catabolique des protéines	19	8%	2,2	2,80E-03	2,02
Régulations des processus basés sur les microtubules	5	2%	7,3	4,80E-03	1,3
Organisation des sous-unités de complexes macromoléculaires	20	8%	2,0	5,00E-03	1,4
Modification de la chromatine	11	4%	2,9	5,20E-03	1,88
Régulation de la transcription	51	20%	1,4	8,90E-03	1,32
Transport vacuolaire	4	2%	8,9	9,90E-03	1,61
Processus métabolique des ARNt	6	2%	3,6	2,50E-02	0,96
Ciblage des protéines	8	3%	2,6	3,20E-02	1,01
Régulation de la traduction	6	2%	3,1	4,30E-02	0,93
Mort cellulaire	17	7%	1,7	4,40E-02	0,96

Tableau III-5 : Appartenance aux groupes fonctionnels (termes GO uniquement) définis avec le logiciel DAVID des gènes dont les transcrits présentent des variations d'épissage après choc thermique. Il s'agit uniquement des groupes fonctionnels pour lesquels les p-values associées aux différents termes GO sont inférieures à 0,01. Le score d'enrichissement DAVID est attribué en fonction des p-values attribuées aux termes composant ce groupe. Plus il est élevé, plus l'enrichissement est significatif.

Afin d'obtenir plus d'informations sur les fonctions affectées par des variations d'épissage alternatif, nous avons déterminé grâce au logiciel DAVID des groupes fonctionnels incluant également les termes GO concernant les fonctions moléculaires, les mots clés associés aux protéines des bases de données INTERPRO ou Pir Superfamily, ainsi que les voies de régulations (bases de données KEGG, BIOCARTA, REACTOME). Au total, 25 groupes fonctionnels ont été définis par cette approche et les p-values associées à chaque terme ont été calculées, qui sont présentés dans l'annexe 4 à la fin de cette partie. En retenant uniquement les termes dont les p-values étaient inférieures à 0,05, le premier groupe fonctionnel présentant le score statistique DAVID le plus élevé (3,24) concernait les propriétés de fixation des protéines à l'ARN et le processus de maturation des ARNm. Avec un score statistique DAVID toutefois moins élevé (de 1,44) nous retrouvions aussi un groupe fonctionnel particulièrement intéressant pour notre étude, puisqu'il concernait la régulation de l'épissage et la formation du spliceosome. Un certain nombre d'autres groupes fonctionnels de gènes sont enrichis en gènes dont les transcrits présentent une variation d'épissage après choc thermique. C'est par exemple le cas du groupe fonctionnel correspondant au cycle cellulaire, dont le déroulement est perturbé par le stress thermique, mais également des groupes fonctionnels correspondant au transport des protéines et au transport vacuolaire, pour lesquels les effets du stress thermique ont été peu étudiés. Au sein de ces groupes fonctionnels, nous nous sommes intéressés en priorité aux protéines ayant une implication dans la maturation des ARNm, en particulier dans l'épissage, qui constitue une spécialité de notre équipe. Toutefois, il pourra être intéressant d'étudier plus en détail les gènes de certains groupes fonctionnels, comme le transport des protéines et les processus cataboliques des protéines. En effet, en se basant sur notre analyse, il est raisonnable de penser que l'épissage alternatif constitue un des moyens de la cellule pour adapter son métabolisme à la protéotoxicité engendrée par le stress thermique.

Une première étape sur le plan expérimental consistait à valider les données de puces obtenues par RT-PCR.

III.4.2. Confirmation par RT-PCR des événements d'épissage alternatif de type exon cassette détectés après un choc thermique

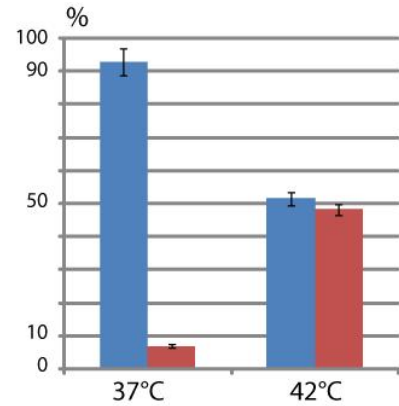
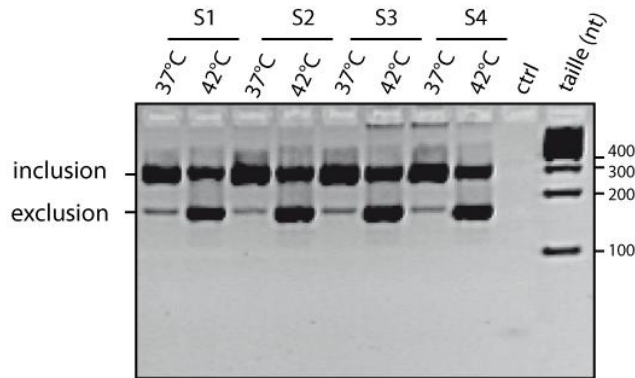
L'analyse des données de puces à ADN implique un traitement statistique d'un nombre conséquent de données soumis à un certain nombre de biais et d'erreurs. Ce traitement est encore plus complexe pour l'analyse des données de puces à jonctions exon-exon, où le nombre de sondes et de paramètres augmente par rapport aux puces à ADN classiques et même aux puces à exons. C'est pourquoi il était nécessaire de confirmer expérimentalement les variations d'épissage alternatif détectées. Nous avons pour cela réalisé des expériences de RT-PCR semi-quantitative sur les ARN totaux utilisés pour l'hybridation sur les puces, mais également sur d'autres préparations d'ARN totaux extraits à partir de cellules HeLa soumises ou non à un choc thermique dans les mêmes conditions. Les ARNm de ces

extraits ont d'abord été rétrotranscrits à l'aide d'un oligonucléotide polydT s'hybridant au niveau de leur queue polyA. L'étape de PCR a ensuite été effectuée en utilisant des oligonucléotides s'hybridant au niveau des exons flanquants l'exon cassette considéré et permettant ainsi la détection des différentes isoformes d'ARNm au sein de la même amplification. Les produits PCR obtenus ont été séparés par électrophorèse en gel d'agarose, puis quantifiés grâce au logiciel SynGene. Ces expériences étant répétées au minimum 3 fois pour chaque événement d'épissage, les signaux quantifiés ont pu être moyennés et une erreur type a pu être calculée.

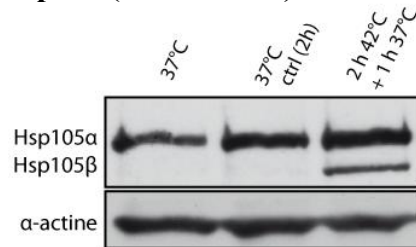
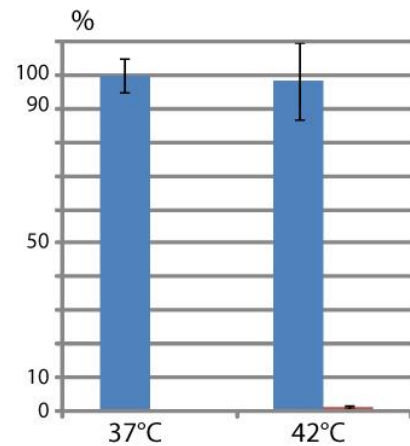
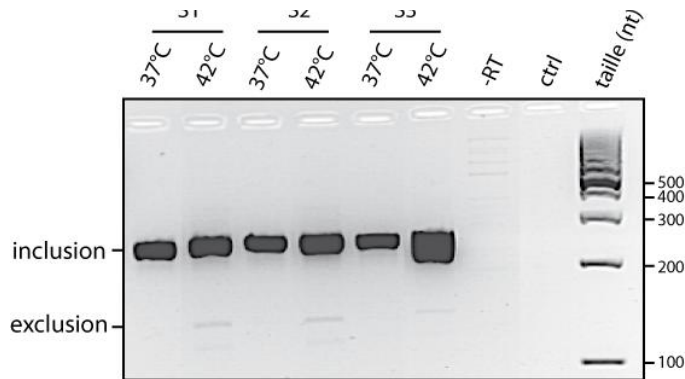
III.4.2.a. Confirmation de la présence des d'événements d'épissage alternatif décrits comme étant associés au stress thermique

Certaines des variations d'épissage alternatif détectées sur nos puces avaient déjà été décrites comme apparaissant après un stress thermique. Ces événements d'épissage alternatif étaient pour nous des contrôles positifs de l'efficacité du choc thermique appliqué aux cellules, nous avons donc quantifié ces événements par RT-PCR semi-quantitative. Un de ces événements d'épissage concerne l'exclusion de l'exon 12 de l'ARNm issu du gène *HSPH1*, qui code la protéine hsp105/110 kDa. Cet ARN est produit principalement sous la forme de deux isoformes α et β conduisant à la production des protéines hsp105 α et hsp105 β . L'isoforme α est une chaperonne cytoplasmique produite constitutivement. L'isoforme hsp105 β est produite en cas de choc thermique modéré, par exclusion de l'exon 12, et est localisée exclusivement dans le noyau, où elle est notamment capable d'activer la transcription du gène de hsp70 (Saito et al., 2009b; Yamagishi et al., 2009b). Nous avons bien observé par RT-PCR, réalisée avec des oligonucléotides spécifiques des exons flanquant l'exon 12, une augmentation de l'exclusion de cet exon après le choc thermique (figure III-3, A.). Le taux d'exclusion, faible au départ (7%), atteint environ 50% après le choc. En accord avec cette observation, l'analyse par western blot d'extraits protéiques issus des mêmes cellules a permis de montrer l'apparition de la forme protéique β après le choc thermique (figure III-3, B.).

Une autre variation d'épissage des ARNm HSPH1 avait été détectée sur nos puces, mais n'avait pas encore été décrite comme associée au stress thermique. Elle correspond à une augmentation de l'exclusion de l'exon 5. Nous avons bien détecté la présence d'une isoforme dépourvue de l'exon 5 après le choc thermique par RT-PCR, mais sa quantité était très faible (moins de 2% des produits d'épissage quantifiés, malgré la saturation des signaux de la bande correspondant à l'isoforme dans lequel l'exon 5 est inclus, ce qui perturbe la quantification) (figure III-3, C.).

A. ARNm hsp105 exon 12/18 (RT-PCR)

% inclusion	93	51,7
% exclusion	7	48,3

B. Protéine hsp105 (western blot)**C. ARNm hsp105 exon 5/18 (RT-PCR)**

% inclusion	100	98,5
% exclusion	0	1,5

Figure III-3 : Confirmation de l'augmentation de l'exclusion de l'exon 12 au sein des ARNm hsp105/hsp110 (HSPH1) après choc thermique. **A.** Des analyses par RT-PCR ont été réalisées sur 4 échantillons biologiques différents (S1 à S4) pour comparer les ARNm hsp105 extraits à 37°C et après 2 h de choc thermique à 42°C suivi d'1 h de régénération à 37°C. La réaction de PCR a été réalisée grâce à des oligonucléotides s'hybridant sur les exons flanquant l'exon 12. Un témoin négatif (ctrl) a été réalisé en absence d'ADN polymérase. **B.** Exemple de western blot réalisé sur les extraits protéiques de cellules non stressées (37°C) ou de cellules soumises au choc thermique dans les mêmes conditions que dans la partie A. Une expérience témoin a également été réalisée en incubant les cellules au bain marie dans les mêmes conditions que pour le choc thermique mais à une température de 37°C (37°C ctrl). **C.** RT-PCR réalisée dans des conditions similaires au panneau A, avec des

oligonucléotides s'hybridant aux exons flanquant l'exon 5. Une expérience témoin a été réalisée en absence de transcriptase inverse (-RT).

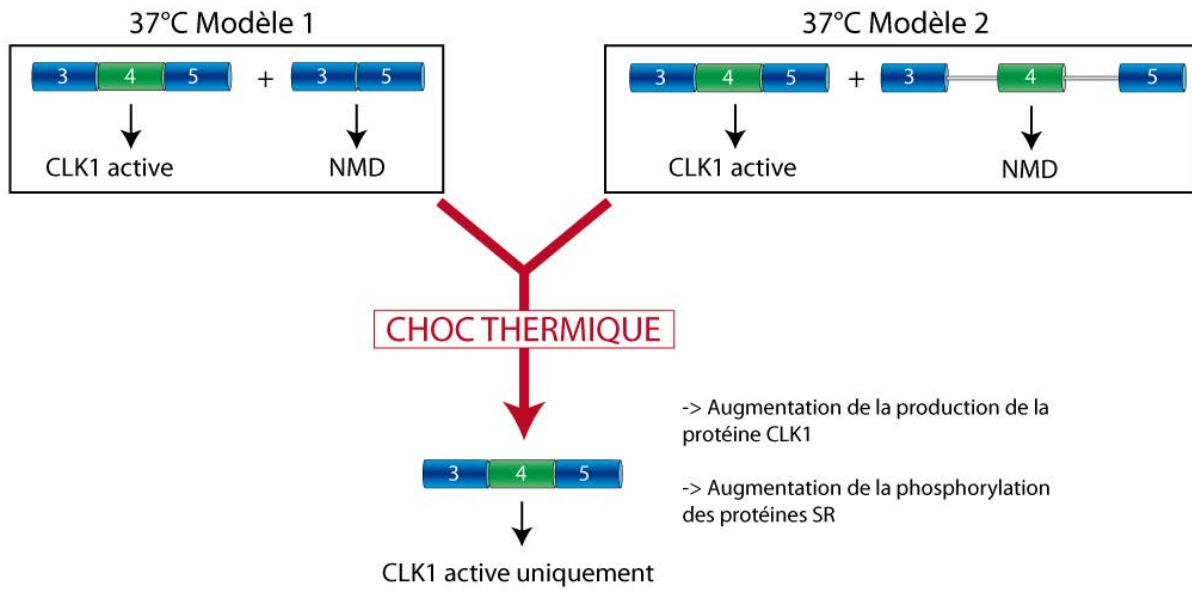
Selon la base de données Ensembl (<http://www.ensembl.org/index.html>), cette isoforme de l'ARNm hsp105 qui est dépourvu de l'exon 4 pourrait être dégradée par le NMD, du fait de la présence d'un codon stop prématuré, ce qui pourrait expliquer la faible quantité observée. La production de cette isoforme pourrait marquer le début de la réduction de la quantité de Hsp105 lors de la régénération des cellules après le choc thermique.

Il est déjà connu que l'exclusion de l'exon 4 de l'ARNm CLK1 provoque la dégradation de cet ARNm par le système NMD du fait de l'inclusion de codons stop prématurés (Hillman et al., 2004b). Une étude a montré qu'une partie des ARNm CLK1 était dégradée par le NMD à 37°C dans les cellules de plusieurs tissus, parce que l'exon 4 et les introns adjacents étaient retenus au sein de l'ARNm. Un choc thermique d'une heure à 43°C provoquait l'élimination rapide (30 min environ) de ce phénomène de rétention d'introns contenant des codons stop prématurés et induisait une production accrue de la protéine CLK1 ainsi que la rephosphorylation des protéines SR dès la fin du choc thermique pendant la période de régénération (Ninomiya et al., 2011).

Nous avons donc réalisé des expériences de RT-PCR pour valider l'exclusion de l'exon 4 au sein des ARNm CLK1 (ou CLK/STY) après choc thermique, observée d'après les données de nos puces (figure III-4, A.). Nos expériences de RT-PCR ont montré en effet qu'à 37°C, environ 20 à 25% des isoformes observées ne contiennent pas l'exon 4 et seraient donc dégradées. Après un choc thermique, la proportion d'ARNm ne contenant pas l'exon 4 devient négligeable, tandis que la quantité d'ARNm contenant cet exon augmente sensiblement, ce qui doit diminuer la dégradation de l'ARNm CLK1 et conduire à une production accrue de protéine CLK1. Les protéines SR, en particulier SRSF4 (SRp75) et SRSF10 (SRp38), qui sont les principales cibles de CLK1, sont rapidement déphosphorylées après le début du choc thermique. L'augmentation de la production de CLK1 pendant la phase de régénération à 37°C pourrait donc contribuer à la rephosphorylation de ces protéines durant cette phase. Il a été proposé en particulier que la rephosphorylation de SRSF10 (SRp38) contribue à la restauration globale de l'épissage dans les cellules.

D'après nos analyses par puces nous détectons seulement une rétention de l'exon 4 durant la période de régénération suivant la période de choc thermique. Nous avons aussi recherché la présence de l'isoforme de l'ARNm CLK1 correspondant à la rétention des introns bordant l'exon 4 déjà décrite dans la littérature, mais à la différence de cette étude, nous n'avons pas réussi à la détecter même en réalisant des électrophorèses plus longue des ADNc (Ninomiya et al., 2011) (figure III-4, B.). Il s'agirait donc d'une régulation avec un résultat identique mais basée sur des événements d'épissage alternatif légèrement.

A.



B. ARNm CLK1 exon 4/13 (RT-PCR)

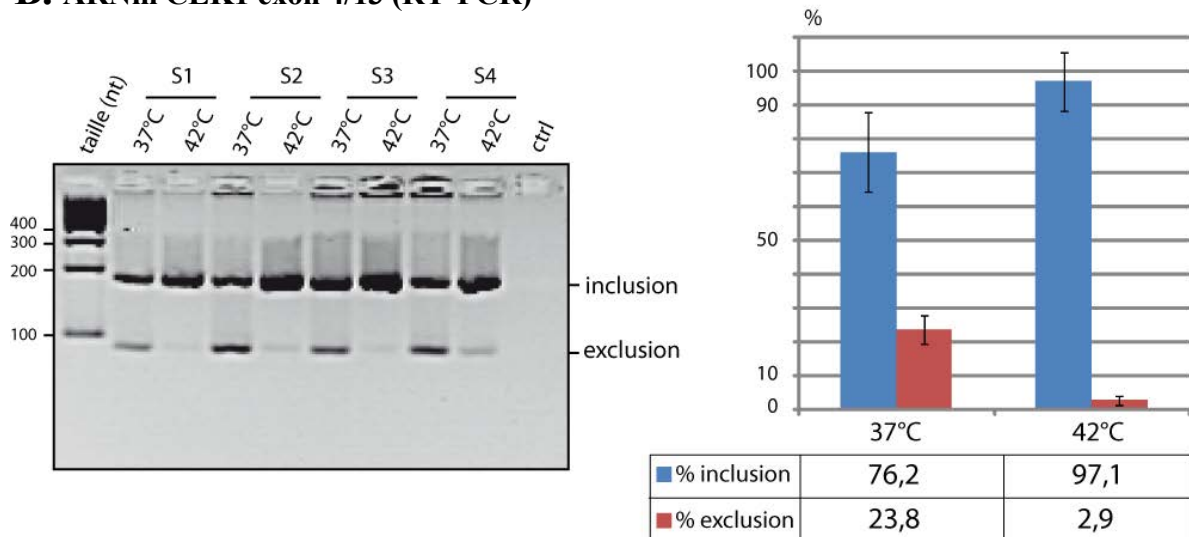


Figure III-4 : Confirmation de l'augmentation de l'inclusion de l'exon 4 au sein des ARNm CLK1 dans des cellules HeLa soumises à un choc thermique. A. Modèles proposés pour expliquer l'épissage alternatif de l'exon 4 de l'ARNm CLK1 et l'augmentation de la phosphorylation des protéines SR par CLK1 après choc thermique. Le modèle 2 est tiré de la publication de (Ninomiya et al., 2011). B. Les expériences de RT-PCR ont été réalisées sur 4 échantillons biologiques différents (S1 à S4) pour comparer les ARNm CLK1 extraits à 37°C ou après 2 h de choc thermique à 42°C puis une heure de régénération à 37°C. La PCR a été réalisée en utilisant des oligonucléotides s'hybridant sur les exons flanquant l'exon 12. Un témoin négatif (ctrl) a été obtenu en omettant l'ajout de l'ADN polymérase lors de la PCR.

III.4.2.b. Validation d'événements d'épissage encore non décrits comme associés à un choc thermique et concernant les ARNm de protéines impliquées dans le transport nucléocytoplasmique

Les protéines Ran (Ras-related small nuclear GTPase), CRM1 (ou encore XPO1 ou exportine 1) et TRN-SR (ou transportine SR ou transportine 3) sont impliquées dans le transport nucléocytoplasmique des ARNm et de certaines des protéines de leur métabolisme, et nous avons détecté grâce aux puces à jonctions exon-exon des variations d'épissage de leurs ARNm après le stress thermique. C'est pourquoi nous avons approfondi l'étude de l'épissage alternatif des ARNm Ran, CRM1 et TRN-SR.

Ces protéines s'associent dans différentes combinaisons pour assurer l'export et l'import des protéines vers le cytoplasme. L'export des protéines vers le cytoplasme est géré par les exportines, qui sont des récepteurs se fixant sur les protéines par le biais des séquences NES, tandis que l'import dans le noyau est géré par les importines, des récepteurs qui se fixent au niveau des séquences NLS ou d'autres séquences spécifiques. La protéine TRN-SR appartient à la famille des importines β nécessaire à l'export des protéines SR dans le cytoplasme (Kataoka et al., 1999). L'association des protéines SR avec le facteur TRN-SR est régulée par le degré de phosphorylation des protéines SR, qui module leur taux de localisation dans le noyau ou dans le cytoplasme (Sanford and Bruzik, 2001; Zhong et al., 2009). La protéine CRM1 assure l'export de très nombreuses protéines du noyau, en particulier des protéines suppresseurs de tumeurs dont hsp90 fait partie. Elle sert également au transport des UsnRNP vers le cytoplasme et des snoRNP vers le nucléole (pour revue, Mao and Yang, 2013; Turner et al., 2012; Verheggen and Bertrand, 2012). La protéine Ran est une GTPase majoritairement associée au GTP dans le noyau et au GDP dans le cytoplasme. Ce gradient nucléo-cytoplasmique de Ran•GTP joue un rôle majeur dans le transit des protéines cargo au travers des pores nucléaires (pour revue, Stewart, 2007). Ces protéines cargo sont importées dans le nucléoplasme en complexe avec un hétérodimère de transporteurs importines α/β . Une fois dans le noyau, Ran•GTP favorise le relargage de la protéine cargo (Imamoto et al., 1995). CRM1 forme un hétérotrimère dans le noyau avec une protéine cargo et Ran•GTP, permettant l'export vers le cytoplasme, où l'hydrolyse du GTP, aidée par les cofacteurs RanGAP et RanBP1, permet le relargage de la protéine cargo (Bischoff et al., 1995a, 1995b; Kutay et al., 1997). Le complexe importine-Ran•GDP est réimporté dans le noyau, où le GDP associé à Ran est remplacé par du GTP, grâce au facteur RCC1 (RanGEF, pour "Ran GTP Exchange Factor") et à un ratio GTP/GDP élevé dans le noyau (Bischoff and Ponstingl, 1991).

En lien direct avec notre étude, il a été montré que la formation des speckles et la localisation des protéines SR dans le noyau sont affectées après l'inhibition des facteurs de transport nucléaires, en particulier RAN, RCC1, RanBP2 qui est nécessaire à la localisation de RanPB1, ou encore TRN-SR (Saitoh et al., 2012). D'autre part, il a été observé que le stress thermique, le stress oxydant, le stress ultraviolet ou le stress hyperosmotique perturbent l'équilibre des concentrations respectives de Ran•GTP dans le noyau et de Ran•GDP dans le cytoplasme. La concentration de Ran•GDP dans le

cytoplasme augmente fortement du fait de la chute de la concentration intracellulaire en ATP ou GTP après l'exposition des cellules aux stress mentionnés ci-dessus (Kelley and Paschal, 2007; Yasuda et al., 2006b). Cela s'accompagne de la relocalisation d'autres facteurs impliqués dans le transport comme les importines α et β ou RanGAP dans le cas du stress hyperosmotique. Il est également intéressant de noter que la production de progérine dans les cellules HGPS est associée à la perturbation du gradient nucléocytoplasmique de Ran•GTP/Ran•GDP (Kelley et al., 2011).

L'analyse des résultats des puces à jonctions exon-exon révélait une augmentation de l'exclusion de l'exon 2 de l'ARNm de Ran. Nous montrons par RT-PCR une augmentation de environ 8% de l'exclusion de l'exon 2 après le choc thermique, et que cet exon est presque totalement inclus en absence de stress (figure III-5).

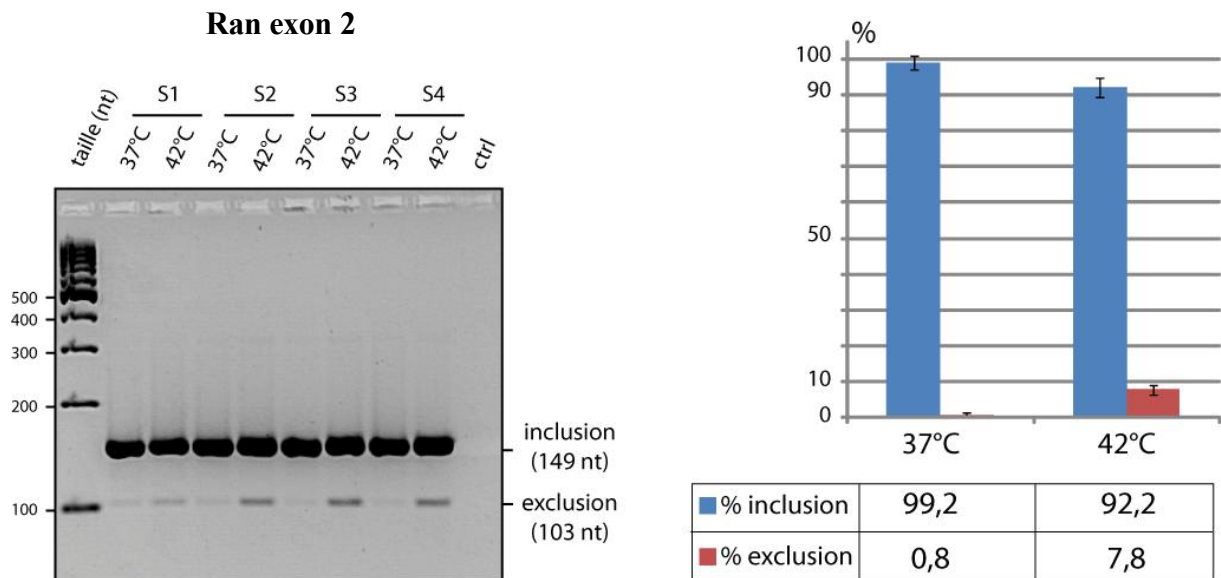


Figure III-5 : Confirmation de l'exclusion de l'exon 2 au sein de certains ARNm Ran après choc thermique. La RT-PCR a été réalisée sur 4 échantillons biologiques différents (S1 à S4) pour comparer les ARNm Ran extraits à 37°C ou après 2 h suivi d'1 h de choc thermique à 42°C puis une heure de régénération à 37°C. La PCR a été réalisée avec des oligonucléotides s'hybridant sur les exons flanquant l'exon 2. Un témoin négatif (ctrl) a été réalisé en absence d'ADN polymérase. La quantification des résultats est présentée à droite de la figure.

L'analyse des EST issues de divers tissus (foie, prostate, poumons, cerveau, carcinome) avait déjà révélé l'existence d'une ou plusieurs isoformes de l'ARNm Ran dépourvues de l'exon 2, (base de données AceView du NCBI : <http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/>). Etant donné que l'exon 2 contient le codon initiateur de la séquence codante, la délétion de cet exon implique l'utilisation d'un codon initiateur situé en aval, au niveau de l'exon 5, et la délétion des 88 premiers résidus sur les 216 que compte la protéine. L'isoforme protéique issue du transcrit dépourvu de l'exon 2 perd alors certains des résidus conservés impliqués dans l'interaction avec les cofacteurs RanGEF et RanGAP, et dans l'interaction avec le GTP (Becker et al., 1995; Renault et al., 2001; Seewald et al., 2002). Cela

entraînerait potentiellement la dérégulation du cycle catalytique de Ran, et pourrait participer à la perte du gradient nucléo-cytoplasmique de Ran•GTP.

En ce qui concerne l'épissage des pré-ARNm TRN-SR, nos expériences de RT-PCR ont confirmé l'augmentation de l'inclusion de l'exon 2 après un choc thermique détecté par l'analyse des puces à jonctions exon-exon. Cependant, cette augmentation n'était que d'environ 4%, ce qui soulève la question de sa signification biologique (figure III-6, A.). De plus, nous observions la présence d'un deuxième produit minoritaire d'une taille très proche, ainsi que d'un autre produit d'une taille inférieure à 100 nt, dont les identités sont inconnues. Les puces à jonctions exon-exon révélaient également l'augmentation de l'exclusion de l'exon 6 au sein des ARNm CRM1, mais nous n'avons pas pu confirmer ce résultat par RT-PCR. Seule l'isoforme résultant de l'inclusion de l'exon 6 était observée, sans que l'intensité relative de détection de cette isoforme ne varie avant ou après le choc thermique (figure III-6, B.). Un produit d'une taille très courte (<100 nt), aspécifique car présent également dans les pistes témoins, était également produit. Sa présence pouvait perturber la détection de l'isoforme résultant de l'exclusion de l'exon 6. Il faut noter que parmi toutes les expériences de RT-PCR que nous avons réalisées, c'est le seul cas où la RT-PCR a infirmé les résultats obtenus par emploi des puces à jonctions exon-exon.

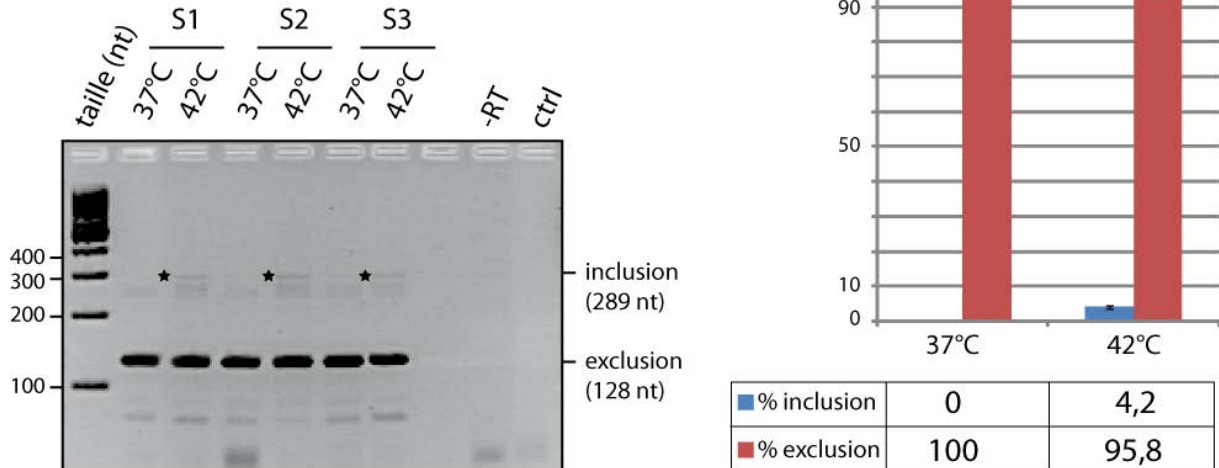
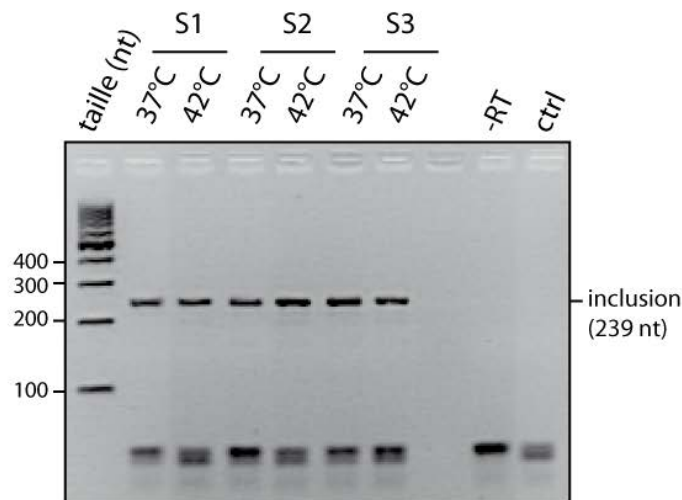
A. Transportine SR (TRN-SR) exon 2**B. Exportine 1 (CRM1), exon 6**

Figure III-6 : Légère augmentation de l'inclusion de l'exon 2 au sein des ARNm TRN-SR (A.) et invalidation de l'exclusion de l'exon 6 au sein des ARNm CRM1 (B.) suite à un choc thermique. **A.** Analyse par RT-PCR du taux d'inclusion de l'exon 2 des ARNm TRN-SR dans 3 échantillons biologiques différents (S1 à S3) préparés à partir de cellules soumises ou non à un choc thermique. La PCR a été réalisée grâce à des oligonucléotides s'hybridant sur les exons flanquant l'exon 2. Un témoin négatif (ctrl) a été réalisé en absence d'ADN polymérase de la PCR, ou en absence de la transcriptase inverse (-RT). **B.** Expérience réalisée dans les mêmes conditions que dans le panneau A., mais avec des oligonucléotides PCR s'hybridant au niveau des exons flanquant l'exon 6 de CRM1. Cette expérience n'a pas permis la détection d'un produit correspondant à l'isoforme résultant de l'exclusion de cet exon.

III.4.2.c. Validation de la variation de l'épissage du pré-ARNm Nop58

Nop58 est une protéine cœur des particules ribonucléoprotéiques snoRNP à boîtes C/D. Ces particules sont impliquées pour la plupart dans la méthylation de riboses au sein des ARN ribosomiques. Certaines de ces particules, telles que la snoRNP U3, sont impliquées dans les clivages nécessaires à la

maturation des ARN ribosomiques. Des défauts de méthylation des ARN ribosomiques ont été observés lors d'un stress thermique, cependant les méthylations concernées étaient catalysées par la méthyltransférase FtsJ et non par des snoRNA à boîtes C/D (Bügl et al., 2000). Les mutations de FtsJ provoquent par ailleurs l'apparition d'un phénotype thermosensible. De manière plus générale, il a été constaté des défauts de la biogenèse des ribosomes chez *E. coli* suite au choc thermique, qui sont accentués en absence des chaperonnes DnaJ et DnaK (Al Refaii and Alix, 2009). Les processus de maturation des ARN ribosomiques sont donc sensibles au choc thermique, et il est possible que des composants des snoRNP, tels que Nop58, soient importants pour la résistance des cellules à ce stress. C'est pourquoi nous avons cherché à confirmer par RT-PCR l'augmentation de l'exclusion de l'exon 6 au sein des ARNm Nop58 détectée par les puces à jonctions exon-exon après un choc thermique. Les expériences de RT-PCR réalisées nous ont permis de confirmer l'apparition d'un produit correspondant à une isoforme dépourvue de l'exon 6, représentant environ 13% des produits PCR obtenus (figure III-7). L'exclusion de cet exon provoque l'apparition de codons stop dans la phase codante portée par les exons en aval, et pourrait provoquer la dégradation de l'ARNm NOP58 par le système NMD, mais cela reste à montrer expérimentalement.

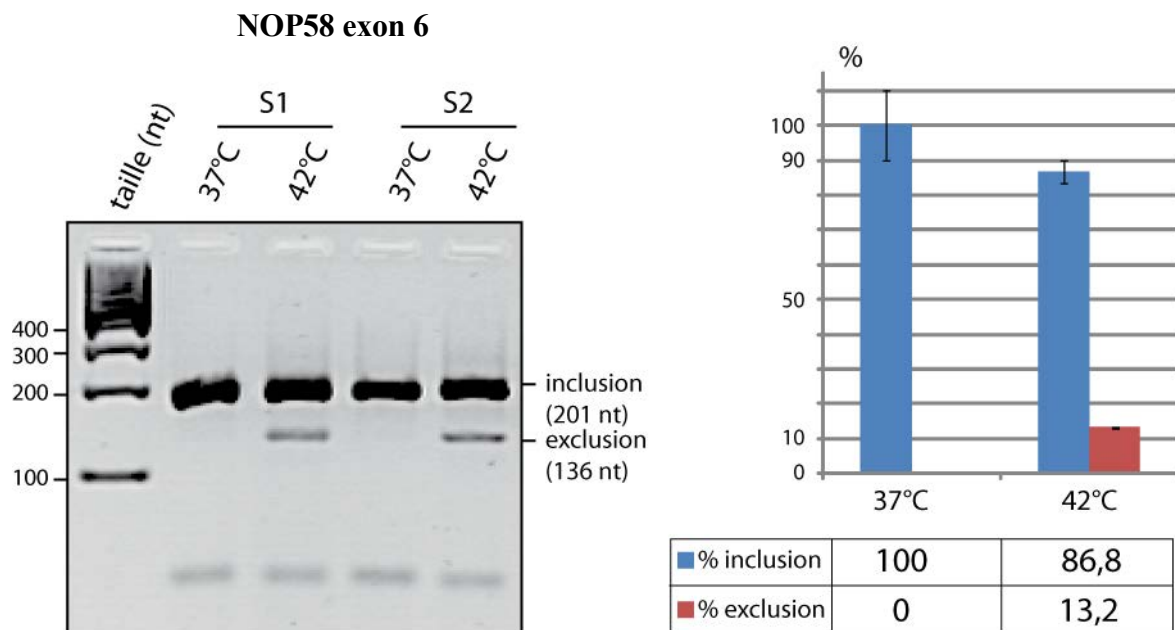


Figure III-7 : Confirmation de l'apparition d'une exclusion de l'exon 6 de l'ARNm de Nop58 par RT-PCR. L'expérience de RT-PCR a été réalisée sur 2 échantillons biologiques différents (S1 et S2) pour comparer les ARNm Nop58 extraits à 37°C ou après 2 h de choc thermique à 42°C puis une heure de régénération à 37°C, ceci avec des oligonucléotides s'hybridant sur les exons flanquant l'exon 6. Un témoin négatif (ctrl) a été obtenu en omettant l'ADN polymérase.

III.4.2.d. Validation de variations d'épissage touchant les pré-ARNm de protéines impliquées dans l'inhibition de la traduction et la dégradation des ARNm

Plusieurs variations d'épissage alternatif conséquentes au choc thermique détectées par les puces à jonctions exon-exon concernaient les ARNm de gènes impliqués dans la formation des granules de stress cytoplasmiques et la formation des P-bodies. Il s'agit des gènes *CNOT2*, *CNOT4*, *CNOT8*, *TNRC6A* (GW182), *EIF4ENIF1* (4E-T), *TIA1* et *EIF4G2*. Le paragraphe suivant décrit les liens fonctionnels qui existent entre les protéines codées par ces gènes.

III.4.2.d.1. Présentation des protéines impliquées dans la formation des P-bodies et des granules de stress cytoplasmiques

Il existe deux voies principales de dégradation des ARNm chez les eucaryotes. La première consiste en une dégradation dans le sens 3'→5' catalysée par un macrocomplexe protéique nommé exosome. La deuxième voie est initiée par le retrait de la queue polyA des ARNm par un complexe de déadénylation, qui peut aussi conduire à la dégradation 3'→5' de l'ARNm mais qui est plus généralement suivie par l'élimination de la coiffe des ARNm et leur dégradation par un complexe d'exoribonucléases, comprenant notamment les enzymes DCP1, DCP2, Xrn1, EDC3 et EDC4 (Parker and Song, 2004). Il existe deux complexes principaux de déadénylation, mais le complexe largement majoritaire chez les mammifères est nommé CCR4/POP2/NOT. L'activité déadénylase est portée principalement par un module formé de deux protéines de ce complexe, CNOT6/CCR4 (*CNOT6*) et CNOT7/CAF1 (*CNOT7*) mais aussi de façon moindre par CNOT8/POP2 (*CNOT8*). L'activité de CCR4 dépend de nombreux co-facteurs, nommés CNOT1 à CNOT5. Ces co-facteurs ont également des fonctions annexes et font le lien avec d'autres processus comme la transcription et l'adressage des ARNm (*CNOT3*), ainsi que le contrôle qualité des protéines (*CNOT4*) (pour revue, Collart et al., 2013). Les complexes CCR4/POP2/NOT, ainsi que les protéines du complexe Dcp1-Dcp2 sont par ailleurs indispensables pour la dégradation des ARNm *via* les microARN et sont aussi retrouvés accumulés dans les P-bodies, qui sont des régions du cytoplasme regroupant plusieurs dizaines de protéines identifiées comme étant impliquées dans l'inhibition de la traduction et la dégradation des ARNm (Behm-Ansmant et al., 2006; Eystathioy et al., 2002)(pour revue, Parker and Sheth, 2007b) (figure III-8). C'est en particulier le cas de la protéine GW182, issue du gène *TNRC6A*. Cette protéine est au centre du mécanisme d'inhibition de la traduction des ARNm. Elle est indispensable à la constitution des P-bodies, initialement nommés GW-bodies en référence à cette protéine (Eulalio et al., 2008, 2009a, 2009b; Yang et al., 2004). Le domaine C-terminal des protéines GW182 est dédié à l'interaction avec les protéines impliquées dans la déadénylation, en particulier le complexe CCR4/POP2/NOT, alors que le domaine N-ter est impliqué dans l'interaction avec le complexe RISC, constitué par les protéines Ago qui sont également présentes dans les P-bodies

(figure III-8)(Castilla-Llorente et al., 2012; Eulalio et al., 2008; Huntzinger et al., 2010, 2012). Les protéines Ago sont recrutées par les microARN qui interagissent avec leurs ARNm cibles, le plus souvent au niveau de la région 3'-UTR, et assurent le recrutement de GW182 ce qui permet l'inhibition de la traduction ou la dégradation des ARNm cibles, *via* le recrutement d'autres complexes, tels que les complexes de déadénylation et d'élimination de la coiffe évoqués ci-dessus (Hock et al., 2007)(pour revue, Huntzinger and Izaurralde, 2011).

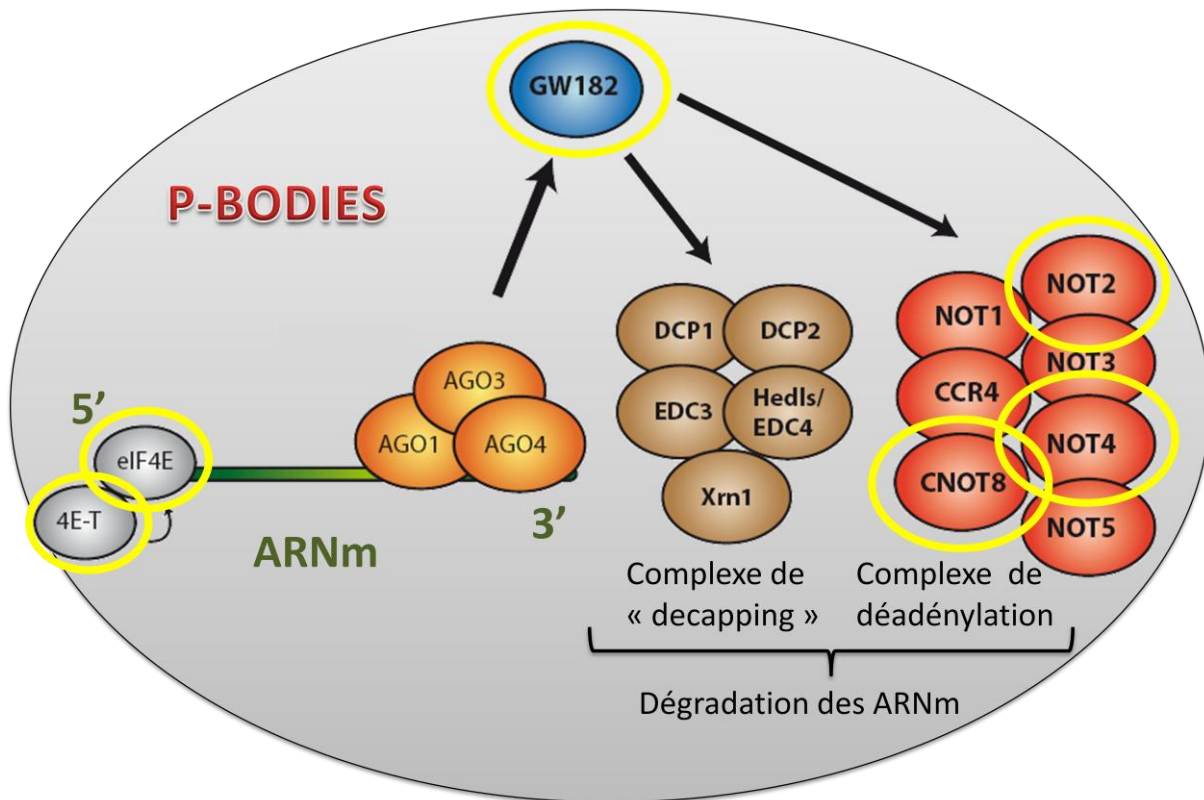


Figure III-8 : Protéines constituant les P-bodies pour lesquelles nous avons détecté des variations d'épissage alternatif de leurs pré-ARNm après un stress thermique par analyse des puces à jonctions exon-exon. Les protéines indiquées représentent une liste non exhaustive de protéines des P-bodies décrites dans le texte. Les protéines pour lesquelles une variation de l'épissage des pré-ARNm a été détectée suite à un stress thermique sont entourées en jaune.

Bien que les protéines qui constituent les P-bodies aient un rôle central dans le contrôle de l'expression des ARNm, le rôle et le mécanisme d'assemblage des P-bodies n'est pas encore connu (pour revue, Ding and Han, 2007). De plus, selon les types cellulaires, les P-bodies peuvent n'être présents que dans une fraction des cellules et non dans la totalité. Par ailleurs, le nombre de P-bodies peut augmenter et leur morphologie varier lors de l'exposition des cellules au stress oxydant (arsenite) et à la surexpression de plusieurs de leurs composants, comme DCP1a ou EDC4 (Kedersha and Anderson, 2007; Kedersha et al., 2005). La protéine 4E-T, pour laquelle les puces révèlent un événement d'épissage alternatif, est également présente et participe à l'assemblage des P-bodies (Cargnello et al., 2012;

Kedersha and Anderson, 2007). Elle fait partie d'un ensemble de protéines nécessaires l'une à l'autre pour leur recrutement dans les P-bodies, parmi lesquelles Lsm1, RCK/p54, et CCR4/CNOT6 (Andrei et al., 2005). Elle est responsable du transport nucléo-cytoplasmique de eIF4E, et sa surexpression cause une augmentation de la présence de eIF4E dans les P-bodies (Ferraiuolo et al., 2005) (figure III-9). Par ailleurs, eIF4E et 4E-T sont également accumulés dans les cSG, suite au blocage de l'initiation de la traduction (Andrei, 2005; Kedersha et al., 2005; Suzuki et al., 2009). Ceci n'est pas surprenant puisque comme mentionné dans l'introduction, il existe de nombreux liens entre les P-bodies et les granules de stress cytoplasmiques (cSG), avec notamment certains composants protéiques communs. L'accumulation d'eIF4E et 4E-T semble directement liée au mécanisme de réponse du choc thermique puisqu'elle est abolie par un inhibiteur de hsp90, la geldamycine (Suzuki et al., 2009). Enfin, lors de la formation des cSG, le complexe 48S* formé sur les ARNm dont la traduction est inhibée provoque l'inclusion et l'agrégation de la protéine TIA-1 dans les granules de stress cytoplasmiques. TIA-1 est une protéine impliquée dans la formation des P-bodies et des cSG qui présente une variation de l'épissage alternatif de ses ARNm dans les résultats des puces à jonctions exon-exon.

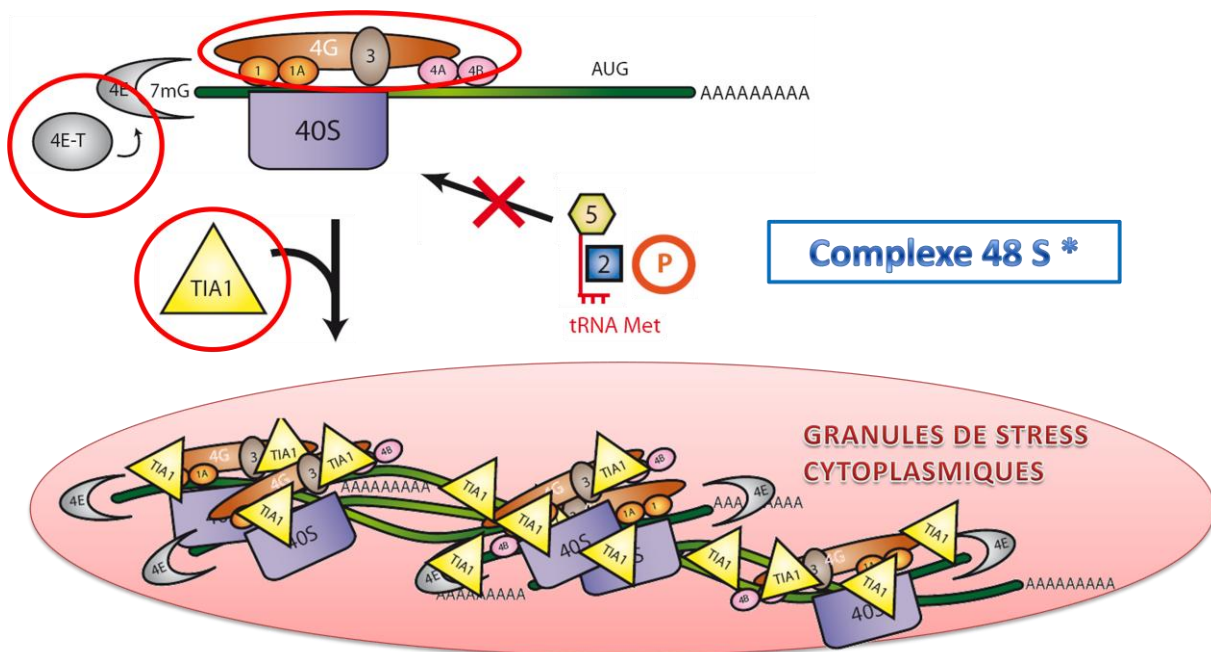
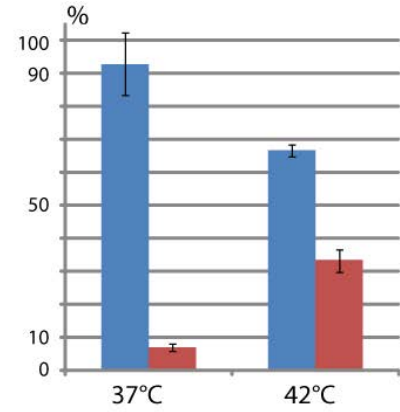
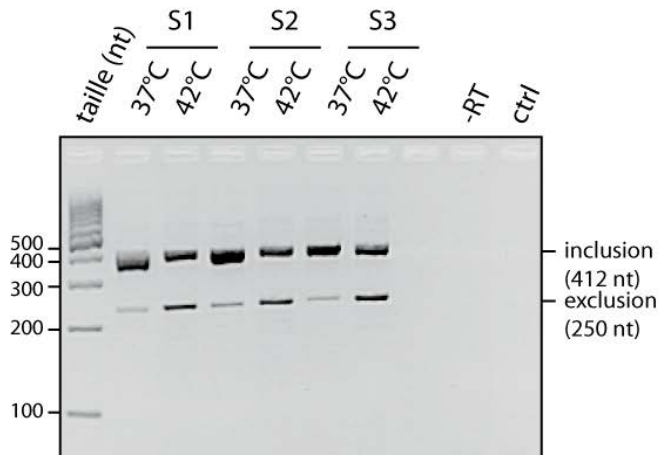


Figure III-9 : Protéines impliquées dans la formation des granules de stress cytoplasmiques pour lesquelles un événement d'épissage alternatif a été décelé dans les résultats des puces à jonctions exon-exon. Cette figure correspond à la figure présentée en introduction concernant la formation des cSG. La phosphorylation (P) de eIF2 par la protéine PKR lors d'un stress thermique arrête l'assemblage du ribosome et provoque la formation d'un complexe 48S inactif (48S*) et le recrutement de la protéine TIA-1. Les ARNm inhibés sont alors relocalisés dans les granules de stress cytoplasmiques en compagnie d'autres facteurs. Les protéines pour lesquelles une variation de l'épissage des ARNm a été détectée après un stress thermique sont entourées en rouge.

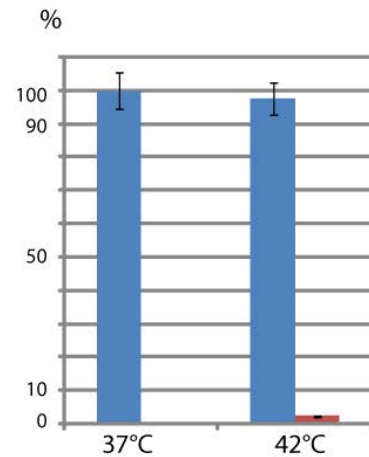
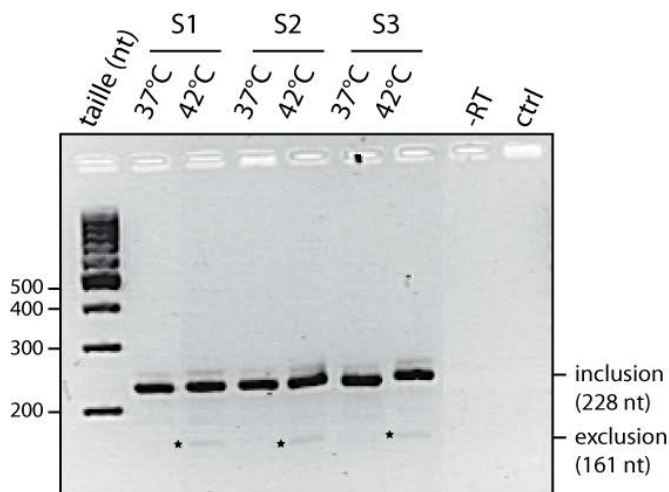
III.4.2.d.2. Validation des variations d'épissage des pré-ARNm des protéines du complexe de déadénylation par RT-PCR

Comme précédemment, nous avons réalisé des expériences de RT-PCR afin de valider les variations d'épissage détectées grâce aux puces. L'analyse des composants du complexe de déadénylation a bien révélé une augmentation d'environ 26% de l'exclusion de l'exon 4 de l'ARNm de CNOT8 (figure III-10). L'exclusion de cet exon conduit à la production d'une protéine de 238aa contre 292aa en présence de l'exon. Cet exon contient une partie du domaine conservé de la famille Caf1, présent également dans CNOT7, mais la littérature ne permet pas de prédire l'effet de la perte de cet exon. Lors du test réalisé sur les ARNm de CNOT2, nous avons également détecté l'apparition d'un produit PCR correspondant à l'exclusion de l'exon 4 après un choc thermique, mais en quantité quasi négligeable (augmentation de 2,4%) (figure III-10). Cependant, de façon similaire à certains cas précédemment décrits, l'exclusion de cet exon provoque l'apparition de codons stop prématurés, ce qui est susceptible de rediriger les isoformes d'ARNm dépourvues de cet exon vers la dégradation par le système NMD ce qui pourrait expliquer le faible taux de produit observés.

En ce qui concerne l'ARNm de la protéine CNOT4, les expériences de RT-PCR, réalisées avec plusieurs couples d'oligonucléotides différents, donnaient toutes de très nombreuses bandes aspécifiques, ce qui a rendu impossible la détection et la quantification des produits PCR correspondant aux isoformes attendues. Le design de nouveaux oligonucléotides de PCR est en cours.

A. CNOT8 exon 4

% inclusion	93	66,7
% exclusion	7	33,3

B. CNOT2 exon 4

% inclusion	100	97,6
% exclusion	0	2,4

Figure III-10 : Confirmation de l'augmentation de l'exclusion de l'exon 4 dans les ARNm CNOT8 et CNOT2. L'expérience de RT-PCR a été réalisée sur 3 échantillons biologiques différents (S1 à S3) pour comparer les ARNm de CNOT8 et CNOT2 extraits à 37°C ou après 2 h de choc thermique à 42°C, puis une heure de régénération à 37°C. L'expérience de PCR a été réalisée grâce à des oligonucléotides s'hybridant sur les exons flanquant les exons candidats. Des témoins négatifs (ctrl) ont été obtenus en omettant l'ADN polymérase ou la transcriptase inverse (-RT).

III.4.2.d.3. Validation des variations d'épissage des pré-ARNm des protéines des P-bodies 4E-T et GW182

Des expériences de RT-PCR réalisées sur les ARNm de 4E-T (*EIF4ENIF1*) ont permis de confirmer l'augmentation de l'exclusion de l'exon 11 indiquée par les puces à jonctions exon-exon. Cette

exclusion est de 4% en absence de stress et atteint un taux d'environ 30% après un stress thermique (figure III-11).

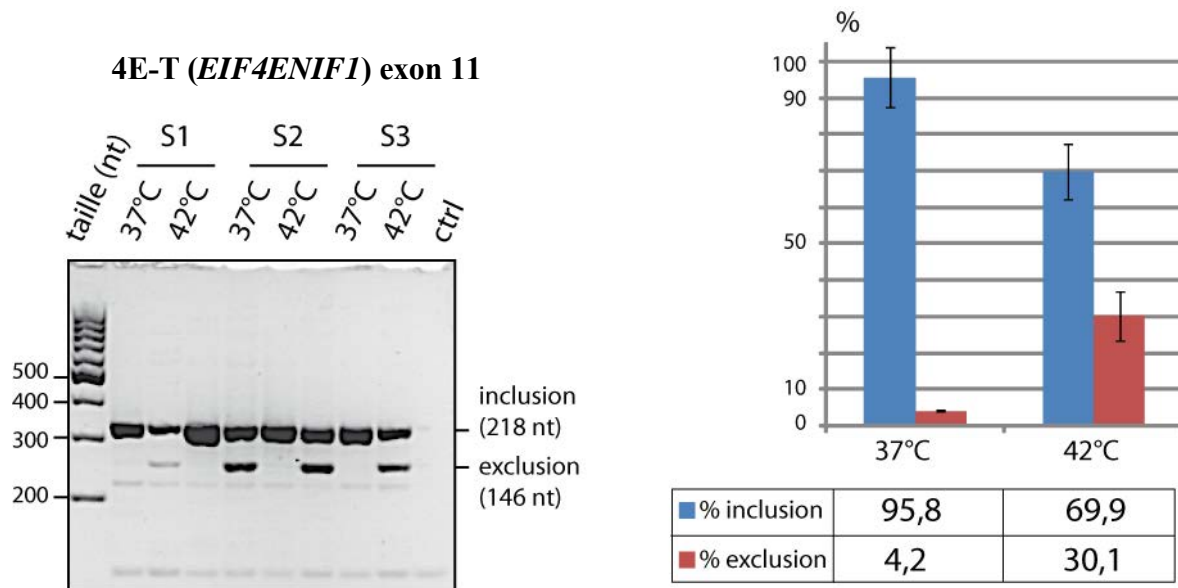


Figure III-11 : Confirmation de l'augmentation de l'exclusion de l'exon 11 dans l'ARNm 4E-T. Les expériences de RT-PCR ont été réalisées sur 3 échantillons biologiques différents (S1 à S3) pour comparer les ARNm de 4E-T extraits à 37°C ou après 2 h de choc thermique à 42°C puis une heure de régénération à 37°C. Les expériences de PCR ont été réalisées grâce à des oligonucléotides s'hybridant sur les exons flanquant l'exon 11. Un témoin négatif (ctrl) a été obtenu en omettant l'ADN polymérase.

Il a été montré qu'en présence d'un stress oxydant induit par l'arsénite, la protéine JNK, une MAPK induite par le stress, est relocalisée dans les P-bodies, où elle catalyse la phosphorylation de la protéine 4E-T au niveau de plusieurs résidus sérine. Or, comme il a été mentionné dans l'introduction, cette MAPK est également activée par phosphorylation lors d'un stress thermique. Cette phosphorylation est nécessaire à la relocalisation de 4E-T dans les P-bodies et est à l'origine d'une augmentation de la taille des P-bodies (Cargnello et al., 2012). De façon intéressante, l'un de ces résidus phosphorylable par JNK (Ser513) est situé au niveau de l'exon 11 épissé alternativement après un stress thermique. Il est donc possible que l'absence de ce résidu impacte la localisation de 4-ET et la formation des P-bodies. Nous avons réalisé des expériences préliminaires d'immunofluorescence grâce à des anticorps dirigés contre la protéine 4E-T endogène. Cet anticorps a permis la visualisation de P-bodies, d'un aspect semblable, que ce soit avant ou après le choc thermique (données non montrées).

Le mécanisme de recrutement de 4E-T dans les P-bodies n'est pas encore connu, mais il impliquerait l'interaction intermoléculaire entre des régions similaires communes à plusieurs protéines des P-bodies. Ce sont des régions protéiques enrichies en résidus Q et N, régulées par phosphorylation (Reijns et al., 2008). Ces régions sont la plupart du temps chevauchantes ou bien situées à proximité d'autres régions enrichies en résidus P. Ces dernières sont flexibles ce qui facilite les interactions

protéine-protéine. La séquence protéique codée par l'exon 11 des ARNm 4E-T correspond à l'une de ces régions, à la fois riche en résidus Q/N et en résidus P. Il est donc possible que cet exon soit important pour l'interaction de 4E-T avec ses partenaires, voire le recrutement de ces partenaires, au sein des P-bodies.

Les expériences de RT-PCR réalisées sur l'ARNm GW182 ont permis de confirmer l'augmentation de l'inclusion de l'exon 13, qui a été détectée lors de l'analyse des puces à jonctions exon-exon. Le produit PCR correspondant à l'inclusion de l'exon 13 n'était pas détecté à 37°C et représentait environ 31% des produits détectés après un choc thermique (figure III-12).

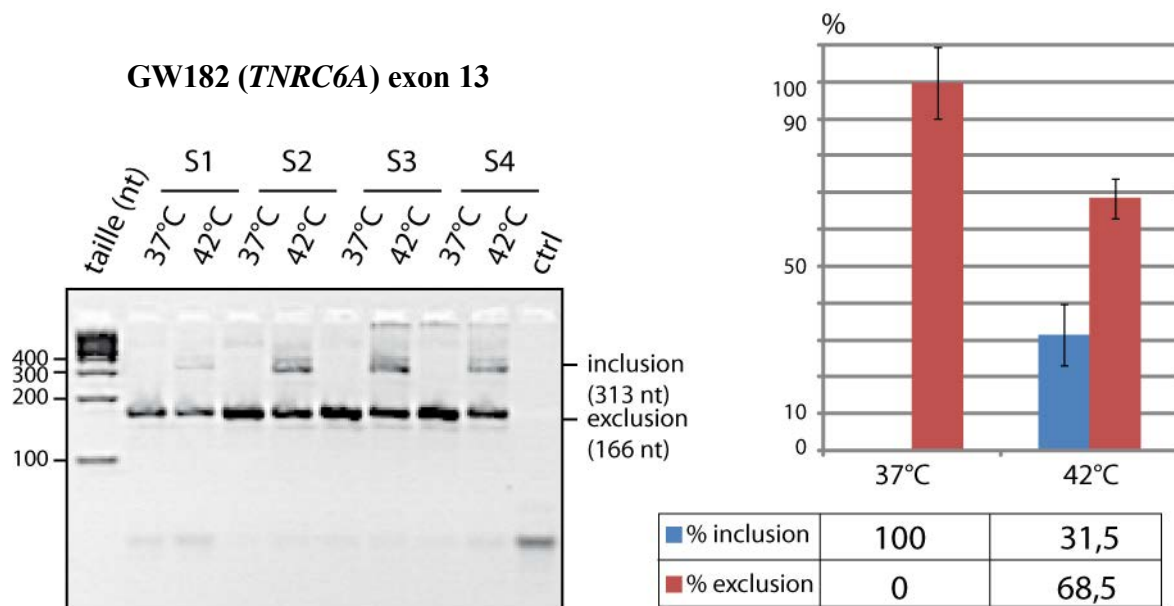


Figure III-12 : Confirmation de l'augmentation de l'inclusion de l'exon 13 dans les ARNm GW182. Les expériences de RT-PCR ont été réalisées sur 3 échantillons biologiques différents (S1 à S3) pour comparer les ARNm de GW182 extraits à 37°C ou après 2 h de choc thermique à 42°C puis une heure de régénération à 37°C. Les expériences de PCR ont été réalisées grâce à des oligonucléotides s'hybridant sur les exons flanquant l'exon 13. Un témoin négatif (ctrl) a été obtenu en omettant l'ADN polymérase.

Des études de structure-fonction menées sur la protéine GW182 ont permis d'établir que la région N-terminale de la protéine est impliquée dans l'interaction avec le complexe Argonaute, et que sa région C-ter est indispensable à l'inhibition des ARNm *via* les microARN et à l'interaction avec les protéines du complexe de déadénylation (figure III-13) (Eulalio et al., 2008, 2009b, 2009b). Deux domaines, nommés UBA-like et Q-rich sont situés entre ces deux régions. Ces domaines sont conservés dans les trois gènes paralogues codant GW182 (TNRC6A, TNRC6B et TNRC6C) et dans la protéine GW182 chez la drosophile, où ils sont impliqués dans la localisation de la protéine dans les P-bodies (Behm-Ansmant et al., 2006; Lazzaretti et al., 2009). Cependant, leur effet sur l'activité de GW182 est moins bien compris.

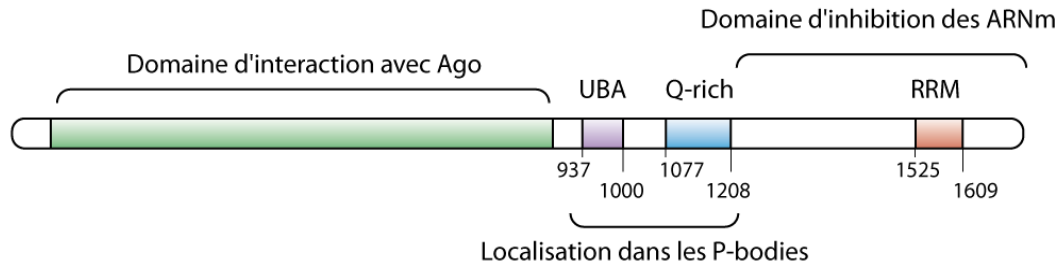
Isoforme protéique GW182 issue du gène *TNRC6A*

Figure III-13 : Les différents domaines de la protéine GW182 (d'après Eulalio et al., 2009c).

L'exon 13 de GW182 dont l'inclusion est augmentée suite à un choc thermique, code une séquence qui au niveau de la protéine GW182 est situé entre les domaines UBA et Q-rich, et n'appartient à aucun domaine fonctionnel connu. Il est à noter que dans les protéines GW182 initialement identifiées dans les cellules HeLa, la séquence codée par cet exon était incluse dans la protéine en absence de stress (Eystathioy et al., 2002). L'utilisation de cellules différentes (HeLa S3) lors de nos expériences de choc thermique pourrait expliquer l'absence de cet exon dans les ARNm GW182 des cellules non stressées. Dans le but d'élucider l'impact de l'inclusion ou l'exclusion de cet exon sur la fonction de la protéine GW182, nous avons réalisé plusieurs expériences. Pour réaliser ces travaux, j'ai été aidé par Christelle AIGUEPERSE et Florence OILLO-BLANLOEIL.

III.4.3. Etude de l'influence de la présence de l'exon 13 dans les ARNm GW182 sur les propriétés de la protéine GW182

III.4.3.a. Influence de l'inclusion de l'exon 13 sur la localisation de GW182

Afin de tester l'influence de la séquence codée par l'exon 13 sur la localisation de la protéine GW182, nous avons transfecté des cellules HeLa avec un plasmide pCIneo portant la séquence codant la protéine GW182 pourvue ou dépourvue de la séquence protéique correspondant à l'exon 13. Dans ce plasmide l'ORF GW182 était placée sous contrôle du promoteur CMV (du CytoMegalovirus) (voir Matériel et méthodes). De plus, pour différencier la protéine transfectée de la protéine endogène, une étiquette HA (séquence : YPYDVPDYA) a été ajoutée au N-terminal de la protéine GW182. Ainsi, l'utilisation d'un anticorps primaire dirigé contre ce peptide permet la localisation spécifique *in cellulo* de la protéine GW182 exprimée à partir du plasmide transfecté. Le plasmide pCIneo-GW182 porteur de la séquence GW182 comprenant l'exon 13 nous a été fourni par E. Izaurralde (Max Planck Institute, Tuebingen). La délétion de l'exon 13 au sein de la séquence de ce plasmide a été réalisée grâce à une adaptation de la technique de mutagenèse dirigée par PCR. Une réaction PCR a été réalisée à l'aide d'un

oligonucléotide antisens s'hybridant en amont de l'exon 13 et d'un oligonucléotide sens s'hybridant en aval de cet exon. Le produit linéaire obtenu correspond à l'intégralité du plasmide dépourvu de l'exon 13. La ligation des extrémités de ce produit a ensuite été réalisée grâce à l'ADN ligase du phage T4, et le produit de ligation a servi à transformer des bactéries compétentes (voir partie Matériel et méthodes). Lors des expériences d'immunofluorescence nous avons utilisé un anticorps anti-HA afin de détecter la protéine transfectée, mais également un anticorps dirigé contre la protéine GW182, afin de détecter à la fois les protéines GW182 endogène et transfectée.

La détection des protéines GW182 endogène ou exprimée à partir du plasmide à l'aide de l'anticorps anti-GW182 utilisé dans les cellules transfectées s'est avérée problématique à cause d'un fort bruit de fond, notamment au niveau du noyau, alors que GW182 est majoritairement cytoplasmique (figure III-14). Nous avons cependant pu constater que la transfection des cellules par le plasmide portant la séquence HA-GW182, incluant ou non l'exon 13, provoquait la formation de plusieurs dizaines de *foci* dans le cytoplasme de certaines cellules correspondant aux cellules transfectées. Ces *foci* étaient détectables avec un anticorps dirigé contre la protéine GW182 ou contre le peptide HA. Dans le cytoplasme des cellules non transfectées, aucun *foci* n'était détectable avec l'anticorps anti-HA, alors que plusieurs *foci* étaient détectables, malgré le bruit de fond, avec l'anticorps dirigé contre la protéine GW182.

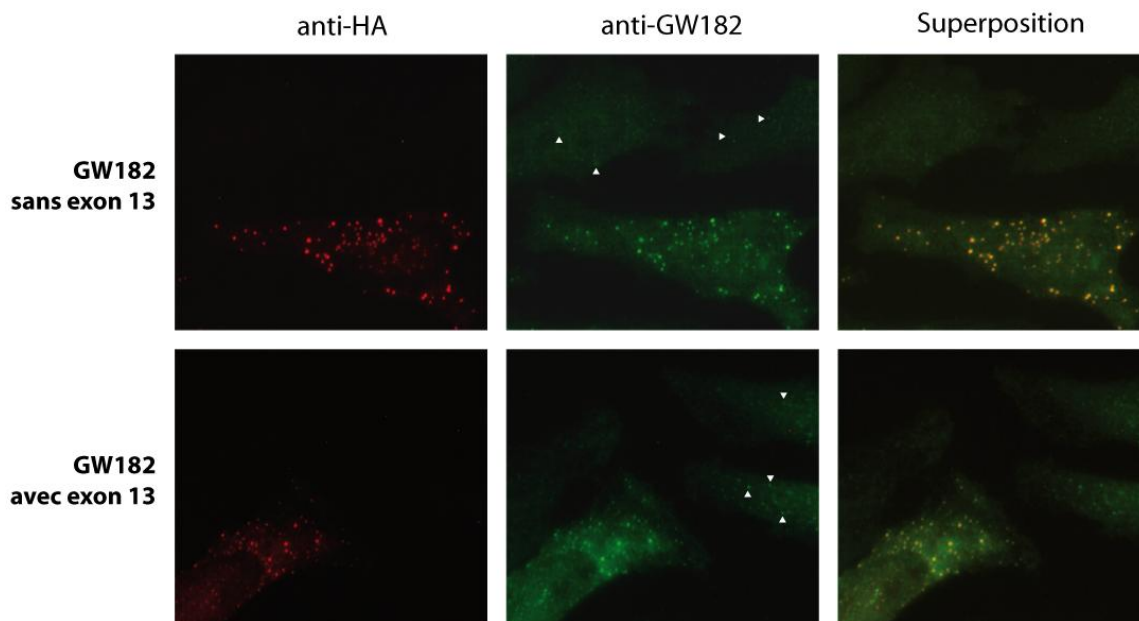


Figure III-14 : Localisation par microscopie à épifluorescence de la protéine GW182 endogène ou exprimée à partir d'un plasmide dans les cellules HeLa S3. Un anticorps dirigé contre le peptide HA permettait la détection de la protéine HA-GW182 exprimée à partir du plasmide pCneo. L'anticorps anti-GW182 employé permettait la détection à la fois de la protéine endogène et de la protéine surexprimée. Les pointes de flèches montrent les granules formés par GW182 endogène dans les cellules où la protéine HA-GW182 surexprimée est indétectable.

Afin de nous assurer que les *foci* détectés correspondaient bien aux P-bodies des expériences de co-localisation de la protéine HA-GW182 (contenant ou non l'exon 13) avec la protéine EDC4, qui est

un marqueur des P-bodies, ont été réalisées. Les résultats obtenus sont résumés dans la figure III-15. Trois cas de figures ont été rencontrés et sont rassemblés de façon fortuite sur la même image obtenue avec la protéine HA-GW182 n'incluant pas l'exon 13. Nous observons que dans de rares cas, les *foci* HA-GW182 observés sont tous co-localisés avec la protéine EDC4, qui outre sa localisation dans le noyau, forme également des *foci* cytoplasmiques nets et en nombre compatible avec le profil des P-bodies observé dans la littérature. Dans d'autre cas, plus communs, les *foci* HA-GW182 ne sont que partiellement co-localisés avec les *foci* EDC4. Enfin, dans un dernier cas, les cellules présentent toujours des *foci* HA-GW182 mais aucun *foci* EDC4. Ces résultats tendent à montrer que la protéine HA-GW182 surexprimée est prioritairement localisée dans les P-bodies, mais forme également dans le cytoplasme des *foci* supplémentaires de nature inconnue. Toutefois, les résultats obtenus sont les mêmes que ce soit en présence ou non de la séquence correspondant à l'exon 13 dans la protéine GW182. Ils ne vont donc pas dans le sens d'un impact de l'inclusion de l'exon 13 sur la localisation de GW182 dans le cytoplasme.

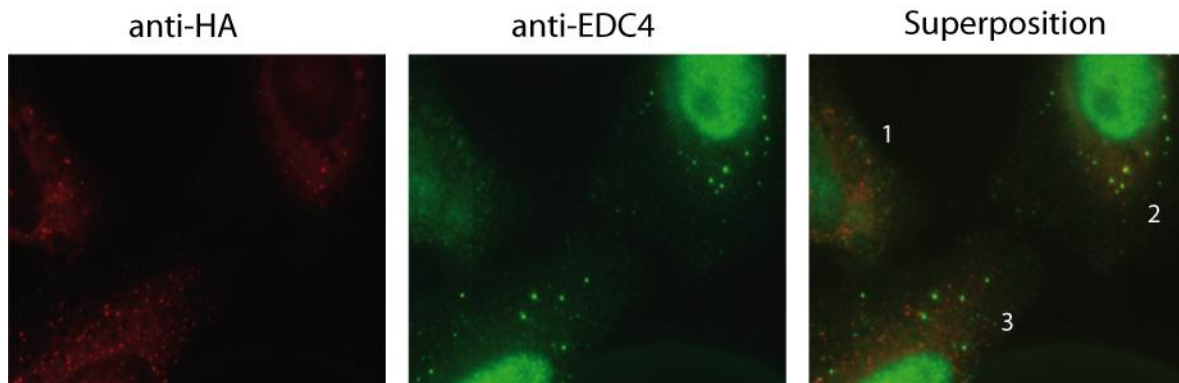


Figure III-15 : Co-localisation partielle entre HA-GW182 et les P-bodies détectés grâce à la protéine EDC4. Un anticorps dirigé contre le peptide HA permettait la détection de la protéine HA-GW182 exprimée à partir du plasmide pCIneo. La protéine EDC4 était utilisée comme marqueur des P-bodies. La superposition des deux images révèle 3 cas de figures différents, numérotés de 1 à 3. 1 : cellule dans laquelle les *foci* EDC4 sont absents. 2 : cellule dans laquelle les *foci* HA-GW182 sont parfaitement co-localisés avec les P-bodies. 3 : cellule dans laquelle les *foci* HA-GW182 ne sont que partiellement co-localisés avec les P-bodies.

III.4.3.b. Influence de la présence de la séquence codée par l'exon 13 sur les fonctions de GW182

Comme mentionné précédemment, la protéine GW182 est impliquée par sa région C-terminale dans la dégradation des ARNm *via* le recrutement du complexe de déadénylation CCR4/POP2/NOT et par sa région N-terminale dans une interaction avec le complexe RISC, constitué par les protéines Ago. Nous avons réalisé différentes expériences afin de tester l'impact de la séquence codée par l'exon 13 sur ces 2 fonctions. Dans un premier temps, des expériences d'ancrage ou « tethering » ont été réalisées. Le principe de ces expériences est de recruter artificiellement la protéine GW182 ou des variants de cette

protéine sur la séquence 3' UTR d'un ARNm cible et de tester la capacité des différentes protéines ainsi recrutées à induire la dégradation de cet ARNm. Ce système avait été utilisé par I. Behm-Ansmant lors de son stage post-doctoral dans le laboratoire d'E. Izaurralde, qui nous a fourni les constructions nécessaires à la réalisation de ces expériences. Le recrutement artificiel de GW182 sur l'ARNm cible est réalisé en utilisant le système λ N-Boîte B. Dans ce système, le peptide λ N du bactériophage lambda (22 aa) se lie spécifiquement à un motif ARN appelé Boîte B (structure tige-boucle de 19 nt) (Gehring et al., 2003). Cinq motifs Boîte B ont été insérés dans la région 3' UTR de l'ARNm cible. La protéine GW182 était exprimée sous la forme d'une protéine de fusion avec le peptide λ N, ce qui permettait son "ancrage" (tethering) au niveau de la région 3'UTR de l'ARNm cible. L'expression d'une protéine GW182 active devait provoquer la dégradation de l'ARNm cible. Cet ARNm cible correspondait à l'ARNm de la luciférase Renilla, qui est une protéine capable de produire de la lumière en présence d'ATP (bioluminescence) (figure III-16, A.). La mesure de l'activité de cette enzyme grâce à un luminomètre permettait d'estimer la quantité de protéine luciférase Renilla présente dans les cellules qui est directement proportionnelle à la quantité d'ARNm cible présent dans l'échantillon. Ceci permettait d'estimer les capacités de la protéine λ N-GW182 à induire la dégradation de l'ARNm cible.

Pour les besoins de cette expérience, les cellules HeLa S3 utilisées ont été co-transfectées avec plusieurs plasmides : le plasmide pCIneo-Rluc-5BoîteB portant la construction luciférase Renilla-5BoîteB, le plasmide pCIneo-HA- λ N-GW182, codant la protéine HA- λ N-GW182 avec ou sans exon 13 et le plasmide pEGFP-N3-FLuc porteur de la séquence codant la luciférase Firefly utilisée pour normaliser les données en fonction de l'efficacité de transfection. Des réactions contrôles ont été réalisées en co-transfectant un plasmide pCIneo-HA- λ N ou pCIneoHA- λ N-MBP codant la protéine HA- λ N seule ou HA- λ N-MBP, ceci afin de nous assurer que le plasmide pCIneo ne contenant pas la séquence codant la protéine GW182 n'avait aucun effet sur l'expression de la luciférase Renilla.

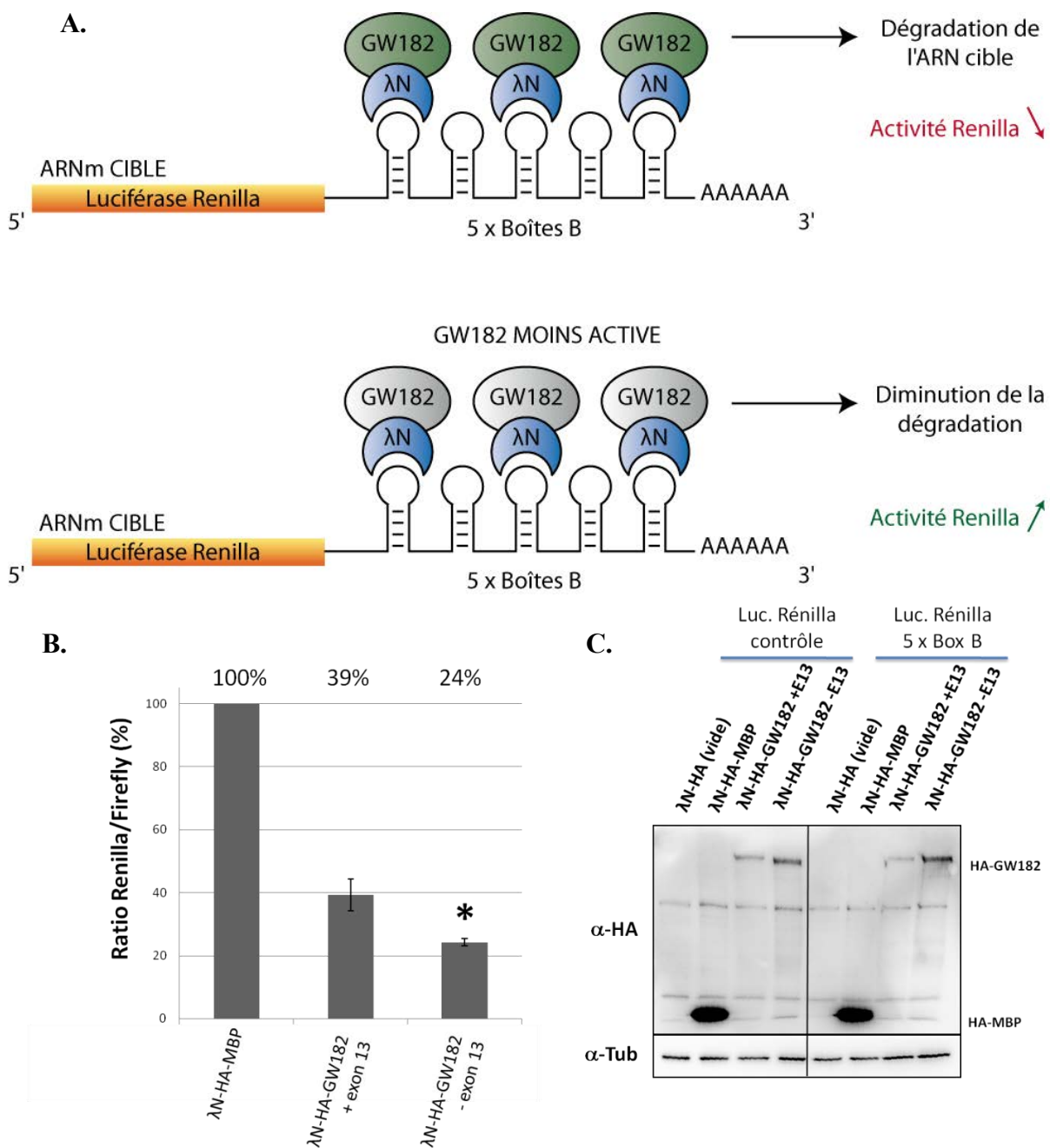


Figure III-16 : L'inclusion de l'exon 13 de GW182 provoque une diminution de l'activité inhibitrice de GW182 imputable à une diminution de sa synthèse. A. Principe de l'expérience de tethering utilisée pour tester l'activité de GW182. Le recrutement de la protéine de fusion λ N-GW182 dans la région 3'UTR de l'ARNm codant la luciférase Renilla a été réalisé grâce à l'ajout de 5 séquences nommées Boîte B, qui ont une très forte affinité pour le peptide λ N. Dans le cas où la protéine GW182 recrutée est moins active, il en résulte en une augmentation de l'activité de la luciférase Renilla. **B.** Mesure de l'activité de la luciférase Renilla synthétisée à partir de l'ARNm cible. L'efficacité de transfection des plasmides contenant l'ARN cible était normalisée par la mesure de l'activité de la luciférase Firefly dont l'ORF était incluse dans un second plasmide utilisé en co-transfection. La mesure était réalisée à une longueur d'onde différente. Les activités relatives des protéines λ N-HA-GW182 avec ou sans exon 13 ont été normalisées par rapport à l'activité obtenue en présence d'une protéine témoin, λ N-HA-MBP, établie arbitrairement à 100%. **C.** Evaluation par western blot du taux des protéines λ N-HA-GW182 avec (+E13) ou sans (-E13) exon 13 synthétisées *in cellulo* à partir des plasmides pCineo, ainsi que des protéines témoin λ N-HA et λ N-HA-MBP. Le taux de ces protéines a été évalué grâce à un anticorps dirigé contre le peptide HA et un anticorps anti-tubuline permet d'estimer la variation de la quantité déposée entre chaque piste.

L'activité luciférase Renilla mesurée en présence de la protéine GW182 a été comparée à celle mesurée en présence de la protéine HA- λ N-MBP, qui n'affecte pas l'activité luciférase, et dont la valeur a été établie à 100% (figure III-16, B.). L'activité Renilla mesurée en présence de la protéine GW182 sans exon 13 représentait ainsi environ 25% de l'activité mesurée en présence de la protéine HA- λ N-MBP, contre 40% en présence de la protéine GW182 avec exon 13. Cela signifie que la présence de la séquence codée par l'exon 13 diminue la capacité de GW182 à induire la dégradation de l'ARNm cible. Cette diminution, d'un facteur d'environ 1,6x, est significative selon le test statistique de Student, et très reproductible dans 3 expériences distinctes avec des échantillons biologiques différents. Des expériences de western blot ont été réalisées à partir des extraits protéiques issus des cellules transfectées afin de vérifier le taux d'expression de la protéine GW182 en présence ou en absence de l'exon 13 (figure III-16, C.). Les protéines HA- λ N-GW182 et HA- λ N-MBP ont été détectées grâce à l'emploi d'un anticorps dirigé contre le peptide HA, et la comparaison de la quantité de protéine déposée dans chaque piste a été réalisée par la détection de la tubuline. Ces expériences montrent que la protéine GW182 était présente en plus grande quantité dans les extraits lorsque l'exon 13 est absent. Notre observation de la plus grande efficacité de la protéine GW182 ne contenant pas l'exon 13 à induire la dégradation de l'ARNm cible pourrait donc n'être due qu'à la présence en plus grande quantité de cette protéine dans les extraits protéiques comparée à la protéine GW182 contenant l'exon 13. Ces données montrent donc que la présence ou non de la séquence codée par l'exon 13 ne semble pas affecter la capacité de GW182 à induire la dégradation d'un ARNm cible et suggèrent un effet possible de la présence de cette séquence sur soit la stabilité de l'ARNm GW182, soit la durée de vie de la protéine GW182.

III.5. Etude des microARN régulés après un choc thermique par analyse haut débit

Puisque nous observons des variations au niveau des protéines des P-bodies, et que les P-bodies sont liés à la dégradation des ARNm *via* les microARN (ou miR), il était intéressant d'utiliser une approche à haut débit pour étudier la variation de l'expression des microARN en réponse au stress thermique. Cette étude a pu être initiée grâce à notre collaboration avec l'équipe de L. Vallar à Luxembourg. L'expression des microARN a été analysée par emploi des puces GeneChip® miRNA 3.0 commercialisées par la société Affymetrix. Ces puces permettent la détection des microARN de la version 17 de la base de données miRbase, soit 1733 microARN, mais également de 2216 snoRNA et scaRNA, et 1658 pre-microARN humains. Le protocole utilisé pour l'hybridation et l'analyse des puces est détaillé dans la partie Matériel et méthodes. Nous avons utilisé les mêmes préparations d'ARN totaux que pour les puces à jonctions exon-exon. Les ARN étaient néanmoins traités avec un kit de purification permettant de conserver les ARN de petite taille dans le mélange d'ARN purifié. Nous avons comparé 3

échantillons issus de cellules non stressées à 3 échantillons de cellules soumises à un choc thermique de 2 h à 42°C suivi d'une période de régénération d'1 h à 37°C.

L'analyse statistique des résultats des puces par un test paramétrique (ANOVA) n'a pas permis de révéler de variation de taux des microARN après le stress. En effet, les variations entre les 3 réplicats biologiques étaient supérieures aux variations dues au choc thermique. Le traitement statistique des données par une méthode non paramétrée, telle que la méthode dite de « rank products », a toutefois permis d'identifier plusieurs microARN dont l'expression était induite ou réprimée de façon significative après le choc thermique (tableau III-6).

microARN	facteur d'induction	RP.PFP
hsa-miR-3127-5p_st	11,64	0
hsa-miR-3198_st	8,90	0
hsa-miR-4793-3p_st	7,93	0
hsa-miR-1972_st	6,45	0
hsa-miR-4440_st	5,97	0
hsa-miR-1275_st	5,56	0
hsa-miR-4448_st	5,17	0
hsa-miR-92a-1-star_st	-4,47	0
hsa-miR-4462_st	3,60	5,00E-04
hsa-miR-4485_st	3,40	1,11E-03
hsa-miR-4458_st	3,20	1,60E-03
hsa-miR-27b-star_st	-2,87	4,00E-03
hsa-miR-665_st	2,88	4,17E-03
hsa-miR-346_st	2,95	4,18E-03
hsa-miR-2276_st	2,52	1,37E-02
hsa-miR-29b-1-star_st	2,32	1,93E-02
hsa-miR-4539_st	2,16	2,24E-02
hsa-miR-212_st	2,38	2,90E-02
hsa-miR-1244_st	-2,17	3,13E-02
hsa-miR-4785_st	1,97	3,26E-02
hp_hsa-mir-4485_st	2,02	3,35E-02

Tableau III-6 : Liste des microARN induits ou réprimés d'après l'exploitation par test statistique "rank product" des données de puces dédiées aux petits ARN. La nomenclature utilisée pour les microARN est celle de la base de données miRbase. Les facteurs d'induction indiqués sont les moyennes des variations de signal observées sur les micropuces à partir des analyses faites sur 3 échantillons biologiques différents. La colonne RP.PFP indique la valeur de probabilité obtenue pour le test statistique rank products (<0,05). Plus elle est proche de 0, plus la variation observée est significative.

La validation expérimentale par RT-qPCR des variations d'expression des 12 premiers microARN du tableau III-9 ayant les meilleurs scores statistiques a été effectuée en utilisant des sondes Taqman selon le protocole décrit dans la partie Matériel et méthodes. Les oligonucléotides Taqman ont la particularité d'être marqués à leur extrémité 5' par un fluorochrome émetteur ("reporter") et à leur extrémité 3' par un fluorochrome suppresseur ("quencher") qui inhibe l'émission du reporter lorsqu'ils sont situés à proximité l'un de l'autre. Au cours de la PCR, si l'oligonucléotide Taqman est hybridé sur sa cible, il est hydrolysé par la Taq polymérase qui possède une activité 5'-exonucléasique. Le reporter est ainsi séparé du quencher et émet un signal, proportionnel au nombre d'oligonucléotides hydrolysés. Les RT-qPCR ont été effectuées sur les extraits d'ARN utilisés pour réaliser les puces. Un échantillon supplémentaire, correspondant à une incubation de 2 h dans un bain marie à 37°C, a été utilisé comme témoin négatif. Il permettait de tester l'effet de l'immersion seule sur l'expression des microARN. La normalisation des signaux obtenus a été réalisée en utilisant un oligonucléotide Taqman spécifique du snRNA U6, dont le signal ne varie pas après un choc thermique.

Des tests préliminaires ont été réalisés pour les 12 microARN du tableau III-9 possédant les meilleurs scores statistiques. Ces expériences n'ont pas été répétées et devront encore être validées. Pour les 12 microARN analysés par RT-PCR nous avons pu observer 4 cas de figures :

- 1- Les microARN ne sont pas détectés avec un signal satisfaisant (c'est à dire des Ct compris entre 25 et 35) pour pouvoir interpréter, (miR 3198, 4793_3p et 4462)
- 2- Les microARN sont détectés mais aucune variation significative de leur expression n'est observée (miR 92a_1*, 4440 et 3127_5p)
- 3- Les microARN sont détectés et une variation significative de leur expression est observée après le choc thermique, et également après une immersion de 2 h à 37°C (miR 27b*, 4448 et 4458) (figure III-17, A.).
- 4- Les microARN sont détectés et une variation significative de leur expression est observée spécifiquement après le choc thermique et pas après immersion à 37°C (miR 1972, 1275 et 4485) (figure III-17, B.). Pour les miR 1972 et 1275 l'analyse a été étendue à une gamme plus large de conditions correspondant à 1 ou 2 h de choc thermique à 42°C, suivi d'une période de régénération à 37°C allant de 1 h à 4 h. Dans ce cas l'induction observée était proportionnelle à la durée du choc thermique, et pour une même durée de choc déterminée l'induction était maximale après 1 h de régénération à 37°C pour le miR 1972 et 2 h de régénération pour le miR 1275. Cette induction est particulièrement importante pour le miR 1972 (environ 15x) et peut tout de même atteindre 3x pour le miR 1275, indiquant que ces deux microARN ont probablement un rôle de régulation durant la période de régénération après un stress thermique.

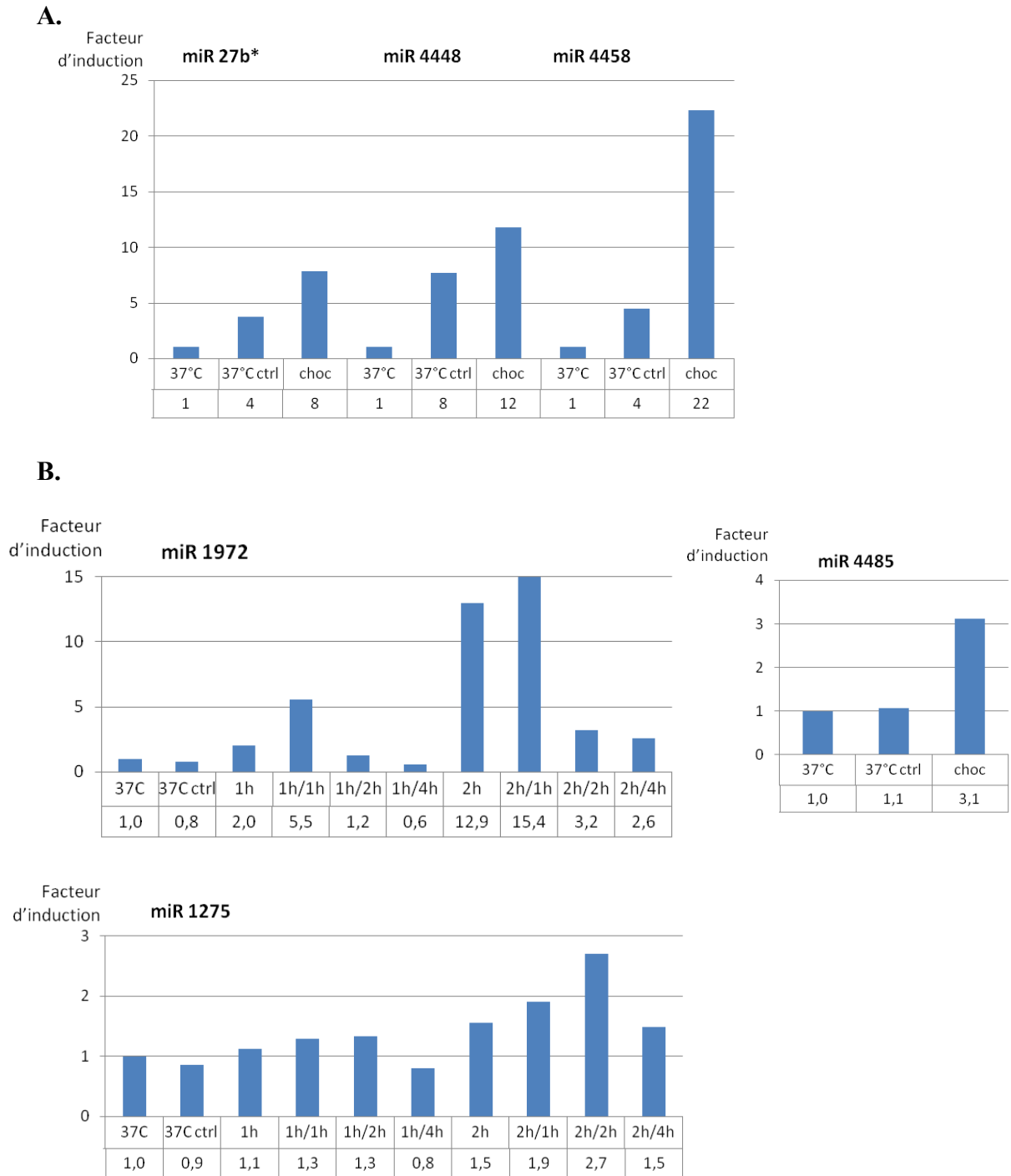


Figure III-17 : Validation par RT-qPCR des taux des variations des microARN suite à un choc thermique. Le témoin à 37°C a servi de référence pour le calcul de l'induction et la valeur correspondante a été arbitrairement fixée à 1. Le témoin "37ctrl" correspond à une immersion de 2 h à 37°C. Les durées des incubations pour les échantillons soumis à un choc thermique sont indiquées sous l'histogramme. Lorsqu'une deuxième durée est indiquée elle correspond à la période de régénération. A. microARN pour lesquels une induction est observée par immersion à 37°C (37°C ctrl). B. microARN pour lesquels une variation d'expression est observée uniquement après un choc thermique.

III.6. Conclusions et perspectives

III.6.1. Analyse globale de l'augmentation ou la diminution de l'expression des gènes après un choc thermique

Notre analyse du taux d'expression des gènes par emploi des puces à jonctions exon-exon a tout d'abord montré l'induction de l'expression de nombreux groupes fonctionnels de gènes connus pour être impliqués dans le choc thermique, notamment les gènes de nombreuses HSP, et ceux des ARN satellite III. Ces données sont en parfait accord avec celles de la littérature ce qui validait les conditions d'expérimentation.

Un premier constat marquant est la place prépondérante, parmi les gènes dont l'expression varie significativement, des régulations de gènes de protéines impliquées dans la régulation de la transcription, et par contre la faible représentation de gènes de protéines impliquées dans l'épissage. Parmi les facteurs de transcriptions, on retrouvait, en plus de HSF1, beaucoup de facteurs déjà décrits comme liés à la résistance à divers types de stress, certains n'ayant pas été reliés au choc thermique. C'est le cas notamment du facteur de transcription NUPR1, ainsi que de la protéine p53. Ces protéines, impliquées dans la régulation de la prolifération et dans l'apoptose avaient été notamment associées à l'induction de systèmes de défense contre le stress génotoxique (Clark et al., 2008; Mirzayans et al., 2013; Vincent et al., 2012). Les dommages à l'ADN causés par le choc thermique induisent donc plusieurs mécanismes de défense qui sont retrouvés en cas de stress génotoxique. On peut également noter la présence dans les résultats des puces, d'événements d'épissage alternatifs comparables ou similaires à ceux déjà observés en cas de stress génotoxique, même s'il ne s'agit pas d'événements identiques. Premièrement, avec un indice de confiance élevé, on note l'exclusion de l'exon 7 de la caspase 8, alors que l'utilisation d'un site alternatif 5' d'épissage a déjà été noté au niveau de l'exon 8 lors de l'exposition des cellules au cisplatine (Edmond et al., 2011). D'autre part, le saut d'un exon dans l'ARNm de MDM4 (6 ou 9 selon le transcrit) présent dans nos résultats de puces, a déjà été noté dans le cas d'un stress aux UV, en compagnie d'autres exons en amont ou en aval (Chandler et al., 2006). Dans la littérature ces événements sont signalés comme étant des événements favorisant l'apoptose ou limitant la prolifération suite à un stress génotoxique. Ainsi l'épissage alternatif de ces transcrits, communs dans la réponse au stress thermique, appuient l'hypothèse que l'épissage alternatif de certains transcrits pourrait être un élément critique dans la réponse à des stress de différentes natures. En effet, bien que nous ayons quantifié la mortalité cellulaire et n'ayons pas observé de mort cellulaire significative de nos cellules soumises au choc thermique (résultats non montrés), il est possible que des processus moléculaires de ce type soient tout de même enclenchés après ce stress, pour à terme mener à l'apoptose si le stress perdure ou devient plus sévère.

Un autre phénomène marquant est la diminution de l'expression des gènes d'histones. Il s'agit majoritairement d'histones appartenant à la famille des histones dites "réplication dépendantes"

(intégrées pendant la réplication de l'ADN). Ces gènes sont situés dans 3 regroupements ("clusters") de gènes d'histones, le plus grand étant appelé HIST1, et les deux autres HIST2 et HIST3. Un lien de proportionnalité entre le taux de réplication de l'ADN et la stabilité des ARNm de ces histones avait déjà pu être établi (DeLisle et al., 1983; Heintz et al., 1983). Ainsi, lors de la phase S, on observe une forte augmentation de la synthèse des ARNm de ces histones, tandis que l'inhibition de la réplication de l'ADN provoque une déstabilisation de ces ARNm. Par ailleurs, l'abondance des histones dites "réplication dépendantes" comme certains variants de l'histone H2A est diminuée dans les cellules de patients atteints de leucémie chronique, et a été associée avec la prolifération et les propriétés tumorales de cellules cancéreuses de la vessie (Singh et al., 2013; Su et al., 2009). Il a également été mis en évidence que la déplétion de l'histone H2A 1C et H2A1 B/E par l'emploi de siRNA peut provoquer l'augmentation du taux de prolifération cellulaire, indiquant un rôle actif des histones dans la régulation de la prolifération (Singh et al., 2013). Cette régulation s'opère *via* les protéines de contrôle de la progression du cycle cellulaire et de l'intégrité du génome, comme ATM ou DNA-PK (Müller et al., 2007; Su et al., 2004). Jusque là, la répression des gènes d'histones n'avait pas été spécifiquement associée à un choc thermique, et il est probable que la répression de ces gènes que nous avons mis en évidence soit en lien direct avec l'arrêt de la réplication de l'ADN ou du cycle cellulaire déjà observés lors d'un choc thermique. La confirmation expérimentale des variations de l'expression des gènes d'histones et des facteurs de transcription NUPR1 et p53 donnera une nouvelle perspective au rôle de la réplication de l'ADN et de la régulation des dommages à l'ADN dans la réponse au stress thermique. La poursuite de l'analyse de nos données transcriptionnelles pourra peut-être révéler l'existence d'autres mécanismes de réponse au choc thermique encore inconnus ou mal connus.

III.6.2. Implications de la synthèse de nouvelles isoformes des protéines des P-bodies et des cSG dans la réponse au stress thermique

L'emploi de puces à jonctions exon-exon nous a par ailleurs permis de détecter de nombreuses variations d'épissage alternatif qui pourraient être impliquées dans la réponse au stress thermique. Un point très intéressant est qu'un nombre élevé de ces variations concerne les ARNm de gènes de protéines des P-bodies et des cSG, qui fonctionnent en collaboration dans le contrôle de la traduction et de la dégradation des ARNm. Il s'agit des gènes *CNOT2*, *CNOT4*, *CNOT8*, *TNRC6A* (GW182), *EIF4ENIF1* (4E-T), *TIA1* et *EIF4G2*. Les variations d'épissage détectées par puces ont pu être vérifiées par RT-PCR. Afin de caractériser l'impact fonctionnel de ces variations, nous avons débuté des expériences, qui sont actuellement poursuivies par des membres de l'équipe.

Les ARNm étant des molécules sensibles à l'augmentation de température, on peut imaginer que des changements profonds dans leur métabolisme ont lieu en cas de stress. Ce constat que nous faisons est donc particulièrement intéressant, car la modification de ces protéines pourrait servir à faciliter le tri

des ARNm qui doivent être inhibés, traduits, voire dégradés. On peut par exemple imaginer que ces modifications dans les protéines des P-bodies et des cSG, servent à protéger certains ARNm de leur dégradation possible à cause de la chaleur, tandis que la dégradation d'une autre catégorie d'ARNm, notamment par les microARN, serait facilitée. La pertinence de ce type d'hypothèse est soulignée par plusieurs éléments décrits dans la littérature. D'une part, une diminution de l'efficacité du système de dégradation des ARNm *via* les microARN a déjà été observée lors d'un choc thermique (Detzer et al., 2011). Selon ces travaux, cette diminution est provoquée par la relocalisation de la protéine Ago2 dans les granules de stress cytoplasmique (Detzer et al., 2011). Ago2 est un composant catalytique majeur du complexe RISC ("RNA induced silencing complex"), capable d'interagir avec les microARN ou les siARN, afin de provoquer l'inhibition de l'expression ou la dégradation d'un ARNm cible, en recrutant d'autres protéines telle que GW182 (Hock et al., 2007). De ce fait, il se pourrait que les changements de l'épissage que nous avons détecté dans l'ARNm GW182 et d'autres interviennent aussi dans la diminution de l'efficacité du système de dégradation des ARNm *via* les microARN après un choc thermique. Afin de tester cette hypothèse, il faudra analyser le temps de demi-vie d'un certain nombre d'ARNm cellulaires après un choc thermique. Pour cela, des cinétiques de traitement à l'actinomycine D devront être réalisées sur des cellules HeLa soumises ou non à un choc thermique et la quantité de différents ARNm cellulaires devra être suivie par RT-qPCR ou par northern blot.

Dans un premier temps, nous avons concentré nos efforts sur l'étude de l'impact du changement du taux d'inclusion de l'exon 13 au sein des pré-ARNm GW182 sur la fonction de la protéine GW182. En effet cette protéine joue un rôle majeur dans la régulation des ARNm par les microARN. D'après nos données, la présence de l'exon 13 dans l'ARNm de GW182 ne semble pas influencer sa localisation. Pour compléter les expériences que nous avons déjà réalisées en vue de tester l'effet de la présence de la région correspondant à l'exon 13 dans la protéine GW182 sur la dégradation des ARNm, nous envisageons de réaliser des expériences telles que celles réalisées dans le groupe d'E. Izaurralde (Eulalio et al. 2009 ; Huntzinger et al., 2012). Dans ces expériences, l'efficacité de la dégradation des ARNm par les microARN sera suivie par le test de la capacité du microARN let-7 à inhiber l'expression d'un minigène rapporteur exprimant un ARNm de la luciférase Renilla porteur dans sa région 3'UTR de plusieurs sites de fixation let-7. L'impact de la présence ou de l'absence de la séquence codée par l'exon 13 dans la protéine GW182 sera testé par déplétion de la protéine GW182 endogène par des siRNA et complémentation des cellules par des variants de GW182 résistants à l'action des siRNA. En réalisant des complémentations avec des protéines GW182 contenant ou non la séquence codée par l'exon 13, nous espérons pouvoir étudier l'effet de l'exon 13 sur la fonction de GW182 dans la dégradation des ARNm *via* les microARN. Les constructions nécessaires à la réalisation de cette expérience nous ont été fournies par E. Izaurralde, et j'ai construit un plasmide permettant l'expression de la protéine GW182 dépourvue de la séquence correspondant à l'exon 13. F. Oillo-Blanloeil et C. Aigueperse réalisent actuellement les tests au laboratoire.

III.6.3. Analyse globale des nouveaux événements d'épissages apparaissant après un choc thermique

En parallèle, il sera intéressant d'essayer de comprendre quels peuvent être les effets d'autres variations d'épissage que nous avons détecté qui concernent d'autres familles de gènes. Par exemple, de nombreux gènes appartenant à la famille des protéines du transport vacuolaire et en particulier trois protéines de la même famille, les protéines "vacuolar protein sorting" (VPS) 52, 4B et 13B, sont retrouvés parmi les ARNm dont l'épissage alternatif est affecté. Ces protéines sont impliquées dans le transport des vésicules lipidiques entre l'appareil de Golgi et les endosomes. On trouve également les ARNm des protéines NECAP-1 et NECAP-2, impliquées dans le revêtement protéique des particules lipidiques de l'endocytose. Il serait donc intéressant de savoir en quoi des protéines reliées à la formation de vésicules et à l'endocytose peuvent être importantes dans la réponse au choc thermique. Ces observations pourraient avoir un lien avec la découverte récente de protéines hsp70 extracellulaires sécrétées par des vésicules (pour revue, De Maio, 2011). Ces vésicules sont impliquées dans la communication et l'activation de voies de signalisation intercellulaires. Leur rôle a été observé récemment dans le cadre de processus inflammatoires mais il est envisageable qu'elles soient également impliquées dans le choc thermique. Des variations d'épissage alternatif sont également détectées pour plusieurs autres familles fonctionnelles de protéines, comme des protéines impliquées dans le cycle cellulaire, le transport et la dégradation des protéines, ou encore les systèmes de réparation de l'ADN. Il nous faudra d'abord valider l'existence de ces variations d'épissage par RT-PCR. Nous espérons que l'exploitation de ces données permettra de révéler l'implication de nouveaux acteurs protéique ou de nouveaux processus de régulation impliqués dans la réponse au choc thermique, et d'une manière générale de préciser l'impact des événements d'épissage alternatif dans la régulation du choc thermique.

Un autre point qu'il sera important d'éclaircir est la contribution exacte de la formation des nSB dans les régulations de l'épissage alternatif observées. En effet, même si nous avons pu clairement établir l'existence de variations d'épissage alternatif dans des conditions où des nSB sont formés, aucune de nos expériences ne permet pour l'instant de relier entre eux ces deux événements. Afin de répondre à cette question nous devons répéter les analyses de transcriptomique en réalisant le choc thermique sur des cellules HeLa préalablement traitées par des siRNA dirigés contre les ARN sat III. Ce traitement a déjà été utilisé avec succès par l'équipe de G. Biamonti et résulte en une très faible sur-expression des ARN sat III suite à un choc thermique du fait de leur dégradation immédiate par ARN interférence (Valgardsdottir et al., 2005a). Nous pourrions ainsi déterminer si l'absence des ARN sat III dans les cellules soumises à un choc thermique affecte les variations d'épissage alternatif des ARNm.

III.6.4. Analyse globale de la variation de l'expression des microARN après un choc thermique

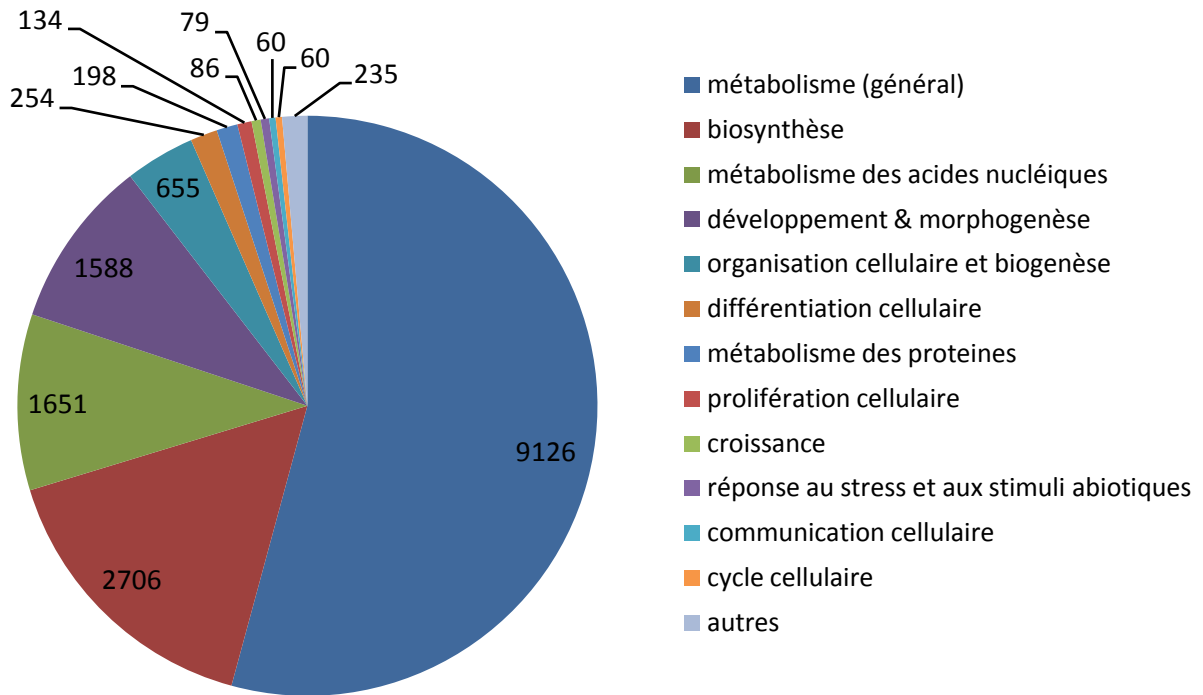
En ce qui concerne l'analyse de la variation de l'expression des microARN après un choc thermique, nos données ne révèlent pas de changements massifs de la production des microARN. Nous avons néanmoins détecté des variations significatives de l'expression d'au moins 4 microARN : miR-1972, miR-1275, miR-4485 et miR-4458. Ces microARN ayant été identifiés récemment, on dispose à l'heure actuelle de peu de données concernant leurs cibles et leurs fonctions. On notera cependant qu'une étude récente a identifié miR-1275 comme étant sur-exprimé 8 heures après l'infection par le virus Influenza A d'une lignée cellulaire humaine de poumon et que l'une des cibles ARNm de miR-1275 est la protéine MAPK3 (ERK1) (Buggele et al., 2012). De façon intéressante MAPK3 intervient dans la régulation de l'activité du facteur de transcription HSF1 qui joue un rôle central dans les mécanismes de réponses au choc thermique. Il est par conséquent concevable, même si cela reste à démontrer, que la régulation de l'expression de miR-1275 puisse jouer un rôle dans la régulation de l'expression de HSF1 au cours de la réponse au choc thermique. Une prédiction informatique des ARNm cibles potentiels de miR-1972, miR-1275, miR-4485 et miR-4458 sera réalisée. Nous analyserons l'expression des ARNm cibles potentiels au sein des données issues de nos puces à jonctions exon-exon. Nous espérons ainsi pouvoir identifier certains couples microARN-ARNm que nous étudierons plus en détails par la suite.

ANNEXE 1 : Résultat de l'analyse grâce à CateGORizer des termes GO associés aux gènes dont l'expression varie après le choc thermique

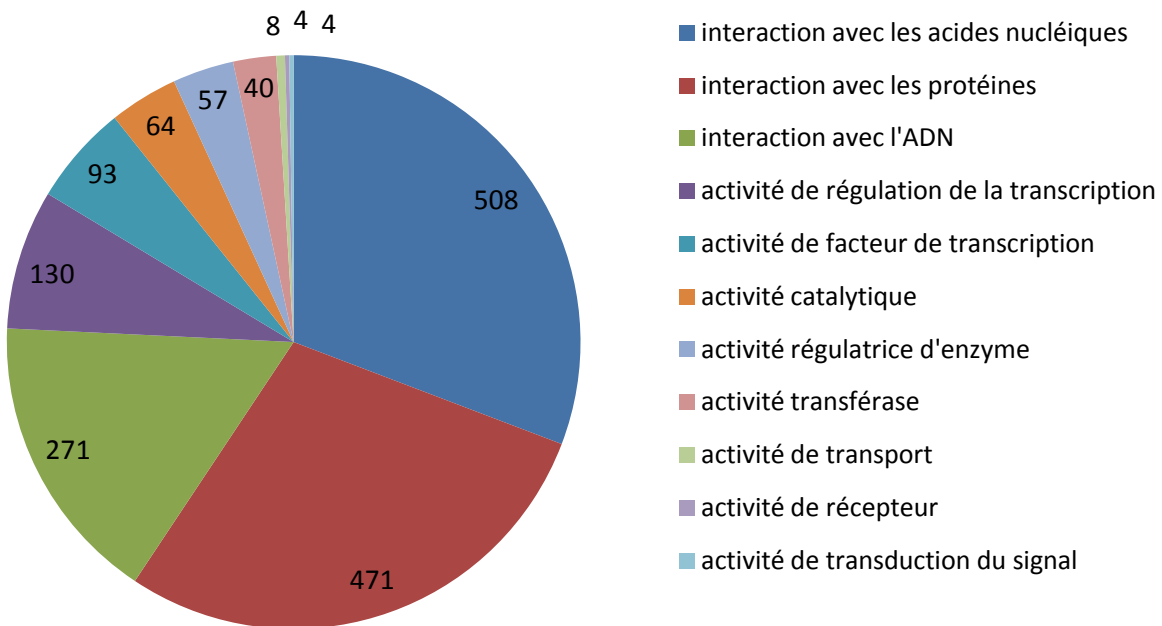
Nous avons utilisé un outil, nommé CateGORizer (<http://www.animalgenome.org/bioinfo/tools/catego/>), permettant, grâce à une classification nommée GOslim, de regrouper des termes GO dans des catégories plus générales. Nous avons appliqué cette catégorisation aux données de puces à jonctions exon-exon, afin d'avoir un aperçu global des grandes fonctions modifiées dans les cellules. Cette technique a été appliquée sur les termes GO reflétant les processus biologiques et ceux reflétant les fonctions moléculaires, qui sont bien moins redondants, et les résultats obtenus sont présentés sous formes de diagramme ci-dessous.

Il faut toutefois garder à l'esprit que cet aperçu ne permet pas de s'affranchir de la redondance des termes GO. Par exemple, le terme GO "métabolisme" utilisé comme terme parent global peut inclure plusieurs termes GO différents pour un seul gène, auquel cas il est compté plusieurs fois pour un même gène, ce qui agrandit considérablement le nombre total de termes GO considérés. D'autre part, un même gène peut être associé à plusieurs catégories de termes génériques simultanément (par exemple "métabolisme" et "métabolisme des acides nucléiques"). Ces graphiques donnent donc surtout une idée du nombre de fois où apparaît un terme GO par rapport aux autres dans la totalité des termes GO référencés pour les gènes ressortant des données de puces. On remarque que les catégories prédominantes de gènes dont l'expression varie concernent le métabolisme et les processus de biosynthèse dans la cellule, probablement parce que ce sont les catégories les plus génériques. Toutefois, ceci est cohérent avec le fait que les fonctions les plus affectées dans le cas du choc thermique sont des fonctions concernant la fixation aux acides nucléiques et l'activité de transcription.

A. Processus biologiques



B. Fonctions moléculaires



Annexe 1 : Représentativité des termes GO associés aux produits de gènes dont l'expression varie selon les données des micropuces. Cette catégorisation des gènes dont l'expression varie est basée sur les catégories fournies par la classification "GOSlim" et a été réalisée grâce au logiciel CateGORizer. Les catégories légendées à droite sont fournies dans l'ordre décroissant de leur représentativité (sauf pour la catégorie "autres"). A. Catégorisation effectuée pour les termes GO concernant les processus biologiques des gènes différentiellement exprimés après le stress thermique. B. Catégorisation effectuée pour les termes GO concernant la fonction moléculaire des gènes différentiellement exprimés après le stress thermique.

ANNEXE 2 : Résultat de l'analyse en amont des données de puces jonctions exon-exon réalisée avec le logiciel Ingenuity (p-value<0,01)

Facteur régulateur identifié	Type de molécule	p-value
NUPR1	régulateur de la transcription	1,95E-16
PDGF BB	complexe	1,69E-15
HSF1	régulateur de la transcription	4,28E-11
ERK	groupe	4,64E-09
F7	peptidase	6,42E-08
TREM1	récepteur transmembranaire	7,03E-07
MAP2K1/2	groupe	6,03E-06
DCLK1	kinase	7,69E-06
SBDS	autre	1,66E-05
P38 MAPK	groupe	3,36E-05
RAE1	autre	3,65E-05
PAF1	autre	1,15E-04
TP53	régulateur de la transcription	1,24E-04
SUPT16H	régulateur de la transcription	2,05E-04
YAP1	régulateur de la transcription	2,21E-04
NR3C1	récepteur nucléaire ligand-dépendant	2,75E-04
HIF1A	régulateur de la transcription	2,80E-04
Mek	groupe	4,24E-04
Creb	groupe	4,62E-04
EGF	facteur de croissance	5,43E-04
RUVBL1	régulateur de la transcription	5,49E-04
GFI1 (1,99)	régulateur de la transcription	6,13E-04
PI3K	complexe	6,74E-04
LNX2	autre	9,15E-04
NDRG1 (1,72)	kinase	9,15E-04
NKX2-1	régulateur de la transcription	9,15E-04
JUN (3,49)	régulateur de la transcription	1,00E-03
BMP10	facteur de croissance	1,48E-03
FBXW7	régulateur de la transcription	1,48E-03
DDX17	enzyme	1,48E-03
DDX5	enzyme	1,48E-03
FOXO3	régulateur de la transcription	1,52E-03
SMARCA4	régulateur de la transcription	1,77E-03
Fgf	groupe	1,89E-03
SMAD1/5	groupe	1,89E-03
RGS3 (-1,71)	autre	1,89E-03
DACH1	régulateur de la transcription	1,89E-03

Facteur régulateur identifié	Type de molécule	p-value
PRKACA	kinase	1,89E-03
Pkc(s)	groupe	1,92E-03
Ap1	complexe	2,28E-03
Smad	complexe	2,51E-03
WWTR1	régulateur de la transcription	2,51E-03
Akt	groupe	2,68E-03
CREBBP	régulateur de la transcription	2,68E-03
Hdac	groupe	2,69E-03
SRF	régulateur de la transcription	3,15E-03
SSRP1	autre	3,27E-03
HLX	régulateur de la transcription	3,89E-03
mir-515	microARN	3,89E-03
SCUBE3	autre	3,89E-03
EGR1 (8,7)	régulateur de la transcription	3,94E-03
ATM	kinase	3,94E-03
STAT3	régulateur de la transcription	4,76E-03
BMPR2	kinase	5,50E-03
ACTL6A	autre	5,50E-03
HOXB1	régulateur de la transcription	5,50E-03
MAP3K3 (-1,61)	kinase	5,50E-03
mir-135	microARN	5,50E-03
LPA	autre	5,50E-03
URI1	régulateur de la transcription	5,50E-03
SMO	Récepteur couplé à la protéine G	5,50E-03
MAP3K11	kinase	5,50E-03
SMAD4	régulateur de la transcription	5,58E-03
Notch	groupe	5,58E-03
FOXL2	régulateur de la transcription	5,59E-03
TP73	régulateur de la transcription	5,64E-03
GDF2	facteur de croissance	5,64E-03
WNT5A	cytokine	5,64E-03
RLIM	enzyme	5,64E-03
mir-373	microARN	5,64E-03
ERK1/2	groupe	6,58E-03
SRC	kinase	6,66E-03
WT1	régulateur de la transcription	6,66E-03
FOXO1 (2,65)	régulateur de la transcription	6,95E-03
SMAD3	régulateur de la transcription	7,19E-03
ATF4	régulateur de la transcription	8,15E-03
Cg	complexe	8,83E-03
SIN3A (-2,11)	régulateur de la transcription	9,84E-03
Pka	complexe	1,04E-02

Facteur régulateur identifié	Type de molécule	p-value
Calcineurin protein(s)	complexe	1,07E-02
BMPR1A	kinase	1,07E-02
FANCA	autre	1,07E-02
NOTCH4	régulateur de la transcription	1,07E-02
NOD1 (-1,68)	autre	1,07E-02
ACVR1	kinase	1,07E-02
HNRNPA1	autre	1,07E-02
RGCC	autre	1,07E-02
EREG	facteur de croissance	1,07E-02
SNRK	kinase	1,07E-02
RNA polymerase II	complexe	1,09E-02
MAP2K1	kinase	1,16E-02
POU5F1	régulateur de la transcription	1,16E-02
PCDH11X/PCDH11Y	autre	1,17E-02
PRKCE	kinase	1,17E-02
BMP7	facteur de croissance	1,17E-02
AR	récepteur nucléaire ligand-dépendant	1,24E-02
HIC1	régulateur de la transcription	1,37E-02
MYB	régulateur de la transcription	1,39E-02
BMP2	facteur de croissance	1,39E-02
CD44	enzyme	1,62E-02
DRAP1	régulateur de la transcription	1,62E-02
MAX	régulateur de la transcription	1,69E-02
CDX2	régulateur de la transcription	1,69E-02
PIM1	kinase	1,69E-02
F2	peptidase	1,69E-02
ATF2	régulateur de la transcription	1,69E-02
caspase	groupe	1,73E-02
KDM3A	régulateur de la transcription	1,73E-02
IRS1	enzyme	1,73E-02
MED12	régulateur de la transcription	1,73E-02
AKT3	kinase	1,73E-02
DUSP1 (1,72)	phosphatase	1,73E-02
SKI	régulateur de la transcription	1,73E-02
SCD	enzyme	1,73E-02
ATG7	enzyme	1,73E-02
MAD2L2	enzyme	1,73E-02
PHB	régulateur de la transcription	1,73E-02
ZBTB7B (-2,02)	autre	1,73E-02
MS4A1	autre	1,73E-02
PELP1	autre	1,73E-02

Facteur régulateur identifié	Type de molécule	p-value
PDX1	régulateur de la transcription	1,89E-02
TGFB1	facteur de croissance	2,07E-02
RAC1	enzyme	2,08E-02
IL12B	cytokine	2,08E-02
AZGP1	transporteur	2,08E-02
GSK3B	kinase	2,16E-02
RBPJ	régulateur de la transcription	2,16E-02
SP1	régulateur de la transcription	2,40E-02
NANOG	régulateur de la transcription	2,43E-02
TWIST2	régulateur de la transcription	2,52E-02
FOXO4	régulateur de la transcription	2,52E-02
TRIB3	kinase	2,52E-02
Hsp70	groupe	2,52E-02
Hedgehog	groupe	2,52E-02
F10	peptidase	2,52E-02
PTPRR	phosphatase	2,52E-02
NOTCH2	régulateur de la transcription	2,52E-02
TAZ	enzyme	2,52E-02
TAF6	régulateur de la transcription	2,52E-02
DEK	régulateur de la transcription	2,52E-02
FSH	complexe	2,54E-02
TERT	enzyme	2,81E-02
ELF4	régulateur de la transcription	3,00E-02
EPAS1	régulateur de la transcription	3,15E-02
HRAS	enzyme	3,16E-02
HDAC1	régulateur de la transcription	3,40E-02
TFRC	transporteur	3,43E-02
mir-181	microARN	3,43E-02
AKT1	kinase	3,43E-02
RB1CC1	autre	3,43E-02
ELANE	peptidase	3,43E-02
TYK2	kinase	3,43E-02
BAG1	autre	3,43E-02
KRAS	enzyme	3,45E-02
RB1	régulateur de la transcription	3,53E-02
ESR1	récepteur nucléaire ligand-dépendant	3,53E-02
SOX2	régulateur de la transcription	3,55E-02
UXT	autre	3,96E-02
NOTCH1 (1,77)	régulateur de la transcription	3,96E-02
A2M	transporteur	4,10E-02
TNFSF10	cytokine	4,10E-02

Facteur régulateur identifié	Type de molécule	p-value
BRAF	enzyme	4,10E-02
EGFR	kinase	4,11E-02
ERBB2	kinase	4,14E-02
MYC (-1,62)	régulateur de la transcription	4,19E-02
SHP	groupe	4,35E-02
Rsk	groupe	4,35E-02
Cofilin	groupe	4,35E-02
MNK1/2	groupe	4,35E-02
RPL21	autre	4,35E-02
CDC42	enzyme	4,35E-02
ZMIZ1	autre	4,35E-02
FGF21	facteur de croissance	4,35E-02
SSBP2	régulateur de la transcription	4,35E-02
CUZD1	autre	4,35E-02
DUOX1	enzyme	4,35E-02
KDM4C	autre	4,35E-02
EGLN2	enzyme	4,35E-02
DNAJC15	autre	4,35E-02
CPEB2	autre	4,35E-02
AFAP1L2	autre	4,35E-02
FANCG	autre	4,35E-02
PTRF	régulateur de la transcription	4,35E-02
FANCL	enzyme	4,35E-02
PRKACB	kinase	4,35E-02
MYH9	transporteur	4,35E-02
FANCF	autre	4,35E-02
TNFRSF10A	récepteur transmembranaire	4,35E-02
LSM10	autre	4,35E-02
SLC2A4	transporteur	4,35E-02
AGR2	autre	4,35E-02
CTNNBIP1	autre	4,35E-02
ACVR2A	kinase	4,35E-02
CDC45	autre	4,35E-02
PCAT6	autre	4,35E-02
PMAIP1 (2,74)	autre	4,35E-02
EEF1A1	translation regulator	4,35E-02
LIMS1	autre	4,35E-02
KSR2	kinase	4,35E-02
miR-185-3p	mature microARN	4,35E-02
mir-580	microARN	4,35E-02
mir-198	microARN	4,35E-02

Facteur régulateur identifié	Type de molécule	p-value
miR-602	mature microARN	4,35E-02
mir-602	microARN	4,35E-02
TOPORS	enzyme	4,35E-02
RPS20 (2,08)	autre	4,35E-02
RNF14	régulateur de la transcription	4,35E-02
ncRNA-a7	autre	4,35E-02
ncRNA-a6	autre	4,35E-02
ELK4	régulateur de la transcription	4,35E-02
NCOA5 (-1,54)	autre	4,35E-02
TRIM63	enzyme	4,35E-02
CKAP4	autre	4,35E-02
EIF2B3	translation regulator	4,35E-02
TGFB1I1	régulateur de la transcription	4,35E-02
DHX9	enzyme	4,35E-02
KLF3 (1,66)	régulateur de la transcription	4,35E-02
TPSD1	peptidase	4,35E-02
CPEB1	autre	4,35E-02
STXBP2	transporteur	4,35E-02
MED26	régulateur de la transcription	4,35E-02
ACVRL1	kinase	4,35E-02
CDK11A/CDK11B	kinase	4,35E-02
FOXD3	régulateur de la transcription	4,35E-02
BMX	kinase	4,35E-02
MAP3K10	kinase	4,35E-02
GPR98	Récepteur couplé à la protéine G	4,35E-02
BMPR1B	kinase	4,35E-02
FANCE	autre	4,35E-02
TP53BP2	autre	4,35E-02
GTF2I	régulateur de la transcription	4,35E-02
PELI2	autre	4,35E-02
ENTPD1	enzyme	4,35E-02
HBB	transporteur	4,35E-02
LMO1	régulateur de la transcription	4,35E-02
RNF2	régulateur de la transcription	4,35E-02
RACGAP1	transporteur	4,35E-02
IL1RL1	récepteur transmembranaire	4,35E-02
ERF	régulateur de la transcription	4,35E-02
TOP2A	enzyme	4,35E-02
TNFRSF10B	récepteur transmembranaire	4,35E-02
STAT5a/b	groupe	4,39E-02
MAP3K7	kinase	4,39E-02

Facteur régulateur identifié	Type de molécule	p-value
UGT2B15	enzyme	4,44E-02
PTGER2 (2,57)	Récepteur couplé à la protéine G	4,44E-02
UGT2B17	enzyme	4,44E-02
SERPINF1	autre	4,44E-02
TLR6	récepteur transmembranaire	4,44E-02
PCGF2	régulateur de la transcription	4,44E-02
TOPBP1	autre	4,44E-02
PPRC1 (-2,31)	régulateur de la transcription	4,47E-02
TGFBR2	kinase	4,54E-02
MYD88	autre	4,72E-02
IL1B	cytokine	4,72E-02
RAB1B	autre	4,86E-02
ESR2	récepteur nucléaire ligand-dépendant	4,86E-02

ANNEXE 3 : Regroupement fonctionnel grâce au logiciel DAVID des termes GO (Biological Process) associés aux variations d'épissage détectées par puces à jonctions exon-exon.

Cette analyse a été effectuée pour mettre en évidence les termes GO-BP uniquement, en utilisant les gènes subissant une variation d'épissage de type exon cassette qui sont détectés avec un indice de confiance moyen à élevé. Avec chaque terme GO-BP répertorié, est fourni le taux d'enrichissement constaté pour les gènes de cette catégorie suivi de la p-value qui lui est associé.

<u>Cycle cellulaire (2,4 ; 8,00E-05)</u>	
ID ENSEMBL	Nom du gène
776427	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11
807700	G-2 and S-phase expressed 1
824595	Holliday junction recognition protein
813217	Mdm4 p53 binding protein homolog (mouse)
787405	RAD1 homolog (S. pombe)
803287	RAN, member RAS oncogene family
826209	SET domain containing (lysine methyltransferase) 8
810642	amyloid beta (A4) precursor protein
805050	anillin, actin binding protein
809850	chromosome 13 open reading frame 3
819748	cyclin B1
799556	cyclin D binding myb-like transcription factor 1
805899	cyclin T2
776371	cyclin-dependent kinase 7
780478	cylindromatosis (turban tumor syndrome)
774797	cytoplasmic linker associated protein 1
782002	eukaryotic translation initiation factor 4 gamma, 2
772426	glycogen synthase kinase 3 beta
774263	inhibitor of growth family, member 4
784977	integrin, beta 1 (fibronectin receptor, beta polypeptide, MSK12)
788987	microtubule-actin crosslinking factor 1
805698	nuclear distribution gene C homolog (A. nidulans)
793195	par-3 partitioning defective 3 homolog (C. elegans)
791255	spindlin 1
798021	suppressor of variegation 3-9 homolog 2 (Drosophila)
783961	vacuolar protein sorting 4 homolog B (S. cerevisiae)
<u>Transport des protéines (2,2 ; 4,00E-04)</u>	
ID ENSEMBL	Nom du gène
773802	API gamma subunit binding protein 1
808723	ATG4 autophagy related 4 homolog A (S. cerevisiae)

816192	Der1-like domain family, member 1
785707	NECAP endocytosis associated 1
804841	NECAP endocytosis associated 2
797262	NOP58 ribonucleoprotein homolog (yeast)
823090	RAB5A, member RAS oncogene family
803287	RAN, member RAS oncogene family
787492	RPA interacting protein
804924	aftiphilin
811536	chromosome 14 open reading frame 133
825289	eukaryotic translation initiation factor 4E nuclear import factor 1
821182	exocyst complex component 7
804942	exportin 1 (CRM1 homolog, yeast)
773845	fracture callus 1 homolog (rat)
772426	glycogen synthase kinase 3 beta
813302	golgi associated, gamma adaptin ear containing, ARF BP 3
773196	lectin, mannose-binding 2-like
788987	microtubule-actin crosslinking factor 1
825103	transportin 3
795850	vacuolar protein sorting 13 homolog B (yeast)
783961	vacuolar protein sorting 4 homolog B (S. cerevisiae)
804203	vacuolar protein sorting 52 homolog (S. cerevisiae)
780037	vesicle-associated membrane protein 7
<u>Maturation des ARN (2,5 ; 6,60E-04)</u>	
ID ENSEMBL	Nom du gène
790898	DIS3 mitotic control homolog (S. cerevisiae)
816714	FUS interacting protein (serine/arginine-rich) 1
797262	NOP58 ribonucleoprotein homolog (yeast)
783520	RNA binding motif protein 23
782499	RNA binding motif protein 5
772310	SFRS protein kinase 2
822994	SYF2 homolog, RNA splicing factor (S. cerevisiae)
810642	amyloid beta (A4) precursor protein
807328	breast carcinoma amplified sequence 2
812773	coiled-coil domain containing 76
821665	dual specificity phosphatase 11 (RNA/RNP complex 1-interacting)
813133	heterogeneous nuclear ribonucleoprotein A1-like 2
774982	heterogeneous nuclear ribonucleoprotein D-like
810407	poly(A) polymerase gamma
825901	polypyrimidine tract binding protein 1
818254	queuine tRNA-ribosyltransferase domain containing 1
785094	ribonuclease P/MRP 14kDa subunit
776781	transformer 2 alpha homolog (Drosophila)
802107	zinc finger, CCHC domain containing 8

Réponse cellulaire au stress (2,4 ; 9,80E-08)	
ID ENSEMBL	Nom du gène
818171	DNA cross-link repair 1B (PSO2 homolog, <i>S. cerevisiae</i>)
810838	DNA cross-link repair 1C (PSO2 homolog, <i>S. cerevisiae</i>)
816192	Der1-like domain family, member 1
807700	G-2 and S-phase expressed 1
787405	RAD1 homolog (<i>S. pombe</i>)
797802	RAD18 homolog (<i>S. cerevisiae</i>)
787492	RPA interacting protein
796017	X-linked inhibitor of apoptosis
806417	alkB, alkylation repair homolog 3 (<i>E. coli</i>)
789252	cerebral cavernous malformation 2
807895	chromosome 11 open reading frame 30
776371	cyclin-dependent kinase 7
822591	denticleless homolog (<i>Drosophila</i>)
772426	glycogen synthase kinase 3 beta
797002	immunoglobulin mu binding protein 2
774263	inhibitor of growth family, member 4
812059	non-SMC element 2, MMS21 homolog (<i>S. cerevisiae</i>)
778167	suppressor of Ty 16 homolog (<i>S. cerevisiae</i>)
782704	trinucleotide repeat containing 6A
Processus catabolique des protéines (2,2 ; 2,80E-03)	
ID ENSEMBL	Nom du gène
808723	ATG4 autophagy related 4 homolog A (<i>S. cerevisiae</i>)
803551	CCR4-NOT transcription complex, subunit 4
816192	Der1-like domain family, member 1
800641	F-box protein 42
797802	RAD18 homolog (<i>S. cerevisiae</i>)
796017	X-linked inhibitor of apoptosis
790347	ankyrin repeat and SOCS box-containing 6
809819	caspase 8, apoptosis-related cysteine peptidase
819748	cyclin B1
780478	cylindromatosis (turban tumor syndrome)
822591	denticleless homolog (<i>Drosophila</i>)
812059	non-SMC element 2, MMS21 homolog (<i>S. cerevisiae</i>)
824685	suppressor of cytokine signaling 4
809467	tripeptidyl peptidase I
779373	ubiquitin specific peptidase 37
815253	ubiquitin specific peptidase 8
814720	ubiquitin-conjugating enzyme E2 variant 1
822361	ubiquitin-conjugating enzyme E2, J2 (UBC6 homolog, yeast)
824316	ubiquitin-like modifier activating enzyme 2

<u>Régulations des processus basés sur les microtubules (7,3 ; 4,80E-03)</u>	
ID ENSEMBL	Nom du gène
816691	McKusick-Kaufman syndrome
809850	chromosome 13 open reading frame 3
774797	cytoplasmic linker associated protein 1
804942	exportin 1 (CRM1 homolog, yeast)
788987	microtubule-actin crosslinking factor 1
<u>Organisation des sous-unités de complexes macromoléculaires (2 ; 5,00E-03)</u>	
ID ENSEMBL	Nom du gène
809021	CREB binding protein
816714	FUS interacting protein (serine/arginine-rich) 1
824595	Holliday junction recognition protein
813217	Mdm4 p53 binding protein homolog (mouse)
782499	RNA binding motif protein 5
772310	SFRS protein kinase 2
821928	TAF12 RNA polymerase II, TBP-associated factor, 20kDa
782508	adenylosuccinate lyase
796205	adhesion regulating molecule 1
805050	anillin, actin binding protein
809819	caspase 8, apoptosis-related cysteine peptidase
776371	cyclin-dependent kinase 7
804942	exportin 1 (CRM1 homolog, yeast)
797002	immunoglobulin mu binding protein 2
807324	mitochondrial ribosome recycling factor
793195	par-3 partitioning defective 3 homolog (C. elegans)
778167	suppressor of Ty 16 homolog (S. cerevisiae)
776899	tubulin, beta 2C
804156	tumor necrosis factor receptor superfamily, member 1A
788097	vesicle-associated membrane protein 3 (cellubrevin)
<u>Modification de la chromatine (2,9 ; 5,20E-03)</u>	
ID ENSEMBL	Nom du gène
808473	BCL6 co-repressor
809021	CREB binding protein
824595	Holliday junction recognition protein
826209	SET domain containing (lysine methyltransferase) 8
821928	TAF12 RNA polymerase II, TBP-associated factor, 20kDa
784778	bromodomain PHD finger transcription factor
775926	bromodomain containing 8
807895	chromosome 11 open reading frame 30
794118	inhibitor of growth family, member 3
774263	inhibitor of growth family, member 4
798021	suppressor of variegation 3-9 homolog 2 (Drosophila)

Régulation de la transcription (1,4 ; 8,90E-03)	
ID ENSEMBL	Nom du gène
808473	BCL6 co-repressor
798333	CCR4-NOT transcription complex, subunit 2
803551	CCR4-NOT transcription complex, subunit 4
788692	CCR4-NOT transcription complex, subunit 8
809021	CREB binding protein
816714	FUS interacting protein (serine/arginine-rich) 1
792983	GC-rich promoter binding protein 1
793576	HMG box domain containing 4
821094	Kruppel-like factor 6
813217	Mdm4 p53 binding protein homolog (mouse)
803287	RAN, member RAS oncogene family
805964	SAP30 binding protein
826209	SET domain containing (lysine methyltransferase) 8
777328	SIN3 homolog A, transcription regulator (yeast)
821928	TAF12 RNA polymerase II, TBP-associated factor, 20kDa
778482	YY1 associated protein 1; gon-4-like (C. elegans)
802433	additional sex combs like 1 (Drosophila)
810642	amyloid beta (A4) precursor protein
784778	bromodomain PHD finger transcription factor
800176	bromodomain adjacent to zinc finger domain, 1A
775926	bromodomain containing 8
805714	calcium regulated heat stable protein 1, 24kDa
782974	catenin (cadherin-associated protein), delta 1
807895	chromosome 11 open reading frame 30
773156	chromosome 17 open reading frame 81
799556	cyclin D binding myb-like transcription factor 1
805899	cyclin T2
776371	cyclin-dependent kinase 7
817937	elongation protein 3 homolog (S. cerevisiae)
815828	forkhead box J3
774982	heterogeneous nuclear ribonucleoprotein D-like
774392	homeodomain interacting protein kinase 3
797002	immunoglobulin mu binding protein 2
794118	inhibitor of growth family, member 3
774263	inhibitor of growth family, member 4
817759	myocardin
776490	nuclear factor of activated T-cells 5, tonicity-responsive
778836	nuclear receptor coactivator 7
785677	recombination signal binding protein for Ig kappa J region
778167	suppressor of Ty 16 homolog (S. cerevisiae)
798021	suppressor of variegation 3-9 homolog 2 (Drosophila)

816847	thymopoietin
804156	tumor necrosis factor receptor superfamily, member 1A
814720	ubiquitin-conjugating enzyme E2 variant 1
802658	v-myb myeloblastosis viral oncogene homolog (avian)-like 2
825813	zinc finger and BTB domain containing 38
804663	zinc finger protein 207
801100	zinc finger protein 317
819156	zinc finger protein 35
807734	zinc finger protein 426
800042	zinc finger protein 548
<u>Transport vacuolaire (8,9 ; 9,90E-03)</u>	
ID ENSEMBL	Nom du gène
808723	ATG4 autophagy related 4 homolog A (S. cerevisiae)
793576	HMG box domain containing 4
811536	chromosome 14 open reading frame 133
780037	vesicle-associated membrane protein 7
<u>Processus métabolique des ARNt (3,6 ; 2,50E-02)</u>	
ID ENSEMBL	Nom du gène
813024	arginyl-tRNA synthetase 2, mitochondrial
812773	coiled-coil domain containing 76
781476	cysteinyl-tRNA synthetase
818254	queuine tRNA-ribosyltransferase domain containing 1
785094	ribonuclease P/MRP 14kDa subunit
814244	tryptophanyl-tRNA synthetase
<u>Ciblage des protéines (2,6 ; 3,20E-02)</u>	
ID ENSEMBL	Nom du gène
808723	ATG4 autophagy related 4 homolog A (S. cerevisiae)
797262	NOP58 ribonucleoprotein homolog (yeast)
803287	RAN, member RAS oncogene family
787492	RPA interacting protein
804942	exportin 1 (CRM1 homolog, yeast)
773845	fracture callus 1 homolog (rat)
772426	glycogen synthase kinase 3 beta
788987	microtubule-actin crosslinking factor 1
<u>Régulation de la traduction (3,1 ; 4,30E-02)</u>	
ID ENSEMBL	Nom du gène
824163	TIA1 cytotoxic granule-associated RNA binding protein
810642	amyloid beta (A4) precursor protein
782002	eukaryotic translation initiation factor 4 gamma, 2
775201	pumilio homolog 2 (Drosophila)
816672	eukaryotic translation initiation factor 4A, isoform 2
782704	trinucleotide repeat containing 6A

<u>Mort cellulaire (1,7 ; 4,40E-02)</u>	
ID ENSEMBL	Nom du gène
798711	A kinase (PRKA) anchor protein 13
813217	Mdm4 p53 binding protein homolog (mouse)
782499	RNA binding motif protein 5
805964	SAP30 binding protein
824163	TIA1 cytotoxic granule-associated RNA binding protein
796017	X-linked inhibitor of apoptosis
810642	amyloid beta (A4) precursor protein
773064	bifunctional apoptosis regulator
809819	caspase 8, apoptosis-related cysteine peptidase
813383	endothelin converting enzyme 1
782002	eukaryotic translation initiation factor 4 gamma, 2
774392	homeodomain interacting protein kinase 3
797002	immunoglobulin mu binding protein 2
774263	inhibitor of growth family, member 4
772977	son of sevenless homolog 1 (Drosophila)
809467	tripeptidyl peptidase I
804156	tumor necrosis factor receptor superfamily, member 1A

ANNEXE 4 : Regroupement fonctionnel grâce au logiciel DAVID des termes associés aux variations d'épissage détectées par puces à jonctions exon-exon.

Cette analyse a été effectuée en utilisant les gènes subissant une variation d'épissage de type exon cassette qui sont détectés avec un indice de confiance moyen à élevé. A chaque groupe fonctionnel répertorié, est fourni le taux d'enrichissement constaté pour les termes de cette catégorie suivi du score d'enrichissement DAVID qui lui est associé. Les groupes fonctionnels indiqués sont ceux qui ont un score d'enrichissement calculé par DAVID supérieur à 1.

Groupe fonctionnel 1 - Score d'enrichissement DAVID : 3,24				
Catégorie	Terme	Nombre	Enrich.	p-value
Pir Superfamily	rna-binding	21	3,1	2,00E-05
Terme GO - MF	RNA binding	24	2,5	6,20E-05
INTERPRO	RNA recognition motif, RNP-1	11	3,9	5,60E-04
INTERPRO	Nucleotide-binding, alpha-beta plait	11	3,8	6,10E-04
Terme GO - BP	mRNA processing	12	2,7	5,50E-03
Terme GO - BP	mRNA metabolic process	12	2,3	1,50E-02
Groupe fonctionnel 2 - Score d'enrichissement DAVID : 2,9				
Catégorie	Terme	Nombre	Enrich.	p-value
INTERPRO	Zinc finger, PHD-finger	7	6,1	9,80E-04
INTERPRO	Zinc finger, PHD-type	7	5,8	1,30E-03
INTERPRO	Zinc finger, PHD-type, conserved site	7	5,6	1,60E-03
Groupe fonctionnel 3 Score d'enrichissement DAVID : 2,81				
Catégorie	Terme	Nombre	Enrich.	p-value
INTERPRO	Bromodomain, conserved site	5	10,0	1,30E-03
Pir Superfamily	bromodomain	5	10,0	1,50E-03
INTERPRO	Bromodomain	5	9,3	2,00E-03
Groupe fonctionnel 4 - Score d'enrichissement DAVID : 2,63				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - CC	intracellular non-membrane-bounded organelle	50	1,6	6,70E-04
Terme GO - CC	non-membrane-bounded organelle	50	1,6	6,70E-04
Terme GO - CC	nucleolus	20	2,3	8,50E-04
Terme GO - CC	membrane-enclosed lumen	38	1,7	1,40E-03
Terme GO - CC	organelle lumen	37	1,7	1,90E-03
Terme GO - CC	intracellular organelle lumen	36	1,6	2,50E-03

Partie 2 - Annexe 4 - groupement fonctionnel DAVID (plusieurs bases de données)

Terme GO - CC	nuclear lumen	30	1,7	4,90E-03
Terme GO - CC	nucleoplasm	17	1,6	7,20E-02
Groupe fonctionnel 5 - Score d'enrichissement DAVID : 2,36				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	response to DNA damage stimulus	15	2,9	7,60E-04
Terme GO - BP	cellular response to stress	19	2,4	9,80E-04
Pir Superfamily	DNA damage	10	3,8	1,30E-03
Pir Superfamily	dna repair	9	3,7	3,10E-03
Terme GO - BP	DNA repair	11	2,8	6,60E-03
Terme GO - BP	DNA metabolic process	15	2,1	1,20E-02
Terme GO - BP	DNA replication	6	2,2	1,30E-01
Groupe fonctionnel 6 - Score d'enrichissement DAVID : 2,05				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	protein catabolic process	19	2,2	2,80E-03
Terme GO - BP	proteolysis involved in cellular protein catabolic process	18	2,1	4,50E-03
Terme GO - BP	cellular protein catabolic process	18	2,1	4,70E-03
Pir Superfamily	ubl conjugation pathway	15	2,3	5,70E-03
Terme GO - BP	modification-dependent protein catabolic process	17	2,1	6,70E-03
Terme GO - BP	modification-dependent macromolecule catabolic process	17	2,1	6,70E-03
Terme GO - BP	cellular macromolecule catabolic process	19	1,9	1,30E-02
Terme GO - BP	macromolecule catabolic process	20	1,8	1,30E-02
Terme GO - BP	proteolysis	20	1,4	1,50E-01
Groupe fonctionnel 7 - Score d'enrichissement DAVID : 2,02				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	cell cycle	26	2,4	8,00E-05
Pir Superfamily	cell cycle	18	3,1	8,80E-05
Pir Superfamily	cell division	11	3,3	2,10E-03
Terme GO - BP	cell cycle process	18	2,3	2,40E-03
Terme GO - BP	cell division	12	2,9	2,90E-03
Terme GO - BP	cell cycle phase	13	2,2	1,40E-02
Pir Superfamily	mitosis	7	3,0	3,00E-02
Terme GO - BP	nuclear division	8	2,6	3,50E-02
Terme GO - BP	mitosis	8	2,6	3,50E-02
Terme GO - BP	mitotic cell cycle	11	2,1	3,50E-02
Terme GO - BP	M phase of mitotic cell cycle	8	2,5	3,80E-02

Partie 2 - Annexe 4 - groupement fonctionnel DAVID (plusieurs bases de données)

Terme GO - BP	M phase	10	2,2	4,20E-02
Terme GO - BP	organelle fission	8	2,5	4,20E-02
Pir Superfamily	cytoskeleton	10	1,2	4,20E-01
Groupe fonctionnel 8 - Score d'enrichissement DAVID : 1,84				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - CC	kinetochore	7	7,4	3,50E-04
Terme GO - CC	chromosome, centromeric region	8	5,3	8,10E-04
Terme GO - CC	chromosome	14	2,5	4,20E-03
Terme GO - CC	condensed chromosome kinetochore	5	7,0	5,40E-03
Terme GO - CC	condensed chromosome, centromeric region	5	6,2	8,60E-03
Terme GO - CC	chromosomal part	11	2,3	2,00E-02
Terme GO - CC	condensed chromosome	5	3,2	7,30E-02
Terme GO - BP	chromosome segregation	4	3,5	1,00E-01
Pir Superfamily	kinetochore	3	3,7	1,90E-01
Terme GO - CC	chromatin	5	2,0	2,30E-01
Groupe fonctionnel 9 - Score d'enrichissement DAVID : 1,83				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - CC	histone acetyltransferase complex	5	8,1	3,20E-03
Terme GO - BP	histone acetylation	5	7,4	4,40E-03
Terme GO - BP	chromatin modification	11	2,9	5,20E-03
Terme GO - BP	protein amino acid acetylation	5	6,8	5,90E-03
Terme GO - BP	protein amino acid acylation	5	5,9	9,80E-03
Terme GO - BP	chromatin organization	12	2,3	1,70E-02
Terme GO - BP	chromosome organization	14	2,1	1,90E-02
Pir Superfamily	chromatin regulator	8	2,9	1,90E-02
Terme GO - BP	histone modification	6	3,5	2,80E-02
Terme GO - BP	histone H4 acetylation	3	11,0	3,10E-02
Terme GO - BP	covalent chromatin modification	6	3,4	3,20E-02
Terme GO - CC	nucleoplasm part	11	1,6	1,40E-01
Groupe fonctionnel 10 - Score d'enrichissement DAVID : 1,82				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - MF	nucleotide binding	46	1,6	1,70E-03
Pir Superfamily	atp-binding	30	1,8	2,90E-03
Pir Superfamily	nucleotide-binding	35	1,6	4,70E-03
Terme GO - MF	ATP binding	30	1,5	1,70E-02
Terme GO - MF	adenyl ribonucleotide binding	30	1,5	2,00E-02
Terme GO - MF	purine ribonucleotide binding	35	1,4	2,20E-02
Terme GO - MF	ribonucleotide binding	35	1,4	2,20E-02

Partie 2 - Annexe 4 - groupement fonctionnel DAVID (plusieurs bases de données)

Terme GO - MF	adenyl nucleotide binding	30	1,4	3,60E-02
Terme GO - MF	purine nucleotide binding	35	1,4	3,80E-02
Terme GO - MF	purine nucleoside binding	30	1,4	4,30E-02
Terme GO - MF	nucleoside binding	30	1,4	4,60E-02
Groupe fonctionnel 11 - Score d'enrichissement DAVID : 1,78				
Catégorie	Terme	Nombre	Enrich.	p-value
Pir Superfamily	protein transport	18	2,9	1,60E-04
Terme GO - BP	protein transport	24	2,2	4,00E-04
Terme GO - BP	establishment of protein localization	24	2,2	4,60E-04
Terme GO - BP	intracellular transport	20	2,2	2,10E-03
Terme GO - BP	protein localization	24	1,9	2,80E-03
Terme GO - BP	intracellular protein transport	12	2,3	1,60E-02
Terme GO - BP	cellular protein localization	12	2,1	3,00E-02
Terme GO - BP	cellular macromolecule localization	12	2,1	3,10E-02
Terme GO - BP	protein targeting	8	2,6	3,20E-02
Terme GO - BP	protein localization in organelle	6	2,9	5,40E-02
Terme GO - BP	nucleocytoplasmic transport	6	2,7	6,80E-02
Terme GO - BP	nuclear transport	6	2,7	7,10E-02
Terme GO - BP	protein import	5	2,7	1,10E-01
Pir Superfamily	transport	28	1,3	1,10E-01
Terme GO - BP	protein import into nucleus	4	3,3	1,20E-01
Terme GO - BP	nuclear import	4	3,2	1,30E-01
Terme GO - BP	protein localization in nucleus	4	3,0	1,40E-01
Groupe fonctionnel 12 - Score d'enrichissement DAVID : 1,61				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	vacuolar transport	4	8,9	9,90E-03
Terme GO - BP	endosome to lysosome transport	3	15,0	1,60E-02
Terme GO - BP	lysosomal transport	3	8,5	4,70E-02
Terme GO - BP	endosome transport	4	4,9	4,70E-02
Groupe fonctionnel 13 - Score d'enrichissement DAVID : 1,44				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	RNA processing	19	2,5	6,60E-04
Pir Superfamily	mrna splicing	9	3,4	5,30E-03
Terme GO - BP	mRNA processing	12	2,7	5,50E-03
Pir Superfamily	mrna processing	10	3,0	6,10E-03
Terme GO - CC	ribonucleoprotein complex	14	2,2	1,00E-02
Terme GO - BP	mRNA metabolic process	12	2,3	1,50E-02
Terme GO - BP	RNA splicing	9	2,3	4,60E-02

Partie 2 - Annexe 4 - groupement fonctionnel DAVID (plusieurs bases de données)

Pir Superfamily	Spliceosome	5	3,3	6,70E-02
Terme GO - BP	spliceosome assembly	3	6,7	7,30E-02
Terme GO - CC	spliceosome	5	3,1	7,80E-02
KEGG_PATHWAY	Spliceosome	5	2,8	1,00E-01
Terme GO - BP	RNA splicing, via transesterification reactions	5	2,3	1,70E-01
Terme GO - BP	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	5	2,3	1,70E-01
Terme GO - BP	nuclear mRNA splicing, via spliceosome	5	2,3	1,70E-01
Terme GO - BP	ribonucleoprotein complex biogenesis	5	2,0	2,40E-01
Terme GO - BP	ribonucleoprotein complex assembly	3	3,1	2,50E-01
Groupe fonctionnel 14 - Score d'enrichissement DAVID : 1,42				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - CC	clathrin coat	5	10,0	1,30E-03
Terme GO - CC	clathrin vesicle coat	4	17,0	1,50E-03
Terme GO - CC	coated vesicle	8	4,1	3,30E-03
Terme GO - CC	clathrin-coated vesicle	7	4,3	5,60E-03
Terme GO - CC	coated membrane	5	6,5	7,30E-03
Terme GO - CC	membrane coat	5	6,5	7,30E-03
Terme GO - CC	vesicle coat	4	8,6	1,10E-02
Terme GO - CC	cytoplasmic membrane-bounded vesicle	14	2,1	1,70E-02
Terme GO - CC	membrane-bounded vesicle	14	2,0	2,20E-02
Terme GO - CC	cytoplasmic vesicle	15	1,9	2,50E-02
Terme GO - CC	clathrin coated vesicle membrane	4	6,1	2,70E-02
Terme GO - CC	cytoplasmic vesicle part	7	3,0	2,70E-02
Terme GO - CC	cytoplasmic vesicle membrane	6	3,5	2,80E-02
Terme GO - BP	vesicle-mediated transport	15	1,9	3,10E-02
Terme GO - CC	vesicle	15	1,8	3,50E-02
Terme GO - CC	vesicle membrane	6	3,2	3,80E-02
Terme GO - CC	Golgi-associated vesicle membrane	3	8,4	4,80E-02
Terme GO - CC	clathrin adaptor complex	3	8,4	4,80E-02
Terme GO - CC	AP-type membrane coat adaptor complex	3	8,1	5,20E-02
Terme GO - CC	coated vesicle membrane	4	4,5	6,00E-02
Terme GO - CC	coated pit	3	5,6	1,00E-01
Pir Superfamily	Endocytosis	4	3,3	1,20E-01
Terme GO - CC	Golgi-associated vesicle	3	4,8	1,30E-01
Terme GO - CC	endomembrane system	14	1,5	1,60E-01
Terme GO - BP	endocytosis	6	1,9	2,00E-01
Terme GO - BP	membrane invagination	6	1,9	2,00E-01

Partie 2 - Annexe 4 - groupement fonctionnel DAVID (plusieurs bases de données)

Terme GO - BP	membrane organization	8	1,5	2,80E-01
Terme GO - CC	Golgi apparatus part	6	1,7	2,90E-01
Pir Superfamily	cytoplasmic vesicle	4	1,3	5,80E-01
Terme GO - CC	Golgi membrane	3	1,3	6,70E-01
Groupe fonctionnel 15 - Score d'enrichissement DAVID : 1,4				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	macromolecular complex subunit organization	20	2,0	5,00E-03
Terme GO - BP	macromolecular complex assembly	18	1,9	1,20E-02
Terme GO - BP	protein complex biogenesis	14	2,0	2,50E-02
Terme GO - BP	protein complex assembly	14	2,0	2,50E-02
Terme GO - BP	cellular macromolecular complex subunit organization	10	2,0	6,30E-02
Terme GO - BP	cellular macromolecular complex assembly	8	1,8	1,60E-01
Terme GO - BP	cellular protein complex assembly	4	1,8	4,00E-01
Groupe fonctionnel 16 - Score d'enrichissement DAVID : 1,4				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - MF	ATPase activity, coupled	11	3,1	3,10E-03
Terme GO - MF	ATPase activity	11	2,5	1,30E-02
Pir Superfamily	helicase	6	3,5	2,80E-02
Terme GO - MF	helicase activity	6	3,3	3,70E-02
Terme GO - MF	DNA-dependent ATPase activity	4	5,3	3,90E-02
Terme GO - MF	purine NTP-dependent helicase activity	5	3,9	4,00E-02
Terme GO - MF	ATP-dependent helicase activity	5	3,9	4,00E-02
INTERPRO	DEAD-like helicase, N-terminal	5	3,4	5,90E-02
INTERPRO	DNA/RNA helicase, ATP-dependent, DEAH-box type, conserved site	3	5,7	9,60E-02
INTERPRO	Helicase, superfamily 1 and 2, ATP-binding	4	2,8	1,70E-01
INTERPRO	DNA/RNA helicase, C-terminal	4	2,8	1,70E-01
Groupe fonctionnel 17 - Score d'enrichissement DAVID : 1,36				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	cell cycle phase	13	2,2	1,40E-02
Terme GO - BP	mitotic cell cycle	11	2,1	3,50E-02
Terme GO - BP	interphase of mitotic cell cycle	5	3,5	5,60E-02
Terme GO - BP	interphase	5	3,4	6,10E-02
Terme GO - BP	regulation of cell cycle	9	1,9	9,40E-02
Groupe fonctionnel 18 - Score d'enrichissement DAVID : 1,3				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	regulation of microtubule-based process	5	7,3	4,80E-03

Partie 2 - Annexe 4 - groupement fonctionnel DAVID (plusieurs bases de données)

Terme GO - BP	regulation of microtubule cytoskeleton organization	4	6,8	2,10E-02
Terme GO - BP	regulation of microtubule polymerization or depolymerization	3	6,9	6,90E-02
Terme GO - BP	regulation of cytoskeleton organization	5	2,6	1,20E-01
Terme GO - BP	regulation of organelle organization	5	1,6	3,60E-01
Groupe fonctionnel 19 - Score d'enrichissement DAVID : 1,25				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - CC	synaptosome	6	5,7	3,80E-03
Terme GO - CC	ruffle	3	3,6	2,00E-01
Terme GO - CC	cell leading edge	4	2,4	2,40E-01
Groupe fonctionnel 20 - Score d'enrichissement DAVID : 1,22				
Catégorie	Terme	Nombre	Enrich.	p-value
Pir Superfamily	transcription regulation	40	1,6	5,40E-03
Pir Superfamily	Transcription	40	1,5	7,70E-03
Terme GO - BP	regulation of transcription	51	1,4	8,90E-03
Terme GO - BP	transcription	41	1,4	2,30E-02
Terme GO - MF	transcription regulator activity	27	1,4	8,80E-02
Terme GO - BP	regulation of RNA metabolic process	33	1,3	9,50E-02
Terme GO - BP	regulation of transcription, DNA-dependent	31	1,2	1,50E-01
Terme GO - MF	DNA binding	37	1,2	1,60E-01
Terme GO - MF	transcription factor activity	14	1,1	5,10E-01
Pir Superfamily	dna-binding	22	0,9	7,50E-01
Groupe fonctionnel 21 - Score d'enrichissement DAVID : 1,21				
Catégorie	Terme	Nombre	Enrich.	p-value
Pir Superfamily	zinc-finger	35	1,6	6,20E-03
Pir Superfamily	zinc	41	1,5	1,10E-02
Terme GO - MF	zinc ion binding	43	1,4	1,40E-02
Terme GO - MF	transition metal ion binding	48	1,3	3,20E-02
Pir Superfamily	metal-binding	46	1,2	1,10E-01
Terme GO - MF	metal ion binding	58	1,1	3,50E-01
Terme GO - MF	cation binding	58	1,1	3,80E-01
Terme GO - MF	ion binding	58	1,0	4,30E-01
Groupe fonctionnel 22 - Score d'enrichissement DAVID : 1,07				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - BP	tRNA metabolic process	6	3,6	2,50E-02
Pir Superfamily	protein biosynthesis	7	2,9	3,30E-02
Terme GO - BP	ncRNA metabolic process	8	2,5	4,30E-02

Partie 2 - Annexe 4 - groupement fonctionnel DAVID (plusieurs bases de données)

INTERPRO	Aminoacyl-tRNA synthetase, class I, conserved site	3	6,6	7,60E-02
Pir Superfamily	Aminoacyl-tRNA synthetase	3	6,0	8,70E-02
INTERPRO	Rossmann-like alpha/beta/alpha sandwich fold	3	5,6	1,00E-01
KEGG_PATHWAY	Aminoacyl-tRNA biosynthesis	3	5,1	1,10E-01
Terme GO - MF	aminoacyl-tRNA ligase activity	3	4,8	1,30E-01
Terme GO - MF	ligase activity, forming carbon-oxygen bonds	3	4,8	1,30E-01
Terme GO - MF	ligase activity, forming aminoacyl-tRNA and related compounds	3	4,8	1,30E-01
Terme GO - BP	tRNA aminoacylation for protein translation	3	4,6	1,40E-01
Terme GO - BP	amino acid activation	3	4,6	1,40E-01
Terme GO - BP	tRNA aminoacylation	3	4,6	1,40E-01
Groupe fonctionnel 23 - Score d'enrichissement DAVID : 1				
Catégorie	Terme	Nombre	Enrich.	p-value
Terme GO - CC	membrane fraction	16	1,6	6,90E-02
Terme GO - CC	insoluble fraction	16	1,6	8,80E-02
Terme GO - CC	cell fraction	18	1,4	1,70E-01

MATERIEL ET METHODES

Matériel et méthodes

I. Matériel

I.1. Souches bactériennes, cellules humaines et plasmides

I.1.1. Souches bactériennes

Souche *E. coli* DH5 α : *supE44*, Δ *lacU169* (Φ 80 *lacZ* Δ M15), *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*. Cette souche est utilisée pour la production et la purification de plasmides.

Souche *E. coli* XL1-Blue : *recA1 endA*, *gyrA9 thi- hsdR1 supE44 relA1 lac* (F' *proAB lacIqZ* Δ M15 *Tn10* (TetR)) (Stratagène). Cette souche a été utilisée pour la production et la purification de plasmides.

Souche *E. coli* BL21-Codon Plus (DE3)-RIL : F- *ompT hsdS* (rB-mB-) *dcm* + TetR *gal* λ (DE3) *endA*Hte [*argU ileY leuW* CamR] (Stratagène). Cette souche a été utilisée pour la production des protéines recombinantes.

I.1.2. Cellules humaines

Cellules HeLa : N°ATCC® : CCL-2™ (Scherer et al., 1953), une lignée cellulaire cancéreuse provenant d'un prélèvement effectué sur une patiente atteinte d'un cancer du col de l'utérus et décédée en 1951, **Henrietta Lacks**.

Fibroblastes : Les lignées de fibroblastes proviennent d'un patient HGPS de sexe féminin (AG11513) et d'un patient féminin non atteint de même âge (AG08470) (Coriell Cell Repositories, Camden, NJ).

I.1.3. Plasmides et constructions génétiques

Plasmides pUC18::T7-LMNA 638 ; pUC18::T7-LMNA 103 ; pUC18::T7-LMNA 48 : Ces plasmides dérivés de pUC18 ont permis de produire les ARN utilisés pour les analyses de structures et les études d'interaction avec SRSF1. La carte du plasmide pUC18 avec les différents segments du gène *LMNA* insérés dans ce plasmide est représentée dans la figure I-1. Ces constructions avaient été réalisées par I. Behm-Ansmant et F. Freyermuth à partir d'un plasmide fourni par J. Tazi. Ce plasmide de départ contenait l'ADN codant l'ARN LMNA 638 WT ou muté 1824C>T.

Plasmide pUC18::T7-MS2 et Plasmide pUC18::T7- MiniHur-MS2 : La construction pADML-M3, renfermant à l'extrémité 3' du minigène ADML le fragment d'ADN correspondant à l'ARN MS2, a été

réalisée par l'équipe de B. Wang et nous a été fournie par l'équipe du Pr. R. Lührmann. Ce fragment d'ADN a été placé sous le contrôle du promoteur T7 (figure I-1) dans le plasmide pUC18::T7 permettant ainsi de transcrire *in vitro* l'ARN MS2. La séquence MiniHur a été insérée entre les sites de restriction *Bam*HI et *Kpn*I.

Plasmide pMBP-MS2 : Ce plasmide a été donné par l'équipe de Lührmann. Nous ne disposons pas de carte de ce plasmide, mais il a été créé par le laboratoire de R. Reed. Il dérive d'un plasmide pMAL (New England Biolabs), qui permet de produire chez *E. coli* des protéines contenant une étiquette MBP en N-ter, et possède une résistance à l'ampicilline. Il permet l'expression de la protéine de fusion MBP-MS2.

Plasmides pCDNA3.1/V5-His TOPO::T7-βGlo LMNA (J. Tazi) : La carte de ce plasmide construit par J. Tazi et les différents plasmides variants que j'ai produits par mutagenèse dirigée (1822G>A, 1868C>G et 1819G>A) sont représentés dans la figure I-2. Ces constructions ont été utilisées pour la production par transcription *in vitro*, d'ARN testés en épissage *in vitro*.

Plasmide pCRII-TOPO::SP6-pHur98-T7 : Ce plasmide a été offert par l'équipe de G. Biamonti. Il contient une construction SP6-pHur98(antisens)-T7(antisens) insérée dans un site de restriction *Eco*RI dans la cassette de clonage du vecteur pCRII-TOPO (Invitrogen). La digestion par *Eco*RI permet la libération de cette construction et la transcription par l'ARN polymérase T7 de l'ARN pHur98 G-rich, et par l'ARN polymérase SP6 de l'ARN pHur98 C-rich (figure I-3).

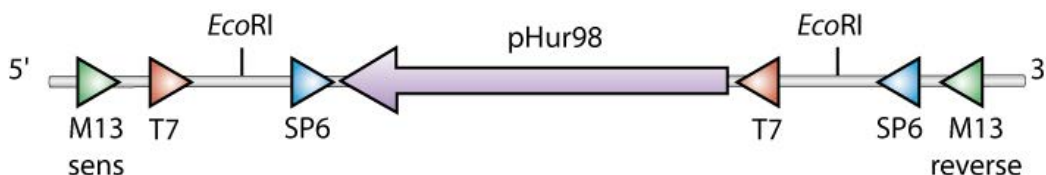


Figure I-3 : Site de clonage du plasmide pCRII-TOPO::SP6-pHur98::T7. La séquence SP6-pHur98-T7 a été insérée dans la cassette de clonage du plasmide au niveau du site de restriction *Eco*RI.

Plasmides pCINeo::λN-HA : Ce plasmide et les plasmides dérivés ont été donnés par l'équipe de E. Izaurralde. La carte du plasmide pCINeo::λN-HA et les plasmides dérivés sont détaillés dans la figure I-4. Ces plasmides ont servi à la réalisation des expériences de "tethering" de la protéine GW182 (TNRC6A). Pour cela, plusieurs dérivés du plasmide ont été utilisés : pCINeo::λN-HA-TNRC6A ou pCINeo::λN-HA-MBP, dans lesquels les séquences codant les protéines GW182 ou la Maltose Binding Protein ont été insérés.

Plasmide pCINeo::Rluc-5BoxB et pEGFP::FLuc : Ce plasmide est pourvu d'un gène rapporteur Renilla luciférase suivi des séquences BoxB nécessaires aux expériences de tethering. Le normalisateur

de l'activité luciférase utilisé dans ces expériences est le plasmide commercial pEGFP dans lequel a été inséré la séquence codant la protéine Fluc (Firefly Luciférase), qui est également un don de l'équipe de E. Izaurralde.

Plasmide pXJ41-ASF, pXJ41-SRp40 : Ces plasmides ont été donnés par l'équipe de J. Stevenin. Nous n'avons pas de carte précise de ces plasmides, ils portent la résistance à l'ampicilline. Ils permettent la production des protéines SR dans les lignées cellulaires humaines. Le pXJ41, qui est notre plasmide contrôle a été produit au laboratoire en retirant la séquence codante pour ASF.

I.2. Oligodésoxyribonucléotides

Les oligodésoxyribonucléotides utilisés lors de ce travail sont décrits dans le tableau I-1. Les sondes TaqMan utilisées pour la détection des microARN ne figurent pas dans ce tableau. Elles sont pré-conçues et commercialisées par la société life technologies.

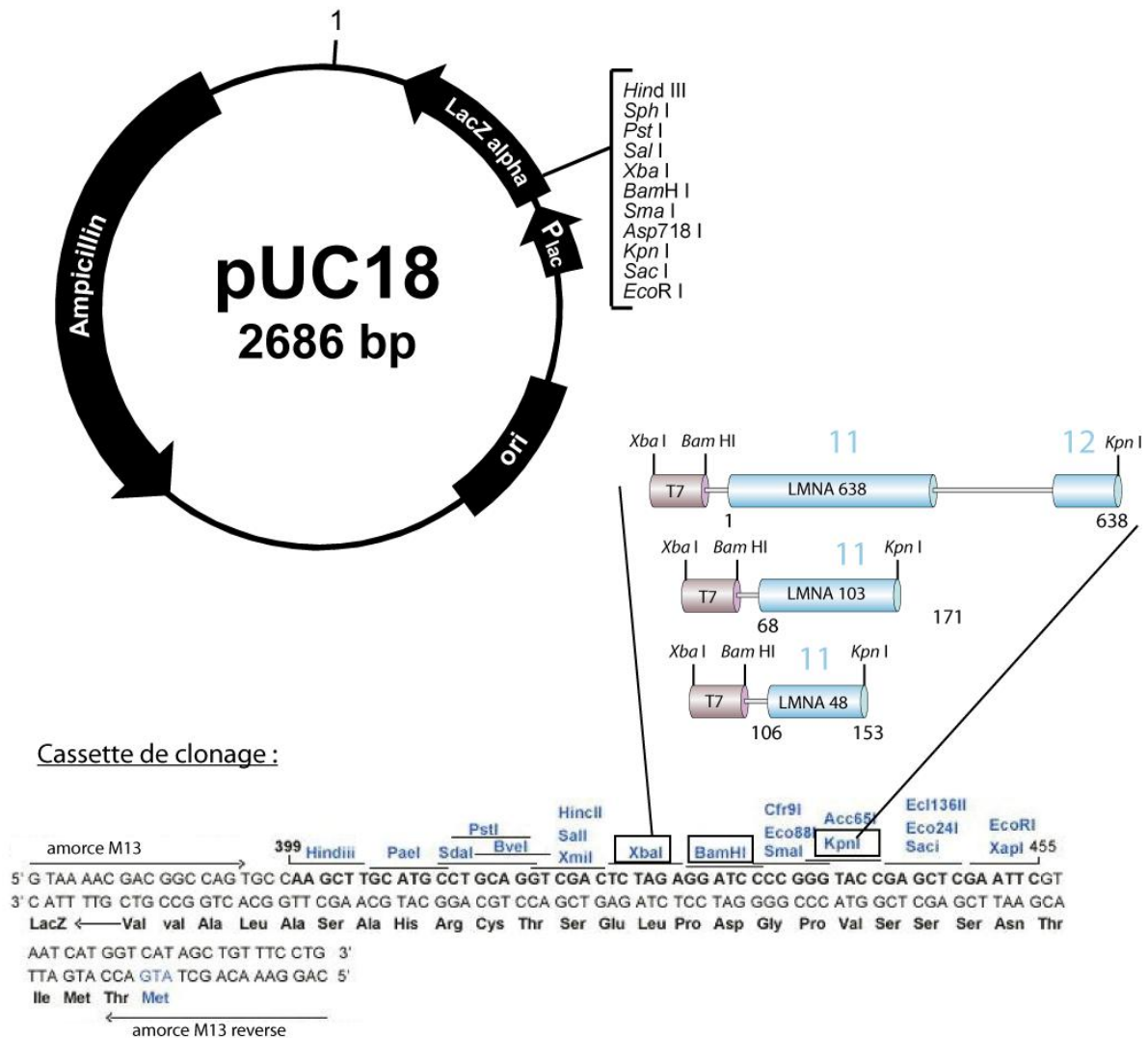
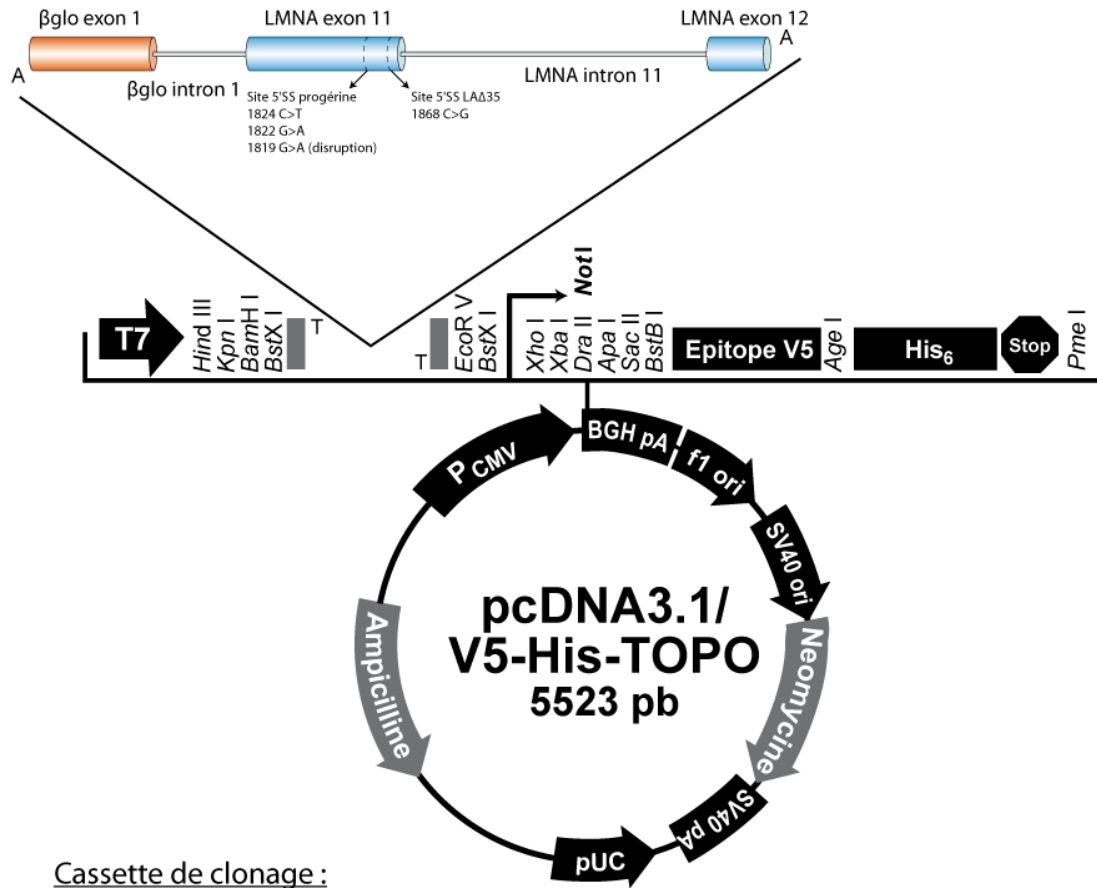


Figure I-1 : Carte du plasmide pUC18. Ce plasmide a été utilisé précédemment au laboratoire pour cloner, entre les sites de restriction BamHI et KpnI, des fragments du gène *LMNA* de différentes tailles (LMNA 48, LMNA 103 et LMNA 638), WT ou porteurs de la mutation 1824C>T. Le promoteur T7 a été inséré en amont de ces fragments entre les sites de restriction XbaI et BamHI. Les constructions obtenues sont indiquées au dessus de la cassette de clonage. Les cylindres bleus correspondent aux exons 11 et 12 du gène *LMNA* (numéros indiqués au dessus des segments ADN insérés), les positions indiquées en dessous correspondent à la numérotation utilisée pour l'ARN *LMNA* 638, sachant que le premier nucléotide de l'ARN 638 correspond au premier nucléotide de l'intron 11 du gène *LMNA*. Les mutations 1822G>A et 1868C>G ont été introduites dans ces constructions par mutagenèse dirigée par PCR.

Le plasmide pUC18 contient :

- La partie lacΔM15 du gène de la β-galactosidase (fragment N-terminal), permettant l'α-complémentation,
- Une cassette de clonage (détaillée en dessous de la carte du plasmide),
- Une origine de répliation (ori),
- Un gène conférant la résistance à l'ampicilline.



Cassette de clonage :

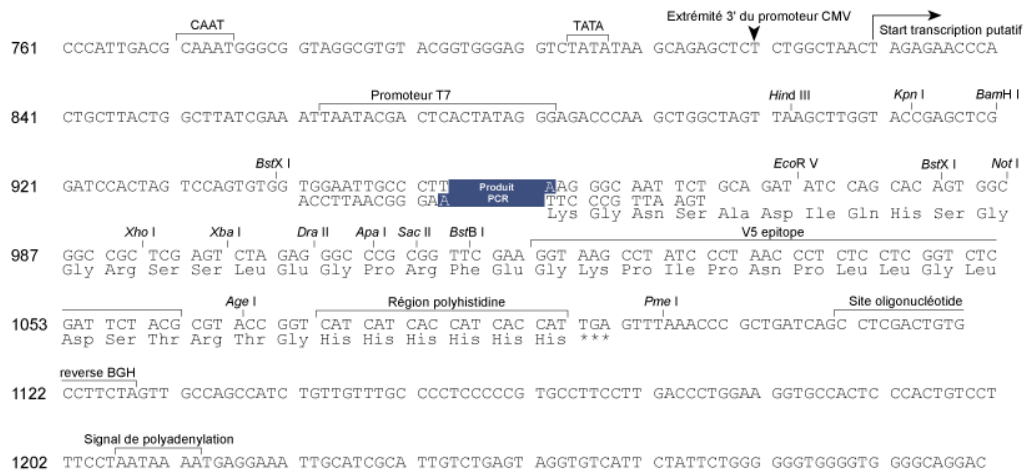
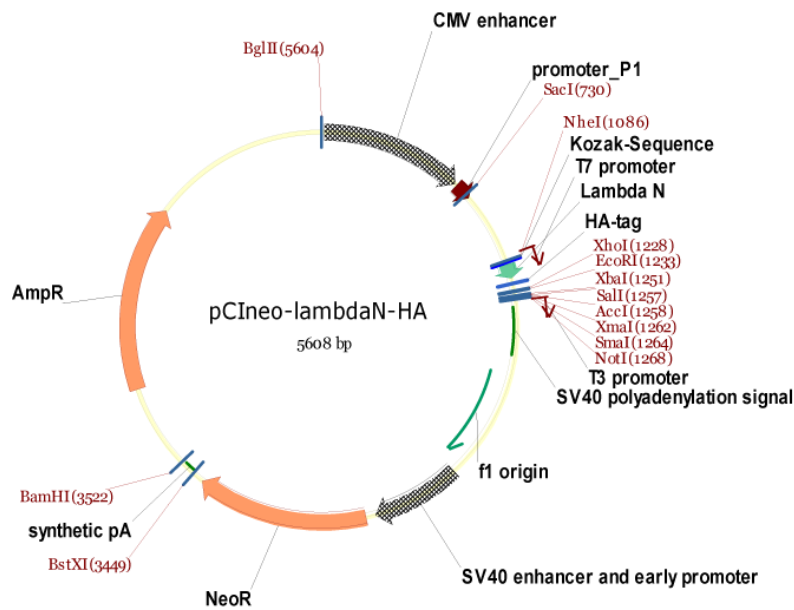


Figure I-2 : Carte de restriction du plasmide pcDNA3.1/V5-His-TOPO (Invitrogen). Ce plasmide a été utilisé pour cloner, grâce à la technique de clonage T/A, le segment d'ADN codant le mini-pré-ARNm hybride LMNA-βGlo indiqué au dessus, qui a été utilisé pour les expériences d'épissage *in vitro*. Les plasmides contenant la séquence WT ou mutée en position 1824C>T dans le fragment du gène LMNA, nous ont été fournis par l'équipe de J.Tazi. J'ai produit par mutagenèse dirigée les plasmides portant les autres mutations indiquées sous l'insert. Le plasmide V5-His-TOPO contient :

- un gène de résistance à l'ampicilline et un gène de résistance à la néomycine,
- Une origine de répliation dans les cellules eucaryotes issue du virus SV40,
- Un signal de polyadénylation aussi issu du virus SV40,
- Deux origines de répliation de type pUC et f1,
- Une séquence promotrice du virus CMV,
- Un site de polyadénylation BGH
- Une cassette de clonage (MCS) détaillée sous la carte du plasmide.



Site de clonage du plasmide pCIneo

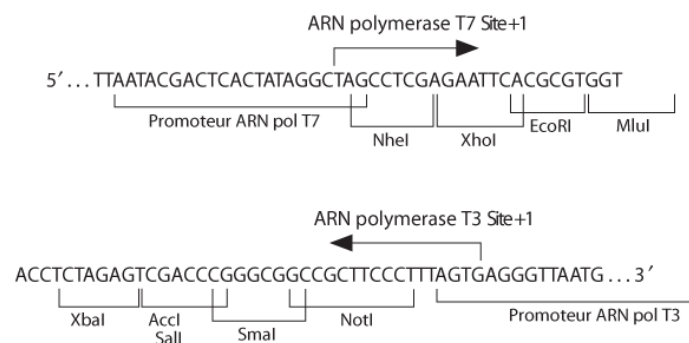


Figure I-4 : Carte du plasmide pCIneo-λN-HA. Ce plasmide est obtenu à partir du clonage des séquences étiquettes λN et HA entre les séquences de restriction NheI et XhoI du plasmide pCIneo. Les séquences codant les protéines GW182 ou MBP ont également été clonées dans ce plasmide entre les sites de restriction XhoI et NotI. Dans le plasmide pCIneo::Rluc-5BoxB, la séquence codant la luciférase Rluc a été clonée entre les sites de restriction NheI et XbaI, tandis que les 5 séquences BoxB ont été clonées entre les sites de restriction XbaI et XhoI. Ce plasmide contient :

- Une origine de réplication dans les cellules eucaryotes issue du virus SV40,
- Une cassette de clonage (MCS) détaillée sous la carte du plasmide,
- Deux origines de réplication bactérienne f1,
- Une séquence promotrice du virus CMV, et un signal de polyadénylation issu du virus SV40, avant et après le MCS, respectivement
- Deux séquences codant les protéines de résistance à l'ampicilline et la néomycine

Numéro	Séquence de 5' vers 3'	Taille (nt)	Description
967	CTAATACGACTCACTATAGGG	21	oligonucléotide universel s'hybridant au niveau du promoteur T7. Surtout utilisé comme oligonucléotide 5' de PCR pour la synthèse de matrices de transcription
4914	ATAAAGAGCTCGAATTCCGTACCCTG	26	s'hybride en 3' de la séquence MS2 et du site de restriction <i>SacI</i> contenu dans le -plasmide pUC18::T7-MS2 (utilisé pour la synthèse de matrices de transcription par PCR)
5478	TGACCAGGGACCCGCCTGCAG	21	oligonucléotide amorce de rétrotranscription de l'ARN LMNA 638 (positions 320-300) (étude de la structure secondaire)
5480	GAAGCTGCCACCCCACTGCCCCC	25	oligonucléotide amorce de rétrotranscription de l'ARN LMNA 638 (positions 213-189) (étude de la structure secondaire)
5576	GGCTCCCACTGCAGCAGCTCG	21	s'hybride en 5' de l'ARN LMNA 638
5725	ATAAAGGATCCGGCTCCCACTGCAGCAGCTCG	32	s'hybride en 5' de l'ARN LMNA 638 et porteur en 5' d'un site <i>Bam</i> HI (utilisé pour le clonage dans pUC18::T7)
5727	ATAAAGGTACCCTGCCTGGCAGGTCCAGATTACATGA	38	s'hybride en 3' de l'ARN LMNA 638 et porteur en 5' d'un site <i>Kpn</i> I (utilisé pour le clonage dans pUC18::T7)
5737	ATAAAGGATCCGGCTCAGGAGCCAGGTGG	40	s'hybride en 5' de l'ARN LMNA 48 et porteur en 5' d'un site <i>Bam</i> HI (utilisé pour le clonage dans pUC18::T7 et la synthèse de matrices de transcription par PCR)
5738	ATAAAGGATCCGGACCTGCGGGCAGCCTGCC	40	s'hybride en 5' de l'ARN LMNA 103 et porteur en 5' d'un site <i>Bam</i> HI (utilisé pour le clonage dans pUC18::T7 et la synthèse de matrices de transcription par PCR)
5739	ATAAAGGTACCGGCAGAAGAGCCAGAGGAGAT	40	s'hybride en 3' de l'ARN LMNA 48 et porteur en 5' d'un site <i>Kpn</i> I (utilisé pour le clonage dans pUC18::T7 et la synthèse de matrices de transcription par PCR)
5740	ATAAAGGTACCGTGACCGTGACACTGGAGGCA	40	s'hybride en 3' de l'ARN LMNA 103 et porteur en 5' d'un site <i>Kpn</i> I (utilisé pour le clonage dans pUC18::T7 et la synthèse de matrices de transcription par PCR)
6547	TCAGGAGCCCAGGTGAGCGGACCCATCTCCT	31	oligonucléotide 5' mutagenèse LMNA G608S
6548	AGGAGATGGGTCCGCTCACCTGGGCTCCTGA	31	oligonucléotide 3' mutagenèse LMNA G608S
6600	CCAGTGTCACGGTCAGTCGCAGCTACCGCAG	31	oligonucléotide 5' mutagenèse LMNA C1868G

6601	CTGCGGTAGCTGCGACTGACCGTGACACTGG	31	oligonucléotide 5' mutagenèse LMNA C1868G
6717	GGCTCAGGAGCCCAGATGGGCGGACCCATCT	31	oligonucléotide 5' mutagenèse LMNA G1819A (destruction site alternatif)
6718	AGATGGGTCCGCCATCTGGGCTCCTGAGCC	31	oligonucléotide 3' mutagenèse LMNA G1819A (destruction site alternatif)
7181	AGCCCCGAACCCAGATGAGTTGTCTCTGCT	31	oligonucléotide 5' mutagenèse LMNA G1969A (destruction site canonique)
7182	AGCAGAGACAACCTCATCTGGGTTCTGGGGGCT	31	oligonucléotide 3' mutagenèse LMNA G1969A (destruction site canonique)
7239	TTCCATTCCATTCTGTACT	20	oligonucléotide reverse pour analyse par RT de la structure secondaire de l'ARN HuR98 (ARN sat III)
7365	GCTTCTGACACAACCTGTGTTCACTAGC	27	amplification PCR de produits d'épissage rétrotranscrits obtenus à partir de la construction β Glo-LMNA (oligonucléotide sens, sur l'exon 11)
7507	GCAGTTCTGGGGGCTCTGG	19	[fourni par I. Lopez de l'équipe de J. Tazi] amplification PCR de produits d'épissage rétrotranscrits obtenus à partir de la construction β Glo-LMNA (oligonucléotide antisens). Chevauche la jonction exon 11-12 obtenue avec le site canonique et avec le site alternatif.
7554	GGATCCAATCAACCCGAGTGCAATCG	26	oligonucléotide sens s'hybridant en 5' de l'ARN Hur98 et porteur d'un site <i>Bam</i> HI pour clonage dans pUC18::T7-MS2
7555	GGTACCTTCCATTCCATTCTGTACT	26	oligonucléotide antisens s'hybridant en 3' de l'ARN Hur98 et porteur d'un site <i>Kpn</i> I pour clonage dans pUC18::T7-MS2
7653	ATAAAGGATCCTTCCATTCCATTCTGTACT	26	oligonucléotide sens s'hybridant en 5' de l'ARN antisens Hur98 (C-rich) et porteur d'un site <i>Bam</i> HI pour clonage dans pUC18::T7-MS2
7654	ATAAAGGTACCAATCAACCCGAGTGCAATCG	26	oligonucléotide antisens s'hybridant en 3' de l'ARN antisens Hur98 (C-rich) et porteur d'un site <i>Kpn</i> I pour clonage dans pUC18::T7-MS2
7741	ATAAAGGATCCGGAATGGCATGGATTGGAATGGAATG GAATGGAATCAACCCGAGTGGTACCATAAA	67	matrice ADN pour PCR, comprenant l'ADN MiniHur flanqué des sites de restriction <i>Bam</i> HI et <i>Kpn</i> I pour le clonage dans le plasmide pUC18::T7-MS2 pour transcription
7742	ATAAAGGATCCGGAATGGCA	20	oligonucléotide 5' pour PCR sur la matrice oligonucléotide 7741, comprenant un site de restriction <i>Bam</i> HI pour le clonage dans le pUC18::T7-MS2
7743	TTTATGGTACCACTCGGGTT	20	oligonucléotide 3' pour PCR sur la matrice oligonucléotide 7741, comprenant un site de restriction <i>Kpn</i> I pour le clonage dans le pUC18::T7-MS2

7795	AATCAACCCGAGTGCAATCG	20	RT-PCR pour la détection d'ARN sat III dans les extraits ARN de cellules HeLa soumises à un stress (oligonucléotide de PCR sens, nommé Hur98R) (Valgardsdottir et al., 2008)
7796	CCGTAAAACGACGGCCAGTTCCTTCCATTCCATTATTA TCC	42	RT-PCR pour la détection d'ARN sat III dans les extraits ARN de cellules HeLa soumises à un stress (oligonucléotide de rétrotranscription, RSM13)
7828	CGTGTTGGAACACCCCCAC	20	amplification par PCR d'un fragment du cDNA hsp70-A1A (oligonucléotide sens) (Valgardsdottir et al., 2008)
7829	TTCCGTTTCCAGCCCCCAATC	21	amplification par PCR d'un fragment du cDNA hsp70-A1A (oligonucléotide antisens) (Valgardsdottir et al., 2008)
7868	GGAGCCAGGTGGGCAGACCCATCTCCTCTG	31	oligonucléotide 5' mutagenèse LMNA G1825A
7869	CAGAGGAGATGGGTCTGCCACCTGGGCTCC	31	oligonucléotide 3' mutagenèse LMNA G1825A
8536	AACAGCCTCAAGATCATCAGC	21	qPCR sur l'ARNm GAPDH rétrotranscrit (Valgardsdottir et al., 2008) (hGP_F)(FW)
8537	GGATGATGTTCTGGAGAGCC	20	qPCR sur l'ARNm GAPDH rétrotranscrit (Valgardsdottir et al., 2008) (hGP_R)(REV)
8538	GGTCGTCAGACTGTCGATGAAGCC	24	séquençage du gène malE codant la Maltose Binding Protein chez E.Coli (contenu dans le plasmide pmal-c2)
8539	CCGTAAAACGACGGCCAG	24	séquence M13 telle qu'utilisée pour la nested qRT-PCR de l'ARN sat III dans Valgardsdottir et al., 2008
8772	GAAGTCACTGTGCCAGCCCA	20	qPCR Rplp0 sur l'ARN rplp0 rétrotranscrit pour normalisation (oligonucléotide sens), (Fugier et al., 2011)
8773	GAAGGTGTAATCCGTCTCCA	20	qPCR Rplp0 sur l'ARN rplp0 rétrotranscrit pour normalisation (oligonucléotide antisens), (Fugier et al., 2011)
9121	GTGATTACAATCGAACGGTC	20	vérification inclusion/exclusion de l'exon 12 TNRC6A par RT-qPCR (oligonucléotide sens, sur l'exon 11)
9122	TGTGCTGGAGGCTGCTGCAT	20	vérification inclusion/exclusion de l'exon 12 TNRC6A par RT-qPCR (oligonucléotide antisens, sur l'exon 13)
9128	AGTACTTCACATCGTCGTTT	20	vérification inclusion/exclusion de l'exon 4 de CLK1 par RT-qPCR (oligonucléotide sens, sur l'exon 3)

9129	ATGATCGATGCACTCCACAA	20	vérification inclusion/exclusion de l'exon 4 de CLK1 par RT-qPCR (oligonucléotide antisens, sur l'exon 5)
9157	TGCGAGTCAACACCCATGGCA	21	vérification inclusion/exclusion de l'exon 13 HSPH1 par RT-qPCR (oligonucléotide sens, sur l'exon 12)
9158	TTAGCTTCTGGAGGCTGGTCA	21	vérification inclusion/exclusion de l'exon 13 HSPH1 par RT-qPCR (oligonucléotide antisens, sur l'exon 14)
9159	GGGCTTGAAGGTTGACCAGCA	21	vérification inclusion/exclusion de l'exon 12 de EIF4ENIF1 par RT-qPCR (oligonucléotide sens, sur l'exon 11)
9160	GCTCCCCATAAGGCCACTCAGA	22	vérification inclusion/exclusion de l'exon 12 de EIF4ENIF1 par RT-qPCR (oligonucléotide antisens, sur l'exon 13)
9161	GCCATCTTTCCAGCCTCAGT	20	vérification inclusion/exclusion de l'exon 2 de RAN par RT-qPCR (oligonucléotide sens, sur l'exon 1)
9162	ATGACGTTTCACGAAGGTCGT	21	vérification inclusion/exclusion de l'exon 2 de RAN par RT-qPCR (oligonucléotide antisens, sur l'exon 3)
9385	AATGGCAATCCCAGTATGTTTGGTGTGGA	30	oligonucléotide 5' de mutagenèse : amplification en délétant l'exon 13 de GW182 du plasmide pCIneo-lambdaN-HA-GW182 (expériences tethering GW182)
9386	CTTATCAAATAAGGATTGCGCTCCATGGA	30	oligonucléotide 3' de mutagenèse : amplification en délétant l'exon 13 de GW182 du plasmide pCIneo-lambdaN-HA-GW182 (expériences tethering GW182)
9504	TTTGTAGAGGGGGTCGACAGT	21	vérification inclusion/exclusion de l'exon 5 de CNOT2 par RT-PCR (oligonucléotide sens, sur l'exon 4)
9505	AGTGCCTTGTGATAAGCTGCGA	22	vérification inclusion/exclusion de l'exon 5 de CNOT2 par RT-PCR (oligonucléotide antisens, sur l'exon 6)
9506	AGGGATTTTAGCCAGTGTGGA	21	vérification inclusion/exclusion de l'exon 13 de CNOT4 par RT-PCR (oligonucléotide antisens, sur l'exon 12)
9507	TGGAGCTTGCAGTGAGCCGAG	21	vérification inclusion/exclusion de l'exon 13 de CNOT4 par RT-PCR (oligonucléotide sens, sur l'exon 14)
9508	ATTTCCAGGTGTTGTGGTGCGA	22	vérification inclusion/exclusion de l'exon 4 de CNOT8 par RT-PCR (oligonucléotide sens, sur l'exon 3)

9509	GGCAAACGAGAATCTGTAAGCA	22	vérification inclusion/exclusion de l'exon 4 de CNOT8 par RT-PCR (oligonucléotide antisens, sur l'exon 5)
9510	TTGACTTGGGTTATCTGTCCAC	22	vérification inclusion/exclusion de l'exon 6 de CRM1 par RT-PCR (oligonucléotide antisens, sur l'exon 5)
9511	ATCAAGACGTCATCTGACCCA	21	vérification inclusion/exclusion de l'exon 6 de CRM1 par RT-PCR (oligonucléotide sens, sur l'exon 7)
9512	TCCACATCCTGCCGGATCTGT	22	vérification inclusion/exclusion de l'exon 2 de Transportin-SR par RT-PCR (oligonucléotide antisens, sur l'exon 1)
9513	CAGGCGCTTACCACGACCCAG	21	vérification inclusion/exclusion de l'exon 2 de Transportin-SR par RT-PCR (oligonucléotide sens, sur l'exon 3)
9514	AACACAGATCGCCTCTCAGCA	21	vérification inclusion/exclusion de l'exon 5 de HSPH1 par RT-PCR (oligonucléotide antisens, sur l'exon 4)
9515	AGATTCATGGCCGAGCATTC	21	vérification inclusion/exclusion de l'exon 5 de HSPH1 par RT-PCR (oligonucléotide sens, sur l'exon 6)
9516	TAATCCCTGGGGTAGAACCACG	22	vérification inclusion/exclusion de l'exon 6 de Nop58 par RT-PCR (oligonucléotide sens, sur l'exon 5)
9517	CAGGGAAATGCCAGCCATACC	21	vérification inclusion/exclusion de l'exon 6 de Nop58 par RT-PCR (oligonucléotide antisens, sur l'exon 7)

Tableau I - 1 : Oligonucléotides utilisés lors de ce travail

I.3. Milieux de culture et tampons

Milieu Luria-Bertani (LB) : tryptone 20 g/L ; extrait de levure 5 g/L ; NaCl 5 g/L. Le pH est ajusté à 7,5 avec du NaOH 10 N.

Milieu LB supplémenté en Ampicilline (LBA) : milieu LB contenant de l'ampicilline à une concentration finale de 100 µg/mL.

Milieu LBAC : milieu LBA complété par du chloramphénicol à une concentration finale de 25 µg/ml.

Milieus LB solides : milieux LB et dérivés dans lequel est ajouté de l'agar à 17,5 g/L.

Les milieux de culture sont autoclavés 20 min à 120°C. Les antibiotiques sont ajoutés après refroidissement du milieu.

Phosphate Buffered Saline (PBS) 10 x : NaCl 1,37 mM, KCl 27 mM, Na₂HPO₄.12H₂O 100 mM, KH₂PO₄ 18 mM. Ajuster le pH à 7,4.

Solution de bleu de dépôt BP pour les protéines : Tris-HCl 80 mM pH 6,8 ; SDS 3% (v/v) ; glycérol 10% (v/v); DTT 100 mM ; Bleu de bromophénol 0,1%.

Solution de bleu de dépôt sur gel d'agarose (6x) : EDTA 125 mM ; SDS 2 % (p/v) ; glycérol 25 % (v/v) ; bleu de bromophénol 0,12 % (p/v).

Solution de bleu dénaturant (bleu de formamide) pour ADN et ARN (2x) : EDTA 20 mM ; bleu de bromophénol 0,05 % (p/v) ; bleu de xylène cyanol 0,05 % (p/v), dans de la formamide désionisée.

Solution de bleu CSB non dénaturant : glycérol 40 % (v/v), bleu de bromophénol 0,03 % (p/v) ; xylène cyanol 0,03 % (p/v) ; HEPES-KOH 200 mM pH 7,9.

Solution SBL (« Sample Buffer Laemmli ») : SDS 2 % (p/v) ; glycérol 10 % (v/v) ; β-mercapto-éthanol 100 mM ; bleu de bromophénol 0,01 % (p/v) ; Tris-HCl 63,65 mM pH 6,8.

Tampon de blocage RNA FISH : BSA 3% ; SSC4x ; tween 0,1%.

Tampon d'hybridation RNA FISH : BSA 1% (p/v) ; SSC4x ; tween 0,1% (v/v) ; Streptavidine Alexa 488 2 µg/mL.

Tampon D : MgCl₂ 1,5 mM ; EDTA 0,2 mM ; KCl 150 mM ; HEPES-KOH 20 mM pH 7,9.

Tampon d'éluion ARN et ADN : NaCl 100 mM ; EDTA 1 mM ; SDS 1 % (p/v) ; Tris-HCl 10 mM pH 7,5. Ce tampon est filtré avec un filtre nitrocellulose (Millipore) 0,22 µm.

Tampon d'hybridation pour la rétrotranscription (10x) : KCl 400 mM ; Tris-HCl 500 mM pH 8,3.

Tampon de lyse protéines / extraction protéines : Tris/HCl 50 mM pH 8, NaCl 150 mM, Igepal 1 % (v/v), EDTA 1 mM, Halt™ protease&phosphatase inhibitor cocktail-EDTA free (Thermo scientific) (10 µL/mL)

Tampon TfbI (préparation de bactéries compétentes) : KAc 30 mM; KCl 100 mM; CaCl₂ 10 mM ; Glycérol 15% (v/v) ; MnCl₂ 50 mM.

Tampon TfbII (préparation de bactéries compétentes) : MOPS Na pH7 10 mM ; KCl 10 mM ; CaCl₂ 75 mM ; Glycérol 15% (v/v).

Tampon de transcription (5x) : DTT 200 mM ; spermidine 10 mM ; MgCl₂ 120 mM ; HEPES-KOH 400 mM pH 7,5.

Tampon de transfert pour western blot : Tris 25 mM ; Glycine 200 mM ; Ethanol 20 % (v/v).

Tampon de sonication : Tris-HCl 20 mM pH 7,6 ; NaCl 200 mM ; inhibiteur de protéase EDTA-free (Roche).

Tampon PBS-Tween 0,1 % : NaH₂PO₄ 10 mM ; KH₂PO₄ 1,8 mM ; NaCl 150 mM ; KCl 2,7 mM ; Tween 0,1 % (v/v). Le pH est ajusté à 7,4.

Tampon TBS-Tween 0,1 % : NaCl 150 mM ; KCl 2,7 mM ; Tween 0,1 % (v/v) ; Tris-HCl 140 mM pH 7,4.

Tampon de transcription US (5x) : MgCl₂ 30 mM ; spermidine 5 mM ; Tris-HCl 200 mM pH 7,5.

Tampon Tris-NaCl réservé à l'épissage : NaCl 200 mM ; Tris-HCl 10 mM pH 7,5.

Tampon protéinase K (2x) : NaCl 0,3 M ; SDS 2 % (v/v) ; EDTA 25 mM pH 7,4 ; Tris-HCl pH 7,5.

I.4. Anticorps

Protéine cible	Fournisseur	Caractéristiques	Organisme	Dilution
hnRNP H	Abcam (ab10374)	polyclonal	lapin	IF : 1/500 WB : 1/20000
hnRNP K	Immuquest (IQ213) Clone 3C2	monoclonal	souris	WB : 1/20000
hnRNP U	Abcam (ab10297), clone 3G6	monoclonal	souris	IF : 1/100 WB : 1/200-1/1000
SRSF2(phosphorylée)	Abcam (ab11826)	monoclonal	souris	IF : 1/1000
SRSF3	Santa cruz (sc-13510)	monoclonal	souris	IF : 1/50
hnRNP A1	Santa cruz (sc-32301)	monoclonal	souris	IF : 1/100 WB : 1/200
hnRNP R	Santa cruz (sc-16541)	polyclonal	chèvre	IF : 1/100 WB : 1/200
hnRNP R	Atlas antibodies (HPA026092)	polyclonal	lapin	IF : 1/50-1/100
hnRNP Q	Santa cruz (sc-16541)	monoclonal	souris	IF : 1/100 WB : 1/200
SRSF1	Santa cruz (sc-73016)	monoclonal	souris	IF : 1/100
SRSF1	Abcam (ab38017)	polyclonal	lapin	IF : 1/20
SRSF2	Abcam (ab11826)	monoclonal	souris	IF : 1/1000 WB : 1/5000
SRSF2	J. Stévenin (1SCF11)	monoclonal	souris	WB : 1/1250
SRSF3	Santa cruz (sc-13510) (clone 7B4)	monoclonal	mouse	WB : 1/200
SRSF4	Santa cruz (sc-51204)	polyclonal	chèvre	WB : 1/200
hnRNP M	Santa cruz (sc-134360)	monoclonal	souris	IF : 1/100 WB : 1/200
hnRNP M (isoformes 3 et 4)	Abcam (ab55486)	monoclonal	souris	IF : 1/100
SAF-B/hnRNP HAP	Santa cruz (sc-135618)	monoclonal	souris	IF : 1/100 WB : 1/200
Sam68	Santa cruz (sc-136062)	monoclonal	souris	IF : 1/100 WB : 1/200
GW182	Santa cruz (sc-56314)	monoclonal	souris	IF : 1/25 - 1/100 WB : 1/2000
EDC4	Sigma (HPA041164)	polyclonal	lapin	IF : 1/50 - 1/300
HSF1	Thermo scientific (MA5- 14632)	monoclonal	rat	IF : 1/502
Anticorps secondaires				
Protéine cible	Fournisseur	Caractéristiques	Organisme	Dilution
Anticorps de souris	Invitrogen	Couplé à l'Alexa555 ou l'Alexa488	Chèvre	IF : 1/100
Anticorps de souris	Pierce	Couplé à la peroxydase		WB : 1/10000
Anticorps de lapin	Invitrogen	Couplé à l'Alexa555 ou l'Alexa488	Chèvre	IF : 1/100
Anticorps de lapin	Pierce	Couplé à la peroxydase		WB : 1/10000
Anticorps de chèvre	Invitrogen	Couplé à l'Alexa555	Âne	IF : 1/100
Anticorps de chèvre	Pierce	Couplé à la peroxydase		WB : 1/10000

Tableau I-2 : Détails des anticorps utilisés lors de ce travail. Pour chaque anticorps est référencé une protéine cible, une référence fournisseur accompagné du nom du clone lorsqu'il est fourni, l'organisme de production et la clonalité ainsi que les spécificités de l'anticorps si il y en a. La dernière colonne indique le facteur de dilution utilisées pour les applications mises en œuvre (WB : western blot, IF : Immunofluorescence).

I.5. Micropuces jonctions exon-exon et microARN

Les puces utilisées pour l'analyse des isoformes d'ARNm produites avant et après choc thermique sont des modèles expérimentaux Affymetrix HJAY non commercialisés. Le descriptif complet des puces est disponible dans la base de données NCBI GEO (Gene Expression Omnibus) à l'adresse : <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL8444>. Ces puces contiennent des probesets dirigés contre 315 137 exons et 260 488 jonctions exon-exon. Cela couvre l'ensemble des isoformes répertoriés dans la version NCBI36/hg18 de la base de données d'assemblage des séquences du génome humain.

Les puces utilisées pour l'analyse de l'expression des microARN sont les puces GeneChip® miRNA 3.0 Array (Affymetrix). Le descriptif complet des puces est disponible auprès du fabricant. Ces puces contiennent notamment 19 931 probesets reconnaissant les miRNA matures de nombreux organismes, dont 1733 microARN spécifiques à l'homme, et dérivés de la version Sanger miRBase v17 (www.mirbase.org). Elles contiennent également 2 999 probesets dirigés contre des pré-miR (1658 pré-miR humains, le restant chez la souris et le rat) 2 216 probesets dirigés contre d'autres ARN non codants chez l'homme comme les snoARN et scaARN.

II. Méthodes d'analyse *in vitro*

II.1. Introduction des mutations par mutagenèse dirigée

Les séquences d'intérêt avaient été précédemment clonées dans le plasmide pUC18. Une amplification PCR est réalisée avec les oligodésoxyribonucléotides permettant d'introduire la mutation. Le détail de la réaction de PCR mise en place pour réaliser les mutagenèses dirigées est décrit dans le tableau II-1 ci-dessous.

Réactifs	Volume
- Tampon Pfu 5x (Fermentas)	5 μ L
- Pfu DNA polymerase (Fermentas ; 2,5 U/ μ L)	1 μ L
- Mélange de dNTP (1,25 mM chaque)	8 μ L
- Matrice plasmidique (200 ng/ μ L)	1 μ L
- Oligodésoxyribonucléotide sens (10 ng/ μ L)	1 μ L
- Oligodésoxyribonucléotide antisens (10 ng/ μ L)	1 μ L
- H ₂ O	qsp 50 μ L

Nombre de cycles	Etape	Température	Durée
1	Dénaturation initiale	94°C	5 min
18	Dénaturation	94°C	1 min
	Hybridation	55°C	1 min
	Elongation	72°C	8 min
1	Elongation finale	72°C	10 min
1	Stockage	20°C	∞

Tableau II-1 : Réaction de PCR pour l'introduction d'une mutation par mutagenèse dirigée. La première partie du tableau indique les quantités de réactifs et la deuxième partie les cycles de la PCR.

Les produits PCR sont ensuite soumis à une digestion par l'enzyme *DpnI* (Fermentas) hydrolysant l'ADN parental méthylé, à raison de 10 UE par réaction. Les produits de la digestion sont utilisés pour transformer par choc thermique 50 μ L de cellules *E. coli* DH5 α compétentes, en incubant le mélange 30 min dans la glace, 45 sec à 42°C puis 5 min dans la glace. Les clones bactériens ayant incorporé le plasmide sont sélectionnés sur gélose en milieu LBA, puis mis en culture dans 2 mL de milieu LBA liquide. Les plasmides sont extraits par lyse alcaline grâce au kit Nucleospin Plasmid

(Macherey-Nagel), puis sont envoyés à la société GATC pour leur séquençage. Les clones d'intérêt sont remis en culture dans environ 150 mL de LBA pour la préparation d'ADN plasmidique à plus grande échelle à l'aide du kit Nucleobond AX (Macherey-Nagel). Les quantités de plasmides obtenues sont estimées par spectrophotométrie à 260 nm (spectrophotomètre SAFAS).

II.2. Production de matrices pour la transcription

Les matrices sont produites par amplification PCR à partir des plasmides recombinants contenant les séquences d'intérêt, en utilisant l'oligodésoxyribonucléotide sens 937 (promoteur T7) combiné à un oligodésoxyribonucléotide antisens dont l'identité dépend de la séquence à amplifier dans le cas des ARN LMNA (tableau I-1). Le protocole de PCR utilisé est indiqué dans le tableau II-2. Les produits PCR obtenus sont fractionnés en gel d'agarose à 1 % contenant du GelRed (Fluoprobes), un agent intercalant utilisé pour la visualisation, dilué selon les spécifications du fabricant. Les bandes visibles par illumination aux UV à 254 nm sont découpées et l'ADN est purifié grâce au kit GeneElute (Fermentas). L'élué est réalisé dans un volume de 20 µL.

Réactifs	Volume
- Tampon DreamTaq 5x (Fermentas)	10 µL
- DreamTaq ADN polymerase (5 U/µL, Fermentas)	0,5 µL
- Mélange de dNTP (1,25 mM chaque)	8 µL
- Oligodésoxyribonucléotide 5' (100 ng/µL)	1 µL
- Oligodésoxyribonucléotide 3' (100 ng/µL)	1 µL
- Matrice plasmidique (10 ng/µL)	1 µL
- H ₂ O	qsp 50 µL

Tableau II-2 : Réaction de PCR pour la synthèse de matrices de transcription.

II.3. Production d'ARN par transcription *in vitro*

Une fraction des produits PCR purifiés (20 µL) est utilisée comme matrice de la transcription par l'ARN polymérase du phage T7. Les réactifs utilisés sont décrits dans le tableau II-3. Le mélange réactionnel est incubé pendant 12 h à température ambiante. Le produit de la transcription est soumis à l'action de l'enzyme DNase I (Fermentas) pendant 30 min à 37°C. Les ARN sont ensuite chauffés pendant 2 min à 96°C, puis purifiés sur gel de polyacrylamide dénaturant à 5 ou 10 % selon la taille de l'ARN à purifier. Les bandes d'ARN sont révélées par « UV shadowing » à 254 nm, découpées et éluées

2 fois dans 200 μL de tampon d'éluion pendant 30 min à 37°C. Les protéines sont retirées par un mélange phénol/chloroforme, et les ARN sont précipités à -80°C en présence de 2,5 volumes d'éthanol 96 % et d'acétate de sodium à une concentration finale de 0,3 M. Les culots sont lavés dans 400 μL d'éthanol 70 %, séchés et repris dans 10 μL d'eau stérile. La quantité d'ARN obtenue est estimée par spectrophotométrie à 260 nm.

Réactifs	Volume
- Tampon de transcription 5x	10 μL
- rNTP (25 mM chaque)	15 μL
- Produit PCR purifié	20 μL
- ARN polymérase T7 (produite au laboratoire)	5 μL

Tableau II-3 : Réaction de transcription *in vitro*.

II.4. Etude de la structure secondaire des ARN par emploi de sondes enzymatiques

II.4.1. Marquage radioactif en 5' d'oligodésoxyribonucléotides ou d'ARN

Marquage en 5' d'oligodésoxyribonucléotides : 100 ng d'oligodésoxyribonucléotides sont marqués radioactivement en mélangeant 1 μL d'ATP [γ - ^{32}P] (Perkin Elmer, 10 mCi/mL ; 3000 Ci/mmol) à 1 μL d'enzyme T4 polynucléotide kinase (Fermentas ; 10 U/ μL). Le mélange est incubé pendant 45 min à 37°C, puis purifié sur gel de polyacrylamide dénaturant à 10 %. Les bandes radioactives sont repérées grâce à un film radiosensible, découpées et éluées 2 fois dans 200 μL de tampon d'éluion pendant 30 min à 65°C. Les oligodésoxyribonucléotides sont extraits par un mélange phénol/chloroforme et soumis à une précipitation à -80°C, comme décrit dans le paragraphe V.2. Les culots sont repris dans de l'eau stérile de manière à obtenir une concentration d'environ 500 cps/ μL .

Déphosphorylation et marquage en 5' d'ARN : 2 μg d'ARN sont déphosphorylés par action de la phosphatase alcaline de veau (1 UE CIAP) (Fermentas) pendant 1 h à 37°C. Ils sont ensuite extraits par un mélange phénol/chloroforme, puis précipités à -80°C et centrifugés. Les culots sont repris dans 14 μL d'eau stérile. La moitié (7 μL) de cette solution est utilisée pour le marquage radioactif, qui se déroule de la même manière que le marquage des oligodésoxyribonucléotides, décrit ci-dessus, sauf que l'éluion est réalisée à 37°C au lieu de 65°C.

II.4.2. Hydrolyse des ARN par les RNases T1, T2 et V1

Tout d'abord, l'ARN est dénaturé pendant 10 min à 65°C puis renaturé par refroidissement progressif à température ambiante pendant 15 min. Pour une seule réaction enzymatique, 300 ng d'ARN renaturé sont soumis à une hydrolyse par de la RNase T1 (Roche), la RNase T2 (Invitrogen) ou la RNase V1 (Kemotech), selon une gamme de quantités 0,1 UE, 0,5 UE et 1 UE pour les enzymes T1 et T2, et 10^{-4} UE, $5 \cdot 10^{-4}$ UE et 10^{-3} UE pour l'enzyme V1. Les enzymes sont employées diluées dans du tampon D. Le mélange est incubé 6 min à température ambiante, la réaction est arrêtée par ajout de : 1) 100 μ L d'une solution comprenant de l'EDTA 4 mM et 20 μ g d'ARNt, 2) un volume de phénol/chloroforme. Les ARN sont ensuite extraits, puis précipités à -80°C en présence de 1 μ L de glycogène et centrifugés. Les culots d'ARN sont repris dans 3 μ L d'une solution de Kborate 25 mM. La composition des différentes solutions utilisées est détaillée dans le tableau II-4 ci-dessous.

Composition du mélange de dénaturation/renaturation	Volume
ARN non marqué (300 ng/ μ L)	1 μ L
ARNt (2 μ g/ μ L)	1 μ L
Tampon D 10x	1 μ L
H ₂ O stérile	qsp 9 μ L

Composition de la solution d'arrêt de la réaction	Volume
EDTA 500 mM	0,8 μ L
ARNt (20 μ g/ μ L)	1 μ L
H ₂ O stérile	100 μ L

Tableau II-4 : Réaction d'hydrolyse enzymatique par les RNases T1, T2 et V1. La deuxième partie du tableau détaille la composition de la solution utilisée pour l'arrêt de la réaction enzymatique.

II.4.3. Identification des positions modifiées par rétro-transcription des ARN non radiomarqués

Cent nanogrammes d'ARN modifiés non radiomarqués sont hybridés pendant 10 min à 65°C avec 100 cps d'oligodésoxyribonucléotide radiomarqué en 5' (voir plus haut) dans le tampon d'hybridation, puis incubés pendant 10 min dans la glace. Les hybrides formés sont ensuite incubés avec une solution d'extension décrite dans tableau II-5. Le mélange réactionnel est incubé pendant 30 min à 42°C. Trois microlitres de bleu de formamide dénaturant sont ajoutés à chaque échantillon. Ceux-ci sont

alors chauffés 2 min à 96°C et fractionnés en gel de séquençage polyacrylamide dénaturant 7 %. Les ARN sont révélés par emploi d'un Phosphorimager Typhoon (GE Healthcare) ou par autoradiographie en présence d'un écran amplificateur. Le séquençage par réverse transcription de l'ARN non modifié est également réalisé en parallèle, grâce au principe de la réaction de Sanger. Pour cela, 100 ng d'ARN sont soumis à une rétrotranscription tel que décrit ci-dessus, en ajoutant en plus 0,5 µL de ddATP, ddCTP, ddGTP ou ddTTP à 0,5 mM permettant l'arrêt de la réaction lors de leur incorporation aléatoire pour visualiser les position A, C, G ou T respectivement. Le protocole utilisé est présenté dans le tableau II-5.

Composition de la solution d'hybridation	Volume
ARN non marqué (100 ng/µL)	1 µL
Tampon d'hybridation 10x	0,25 µL
Oligodésoxyribonucléotide radiomarqué (100 cps/µL)	1 µL
H ₂ O stérile	qsp 2,5 µL
Composition de la solution d'extension	Volume
dNTP MIX (5mM chaque)	0,1 µL
Tampon de rétrotranscription 10x (Q-BIOgene)	0,275 µL
Rétro-transcriptase AMV (Q-BIOgene ; 20 U/µL)	0,025 µL
H ₂ O stérile	qsp 2,5 µL

Volumes utilisés pour une réaction de séquençage				
	A	C	G	T
ddATP 0,5 mM	0,5 µL	-	-	-
ddCTP 0,5 mM	-	0,5 µL	-	-
ddGTP 0,5 mM	-	-	0,5 µL	-
ddTTP 0,5 mM	-	-	-	0,5 µL
Solution d'extension	2,5 µL	2,5 µL	2,5 µL	2,5 µL
Solution d'hybridation	2,5 µL	2,5 µL	2,5 µL	2,5 µL

Tableau II-5 : Réaction de rétrotranscription des produits ARN d'hydrolyse enzymatique par les RNases T1, T2 et V1. Les différentes parties du tableau correspondent au détail des réactifs utilisés lors des deux étapes de cette réaction, l'étape d'hybridation de l'amorce à l'ARN cible et l'étape d'extension par la rétrotranscriptase. La dernière partie du tableau correspond aux volumes de ddNTP utilisés pour permettre le séquençage par la réaction de Sanger.

II.5. Etude de l'interaction des protéines avec les ARN

II.5.1. Purification par chromatographie d'affinité MBP-MS2

II.5.1.a. Production et purification de la protéine de fusion MBP-MS2

Des bactéries *E. coli* BL21 (λ DE3) sont transformées par le plasmide d'expression codant la protéine de fusion MBP-MS2 (pMBP-MS2) donné par le laboratoire de Luhrmann. Une pré-culture de 4 ml réalisée à 37°C sous agitation à partir d'un clone transformé, est transférée dans 1 L de milieu LB complété avec l'antibiotique adéquat. La culture est incubée à 37°C sous agitation jusqu'à une absorbance de 0,8 UA_{600nm}/ml. La production de la protéine est induite par ajout d'IPTG à 0,1 mM final et la culture est poursuivie pendant 4 h à 37°C. Les bactéries sont ensuite collectées par centrifugation à 4000 rpm pendant 15 min à 4°C. Le culot cellulaire est repris dans du tampon de sonication et les cellules sont lysées par 2 traitements de 2 min aux ultrasons, interrompus par 2 min d'incubation dans la glace. Le surnageant de sonication contenant la protéine MBP-MS2 est mis en contact pendant 30 à 45 min sur roue (agitation lente) à 4°C avec 1,8 mL de résine amylose (New England Biolabs) préalablement lavée 2 fois avec 10 mL de tampon de sonication. La protéine de fusion MBP-MS2 est retenue sur les billes par l'intermédiaire de son domaine MBP. Après 3 lavages de la résine avec du tampon PBS (L1, L2 et L3), la protéine MBP-MS2 est éluée en 2 fractions de 1 mL (E1 et E2) par l'application d'une solution contenant du maltose à 10 mM.

II.5.1.b. Chromatographie d'affinité MBP-MS2

L'intégralité de cette chromatographie est réalisée dans des conditions minimisant la contamination à la kératine afin de ne pas perturber l'identification ultérieure des protéines par spectrométrie de masse. Les ARN de fusion portant les séquences MS2 (300 pmoles) sont dilués dans 50 μ l de tampon D, dénaturés et renaturés comme décrit précédemment en présence de 3 mM de MgCl₂. Après cela, un aliquot de 5 μ L (fraction C1, "input") est prélevé pour une analyse ultérieure. Les ARN sont ensuite incubés en présence d'un excès molaire de 5 fois de protéine recombinante MBP-MS2 pendant 30 min dans la glace, suite à quoi un aliquot (C2) est prélevé. A ce stade, l'interaction des protéines MBP-MS2 avec leur ARN cible est vérifiée par expérience de retard sur gel grâce à l'électrophorèse des fractions C1 et C2 sur un gel de polyacrylamide natif (voir aussi § II.5.3., retard sur gel) (figure II-1). Les ARN sont révélés par exposition aux UV en présence de révélateur GelRed (FluoProbes).

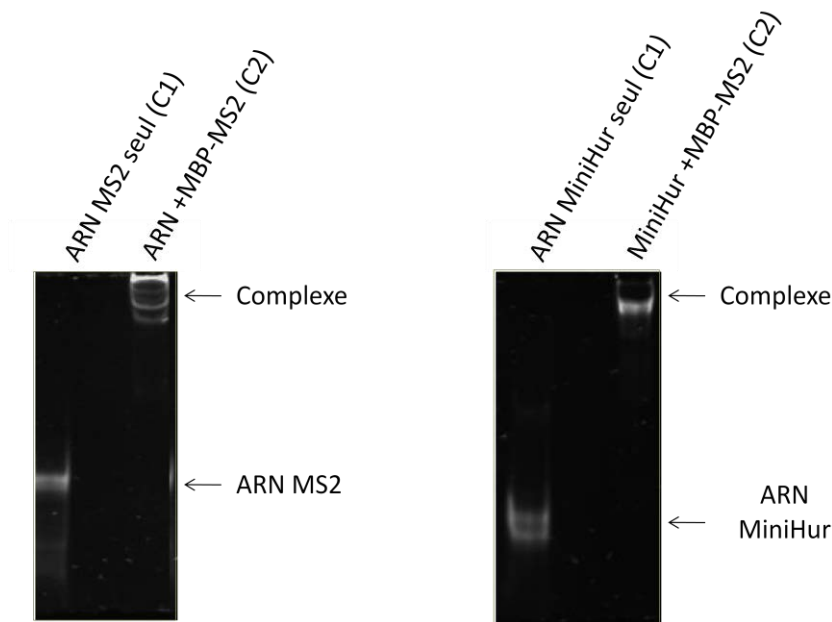


Figure II-1 : Vérification par retard sur gel de l'interaction entre les ARN MS2 et la protéine MS2 durant la chromatographie d'affinité. Pour cela, les fractions C1 et C2 obtenues avec l'ARN témoin MS2 ou avec l'ARN MiniHur-MS2 sont déposées sur gel de polyacrylamide natif 10 ou 8%, respectivement. Le retard de migration témoigne de la formation d'un complexe ARN-protéine.

Le mélange ARN-MS2/MBP-MS2 est ensuite mis en contact avec 40 μ l d'une suspension de résine amylose (New England Biolabs), préalablement lavée 3 fois avec 500 μ l de tampon D, pendant 1 h sous agitation lente à 4°C. Cette étape permet de fixer le complexe sur les billes d'amylose par l'intermédiaire de la protéine MBP. Les billes sont centrifugées 1 min à 1000 rpm puis sont lavées 3 fois avec du tampon D avant d'ajouter 250 μ l d'extrait nucléaire ou cytoplasmique de cellules HeLa (CilBiotech) et des ARNt en excès 50x molaire en présence de 3 mM de $MgCl_2$ et de 0,1% de BSA. L'ensemble est incubé sous agitation lente à 4°C pendant 40 min. Les billes sont centrifugées 1 min à 1000 rpm, et un aliquot du surnageant est conservé pour analyse (fraction Non Retenue NR). Les billes sont ensuite lavées 3 fois pendant 15 min sous agitation lente à 4°C avec 700 μ l de tampon D. A chaque lavage est prélevé un aliquot pour analyse (fractions Lavage 1 à Lavage 3). Enfin, les complexes RNP sont élués par addition de 100 μ l d'une solution de maltose à 10 mM pendant 30 min sous agitation lente à température ambiante (fraction Elution 1). L'opération est répétée une seconde fois pour une élution plus totale (fraction Elution 2). Un aperçu du contenu en protéines, notamment MBP-MS2, est obtenu grâce à l'électrophorèse des différentes fractions collectées au cours de cette purification sur gel SDS-PAGE 10% (figure II-2).

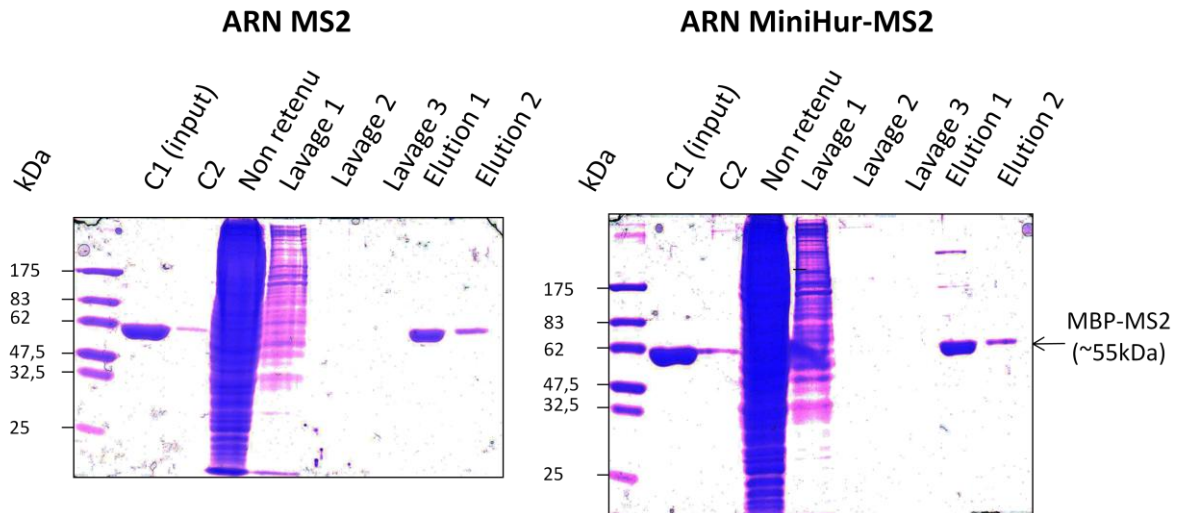


Figure II-2 : Vérification de la présence des protéines MBP-MS2 dans les fractions récoltées lors de la chromatographie MBP-MS2. La détection des protéines est effectuée par électrophorèse des différentes fractions sur gel dénaturant SDS-PAGE 10%. Ces fractions sont déposées dans l'ordre chronologique du déroulement de la chromatographie (voir texte), effectuée en présence de l'ARN témoin MS2 ou de l'ARN d'intérêt MiniHur fusionné aux séquences MS2.

Pour vérifier que les ARN MS2 et MiniHur-MS2 ont bien été récupérés lors des étapes d'éluion, les fractions de purification sont également déposées sur gel de polyacrylamide-urée dénaturant au pourcentage adapté (figure II-3).

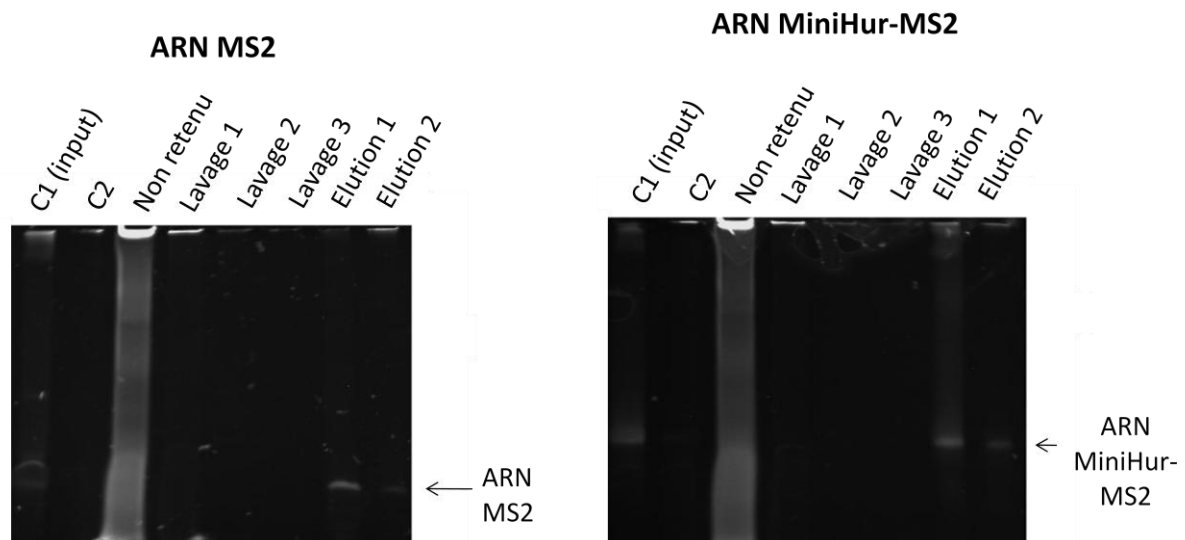


Figure II-2 : Vérification de la présence des ARN MS2 et MiniHur-MS2 dans les fractions récoltées lors de la chromatographie MBP-MS2. La détection de ces ARN est effectuée par électrophorèse des différentes fractions sur gel dénaturant polyacrylamide urée à 8 ou 10%. Ces fractions sont déposées dans l'ordre chronologique du déroulement de la chromatographie (voir texte), effectuée en présence de l'ARN témoin MS2 (à gauche) ou de l'ARN d'intérêt MiniHur fusionné aux séquences MS2 (à droite).

Pour l'analyse du contenu en protéines par spectrométrie de masse, une première étape a consisté à séparer les protéines de la fraction d'éluion (environ 50%) sur gel d'électrophorèse SDS-PAGE (voir section résultats pour l'image du gel). Le gel est fixé par incubation dans 3 bains d'acide acétique 10 % - éthanol 20 % pendant 10 min, puis coloré au bleu colloïdal toute la nuit. Les bandes du gel sont alors découpées de manière systématique et leur contenu en protéines est analysé par spectrométrie de masse en collaboration avec l'équipe d'Alain Vandoersaeler (LSMBO, Strasbourg). Pour cela, les protéines contenues dans les bandes du gel ont été digérées à la trypsine, extraites du gel par un solvant ACN/H₂O/HCOOH (50/50/1), puis ionisées selon la technique d'électrospray et enfin analysées par nanoLC MS/MS (Miguet et al., 2006).

II.5.2. Estimation des quantités de protéines par « western blot »

Un microlitre d'extrait nucléaire ou 1 μ L d'éluat sont fractionnés en gel SDS-PAGE dénaturant à 10 %. Parallèlement, la migration d'un marqueur de taille (Pre-stained Biorad marker ; Biorad) permet de contrôler la masse moléculaire des protéines étudiées. Après migration, le transfert a lieu sur membrane de nitrocellulose 0.45 μ m pendant 1 h à 100 V dans du tampon de transfert refroidi. La membrane est ensuite saturée avec 5 mL de lait écrémé 5 % (p/v), pendant 1 h à température ambiante sous agitation, puis lavée 3 fois dans du tampon PBS-Tween 0,1 % (v/v), et incubée avec la dilution adéquate d'anticorps primaire pendant environ 12 h à 4°C sous agitation. La membrane est à nouveau lavée avec du PBS-T et incubée avec la dilution adéquate d'anticorps secondaire pendant 45 min à température ambiante. La membrane est une dernière fois lavée et incubée pendant 5 min avec une solution de révélation contenant un substrat peroxydase et du luminol, fournie dans le kit Immobilon Western (Millipore). La révélation se fait par chimioluminescence sur un film photosensible, ou grâce à un scanner Fusion Solo (Vilber Lourmat) et au logiciel d'acquisition Fusion (pour les western blots du manuscrit de la publication N°2 concernant la mutation de l'ARN LMNA 1868C>G).

II.5.3. Expériences de retard sur gel

Ces expériences sont réalisées avec 47 fmoles d'ARN non radiomarqué, en présence de 5 cps du même ARN radiomarqué (voir plus haut), de MgCl₂ 2,8 mM et d'ARNt en excès 100 molaire par rapport à l'ARN non radiomarqué. Les ARN sont dénaturés pendant 10 min à 65°C puis renaturés pendant 15 min par refroidissement progressif à température ambiante. Deux microlitres de différentes solutions de protéines à des concentrations croissantes spécifiées sont additionnés au mélange, qui est alors incubé pendant 30 min à température ambiante. La protéine est diluée dans du tampon D contenant 10 % de glycérol, selon une gamme allant de 0 à 1000 fois ou 0 à 10 000 fois la quantité molaire d'ARN non radiomarqué. Trois microlitres et demi de bleu CSB non dénaturant sont ajoutés à chaque échantillon.

Ceux-ci sont ensuite fractionnés sur gel de polyacrylamide 19 :1(acrylamide : bis-acrylamide) natif de 3,5 à 8 %, selon la taille de l'ARN non radiomarké utilisé. La révélation des ARN se fait par autoradiographie du gel sur un film radiosensible.

II.6. Epissage *in vitro* de l'ARN LMNA et de ses variants

II.6.1. Préparation des matrices de transcription par hydrolyse de plasmide

Cinq microgrammes de plasmides pCDNA3.1/V5-His TOPO::T7-βGlo LMNA sont linéarisés par digestion avec l'enzyme *Not* I utilisée dans du tampon O (Fermentas), à raison de 5 UE/μg d'ADN, dans un volume total de 10 μL. L'efficacité d'hydrolyse est vérifiée par fractionnement des produits de la réaction en gel d'agarose à 1 %. Les protéines sont retirées par un mélange de phénol/chloroforme et les plasmides sont précipités à -80°C en présence de 1 μL de glycogène puis centrifugés. Les culots sont repris dans un volume de 10 μL d'eau stérile.

II.6.2. Production *in vitro* de transcrits uniformément radiomarqués

Cinq cent nanogrammes de plasmide linéarisé sont utilisés comme matrice pour la réaction de transcription *in vitro*. Le radiomarquage est effectué grâce à l'incorporation de rUTP [α -³²P] lors de la transcription. La composition en réactifs pour ce radiomarquage est détaillée dans le tableau II-6.

Réactifs	Volume
- Tampon de transcription US (5x)	2 μL
- DTT 0,1 M	1 μL
- Tampon RiboLock (Fermentas)	0,25 μL
- Inhibiteur de RNases : RiboLock (40 U/μL - Fermentas)	0,25 μL
- rNTP A,C,G (5,5,3 mM)	1 μL
- rUTP 1 mM	0,5 μL
- BSA (2 μg/μL)	0,5 μL
- Cap (m ⁷ GpppGp) 10 mM	1,25 μL
- rUTP [α - ³² P] (Perkin-Elmer, 800 Ci/mmol, 20 mCi/mL)	0,9 μL
- H ₂ O	0,85 μL
- ARN polymérase T7 (produite au laboratoire)	0,5 μL
- Matrice : plasmide linéarisé (environ 500 ng/μL)	1 μL

Tableau II-6 : Réactifs utilisés pour la production de transcrits uniformément radiomarqués *in vitro* par incorporation de rUTP [α -³²P].

II.6.3. Épissage *in vitro* des transcrits uniformément radiomarqués

La réaction d'épissage *in vitro* est réalisée dans un extrait nucléaire de cellules HeLa (EN 2303 ; CilBiotech) avec 10 fmoles d'ARN uniformément radiomarqués. Les réactifs utilisés pour cette réaction sont indiqués dans le tableau II-7.

Solution contenant l'ARN	Volume
- MgCl ₂ (31,25 mM)	1 µL
- Créatine phosphate 625 mM	1 µL
- rATP dilué extemporanément (25 mM)	1 µL
- Tampon RiboLock (Fermentas)	0,2 µL
- Inhibiteur de RNases : RiboLock (40 U/µL - Fermentas)	0,2 µL
- ARN uniformément radiomarqué (10 fmol/µL)	1 µL
- PVA 13 % préchauffé 10 min à 30°C	6 µL
- Solution contenant l'extrait nucléaire <ul style="list-style-type: none"> o Extrait nucléaire préchauffé 10 min à 30°C (8 µL) o Tampon D 1x (3,6 µL) 	11,6 µL

Solution pour le traitement à la protéinase K	Volume
- Tampon protéinase K 2x	100 µL
- H ₂ O	80 µL
- Protéinase K (20 µg/µL)	1 µL

Tableau II-7 : Réaction d'épissage et solution de protéinase K utilisée pour l'arrêt de cette réaction.

Le mélange est incubé pendant 5 min dans la glace puis 2 h à 30°C (sauf dans le cas d'expériences de cinétique d'épissage où le temps d'incubation à 30°C varie de 30 à 150 min). Les réactions d'épissage sont réalisées avec ou sans protéines SR (200 ou 400 ng de protéines SRSF1, SRSF7, SRSF2, SRSF5 ou SRSF6, produites en système baculovirus par l'équipe de J. Stévenin (IGBMC, Strasbourg). Les protéines contenues dans les produits de la réaction d'épissage sont ensuite hydrolysées par incubation 30 min à 30°C en présence de protéinase K (tableau II-7). Après extraction par du phénol/chloroforme, les ARN sont précipités à -80°C dans 2,5 volumes d'éthanol sans ajout de sels. Les culots sont lavés à l'éthanol 70 %, et la radioactivité de chaque échantillon est quantifiée au compteur à scintillation (Cerenkov – Beckman). Ceux-ci sont repris dans du bleu de formamide dénaturant de manière à obtenir une concentration finale de 1000 cpm/µL. Une quantité d'ARN correspondant à deux milles cpm est ensuite chauffée pendant 2 min à 96°C et fractionnée en gel de séquençage de polyacrylamide dénaturant 5 %. Les ARN sont révélés au PhosphorImager Typhoon (GE

Healthcare) ou par autoradiographie sur un film radiosensible en présence d'un écran amplificateur. La quantification de la radioactivité dans les bandes de gel se fait à l'aide du logiciel ImageQuant (MolecularDynamics).

III. Méthodes d'analyse de cellules humaines en culture

III.1. Culture et transfection des cellules

III.1.1. Culture à 37°C et déroulement du choc thermique

Les cellules HeLa sont cultivées dans le milieu DMEM (pour Dulbecco's Modified Eagle Medium) supplémenté avec 10% de Sérum de Veau Foetal (FBS, Invitrogen), 100 U/ml de pénicilline, 100µg/ml de streptomycine et 4mM de glutamine (Sigma).

L'entretien et l'amplification des cellules s'effectuent par culture dans des boîtes de 10 cm de diamètre, placées dans un incubateur à 37°C, sous une atmosphère humide et à 5% CO₂. La confluence des cellules est vérifiée par microscopie. Une fois à 70-90%, les cellules sont rincées avec du PBS et puis incubées avec 1 ml de trypsine (Invitrogen) à 37°C environ 2 min pour décrocher des cellules. L'activité enzymatique de la trypsine est arrêtée par ajout de 1 ml de milieu. Une nouvelle boîte contenant 10 ml de milieu estensemencée avec 200 µl de cette suspension. Les cellules sont mises en culture jusqu'à confluence (2 à 3 jours pour les cellules HeLa et les fibroblastes, plus d'une semaine pour les fibroblastes de patient HGPS). Pour un choc thermique optimal, les boîtes de cultures sont enveloppées dans plusieurs couches de parafilm afin d'être totalement hermétiques, puis plongées dans un bain marie à 42°C. Après le temps d'immersion choisi (souvent 2 h), les boîtes sont transférées en étuves à 37°C après avoir enlevé le parafilm, pour la période de régénération voulue.

III.1.2. Transfection des cellules

Un nombre nécessaire de plaques de 6 puits estensemencé avec 300 000 cellules par puits contenant 4 ml de DMEM. Le lendemain, les cellules sont transfectées avec des constructions plasmidiques aux concentrations spécifiées à l'aide du kit JetPEI (PolyPlus Transfection) ou à l'aide du kit LipofectamineTM 2000 (Invitrogen), selon les indications du fournisseur. Après 48h, les cellules sont rincées avec du PBS 1X, puis récoltées à l'aide d'une spatule en plastique stérile dans 500 µl de PBS1X.

III.2. Extraction des ARN totaux

Après récupération des cellules ou après l'arrêt de la réaction de la trypsine lors du passage des cellules, les cellules sont comptées au microscope en cellule de Malassez, puis centrifugées pendant 2 min à 200 g et lavées 2 fois dans le PBS et recentrifugées. Le réactif trizol (Invitrogen) est ajouté sur

le culot de cellules à raison de 1 mL par $10 \cdot 10^6$ cellules, qui sont alors incubées 5 min à température ambiante, puis mélangées à 0,2 mL de chloroforme ajoutés pour chaque mL de Trizol. Le mélange est vortexé, incubé 3 min à température ambiante, et centrifugé à 12000 g à 4°C pendant 15 min. La phase aqueuse est récupérée puis les ARN sont précipités par ajout de 0,5 mL d'isopropanol par mL de trizol initial et incubés à température ambiante pendant 10 min. Les ARN sont récupérés par centrifugation à 12000 g à 4°C pendant 10 min, le culot est lavé à l'éthanol 75%, puis repris dans 11 μ L de H₂O, dont 1 μ L sert au dosage par absorption optique à 260 nm grâce à un appareil nanodrop.

III.3. Analyses transcriptomiques par puces à jonctions exon-exon et puces microARN

III.3.1. Marquage des ARN totaux et hybridation sur les puces

Les échantillons destinés à l'hybridation subissent une étape de purification supplémentaire grâce au kit "miRCURY RNA Isolation kit" (Exiqon). Afin de vérifier l'intégrité des ARN totaux, ils sont également soumis à une analyse sur des puces d'appareil "bioanalyser" (Eukaryote Total RNA Nano 2.6, Agilent technologies). Le RIN (RNA Integrity Number) obtenu était systématiquement compris entre 9 et 10, ce qui reflète l'absence de dégradation des ARN purifiés.

III.3.1.a. Hybridation sur les puces à jonctions exon-exon HJAY

La préparation des ARN totaux pour l'hybridation sur les puces jonctions exon-exon HJAY (Affymetrix) est réalisée selon le protocole et avec le matériel préconisé par le fabricant pour les puces à exon. Il s'agit du protocole intitulé "Ambion Whole Transcript expression kit for Affymetrix GeneChip protocole". Rapidement, les ARN totaux sont rétrotranscrits de manière à ajouter en étiquette la séquence du promoteur T7 en 3' de la séquence des ADNc. Après la synthèse du second brin grâce à une ADN polymérase ADN-dépendante et la dégradation des ARN par la RNase H, un ARNc antisens est synthétisé grâce à une ARN polymérase et la présence du promoteur T7 en 3' de chaque fragment d'ADNc double brin. Cet ARNc est à nouveau rétrotranscrit en ADNc et éliminé par la RNase H. Cette rétrotranscription a lieu grâce à des oligonucléotides de séquences dégénérées, et en présence de dUTP, ce qui permet l'incorporation de résidus U dans l'ADNc obtenu. Après une étape de purification, ces ADNc sens sont fragmentés grâce à la combinaison d'enzymes Uracile DNA glycosylase (UDG) et "apurinic/apyrimidinic endonuclease 1" (APE1) qui hydrolyse l'ADN au niveau des résidus U précédemment incorporés. Ces fragments sont marqués en 3' par un fluorophore propriétaire Affymetrix (DLR pour DNA Labelling Reagent) grâce à l'action de la TdT ("Terminal desoxynucleotide Transférase"). Ce sont ces fragments marqués qui sont hybridés sur les puces jonctions exon-exon

HJAY. Nous disposions au total de 6 puces, 3 destinées à l'hybridation des échantillons de cellules non stressées, et 3 pour les échantillons de cellules soumises à un choc thermique.

III.3.2.b. Hybridation sur les puces microARN GeneChip® miRNA 3.0 Array

La préparation des ARN totaux pour l'hybridation sur les puces microARN GeneChip® miRNA 3.0 Array (Affymetrix) est réalisée avec le matériel et selon le protocole préconisés par le fabricant grâce aux kits "FlashTag™ HSR Biotin RNA Labeling Assay", GeneTitan® Hybridization, Wash and Stain Kit for miRNA Array Plates (P/N 902276) et GeneChip® Hybridization Control Kit (P/N 900457).

III.3.2. Protocole de traitement du signal et d'analyse des données des puces

III.3.2.a. Analyse des données des puces à jonctions exon-exon HJAY

L'analyse des puces à jonctions exon-exon est réalisée selon le protocole décrit dans la publication : (Shen et al., 2013). Pour l'analyse, le signal des puces a été normalisé en utilisant la méthode dite de "probe scaling". Le bruit de fond a été corrigé en utilisant le programme GeneBase utilisé avec le logiciel ProbeEffect (Kapur et al., 2008). L'index d'expression des gènes a été calculé à partir de sondes sélectionnées depuis le programme GeneBase utilisé avec le logiciel ProbeSelect (Kapur et al., 2008). Les signaux d'expression des gènes ont été traités en utilisant ces sondes. Les gènes ont été considérés comme exprimés si l'intensité moyenne était supérieure ou égale à 500. Les gènes étaient considérés comme ayant une expression qui varie si (1) ils étaient exprimés dans au moins une condition, (2) le taux de variation était supérieur ou égal à 1,5 et (3) la p-value obtenue par le test statistique "unpaired *t*-test" entre les intensités de gènes était inférieure ou égale à 0,05. Pour chaque sonde, un indice d'épissage a été calculé. Les tests "unpaired *t*-test" ont été réalisés pour déterminer la différence dans le signal des sondes entre deux échantillons en utilisant le programme MADS+ (Shen et al., 2010). Les p-values pour chaque probeset (set de sondes) ont été regroupées en utilisant la méthode de Fisher. La différence entre les isoformes des profils d'épissages (exons cassettes, sites 5'/3' alternatifs et exons mutuellement exclusifs) a été déterminée en utilisant les fichiers d'annotation de la puce HJAY, et en sélectionnant les isoformes en compétitions qui présentaient un nombre minimum de probeset dont les signaux qui variaient (avec une $p\text{-value} \leq 0,01$). Au moins 1 probeset sur 3 du profil "exon inclus" devaient être significatifs, et au moins 1 probeset sur 3 du profil "exon exclus" devaient être significatifs et montrer une variation d'expression opposée aux probesets du profil "exon exclus" pour l'index d'épissage. Par exemple, pour un seul exon cassette, la jonction du profil "exon exclus" et au moins un des trois probeset du profil "exon inclus" (c'est à dire un probeset correspondant à l'exon et deux

probesets correspondant aux deux jonctions exon-exon aux extrémités) devaient être significatif et devaient avoir une variation d'expression opposée pour l'index d'épissage.

III.3.2.b. Analyse des puces microARN GeneChip® miRNA 3.0 Array

Une fois les données de signal obtenues, elles ont été traitées une première fois avec le logiciel Partek (Partek, Inc.), ce qui a permis de corriger le bruit de fond, de normaliser et de moyenner les signaux des sondes. Le contrôle qualité a été effectué avec plusieurs méthodes : RLE ("Relative Log Expression Values"), BoxPlot et HC grâce à l'environnement logiciel R, et PCA *via* le logiciel Partek (Partek, Inc.). La détection des gènes différentiellement exprimés a été réalisée par le test statistique ANOVA grâce au logiciel Partek, puis Rank Product grâce à l'environnement logiciel R (Breitling et al., 2004; Eisinga et al., 2013).

III.4. RT-PCR et RT-qPCR

III.4.1. Digestion des extraits d'ARN à la DNase RQ1

Avant chaque réaction de transcription inverse, 10 µg d'ARN totaux sont soumis à digestion par 10U de DNase RQ1 (Promega) en présence de son tampon commercial (Promega) dans un volume final de 100 µL. Cette enzyme permet d'éliminer les acides désoxyribonucléiques éventuellement présents dans l'extrait d'ARN. La réaction de digestion est stoppée par l'ajout de 10 µL d'un tampon stop (Promega) et l'incubation pendant 10 min à 65°C. Les ARN sont ensuite dosés par mesure de l'absorption optique à 260 nm grâce à un appareil nanodrop.

III.4.2. Réaction de rétrotranscription

Avant la rétrotranscription, 1 µg d'ARN totaux traités à la DNase sont préalablement dénaturés à 65°C pendant 10 min puis refroidis rapidement dans la glace, en présence de 100 pmol de l'oligonucléotide amorce. Cet oligonucléotide est soit l'oligonucléotide oligodT¹⁶, soit l'oligonucléotide spécifique de l'ARN sat III (7796), soit un oligonucléotide hexamérique aléatoirement dégénéré. Un mélange de chaque dNTP (25 mM), 200 U d'enzyme MMLV (Promega) ou l'enzyme RevertAid Retrotranscriptase, et le tampon commercial de l'enzyme MMLV (Tris HCl 50 mM pH 8.3, KCl 75 mM, MgCl₂ 3 mM, DTT 10 mM final) sont ensuite ajoutés. L'élongation est effectuée à 42°C pendant 60 minutes, sauf dans le cas de l'oligonucléotide hexamérique où elle est effectuée à 37°C pendant 30 minutes puis à 42°C pendant 30 minutes. La réaction de transcription inverse est stoppée par incubation pendant 10 min à 65°C. Les ADNc obtenus sont ensuite soumis à une détection de produits spécifiques

grâce à la qPCR (pour les validations du choc thermique avec l'ARNm hsp70 et l'ARN sat III) ou par PCR "point final" (pour la détection de produits d'épissage alternatifs).

III.4.3. Détection des produits rétrotranscrits par PCR point final

Deux microlitres de la solution d'ADNc produits par RT-PCR est utilisé comme matrice pour une réaction de PCR. Le milieu réactionnel utilisé est le suivant :

Réaction PCR	Volume
- ADNc	2 μ L
- mélange de dNTP (1,25 mM chaque)	8 μ L
- Oligonucléotide 5' (20 μ M)	1 μ L
- Oligonucléotide 3' (20 μ M)	1 μ L
- DreamTaq ADN polymérase (5 U/ μ L, Fermentas)	0,25 μ L
- DreamTaq buffer green 10X (Fermentas)	5 μ L
- H ₂ O	2 μ L

Nombre de cycles	Etape	Température	Durée
1	Dénaturation initiale	95°C	5 min
35 (amplification)	Dénaturation	95°C	30 sec
	Hybridation	58°C	30 sec
	Elongation	72°C	30 sec
1	Elongation finale	72°C	10 min

Tableau III-1 : Réaction de qPCR pour la détection de produits spécifiques à partir d'ARN totaux rétrotranscrits.

Les produits PCR obtenus sont fractionnés en gel d'agarose 3 à 4 % (NUSEVE GTG Agarose, Lonza) en présence de GelRed (Fluoprobes), un agent intercalant utilisé pour la visualisation, dilué selon les spécifications du fabricant. Les bandes visibles par illumination aux UV à 254 nm sont quantifiées grâce au logiciel GeneTools (Syngene).

III.4.4. Détection des produits rétrotranscrits par qPCR

Un microlitre de la solution d'ADNc produits par RT-PCR est utilisé comme matrice pour une réaction de qPCR. Selon la nature du produit à détecter, l'ADNc matrice peut être dilué de 1 à 300fois, afin que la détection des produits d'intérêt apparaissent entre 25 et 30 Ct pour optimiser la quantification. Lorsque les ARN totaux sont rétrotranscrits grâce à l'oligonucléotide oligodT16, les ADNc obtenus sont dilués 20fois pour la détection de l'ARNm hsp70, et 300fois pour la détection de l'ARNm de la protéine ribosomique Rplp0 servant de normalisateur lors de la quantification. Lorsque les ARN totaux sont rétrotranscrits grâce à l'oligonucléotide 7796 pour la détection spécifique de l'ARN sat III, les ADNc matrices ne sont pas dilués. Le milieu réactionnel est composé d'un couple d'oligonucléotides spécifique (0,5 μ M chacun) et de SYBR Green I Master mix (Roche) dans un volume final de 10 μ l. Le mélange réactionnel et le programme utilisés pour la qPCR sont détaillés dans le tableau III-2.

Réaction qPCR	Volume
- ADNc dilué ou non	1 μ L
- Master MIX 2x	5 μ L
- Oligonucléotide 5' (5 μ M)	1 μ L
- Oligonucléotide 3' (5 μ M)	1 μ L
- H ₂ O	2 μ L

Nombre de cycles	Etape	Température	Durée	Vitesse de transition
1	Dénaturation initiale	95°C	10 min	-
45 (amplification)	Dénaturation	95°C	10 sec	20°C/sec
	Hybridation	58°C	10 sec	20°C/sec
	Elongation	72°C	10 sec	5°C/sec
1	Courbe de fusion	95°C	10 sec	20°C/sec
		60°C	5 sec	20°C/sec
		95°C	-	0,1°C/sec

Tableau III-2 : Réaction de qPCR pour la détection de produits spécifiques à partir d'ARN totaux rétrotranscrits.

Ce programme contient en étape finale un cycle permettant l'établissement d'une courbe de fusion afin de vérifier qu'un seul produit de PCR est amplifié. Les réactions de qPCR sont réalisées dans un appareil Light Cycler Carousel (Roche). Le calcul des Ct à partir de la fluorescence du SYBR Green et la vérification de la courbe de fusion sont effectués grâce au logiciel Light Cycler 3(v.5.32) (Roche). La détection des produits d'intérêt est normalisée entre chaque condition grâce à la détection parallèle d'un ARNm invariant, celui de la protéine ribosomique RPLP0. La technique de quantification relative

utilisée, nommée " $2^{(-\Delta\Delta Ct)}$ " est détaillée dans la publication (Livak and Schmittgen, 2001). Elle consiste succinctement à normaliser les Ct obtenus pour chaque condition d'intérêt (ici pour les ADNc obtenus après l'application du choc thermique) par rapport à une condition référence (ici pour les ADNc obtenus à partir de cellules non stressées), puis de normaliser le résultat grâce aux Ct obtenus à partir d'un produit invariant (ici l'ARNm de Rplp0). Cette technique est applicable uniquement avec un certain degré d'efficacité de la réaction de qPCR, qui doit être assez proche de l'efficacité théorique maximale. Cette efficacité varie d'une condition à l'autre principalement à cause des différents couples d'oligonucléotides utilisés. Elle est calculable grâce au logiciel de qPCR Light Cycler 3, à partir des Ct obtenus pour une gamme de différentes dilutions du même produit de départ. Nous avons donc au préalable vérifié la compatibilité de chaque couple d'oligonucléotide utilisé.

III.4.5. Détection des microARN par RT-qPCR TaqMan

NB1 : Les sondes TaqMan utilisées avec ce protocole sont préconçues et commercialisées par la société life technologies.

NB2 : Un oligonucléotide propre à chaque miRNA étudié permet d'amorcer la réaction transcription inverse. La méthode TaqMan consiste à utiliser amorce de transcription inverse permettant de rallonger spécifiquement les miRNA matures et donc les ADNc produits. Ainsi cette amorce d'environ 60 nucléotides se structure en tige boucle et son extrémité 3' est complémentaire de l'extrémité 5' du miRNA mature étudié. Cela rend l'appariement entre l'amorce et le miRNA mature plus spécifique et de différencier plus facilement les miRNA.

Les ARN totaux sont extraits grâce au kit miRCury™-RNA isolation kit (Tissue). Le kit a été utilisé selon les recommandations du fournisseur Exiqon. 20 ng d'extraits d'ARN totaux sont mis en présence de l'amorce tige-boucle (5X), d'un mélange de dNTP (100 mM), d'un mix d'enzyme MRT, de tampon de transcription inverse (10X), d'inhibiteur de RNase, dans un volume final de 7 µl (tableau III-3). La réaction de transcription inverse est incubée 30 min à 16°C puis 30 min à 42°C. L'enzyme est dénaturée 5 min à 80°C puis le mélange est placé dans la glace.

Les ADNc produits sont dilués 15 fois et servent de matrice pour la réaction de qPCR. Le mélange réactionnel est constitué des ADNc, d'une sonde TaqMan (5x ou 20x, spécifique du miRNA mature étudié) et de tampon TaqMan® Universal Master Mix II (2x) (life technologies), dans un volume final de 20 µl (tableau III-3). Le programme de qPCR est exposé dans le tableau III-3.

Les résultats sont exprimés en Ct et sont normalisés à partir de la détection de miRNA ou du snRNA U6, invariants dans les conditions étudiées. La méthode de quantification employée est la méthode dite du " $2^{(-\Delta\Delta Ct)}$ " décrite dans le paragraphe précédent.

Mélange RT TaqMan	
dNTP	0,15 µL
MRT (mélange de RT)	1 µL
buffer MRT 10x	1,5 µL
RNase inhibitor	0,19 µL
H ₂ O qsp 7 µL	4,16 µL

Mélange qPCR TaqMan	
Mélange de sondes TaqMan	1 µL
MasterMix Taqman 2x	10 µL
H ₂ O	7,67 µL

Nombre de cycles	Température	Durée
1	50°C	2 min
1	95°C	10 sec
40	95°C	15 sec
	60°C	1 min

Tableau III-3 : Réaction de RT-qPCR pour la détection de microARN grâce aux oligonucléotides TaqMan

III.5. Microscopie de fluorescence

III.5.1. Visualisation des ARN satellite III après un stress thermique par RNA-FISH

III.5.1.a. Préparation des sondes biotinylées

Deux types de sondes biotinylées nécessaires à la détection des ARN cellulaires par RNA-FISH ont été utilisées au cours de ce travail. Le premier type de sonde biotinylée a été directement produit à partir d'une matrice ADN grâce à un kit de marquage par "random priming" (voir ci-dessous). Nous avons ensuite utilisé des sondes LNA biotinylées de la société Exiqon, qui se sont avérées plus efficaces pour la détection des ARN sat III.

Synthèse des sondes biotinylées à partir d'une matrice ADN par "random priming"

Les sondes biotinylées sont synthétisées à partir d'une matrice ADN plasmidique porteur du clone pHur98, qui a été donné par l'équipe de G. Biamonti. Cette synthèse est effectuée à l'aide d'un kit Bioprime DNA labeling systems (Invitrogen) combinant des oligonucléotides octomériques aléatoirement dégénérés, un mélange à façon de dNTP biotinylés et le fragment de Klenow pour un marquage de type "random priming". Les solutions d'oligonucléotides dégénérés (solution 2,5x : [125 mM Tris-HCl (pH 6,8), 12,5 mM MgCl₂, 25 mM 2-mercaptoethanol, 750 µg/ml d'octamères aléatoires]) et de dNTP biotinylés (solution 10x : [1 mM biotine-14-dCTP, 1 mM dCTP, 2 mM dATP, 2 mM dGTP, 2 mM dTTP in 10 mM Tris-HCl (pH 7,5), 1 mM Na₂EDTA]), ainsi que le fragment de Klenow sont

fournis dans le kit. Trois cent nanogrammes de plasmide matrice sont dénaturés 5 min à 96°C en présence de la solution d'oligonucléotides aléatoires à 1x, puis refroidis 5 min sur glace. La réaction de synthèse a lieu en présence du mélange de dNTP 1x et de 40U du fragment de Klenow pendant 2 h minimum à 37°C. La réaction est stoppée grâce à une solution de Na₂EDTA 0,5 M pH 8,0 également fournie dans le kit. Juste avant de procéder à l'hybridation avec les cellules, 100 ng de sondes biotinylées par condition d'hybridation sont précipitées à -80°C dans 100 µL d'éthanol absolu en présence d'acétate de sodium et d'ADN de sperme de saumon. Après centrifugation les culots d'ADN biotinylés sont lavés à l'éthanol 70% et sont repris dans une solution [formamide 50% (v/v); SSC2x; Dextran 10% (v/v)]. L'ensemble est ensuite dénaturé 5 min à 75°C et refroidi 2 min dans la glace avant leur incubation avec les cellules fixées et perméabilisées.

Préparation des sondes LNA biotinylées de la société Exiqon

Les sondes LNA fournies par la société Exiqon sont diluées à 30 nM dans une solution [SSC 2x; Dextran sulfate 10% (v/v); Formamide 50% (v/v); Sodium phosphate 50 mM pH7]. Les sondes sont ensuite dénaturées 5 min à 75°C et refroidies 2 min dans la glace avant leur incubation avec les cellules fixées et perméabilisées.

III.5.1.b. Fixation des cellules et incubation avec les sondes

Les cellules sont mises en culture comme décrit dans le paragraphe III.1., sur une lame de microscope plongée dans les boîtes de culture (lames superfrost, Thermo scientific), afin de permettre leur visualisation au microscope ultérieurement. Après choc thermique, le milieu de culture est retiré et les cellules sont lavées avec du PBS. Les lames portant les cellules sont ensuite incubées avec une solution de formaldéhyde 4% (v/v)- PBS pendant 10 min afin de fixer les cellules par pontage covalent. Les cellules sont ensuite lavées 2 fois avec du PBS puis plongées dans une solution de Tris 0,1 M pH 7,5 pendant 5 min. Suite à cela elles sont incubées pendant 20 min à 1 h dans une solution de glycérol 20% (v/v)- PBS, puis congelées et décongelées 3 fois successivement dans de l'azote liquide. Une alternative utilisée pour la fixation et la perméabilisation des cellules est l'incubation 5 à 10 min des cellules dans du méthanol froid à -20°C (voir partie résultats). Suite à cela les cellules sont déshydratées par 3 bains successifs d'éthanol à 70, 90 et 100%.

Les cellules sont ensuite incubées pendant environ 16 à 18h avec la solution contenant les sondes biotinylées (synthétisée grâce à un kit ou les sondes LNA - voir paragraphe précédent), dans un montage hermétique en chambre humide pour éviter l'assèchement des lames. Ces lames sont ensuite lavées 3x5 min à 45°C dans une solution 60% formamide (v/v)- SSC2x, puis 3x5 min à température ambiante dans une solution de SSC 2x. Suite à cela elles sont incubées avec un tampon de blocage (3% BSA (p/v); SSC4x; 0,1% tween (v/v)) pendant 30 min à 37°C. Le tampon de blocage est ensuite

remplacé par un tampon d'hybridation (1% BSA (p/v); SSC4x ; 0,1% tween (v/v)) qui contient également de la streptavidine couplée au fluorophore Alexa 488 (Fischer Scientific). La durée de l'incubation est de 30 min minimum environ si l'expérience est réalisée avec les sondes LNA (Exiqon) ou de 2 h minimum dans le cas contraire. Les lames sont ensuite lavées 3x dans un tampon de lavage SSC4x - tween 0,1% à 45°C. A ce stade, les lames obtenues peuvent être utilisées pour le protocole de visualisation de protéines par immunofluorescence (voir paragraphe suivant III.5.2). Dans le cas contraire, des lamelles sont directement montées dessus pour la visualisation au microscope à épifluorescence ou au microscope confocal (voir paragraphe III.5.3).

III.5.2. Visualisation des protéines par immunofluorescence

Les cellules sont mises en culture comme décrit dans le paragraphe III.1.1., sur une lame de microscope plongée dans les boîtes de culture (lames superfrost, Thermo scientific), afin de permettre leur visualisation au microscope ultérieurement. Après choc thermique, le milieu de culture est retiré et les cellules sont lavées avec du PBS. Les lames portant les cellules sont ensuite incubées avec une solution de formaldéhyde 4% (v/v) - PBS pendant 10 min afin de fixer les cellules par pontage covalent. Les cellules sont ensuite rendues perméables par incubation avec une solution de PBS -Triton 0,5 % (v/v) pendant 5 min. Après la fixation et la perméabilisation des cellules, 3 lavages au PBS sont réalisés afin d'éliminer le détergent. Les cellules sont ensuite incubées avec l'anticorps primaire à la dilution désirée (cf. tableau des anticorps, §I.4.) pendant une heure. Après 3 nouveaux lavages au PBS, les cellules sont incubées avec un anticorps secondaire couplé à un fluorophore pendant 1h puis de nouveau lavées 3 fois avec une solution de PBS. Les lames sont ensuite prêtes pour le montage des lamelles décrit dans le paragraphe suivant et l'observation au microscope.

III.5.3. Montage des lamelles

Des lamelles de microscopie (18x18) sont montées sur les lames obtenues à l'aide d'un milieu de montage contenant du 4,6-diamidino-2-phénylindole (DAPI) (Duolink Mounting Media). Lorsque les lames sont destinées à une observation au microscope confocal que nous avons utilisé, ce milieu est complété avec le fluorophore intercalant Topro-3 à une concentration finale de 10 μ M, ce qui permet la visualisation des noyaux sur cet appareil qui ne possède pas de laser d'excitation dans la longueur d'onde UV.

III.6. Expériences de tethering GW182

Un nombre nécessaire de plaques de 6 puits estensemencé avec 300 000 cellules par puits contenant 4 ml de DMEM. Le lendemain, les cellules sont co-transfectées avec 0,3 µg du plasmide pEGFP::FLuc et 0,2 µg de plasmide pCINeo::Rluc-5BoxB ou bien pCINeo::Rluc contrôle. Après 48 h, les cellules sont récoltées dans 1 mL de PBS. Un autre aliquot de 100 µL est utilisé pour les tests luciférase. Cet aliquot est centrifugé à 3500 rpm et le culot de cellules est resuspendu dans 100 µL du tampon PLB 1x du kit "Dual luciférase® reporter" (Promega). Le mélange est incubé 20 min à température ambiante puis est soumis à 3 cycles de congélation/décongélation pour la lyse des cellules. 10 µL sont ensuite utilisés pour doser l'activité luciférase grâce à l'appareil luminomètre Xenius (SAFAS) et au logiciel SAFAS Sp2000.

L'extraction des protéines est réalisée sur un aliquot de 300 µL de cellules en ajoutant 60 µL de tampon de lyse des protéines suivi d'une incubation 30 min dans la glace. Le mélange est centrifugé à 4000 rpm, et le surnageant est récupéré. Les protéines sont dosées grâce à l'appareil Nanodrop 2000c (thermo scientific) et déposée sur gel SDS-PAGE pour analyse des protéines présentes par western blot (voir II.5.2 de cette partie).

BIBLIOGRAPHIE

Bibliographie

A.

Aaronson, R.P., and Blobel, G. (1975). Isolation of nuclear pore complexes in association with a lamina. *Proc. Natl. Acad. Sci. U. S. A.* *72*, 1007–1011.

Abdel-Halim, H.I., Mullenders, L.H.F., and Boei, J.J.W.A. (2006). Pairing of heterochromatin in response to cellular stress. *Exp. Cell Res.* *312*, 1961–1969.

Abravaya, K., Phillips, B., and Morimoto, R.I. (1991). Heat shock-induced interactions of heat shock transcription factor and the human hsp70 promoter examined by in vivo footprinting. *Mol. Cell. Biol.* *11*, 586.

Achsel, T., Brahm, H., Kastner, B., Bachi, A., Wilm, M., and Lührmann, R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J.* *18*, 5789–5802.

Adhikari, A.S., Sridhar Rao, K., Rangaraj, N., Parnaik, V.K., and Mohan Rao, C. (2004). Heat stress-induced localization of small heat shock proteins in mouse myoblasts: intranuclear lamin A/C speckles as target for alphaB-crystallin and Hsp25. *Exp. Cell Res.* *299*, 393–403.

Aebi, U., Cohn, J., Buhle, L., and Gerace, L. (1986). The nuclear lamina is a meshwork of intermediate-type filaments. *Nature* *323*, 560–564.

Aikawa, R., Komuro, I., Yamazaki, T., Zou, Y., Kudoh, S., Tanaka, M., Shiojima, I., Hiroi, Y., and Yazaki, Y. (1997). Oxidative stress activates extracellular signal-regulated kinases through Src and Ras in cultured cardiac myocytes of neonatal rats. *J. Clin. Invest.* *100*, 1813–1821.

Alastalo, T.-P., Hellesuo, M., Sandqvist, A., Hietakangas, V., Kallio, M., and Sistonen, L. (2003). Formation of nuclear stress granules involves HSF2 and coincides with the nucleolar localization of Hsp70. *J. Cell Sci.* *116*, 3557–3570.

Alexander, R., and Beggs, J.D. (2010). Cross-talk in transcription, splicing and chromatin: who makes the first call? *Biochem. Soc. Trans.* *38*, 1251–1256.

Allemand, E., Guil, S., Myers, M., Moscat, J., Cáceres, J.F., and Krainer, A.R. (2005). Regulation of heterogeneous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 3605–3610.

Ameur, A., Zaghlool, A., Halvardson, J., Wetterbom, A., Gyllenstein, U., Cavelier, L., and Feuk, L. (2011). Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat. Struct. Mol. Biol.* *18*, 1435–1440.

Amin, J., Ananthan, J., and Voellmy, R. (1988). Key features of heat shock regulatory elements. *Mol. Cell. Biol.* *8*, 3761–3769.

Amor, D.J., Kalitsis, P., Sumer, H., and Choo, K.H.A. (2004). Building the centromere: from foundation proteins to 3D organization. *Trends Cell Biol.* *14*, 359–368.

Anckar, J., and Sistonen, L. (2011). Regulation of HSF1 Function in the Heat Stress Response: Implications in Aging and Disease. *Annu. Rev. Biochem.* *80*, 1089–1115.

- Anderson, P., and Kedersha, N. (2002). Stressful initiations. *J. Cell Sci.* *115*, 3227–3234.
- Anderson, P., and Kedersha, N. (2006). RNA granules. *J. Cell Biol.* *172*, 803–808.
- Andersson, M.K., Ståhlberg, A., Arvidsson, Y., Olofsson, A., Semb, H., Stenman, G., Nilsson, O., and Aman, P. (2008). The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. *BMC Cell Biol.* *9*, 37.
- Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C., and Komorowski, J. (2009). Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res.* *19*, 1732–1741.
- Andreadis, A. (2005). Tau gene alternative splicing: expression patterns, regulation and modulation of function in normal brain and neurodegenerative diseases. *Biochim. Biophys. Acta* *1739*, 91–103.
- Andrei, M.A. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA* *11*, 717–727.
- Andrei, M.A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R., and Lührmann, R. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA N. Y. N* *11*, 717–727.
- Andressoo, J.O., and Hoeijmakers, J.H.J. (2005). Transcription-coupled repair and premature ageing. *Mutat. Res.* *577*, 179–194.
- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* *25*, 25–29.
- Ayane, M., Preuss, U., Köhler, G., and Nielsen, P.J. (1991). A differentially expressed murine RNA encoding a protein with similarities to two types of nucleic acid binding motifs. *Nucleic Acids Res.* *19*, 1273–1278.
- Back, R., Dominguez, C., Rothé, B., Bobo, C., Beaufils, C., Moréra, S., Meyer, P., Charpentier, B., Branlant, C., Allain, F.H.-T., et al. (2013). High-Resolution Structural Analysis Shows How Tah1 Tethers Hsp90 to the R2TP Complex. *Struct. Lond. Engl.* *1993* *21*, 1834–1847.
- Bang, O.S., Ha, B.G., Park, E.K., and Kang, S.S. (2000). Activation of Akt is induced by heat shock and involved in suppression of heat-shock-induced apoptosis of NIH3T3 cells. *Biochem. Biophys. Res. Commun.* *278*, 306–311.
- Barascu, A., Le Chalony, C., Pennarun, G., Genet, D., Zaarour, N., and Bertrand, P. (2012). Oxidative stress alters nuclear shape through lamins dysregulation: a route to senescence. *Nucl. Austin Tex* *3*, 411–417.
- Barbier, J., Dutertre, M., Bittencourt, D., Sanchez, G., Gratadou, L., de la Grange, P., and Auboeuf, D. (2007). Regulation of H-ras splice variant expression by cross talk between the p53 and nonsense-mediated mRNA decay pathways. *Mol. Cell. Biol.* *27*, 7315–7333.
- Baricheva, E.A., Berrios, M., Bogachev, S.S., Borisevich, I.V., Lapik, E.R., Sharakhov, I.V., Stuurman, N., and Fisher, P.A. (1996). DNA from *Drosophila melanogaster* beta-heterochromatin binds specifically to nuclear lamins in vitro and the nuclear envelope in situ. *Gene* *171*, 171–176.
- Barlat, I., Maurier, F., Duchesne, M., Guitard, E., Tocque, B., and Schweighoffer, F. (1997). A role for Sam68 in cell cycle progression antagonized by a spliced variant within the KH domain. *J. Biol. Chem.* *272*, 3129–3132.

- Barta, A., Kalyna, M., and Reddy, A.S.N. (2010). Implementing a rational and consistent nomenclature for serine/arginine-rich protein splicing factors (SR proteins) in plants. *Plant Cell* 22, 2926–2929.
- Batsché, E., Yaniv, M., and Muchardt, C. (2006). The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat. Struct. Mol. Biol.* 13, 22–29.
- Beck, L.A., Hosick, T.J., and Sinensky, M. (1990). Isoprenylation is required for the processing of the lamin A precursor. *J. Cell Biol.* 110, 1489–1499.
- Becker, J., Melchior, F., Gerke, V., Bischoff, F.R., Ponstingl, H., and Wittinghofer, A. (1995). RNA1 encodes a GTPase-activating protein specific for Gsp1p, the Ran/TC4 homologue of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 270, 11860–11865.
- Bedard, K.M., Daijogo, S., and Semler, B.L. (2007). A nucleo-cytoplasmic SR protein functions in viral IRES-mediated translation initiation. *EMBO J.* 26, 459–467.
- Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* 20, 1885–1898.
- Belmont, A.S., Zhai, Y., and Thilenius, A. (1993). Lamin B distribution and association with peripheral chromatin revealed by optical sectioning and electron microscopy tomography. *J. Cell Biol.* 123, 1671–1685.
- Beyer, A.L., Christensen, M.E., Walker, B.W., and LeStourgeon, W.M. (1977). Identification and characterization of the packaging proteins of core 40S hnRNP particles. *Cell* 11, 127–138.
- Bhatt, D.M., Pandya-Jones, A., Tong, A.-J., Barozzi, I., Lissner, M.M., Natoli, G., Black, D.L., and Smale, S.T. (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* 150, 279–290.
- Biamonti, G., and Caceres, J.F. (2009). Cellular stress and RNA splicing. *Trends Biochem. Sci.* 34, 146–153.
- Bieberstein, N.I., Carrillo Oesterreich, F., Straube, K., and Neugebauer, K.M. (2012). First exon length controls active chromatin signatures and transcription. *Cell Rep.* 2, 62–68.
- Bijur, G.N., and Jope, R.S. (2000). Opposing actions of phosphatidylinositol 3-kinase and glycogen synthase kinase-3beta in the regulation of HSF-1 activity. *J. Neurochem.* 75, 2401–2408.
- Birney, E., Kumar, S., and Krainer, A.R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.* 21, 5803–5816.
- Bischoff, F.R., and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature* 354, 80–82.
- Bischoff, F.R., Krebber, H., Kempf, T., Hermes, I., and Ponstingl, H. (1995a). Human RanGTPase-activating protein RanGAP1 is a homologue of yeast Rna1p involved in mRNA processing and transport. *Proc. Natl. Acad. Sci. U. S. A.* 92, 1749–1753.
- Bischoff, F.R., Krebber, H., Smirnova, E., Dong, W., and Ponstingl, H. (1995b). Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1. *EMBO J.* 14, 705–715.

- Björk, P., and Wieslander, L. (2011). Nucleocytoplasmic mRNP export is an integral part of mRNP biogenesis. *Chromosoma* *120*, 23–38.
- Black, D.L., and Grabowski, P.J. (2003). Alternative pre-mRNA splicing and neuronal function. *Prog. Mol. Subcell. Biol.* *31*, 187–216.
- Blake, M.J., Fagnoli, J., Gershon, D., and Holbrook, N.J. (1991). Concomitant decline in heat-induced hyperthermia and HSP70 mRNA expression in aged rats. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* *260*, R663–R667.
- Blaustein, M., Quadrana, L., Risso, G., Mata, M. de la, Pelisch, F., and Srebrow, A. (2009). SF2/ASF regulates proteomic diversity by affecting the balance between translation initiation mechanisms. *J. Cell. Biochem.* *107*, 826–833.
- Blencowe, B.J., Nickerson, J.A., Issner, R., Penman, S., and Sharp, P.A. (1994). Association of nuclear matrix antigens with exon-containing splicing complexes. *J. Cell Biol.* *127*, 593–607.
- Blencowe, B.J., Baurén, G., Eldridge, A.G., Issner, R., Nickerson, J.A., Rosonina, E., and Sharp, P.A. (2000). The SRm160/300 splicing coactivator subunits. *RNA N. Y. N* *6*, 111–120.
- Bonne, G., Di Barletta, M.R., Varnous, S., Bécane, H.M., Hammouda, E.H., Merlini, L., Muntoni, F., Greenberg, C.R., Gary, F., Urtizbera, J.A., et al. (1999). Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. *Nat. Genet.* *21*, 285–288.
- Borkovich, K.A., Farrelly, F.W., Finkelstein, D.B., Taulien, J., and Lindquist, S. (1989). hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. *Mol. Cell. Biol.* *9*, 3919.
- Boulon, S., Marmier-Gourrier, N., Pradet-Balade, B., Wurth, L., Verheggen, C., Jády, B.E., Rothé, 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.
- Boulon, S., Westman, B.J., Hutten, S., Boisvert, F.-M., and Lamond, A.I. (2010a). The Nucleolus under Stress. *Mol. Cell* *40*, 216–227.
- Boulon, S., Pradet-Balade, B., Verheggen, C., Molle, D., Boireau, S., Georgieva, M., Azzag, K., Robert, M.-C., Ahmad, Y., Neel, H., et al. (2010b). HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol. Cell* *39*, 912–924.
- Boulon, S., Bertrand, E., and Pradet-Balade, B. (2012). HSP90 and the R2TP co-chaperone complex: building multi-protein machineries essential for cell growth and gene expression. *RNA Biol.* *9*, 148–154.
- Bourgeois, C.F., Lejeune, F., and Stévenin, J. (2004). Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. *Prog. Nucleic Acid Res. Mol. Biol.* *78*, 37–88.
- Bracken, A.P., and Bond, U. (1999). Reassembly and protection of small nuclear ribonucleoprotein particles by heat shock proteins in yeast cells. *RNA N. Y. N* *5*, 1586–1596.
- Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Gallinaro, H., Jacob, M., Sri-Widada, J., and Jeanteur, P. (1980). Nucleotide sequences of nuclear U1A RNAs from chicken, rat and man. *Nucleic Acids Res.* *8*, 4143–4154.
- 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.

- Branlant, C., Krol, A., Lazar, E., Haendler, B., Jacob, M., Galego-Dias, L., and Pousada, C. (1983). High evolutionary conservation of the secondary structure and of certain nucleotide sequences of U5 RNA. *Nucleic Acids Res.* *11*, 8359–8367.
- Braunschweig, U., Gueroussov, S., Plocik, A.M., Graveley, B.R., and Blencowe, B.J. (2013). Dynamic integration of splicing within gene regulatory pathways. *Cell* *152*, 1252–1269.
- Breitkreutz, D., Braiman-Wiksmann, L., Daum, N., Denning, M.F., and Tennenbaum, T. (2007). Protein kinase C family: on the crossroads of cell signaling in skin and tumor epithelium. *J. Cancer Res. Clin. Oncol.* *133*, 793–808.
- Breitling, R., Armengaud, P., Amtmann, A., and Herzyk, P. (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.* *573*, 83–92.
- Bridger, J.M., Kill, I.R., O’Farrell, M., and Hutchison, C.J. (1993). Internal lamin structures within G1 nuclei of human dermal fibroblasts. *J. Cell Sci.* *104 (Pt 2)*, 297–306.
- Bridger, J.M., Foeger, N., Kill, I.R., and Herrmann, H. (2007). The nuclear lamina. *Febs J.* *274*, 1354–1361.
- Brocchieri, L., Conway de Macario, E., and Macario, A.J.L. (2008). hsp70 genes in the human genome: Conservation and differentiation patterns predict a wide array of overlapping and specialized functions. *BMC Evol. Biol.* *8*, 19.
- Bromberg, Z., Goloubinoff, P., Saidi, Y., and Weiss, Y.G. (2013). The Membrane-Associated Transient Receptor Potential Vanilloid Channel Is the Central Heat Shock Receptor Controlling the Cellular Heat Shock Response in Epithelial Cells. *PLoS ONE* *8*, e57149.
- Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T., et al. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3’ end of a transcript encoding a protein kinase family member. *Cell* *69*.
- Brunet Simioni, M., De Thonel, A., Hammann, A., Joly, A.L., Bossis, G., Fourmaux, E., Bouchot, A., Landry, J., Piechaczyk, M., and Garrido, C. (2009). Heat shock protein 27 is involved in SUMO-2/3 modification of heat shock factor 1 and thereby modulates the transcription factor activity. *Oncogene* *28*, 3332–3344.
- Buchan, J.R., and Parker, R. (2009). Eukaryotic Stress Granules: The Ins and Out of Translation. *Mol. Cell* *36*.
- Buckanovich, R.J., Posner, J.B., and Darnell, R.B. (1993). Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. *Neuron* *11*, 657–672.
- Buggele, W.A., Johnson, K.E., and Horvath, C.M. (2012). Influenza A virus infection of human respiratory cells induces primary microRNA expression. *J. Biol. Chem.* *287*, 31027–31040.
- Bügl, H., Fauman, E.B., Staker, B.L., Zheng, F., Kushner, S.R., Saper, M.A., Bardwell, J.C.A., and Jakob, U. (2000). RNA Methylation under Heat Shock Control. *Mol. Cell* *6*, 349–360.
- Buratti, E., and Baralle, F.E. (2004). Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol. Cell. Biol.* *24*, 10505–10514.

- Burgman, P., Ouyang, H., Peterson, S., Chen, D.J., and Li, G.C. (1997). Heat Inactivation of Ku Autoantigen: Possible Role in Hyperthermic Radiosensitization. *Cancer Res.* *57*, 2847–2850.
- Burke, B., and Gerace, L. (1986). A cell free system to study reassembly of the nuclear envelope at the end of mitosis. *Cell* *44*, 639–652.
- Burke, B., and Roux, K.J. (2009). Nuclei Take a Position: Managing Nuclear Location. *Dev. Cell* *17*, 587–597.
- Busch, A., and Hertel, K.J. (2012). Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip. Rev. RNA* *3*, 1–12.
- Caceres, J.F., Stamm, S., Helfman, D.M., and Krainer, A.R. (1994). Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science* *265*, 1706–1709.
- Caceres, J.F., Misteli, T., Sreaton, G.R., Spector, D.L., and Krainer, A.R. (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol.* *138*, 225–238.
- Caceres, J.F., Sreaton, G.R., and Krainer, A.R. (1998). A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes Dev.* *12*, 55–66.
- Calabrese, V., Cornelius, C., Cuzzocrea, S., Iavicoli, I., Rizzarelli, E., and Calabrese, E.J. (2011). Hormesis, cellular stress response and vitagenes as critical determinants in aging and longevity. *Mol. Aspects Med.* *32*, 279–304.
- Calabrese, V., Cornelius, C., Dinkova-Kostova, A.T., Iavicoli, I., Di Paola, R., Koverech, A., Cuzzocrea, S., Rizzarelli, E., and Calabrese, E.J. (2012). Cellular stress responses, hormetic phytochemicals and vitagenes in aging and longevity. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* *1822*, 753–783.
- Cano, C.E., Hamidi, T., Garcia, M.N., Grasso, D., Loncle, C., Garcia, S., Calvo, E., Lomberk, G., Dusetti, N., Bartholin, L., et al. (2013). Genetic inactivation of Nupr1 acts as a dominant suppressor event in a two-hit model of pancreatic carcinogenesis. *Gut*.
- Cao, H., and Hegele, R.A. (2003). LMNA is mutated in Hutchinson-Gilford progeria (MIM 176670) but not in Wiedemann-Rautenstrauch progeroid syndrome (MIM 264090). *J. Hum. Genet.* *48*, 271–274.
- Cao, K., Capell, B.C., Erdos, M.R., Djabali, K., and Collins, F.S. (2007). A lamin A protein isoform overexpressed in Hutchinson-Gilford progeria syndrome interferes with mitosis in progeria and normal cells. *Proc. Natl. Acad. Sci.* *104*, 4949–4954.
- Cao, K., Blair, C.D., Faddah, D.A., Kieckhafer, J.E., Olive, M., Erdos, M.R., Nabel, E.G., and Collins, F.S. (2011). Progerin and telomere dysfunction collaborate to trigger cellular senescence in normal human fibroblasts. *J. Clin. Invest.* *121*, 2833–2844.
- Capell, B.C., Erdos, M.R., Madigan, J.P., Fiordalisi, J.J., Varga, R., Conneely, K.N., Gordon, L.B., Der, C.J., Cox, A.D., and Collins, F.S. (2005). Inhibiting farnesylation of progerin prevents the characteristic nuclear blebbing of Hutchinson-Gilford progeria syndrome. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 12879–12884.
- Caputi, M., Mayeda, A., Krainer, A.R., and Zahler, A.M. (1999). hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO J.* *18*, 4060–4067.
- Cargnello, M., Tcherkezian, J., Dorn, J.F., Huttlin, E.L., Maddox, P.S., Gygi, S.P., and Roux, P.P. (2012). Phosphorylation of the eukaryotic translation initiation factor 4E-transporter (4E-T) by c-Jun N-terminal kinase promotes stress-dependent P-body assembly. *Mol. Cell. Biol.* *32*, 4572–4584.

- Carmo-Fonseca, M., Ferreira, J., and Lamond, A.I. (1993). Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis—evidence that the coiled body is a kinetic nuclear structure. *J. Cell Biol.* *120*, 841–852.
- Carrier, L., Schlossarek, S., Willis, M.S., and Eschenhagen, T. (2010). The ubiquitin-proteasome system and nonsense-mediated mRNA decay in hypertrophic cardiomyopathy. *Cardiovasc. Res.* *85*, 330–338.
- Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., and Krainer, A.R. (2003). ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* *31*, 3568–3571.
- Casas-Finet, J.R., Smith, J.D., Jr, Kumar, A., Kim, J.G., Wilson, S.H., and Karpel, R.L. (1993). Mammalian heterogeneous ribonucleoprotein A1 and its constituent domains. Nucleic acid interaction, structural stability and self-association. *J. Mol. Biol.* *229*, 873–889.
- Castilla-Llorente, V., Spraggon, L., Okamura, M., Naseeruddin, S., Adamow, M., Qamar, S., and Liu, J. (2012). Mammalian GW220/TNGW1 is essential for the formation of GW/P bodies containing miRISC. *J. Cell Biol.* *198*, 529–544.
- Cavaloc, Y., Popielarz, M., Fuchs, J.P., Gattoni, R., and Stévenin, J. (1994). Characterization and cloning of the human splicing factor 9G8: a novel 35 kDa factor of the serine/arginine protein family. *EMBO J.* *13*, 2639–2649.
- Cazalla, D., Zhu, J., Manche, L., Huber, E., Krainer, A.R., and Cáceres, J.F. (2002). Nuclear export and retention signals in the RS domain of SR proteins. *Mol. Cell. Biol.* *22*, 6871–6882.
- Chandler, D.S., Singh, R.K., Caldwell, L.C., Bitler, J.L., and Lozano, G. (2006). Genotoxic stress induces coordinately regulated alternative splicing of the p53 modulators MDM2 and MDM4. *Cancer Res.* *66*, 9502–9508.
- Chang, L., and Karin, M. (2001). Mammalian MAP kinase signalling cascades. *Nature* *410*, 37–40.
- Chang, T.-H., Huang, H.-Y., Hsu, J.B.-K., Weng, S.-L., Horng, J.-T., and Huang, H.-D. (2013). An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. *BMC Bioinformatics* *14 Suppl 2*, S4.
- Charlier, C., Lamaye, F., Thelen, N., and Thiry, M. (2009). Ultrastructural detection of nucleic acids within heat shock-induced perichromatin granules of HeLa cells by cytochemical and immunocytological methods. *J. Struct. Biol.* *166*, 329–336.
- Chaudhary, N., McMahon, C., and Blobel, G. (1991). Primary structure of a human arginine-rich nuclear protein that colocalizes with spliceosome components. *Proc. Natl. Acad. Sci. U. S. A.* *88*, 8189–8193.
- Chen, H.-C., and Cheng, S.-C. (2012). Functional roles of protein splicing factors. *Biosci. Rep.* *32*, 345–359.
- Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* *10*, 741–754.
- Chen, T., Damaj, B.B., Herrera, C., Lasko, P., and Richard, S. (1997). Self-association of the single-KH-domain family members Sam68, GRP33, GLD-1, and Qk1: role of the KH domain. *Mol. Cell. Biol.* *17*, 5707–5718.
- Chen, T., Boisvert, F.M., Bazett-Jones, D.P., and Richard, S. (1999). A role for the GSG domain in localizing Sam68 to novel nuclear structures in cancer cell lines. *Mol. Biol. Cell* *10*, 3015–3033.

- Cheng, H.-L., Mostoslavsky, R., Saito, S., Manis, J.P., Gu, Y., Patel, P., Bronson, R., Appella, E., Alt, F.W., and Chua, K.F. (2003). Developmental defects and p53 hyperacetylation in Sir2 homolog (SIRT1)-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 10794–10799.
- Chiodi, I., Biggiogera, M., Denegri, M., Corioni, M., Weighardt, F., Cobiانchi, F., Riva, S., and Biamonti, G. (2000). Structure and dynamics of hnRNP-labelled nuclear bodies induced by stress treatments. *J. Cell Sci.* *113*, 4043–4053.
- Chiodi, I., Corioni, M., Giordano, M., Valgardsdottir, R., Ghigna, C., Cobiانchi, F., Xu, R.-M., Riva, S., and Biamonti, G. (2004). RNA recognition motif 2 directs the recruitment of SF2/ASF to nuclear stress bodies. *Nucleic Acids Res.* *32*, 4127–4136.
- Choi, H.S., Lin, Z., Li, B.S., and Liu, A.Y. (1990). Age-dependent decrease in the heat-inducible DNA sequence-specific binding activity in human diploid fibroblasts. *J. Biol. Chem.* *265*, 18005–18011.
- Chu, B., Soncin, F., Price, B.D., Stevenson, M.A., and Calderwood, S.K. (1996). Sequential phosphorylation by mitogen-activated protein kinase and glycogen synthase kinase 3 represses transcriptional activation by heat shock factor-1. *J. Biol. Chem.* *271*, 30847–30857.
- Clark, D.W., Mitra, A., Fillmore, R.A., Jiang, W.G., Samant, R.S., Fodstad, O., and Shevde, L.A. (2008). NUPR1 interacts with p53, transcriptionally regulates p21 and rescues breast epithelial cells from doxorubicin-induced genotoxic stress. *Curr. Cancer Drug Targets* *8*, 421–430.
- Clemson, C.M., McNeil, J.A., Willard, H.F., and Lawrence, J.B. (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* *132*, 259–275.
- Close, P., East, P., Dirac-Svejstrup, A.B., Hartmann, H., Heron, M., Maslen, S., Chariot, A., Söding, J., Skehel, M., and Svejstrup, J.Q. (2012). DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation. *Nature* *484*, 386–389.
- Clouet d’Orval, B., d’ Aubenton Carafa, Y., Sirand-Pugnet, P., Gallego, M., Brody, E., and Marie, J. (1991a). RNA secondary structure repression of a muscle-specific exon in HeLa cell nuclear extracts. *Science* *252*, 1823–1828.
- Clouet d’Orval, B., d’ Aubenton-Carafa, Y., Brody, J.M., and Brody, E. (1991b). Determination of an RNA structure involved in splicing inhibition of a muscle-specific exon. *J. Mol. Biol.* *221*, 837–856.
- Cohen, E., Bieschke, J., Perciavalle, R.M., Kelly, J.W., and Dillin, A. (2006). Opposing activities protect against age-onset proteotoxicity. *Science* *313*, 1604–1610.
- Cohen, E., Paulsson, J.F., Blinder, P., Burstyn-Cohen, T., Du, D., Estepa, G., Adame, A., Pham, H.M., Holzenberger, M., Kelly, J.W., et al. (2009). Reduced IGF-1 signaling delays age-associated proteotoxicity in mice. *Cell* *139*, 1157–1169.
- Cohen, M., Lee, K.K., Wilson, K.L., and Gruenbaum, Y. (2001). Transcriptional repression, apoptosis, human disease and the functional evolution of the nuclear lamina. *Trends Biochem. Sci.* *26*, 41–47.
- Collart, M.A., Panasenko, O.O., and Nikolaev, S.I. (2013). The Not3/5 subunit of the Ccr4-Not complex: a central regulator of gene expression that integrates signals between the cytoplasm and the nucleus in eukaryotic cells. *Cell. Signal.* *25*, 743–751.
- Collins, C.A., and Guthrie, C. (2000). The question remains: is the spliceosome a ribozyme? *Nat. Struct. Biol.* *7*, 850–854.

- Collins, P.L., and Hightower, L.E. (1982). Newcastle disease virus stimulates the cellular accumulation of stress (heat shock) mRNAs and proteins. *J. Virol.* *44*, 703–707.
- Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J.L., Bell, J.C., and Duncan, P.I. (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J.* *15*, 265–275.
- Cook, P.R. (1991). The nucleoskeleton and the topology of replication. *Cell* *66*, 627–635.
- Cooke, H.J., and Hindley, J. (1979). Cloning of human satellite III DNA: different components are on different chromosomes. *Nucleic Acids Res.* *6*, 3177.
- Cooper, T.A. (1998). Muscle-specific splicing of a heterologous exon mediated by a single muscle-specific splicing enhancer from the cardiac troponin T gene. *Mol. Cell. Biol.* *18*, 4519–4525.
- Cooper, T.A., Wan, L., and Dreyfuss, G. (2009). RNA and Disease. *Cell* *136*, 777–793.
- Coote, P.J., Cole, M.B., and Jones, M.V. (1991). Induction of increased thermotolerance in *Saccharomyces cerevisiae* may be triggered by a mechanism involving intracellular pH. *J. Gen. Microbiol.* *137*, 1701–1708.
- Corcos, L., and Solier, S. (2005). Épissage alternatif, pathologie et thérapeutique moléculaire. *MS Médecine Sci.* *21*.
- Corden, J.L., and Patturajan, M. (1997). A CTD function linking transcription to splicing. *Trends Biochem. Sci.* *22*, 413–416.
- Corrigan, D., Kuszczak, D., Rusinol, A., Thewke, D., Hrycyna, C., Michaelis, S., and Sinensky, M. (2005). Prelamin A endoproteolytic processing in vitro by recombinant Zmpste24. *Biochem J* *387*, 129–138.
- Corton, J.M., Gillespie, J.G., and Hardie, D.G. (1994). Role of the AMP-activated protein kinase in the cellular stress response. *Curr. Biol.* *4*, 315–324.
- Cotto, J., Fox, S., and Morimoto, R. (1997). HSF1 granules: a novel stress-induced nuclear compartment of human cells. *J. Cell Sci.* *110*, 2925–2934.
- Craig, J.M. (2005). Heterochromatin--many flavours, common themes. *BioEssays News Rev. Mol. Cell. Dev. Biol.* *27*, 17–28.
- Cramer, P., Pesce, C.G., Baralle, F.E., and Kornblihtt, A.R. (1997). Functional association between promoter structure and transcript alternative splicing. *Proc. Natl. Acad. Sci. U. S. A.* *94*, 11456–11460.
- Csoka, A.B., English, S.B., Simkevich, C.P., Ginzinger, D.G., Butte, A.J., Schatten, G.P., Rothman, F.G., and Sedivy, J.M. (2004). Genome-scale expression profiling of Hutchinson-Gilford progeria syndrome reveals widespread transcriptional misregulation leading to mesodermal/mesenchymal defects and accelerated atherosclerosis. *Aging Cell* *3*, 235–243.
- Cuesta, R., Laroia, G., and Schneider, R.J. (2000). Chaperone hsp27 inhibits translation during heat shock by binding eIF4G and facilitating dissociation of cap-initiation complexes. *Genes Dev.* *14*, 1460–1470.
- Dahl, K.N., Scaffidi, P., Islam, M.F., Yodh, A.G., Wilson, K.L., and Misteli, T. (2006). Distinct structural and mechanical properties of the nuclear lamina in Hutchinson–Gilford progeria syndrome. *Proc. Natl. Acad. Sci.* *103*, 10271–10276.

- Das, R., Yu, J., Zhang, Z., Gygi, M.P., Krainer, A.R., Gygi, S.P., and Reed, R. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol. Cell* 26, 867–881.
- Daskal, I., Merski, A., Hughes, J.B., and Busch, H. (1975). The effects of cycloheximide on the ultrastructure of rat liver cells. *Exp. Cell Res.* 93, 395–401.
- Datar, K.V., Dreyfuss, G., and Swanson, M.S. (1993). The human hnRNP M proteins: identification of a methionine/arginine-rich repeat motif in ribonucleoproteins. *Nucleic Acids Res.* 21, 439–446.
- David, C.J., and Manley, J.L. (2010). Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev.* 24, 2343–2364.
- David, C.J., Boyne, A.R., Millhouse, S.R., and Manley, J.L. (2011). The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev.* 25, 972–983.
- Davies, B.S.J., Barnes, R.H., 2nd, Tu, Y., Ren, S., Andres, D.A., Spielmann, H.P., Lammerding, J., Wang, Y., Young, S.G., and Fong, L.G. (2010). An accumulation of non-farnesylated prelamin A causes cardiomyopathy but not progeria. *Hum. Mol. Genet.* 19, 2682–2694.
- Davies, P.F., Zilberberg, J., and Helmke, B.P. (2003). Spatial microstimuli in endothelial mechanosignaling. *Circ. Res.* 92, 359–370.
- Dechat, T., Korbei, B., Vaughan, O.A., Vlcek, S., Hutchison, C.J., and Foisner, R. (2000). Lamina-associated polypeptide 2alpha binds intranuclear A-type lamins. *J. Cell Sci.* 113 Pt 19, 3473–3484.
- Dechat, T., Shimi, T., Adam, S.A., Rusinol, A.E., Andres, D.A., Spielmann, H.P., Sinensky, M.S., and Goldman, R.D. (2007). Alterations in mitosis and cell cycle progression caused by a mutant lamin A known to accelerate human aging. *Proc. Natl. Acad. Sci.* 104, 4955–4960.
- DeLisle, A.J., Graves, R.A., Marzluff, W.F., and Johnson, L.F. (1983). Regulation of histone mRNA production and stability in serum-stimulated mouse 3T6 fibroblasts. *Mol. Cell. Biol.* 3, 1920–1929.
- Denegri, M., Chiodi, I., Corioni, M., Cobianchi, F., Riva, S., and Biamonti, G. (2001). Stress-induced nuclear bodies are sites of accumulation of pre-mRNA processing factors. *Mol. Biol. Cell* 12, 3502–3514.
- Denegri, M., Moralli, D., Rocchi, M., Biggiogera, M., Raimondi, E., Cobianchi, F., De Carli, L., Riva, S., and Biamonti, G. (2002). Human chromosomes 9, 12, and 15 contain the nucleation sites of stress-induced nuclear bodies. *Mol. Biol. Cell* 13, 2069–2079.
- Desbiens, K.M., Deschesnes, R.G., Labrie, M.M., Desfossés, Y., Lambert, H., Landry, J., and Bellmann, K. (2003). c-Myc potentiates the mitochondrial pathway of apoptosis by acting upstream of apoptosis signal-regulating kinase 1 (Ask1) in the p38 signalling cascade. *Biochem. J.* 372, 631–641.
- Detzer, A., Engel, C., Wünsche, W., and Sczakiel, G. (2011). Cell stress is related to re-localization of Argonaute 2 and to decreased RNA interference in human cells. *Nucleic Acids Res.* 39, 2727–2741.
- Dezwaan, D.C., and Freeman, B.C. (2008). HSP90: the Rosetta stone for cellular protein dynamics? *Cell Cycle Georget. Tex* 7, 1006–1012.
- Ding, L., and Han, M. (2007). GW182 family proteins are crucial for microRNA-mediated gene silencing. *Trends Cell Biol.* 17, 411–416.

- Ding, X.Z., Smallridge, R.C., Galloway, R.J., and Kiang, J.G. (1996). Increases in HSF1 translocation and synthesis in human epidermoid A-431 cells: role of protein kinase C and $[Ca^{2+}]_i$. *J. Investig. Med. Off. Publ. Am. Fed. Clin. Res.* *44*, 144–153.
- Ding, X.Z., Tsokos, G.C., and Kiang, J.G. (1998). Overexpression of HSP-70 inhibits the phosphorylation of HSF1 by activating protein phosphatase and inhibiting protein kinase C activity. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *12*, 451–459.
- Dirksen, W.P., Li, X., Mayeda, A., Krainer, A.R., and Rottman, F.M. (2000). Mapping the SF2/ASF binding sites in the bovine growth hormone exonic splicing enhancer. *J. Biol. Chem.* *275*, 29170–29177.
- Dittmer, T.A., and Misteli, T. (2011). The lamin protein family. *Genome Biol.* *12*, 222.
- Dorion, S., Lambert, H., and Landry, J. (2002). Activation of the p38 signaling pathway by heat shock involves the dissociation of glutathione S-transferase Mu from Ask1. *J. Biol. Chem.* *277*, 30792–30797.
- Dostie, J., and Dreyfuss, G. (2002). Translation is required to remove Y14 from mRNAs in the cytoplasm. *Curr. Biol. CB* *12*, 1060–1067.
- Le Dour, C., Schneebeli, S., Bakiri, F., Darcel, F., Jacquemont, M.-L., Maubert, M.-A., Auclair, M., Jeziorowska, D., Reznik, Y., Béréziat, V., et al. (2011). A homozygous mutation of prelamin-A preventing its farnesylation and maturation leads to a severe lipodystrophic phenotype: new insights into the pathogenicity of nonfarnesylated prelamin-A. *J. Clin. Endocrinol. Metab.* *96*, E856–862.
- Dredge, B.K., and Darnell, R.B. (2003). Nova regulates GABA(A) receptor gamma2 alternative splicing via a distal downstream UCAU-rich intronic splicing enhancer. *Mol. Cell. Biol.* *23*, 4687–4700.
- Dreyfuss, G., Matunis, M.J., Piñol-Roma, S., and Burd, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* *62*, 289–321.
- Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* *3*, 195–205.
- Du, P., Ye, L., Yang, Y., and Jiang, W.G. (2013). Candidate of metastasis 1 regulates in vitro growth and invasion of bladder cancer cells. *Int. J. Oncol.* *42*, 1249–1256.
- Duband-Goulet, I., and Courvalin, J.C. (2000). Inner nuclear membrane protein LBR preferentially interacts with DNA secondary structures and nucleosomal linker. *Biochemistry (Mosc.)* *39*, 6483–6488.
- Dutertre, M., Sanchez, G., Barbier, J., Corcos, L., and Auboeuf, D. (2011). The emerging role of pre-messenger RNA splicing in stress responses: sending alternative messages and silent messengers. *RNA Biol.* *8*, 740–747.
- Dwyer, N., and Blobel, G. (1976). A modified procedure for the isolation of a pore complex-lamina fraction from rat liver nuclei. *J Cell Biol* *70*, 581–591.
- Edmond, V., Moysan, E., Khochbin, S., Matthias, P., Brambilla, C., Brambilla, E., Gazzeri, S., and Eymine, B. (2011). Acetylation and phosphorylation of SRSF2 control cell fate decision in response to cisplatin. *EMBO J.* *30*, 510–523.
- Eisinga, R., Breitling, R., and Heskes, T. (2013). The exact probability distribution of the rank product statistics for replicated experiments. *FEBS Lett.* *587*, 677–682.
- Eldridge, A.G., Li, Y., Sharp, P.A., and Blencowe, B.J. (1999). The SRm160/300 splicing coactivator is required for exon-enhancer function. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 6125–6130.

- Ellis, D.J., Jenkins, H., Whitfield, W.G., and Hutchison, C.J. (1997). GST-lamin fusion proteins act as dominant negative mutants in *Xenopus* egg extract and reveal the function of the lamina in DNA replication. *J. Cell Sci.* *110* (Pt 20), 2507–2518.
- Eukashvily, N.I., and Ponomartsev, N.V. (2013). Mammalian satellite DNA: a speaking dumb. *Adv. Protein Chem. Struct. Biol.* *90*, 31–65.
- Eukashvily, N.I., Donev, R., Waisertreiger, I.S.-R., and Podgornaya, O.I. (2007). Human chromosome 1 satellite 3 DNA is decondensed, demethylated and transcribed in senescent cells and in A431 epithelial carcinoma cells. *Cytogenet. Genome Res.* *118*, 42–54.
- Eperon, I.C., Makarova, O.V., Mayeda, A., Munroe, S.H., Cáceres, J.F., Hayward, D.G., and Krainer, A.R. (2000). Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1. *Mol. Cell. Biol.* *20*, 8303–8318.
- Eriksson, M., Brown, W.T., Gordon, L.B., Glynn, M.W., Singer, J., Scott, L., Erdos, M.R., Robbins, C.M., Moses, T.Y., and Berglund, P. (2003). Recurrent de novo point mutations in lamin A cause Hutchinson–Gilford progeria syndrome. *Nature* *423*, 293–298.
- Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008). GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nat. Struct. Mol. Biol.* *15*, 346–353.
- Eulalio, A., Helms, S., Fritsch, C., Fauser, M., and Izaurralde, E. (2009a). A C-terminal silencing domain in GW182 is essential for miRNA function. *RNA* *15*, 1067–1077.
- Eulalio, A., Tritschler, F., Buttner, R., Weichenrieder, O., Izaurralde, E., and Truffault, V. (2009b). The RRM domain in GW182 proteins contributes to miRNA-mediated gene silencing. *Nucleic Acids Res.* *37*, 2974–2983.
- Eulalio, A., Tritschler, F., and Izaurralde, E. (2009c). The GW182 protein family in animal cells: new insights into domains required for miRNA-mediated gene silencing. *Rna* *15*, 1433–1442.
- Evsyukova, I., Bradrick, S.S., Gregory, S.G., and Garcia-Blanco, M.A. (2013). Cleavage and polyadenylation specificity factor 1 (CPSF1) regulates alternative splicing of interleukin 7 receptor (IL7R) exon 6. *RNA N. Y. N* *19*, 103–115.
- Eymery, A., Horard, B., Atifi-Borel, M.E., Fourel, G., Berger, F., Vitte, A.-L., Van den Broeck, A., Brambilla, E., Fournier, A., Callanan, M., et al. (2009a). A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells. *Nucleic Acids Res.* *37*, 6340–6354.
- Eymery, A., Callanan, M., and Vourc'h, C. (2009b). The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *Int. J. Dev. Biol.* *53*, 259–268.
- Eymery, A., Souchier, C., Vourc'h, C., and Jolly, C. (2010). Heat shock factor 1 binds to and transcribes satellite II and III sequences at several pericentromeric regions in heat-shocked cells. *Exp. Cell Res.* *316*, 1845–1855.
- Eystathioy, T., Chan, E.K.L., Tenenbaum, S.A., Keene, J.D., Griffith, K., and Fritzler, M.J. (2002). A phosphorylated cytoplasmic autoantigen, gw182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. *Mol. Biol. Cell* *13*, 1338–1351.
- Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., and Lührmann, R. (2009). The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Mol. Cell* *36*, 593–608.

- Fairbrother, W.G., and Chasin, L.A. (2000). Human genomic sequences that inhibit splicing. *Mol. Cell Biol.* *20*, 6816–6825.
- Fakan, S., Leser, G., and Martin, T.E. (1984). Ultrastructural distribution of nuclear ribonucleoproteins as visualized by immunocytochemistry on thin sections. *J. Cell Biol.* *98*, 358–363.
- Fang, L., Zhan, S., Huang, C., Cheng, X., Lv, X., Si, H., and Li, J. (2013). TRPM7 channel regulates PDGF-BB-induced proliferation of hepatic stellate cells via PI3K and ERK pathways. *Toxicol. Appl. Pharmacol.* *272*, 713–725.
- Faustino, N.A., and Cooper, T.A. (2003). Pre-mRNA splicing and human disease. *Genes Dev.* *17*, 419–437.
- Feng, Y., Chen, M., and Manley, J.L. (2008). Phosphorylation switches the general splicing repressor SRp38 to a sequence-specific activator. *Nat. Struct. Mol. Biol.* *15*, 1040–1048.
- Ferraiuolo, M.A., Basak, S., Dostie, J., Murray, E.L., Schoenberg, D.R., and Sonenberg, N. (2005). A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. *J. Cell Biol.* *170*, 913–924.
- Ferraris, J.D., Williams, C.K., Persaud, P., Zhang, Z., Chen, Y., and Burg, M.B. (2002). Activity of the TonEBP/OREBP transactivation domain varies directly with extracellular NaCl concentration. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 739–744.
- Ferris, D.K., Harel-Bellan, A., Morimoto, R.I., Welch, W.J., and Farrar, W.L. (1988). Mitogen and lymphokine stimulation of heat shock proteins in T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* *85*, 3850.
- Finelli, P., Antonacci, R., Marzella, R., Lonoce, A., Archidiacono, N., and Rocchi, M. (1996). Structural organization of multiple alphoid subsets coexisting on human chromosomes 1, 4, 5, 7, 9, 15, 18, and 19. *Genomics* *38*, 325–330.
- Fisher, D.Z., Chaudhary, N., and Blobel, G. (1986). cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins. *Proc. Natl. Acad. Sci. U. S. A.* *83*, 6450–6454.
- Fong, L.G., Ng, J.K., Meta, M., Coté, N., Yang, S.H., Stewart, C.L., Sullivan, T., Burghardt, A., Majumdar, S., Reue, K., et al. (2004). Heterozygosity for *Lmna* deficiency eliminates the progeria-like phenotypes in *Zmpste24*-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 18111–18116.
- Franklin, A.E., and Cande, W.Z. (1999). Nuclear organization and chromosome segregation. *Plant Cell* *11*, 523–534.
- Fu, X.D., and Maniatis, T. (1992). Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science* *256*, 535–538.
- Fu, J., Bian, L., Zhao, L., Dong, Z., Gao, X., Luan, H., Sun, Y., and Song, H. (2010). Identification of genes for normalization of quantitative real-time PCR data in ovarian tissues. *Acta Biochim. Biophys. Sin.* *42*, 568–574.
- Fuchs, E., and Weber, K. (1994). Intermediate filaments: structure, dynamics, function and disease. *Annu. Rev. Biochem.* *63*, 345–382.
- Fugier, C., Klein, A.F., Hammer, C., Vassilopoulos, S., Ivarsson, Y., Toussaint, A., Tosch, V., Vignaud, A., Ferry, A., Messaddeq, N., et al. (2011). Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat. Med.* *17*, 720–725.

- Fujita, N., Sato, S., Ishida, A., and Tsuruo, T. (2002). Involvement of Hsp90 in signaling and stability of 3-phosphoinositide-dependent kinase-1. *J. Biol. Chem.* *277*, 10346–10353.
- Fukuchi, K., Katsuya, T., Sugimoto, K., Kuremura, M., Kim, H.D., Li, L., and Ogihara, T. (2004). LMNA mutation in a 45 year old Japanese subject with Hutchinson-Gilford progeria syndrome. *J. Med. Genet.* *41*, e67.
- Furger, A., O'Sullivan, J.M., Binnie, A., Lee, B.A., and Proudfoot, N.J. (2002). Promoter proximal splice sites enhance transcription. *Genes Dev.* *16*, 2792–2799.
- Furukawa, K., and Hotta, Y. (1993). cDNA cloning of a germ cell specific lamin B3 from mouse spermatocytes and analysis of its function by ectopic expression in somatic cells. *EMBO J.* *12*, 97.
- Furukawa, K., and Kondo, T. (1998). Identification of the lamina-associated-polypeptide-2-binding domain of B-type lamin. *Eur. J. Biochem. FEBS* *251*, 729–733.
- Furukawa, K., Fritze, C.E., and Gerace, L. (1998). The major nuclear envelope targeting domain of LAP2 coincides with its lamin binding region but is distinct from its chromatin interaction domain. *J. Biol. Chem.* *273*, 4213–4219.
- Furusawa, Y., Tabuchi, Y., Takasaki, I., Wada, S., Ohtsuka, K., and Kondo, T. (2009). Gene networks involved in apoptosis induced by hyperthermia in human lymphoma U937 cells. *Cell Biol. Int.* *33*, 1253–1262.
- Gallinaro, H., and Jacob, M. (1981). The status of small nuclear RNA in the ribonucleoprotein fibrils containing heterogeneous nuclear RNA. *Biochim. Biophys. Acta* *652*, 109–120.
- Gama-Carvalho, M., and Carmo-Fonseca, M. (2001). The rules and roles of nucleocytoplasmic shuttling proteins. *FEBS Lett.* *498*, 157–163.
- Ganot, P., Jády, B.E., Bortolin, M.L., Darzacq, X., and Kiss, T. (1999). Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. *Mol. Cell. Biol.* *19*, 6906–6917.
- Garrido, C., Bruey, J.-M., Fromentin, A., Hammann, A., Arrigo, A.P., and Solary, E. (1999). HSP27 inhibits cytochrome c-dependent activation of procaspase-9. *FASEB J.* *13*, 2061–2070.
- Garrido, C., Paul, C., Seigneuric, R., and Kampinga, H.H. (2012). The small heat shock proteins family: the long forgotten chaperones. *Int. J. Biochem. Cell Biol.* *44*, 1588–1592.
- Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. *Mol. Biol. Cell* *11*, 4241–4257.
- Gatfield, D., Le Hir, H., Schmitt, C., Braun, I.C., Köcher, T., Wilm, M., and Izaurralde, E. (2001). The DEXH/D box protein HEL/UAP56 is essential for mRNA nuclear export in *Drosophila*. *Curr. Biol. CB* *11*, 1716–1721.
- Gattoni, R., Chebli, K., Himmelsbach, M., and Stévenin, J. (1991). Modulation of alternative splicing of adenoviral E1A transcripts: factors involved in the early-to-late transition. *Genes Dev.* *5*, 1847–1858.
- Gattoni, R., Mahé, D., Mähl, P., Fischer, N., Mattei, M.-G., Stévenin, J., and Fuchs, J.-P. (1996). The human hnRNP-M proteins: structure and relation with early heat shock-induced splicing arrest and chromosome mapping. *Nucleic Acids Res.* *24*, 2535–2542.
- Ge, H., Zuo, P., and Manley, J.L. (1991). Primary structure of the human splicing factor ASF reveals similarities with *Drosophila* regulators. *Cell* *66*, 373–382.

- Gems, D., and Doonan, R. (2009). Antioxidant defense and aging in *C. elegans*: is the oxidative damage theory of aging wrong? *Cell Cycle Georget. Tex* 8, 1681–1687.
- Gerace, L., and Blobel, G. (1980). The nuclear envelope lamina is reversibly depolymerized during mitosis. *Cell* 19, 277–287.
- Ghetti, A., Piñol-Roma, S., Michael, W.M., Morandi, C., and Dreyfuss, G. (1992). hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Res.* 20, 3671–3678.
- Giacinti, C., and Giordano, A. (2006). RB and cell cycle progression. *Oncogene* 25, 5220–5227.
- Gilford, H. (1897). On a condition of mixed premature and immature development. *Med Chir Trans* 80, 17–46 25.
- Girard, C., Will, C.L., Peng, J., Makarov, E.M., Kastner, B., Lemm, I., Urlaub, H., Hartmuth, K., and Lührmann, R. (2012). Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat. Commun.* 3, 994.
- Giuliodori, A.M., Di Pietro, F., Marzi, S., Masquida, B., Wagner, R., Romby, P., Gualerzi, C.O., and Pon, C.L. (2010). The *cspA* mRNA is a thermosensor that modulates translation of the cold-shock protein CspA. *Mol. Cell* 37, 21–33.
- Glass, J.R., and Gerace, L. (1990). Lamins A and C bind and assemble at the surface of mitotic chromosomes. *J. Cell Biol.* 111, 1047–1057.
- Glynn, M.W., and Glover, T.W. (2005). Incomplete processing of mutant lamin A in Hutchinson-Gilford progeria leads to nuclear abnormalities, which are reversed by farnesyltransferase inhibition. *Hum. Mol. Genet.* 14, 2959–2969.
- Goldberg, M., Jenkins, H., Allen, T., Whitfield, W.G., and Hutchison, C.J. (1995). *Xenopus* lamin B3 has a direct role in the assembly of a replication competent nucleus: evidence from cell-free egg extracts. *J. Cell Sci.* 108 (Pt 11), 3451–3461.
- Goldberg, M., Harel, A., Brandeis, M., Rechsteiner, T., Richmond, T.J., Weiss, A.M., and Gruenbaum, Y. (1999). The tail domain of lamin Dm0 binds histones H2A and H2B. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2852–2857.
- Goldman, A.E., Moir, R.D., Montag-Lowy, M., Stewart, M., and Goldman, R.D. (1992). Pathway of incorporation of microinjected lamin A into the nuclear envelope. *J. Cell Biol.* 119, 725–735.
- Goldman, R.D., Gruenbaum, Y., Moir, R.D., Shumaker, D.K., and Spann, T.P. (2002). Nuclear lamins: building blocks of nuclear architecture. *Genes Dev.* 16, 533–547.
- Goldman, R.D., Shumaker, D.K., Erdos, M.R., Eriksson, M., Goldman, A.E., Gordon, L.B., Gruenbaum, Y., Khuon, S., Mendez, M., Varga, R., et al. (2004). Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. *Proc. Natl. Acad. Sci. U. S. A.* 101, 8963–8968.
- Gordon, L.B., Kleinman, M.E., Miller, D.T., Neuberger, D.S., Giobbie-Hurder, A., Gerhard-Herman, M., Smoot, L.B., Gordon, C.M., Cleveland, R., Snyder, B.D., et al. (2012). Clinical trial of a farnesyltransferase inhibitor in children with Hutchinson-Gilford progeria syndrome. *Proc. Natl. Acad. Sci. U. S. A.* 109, 16666–16671.
- Gorlach, M., Wittekind, M., Beckman, R.A., Mueller, L., and Dreyfuss, G. (1992). Interaction of the RNA-binding domain of the hnRNP C proteins with RNA. *EMBO J.* 11, 3289–3295.

- Grady, D.L., Ratliff, R.L., Robinson, D.L., McCanlies, E.C., Meyne, J., and Moyzis, R.K. (1992). Highly conserved repetitive DNA sequences are present at human centromeres. *Proc. Natl. Acad. Sci. U. S. A.* *89*, 1695.
- Graveley, B.R. (2000). Sorting out the complexity of SR protein functions. *Rna* *6*, 1197–1211.
- Grishin, N.V. (2001). KH domain: one motif, two folds. *Nucleic Acids Res.* *29*, 638–643.
- Gruenbaum, Y., Goldman, R.D., Meyuhass, R., Mills, E., Margalit, A., Fridkin, A., Dayani, Y., Prokocimer, M., and Enosh, A. (2003). The nuclear lamina and its functions in the nucleus. *Int. Rev. Cytol.* *226*, 1–62.
- GuhaThakurta, D., Palomar, L., Stormo, G.D., Tedesco, P., Johnson, T.E., Walker, D.W., Lithgow, G., Kim, S., and Link, C.D. (2002). Identification of a novel cis-regulatory element involved in the heat shock response in *Caenorhabditis elegans* using microarray gene expression and computational methods. *Genome Res.* *12*, 701–712.
- Guil, S., Long, J.C., and Cáceres, J.F. (2006). hnRNP A1 relocalization to the stress granules reflects a role in the stress response. *Mol. Cell. Biol.* *26*, 5744–5758.
- Haddad, N., and Paulin-Levasseur, M. (2008). Effects of heat shock on the distribution and expression levels of nuclear proteins in HeLa S3 cells. *J. Cell. Biochem.* *105*, 1485–1500.
- Hahn, M., Dambacher, S., Dulev, S., Kuznetsova, A.Y., Eck, S., Wörz, S., Sadic, D., Schulte, M., Mallm, J.-P., Maiser, A., et al. (2013). Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes Dev.* *27*, 859–872.
- Hallay, H., Locker, N., Ayadi, L., Ropers, D., Guittet, E., and Branlant, C. (2006). Biochemical and NMR study on the competition between proteins SC35, SRp40, and heterogeneous nuclear ribonucleoprotein A1 at the HIV-1 Tat exon 2 splicing site. *J. Biol. Chem.* *281*, 37159–37174.
- Han, K., Yeo, G., An, P., Burge, C.B., and Grabowski, P.J. (2005). A combinatorial code for splicing silencing: UAGG and GGGG motifs. *PLoS Biol.* *3*, e158.
- Han, S.P., Tang, Y.H., and Smith, R. (2010a). Functional diversity of the hnRNPs: past, present and perspectives. *Biochem. J.* *430*, 379–392.
- Han, S.P., Kassahn, K.S., Skarszewski, A., Ragan, M.A., Rothnagel, J.A., and Smith, R. (2010b). Functional implications of the emergence of alternative splicing in hnRNP A/B transcripts. *RNA N. Y. N* *16*, 1760–1768.
- Hanamura, A., Cáceres, J.F., Mayeda, A., Franza, B.R., Jr, and Krainer, A.R. (1998). Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. *RNA N. Y. N* *4*, 430–444.
- Harborth, J., Elbashir, S.M., Bechert, K., Tuschl, T., and Weber, K. (2001). Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J. Cell Sci.* *114*, 4557–4565.
- Harel, J., Hanania, N., Tapiero, H., and Harel, L. (1968). RNA replication by nuclear satellite DNA in different mouse cells. *Biochem. Biophys. Res. Commun.* *33*, 696–701.
- Harman, D. (1956). Aging: a theory based on free radical and radiation chemistry. *J. Gerontol.* *11*, 298–300.
- Hasegawa, Y., Brockdorff, N., Kawano, S., Tsutui, K., Tsutui, K., and Nakagawa, S. (2010). The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell* *19*, 469–476.

- Hautbergue, G.M., Hung, M.-L., Golovanov, A.P., Lian, L.-Y., and Wilson, S.A. (2008). Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 5154–5159.
- Hegde, M.L., Banerjee, S., Hegde, P.M., Bellot, L.J., Hazra, T.K., Boldogh, I., and Mitra, S. (2012). Enhancement of NEIL1 protein-initiated oxidized DNA base excision repair by heterogeneous nuclear ribonucleoprotein U (hnRNP-U) via direct interaction. *J. Biol. Chem.* *287*, 34202–34211.
- Heinrichs, V., Bach, M., Winkelmann, G., and Lührmann, R. (1990). U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site. *Science* *247*, 69–72.
- Heintz, N., Sive, H.L., and Roeder, R.G. (1983). Regulation of human histone gene expression: kinetics of accumulation and changes in the rate of synthesis and in the half-lives of individual histone mRNAs during the HeLa cell cycle. *Mol. Cell. Biol.* *3*, 539–550.
- Heldin, C.-H., and Lennartsson, J. (2013). Structural and functional properties of platelet-derived growth factor and stem cell factor receptors. *Cold Spring Harb. Perspect. Biol.* *5*, a009100.
- Henzel, M.J., Delcuve, G.P., and Davie, J.R. (1991). Histone deacetylase is a component of the internal nuclear matrix. *J. Biol. Chem.* *266*, 21936–21942.
- Hernandez, L., Roux, K.J., Wong, E.S.M., Mounkes, L.C., Mutalif, R., Navasankari, R., Rai, B., Cool, S., Jeong, J.-W., Wang, H., et al. (2010). Functional Coupling between the Extracellular Matrix and Nuclear Lamina by Wnt Signaling in Progeria. *Dev. Cell* *19*, 413–425.
- Heydari, A.R., Wu, B., Takahashi, R., Strong, R., and Richardson, A. (1993). Expression of heat shock protein 70 is altered by age and diet at the level of transcription. *Mol. Cell. Biol.* *13*, 2909–2918.
- Hietakangas, V., Ahlskog, J.K., Jakobsson, A.M., Hellesuo, M., Sahlberg, N.M., Holmberg, C.I., Mikhailov, A., Palvimo, J.J., Pirkkala, L., and Sistonen, L. (2003). Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. *Mol. Cell. Biol.* *23*, 2953–2968.
- Hillman, R.T., Green, R.E., and Brenner, S.E. (2004a). An unappreciated role for RNA surveillance. *Genome Biol.* *5*.
- Hillman, R.T., Green, R.E., and Brenner, S.E. (2004b). An unappreciated role for RNA surveillance. *Genome Biol.* *5*, R8.
- Le Hir, H., Moore, M.J., and Maquat, L.E. (2000a). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes Dev.* *14*, 1098–1108.
- Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000b). The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* *19*, 6860–6869.
- Hock, J., Weinmann, L., Ender, C., Rüdell, S., Kremmer, E., Raabe, M., Urlaub, H., and Meister, G. (2007). Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep.* *8*, 1052–1060.
- Holmberg, C.I., Hietakangas, V., Mikhailov, A., Rantanen, J.O., Kallio, M., Meinander, A., Hellman, J., Morrice, N., MacKintosh, C., Morimoto, R.I., et al. (2001). Phosphorylation of serine 230 promotes inducible transcriptional activity of heat shock factor 1. *EMBO J.* *20*, 3800–3810.
- Holmer, L., Pezhman, A., and Worman, H.J. (1998). The human lamin B receptor/sterol reductase multigene family. *Genomics* *54*, 469–476.

- Holt, I., Mittal, S., Furling, D., Butler-Browne, G.S., Brook, J.D., and Morris, G.E. (2007). Defective mRNA in myotonic dystrophy accumulates at the periphery of nuclear splicing speckles. *Genes Cells* *12*, 1035–1048.
- Hori, T.S., Gamperl, A.K., Afonso, L.O., Johnson, S.C., Hubert, S., Kimball, J., Bowman, S., and Rise, M.L. (2010). Heat-shock responsive genes identified and validated in Atlantic cod (*Gadus morhua*) liver, head kidney and skeletal muscle using genomic techniques. *BMC Genomics* *11*, 72.
- Hoskins, A.A., and Moore, M.J. (2012). The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.* *37*, 179–188.
- Hovhannisyann, R.H., and Carstens, R.P. (2007). Heterogeneous ribonucleoprotein m is a splicing regulatory protein that can enhance or silence splicing of alternatively spliced exons. *J. Biol. Chem.* *282*, 36265–36274.
- Huang, Y., and Steitz, J.A. (2001). Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol. Cell* *7*, 899–905.
- Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* *4*, 44–57.
- Huang, S., Chen, L., Libina, N., Janes, J., Martin, G.M., Campisi, J., and Oshima, J. (2005). Correction of cellular phenotypes of Hutchinson-Gilford Progeria cells by RNA interference. *Hum. Genet.* *118*, 444–450.
- Huang, S., Risques, R.A., Martin, G.M., Rabinovitch, P.S., and Oshima, J. (2008). Accelerated telomere shortening and replicative senescence in human fibroblasts overexpressing mutant and wild-type lamin A. *Exp. Cell Res.* *314*, 82–91.
- Huang, Y., Gattoni, R., Stévenin, J., and Steitz, J.A. (2003). SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol. Cell* *11*, 837–843.
- Humphray, S.J., Oliver, K., Hunt, A.R., Plumb, R.W., Loveland, J.E., Howe, K.L., Andrews, T.D., Searle, S., Hunt, S.E., Scott, C.E., et al. (2004). DNA sequence and analysis of human chromosome 9. *Nature* *429*, 369–374.
- Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat. Rev. Genet.* *12*, 99–110.
- Huntzinger, E., Braun, J.E., Heimstädt, S., Zekri, L., and Izaurralde, E. (2010). Two PABPC1-binding sites in GW182 proteins promote miRNA-mediated gene silencing. *EMBO J.* *29*, 4146–4160.
- Huntzinger, E., Kuzuoglu-Ozturk, D., Braun, J.E., Eulalio, A., Wohlbold, L., and Izaurralde, E. (2012). The interactions of GW182 proteins with PABP and deadenylases are required for both translational repression and degradation of miRNA targets. *Nucleic Acids Res.* *41*, 978–994.
- Huot, M.-É., Vogel, G., Zabarauskas, A., Ngo, C.T.-A., Coulombe-Huntington, J., Majewski, J., and Richard, S. (2012). The Sam68 STAR RNA-binding protein regulates mTOR alternative splicing during adipogenesis. *Mol. Cell* *46*, 187–199.
- Hutchinson, J. (1886). Congenital Absence of Hair and Mammary Glands with Atrophic Condition of the Skin and its Appendages, in a Boy whose Mother had been almost wholly Bald from Alopecia Areata from the age of Six. *Med Chir Trans* *69*, 473–477.

- Iborra, F.J., Jackson, D.A., and Cook, P.R. (1998). The path of transcripts from extra-nucleolar synthetic sites to nuclear pores: transcripts in transit are concentrated in discrete structures containing SR proteins. *J. Cell Sci.* *111* (Pt 15), 2269–2282.
- Ibrahim, E.C., Schaal, T.D., Hertel, K.J., Reed, R., and Maniatis, T. (2005). Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 5002–5007.
- Imamoto, N., Tachibana, T., Matsubae, M., and Yoneda, Y. (1995). A karyophilic protein forms a stable complex with cytoplasmic components prior to nuclear pore binding. *J. Biol. Chem.* *270*, 8559–8565.
- Isken, O., Kim, Y.K., Hosoda, N., Mayeur, G.L., Hershey, J.W.B., and Maquat, L.E. (2008). Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. *Cell* *133*, 314–327.
- Jacquet, S., Ropers, D., Bilodeau, P.S., Damier, L., Mougin, A., Stoltzfus, C.M., and Branlant, C. (2001). Conserved stem-loop structures in the HIV-1 RNA region containing the A3 3' splice site and its cis-regulatory element: possible involvement in RNA splicing. *Nucleic Acids Res.* *29*, 464–478.
- Jagatheesan, G., Thanumalayan, S., Muralikrishna, B., Rangaraj, N., Karande, A.A., and Parnaik, V.K. (1999). Colocalization of intranuclear lamin foci with RNA splicing factors. *J. Cell Sci.* *112* (Pt 24), 4651–4661.
- Jahn, D., Schramm, S., Benavente, R., and Alsheimer, M. (2010). Dynamic properties of meiosis-specific lamin C2 and its impact on nuclear envelope integrity. *Nucleus* *1*, 273–283.
- Jensen, T.H., Boulay, J., Rosbash, M., and Libri, D. (2001). The DECD box putative ATPase Sub2p is an early mRNA export factor. *Curr. Biol. CB* *11*, 1711–1715.
- Jiang, W.G., Davies, G., Martin, T.A., Kynaston, H., Mason, M.D., and Fodstad, O. (2006). Com-1/p8 acts as a putative tumour suppressor in prostate cancer. *Int. J. Mol. Med.* *18*, 981–986.
- Jin, H.-O., Seo, S.-K., Woo, S.-H., Choe, T.-B., Hong, S.-I., Kim, J.-I., and Park, I.-C. (2009). Nuclear protein 1 induced by ATF4 in response to various stressors acts as a positive regulator on the transcriptional activation of ATF4. *IUBMB Life* *61*, 1153–1158.
- Jolly, C. (2004). Stress-induced transcription of satellite III repeats. *J. Cell Biol.* *164*, 25–33.
- Jolly, C., Morimoto, R., Robert-Nicoud, M., and Vourc'h, C. (1997a). HSF1 transcription factor concentrates in nuclear foci during heat shock: relationship with transcription sites. *J. Cell Sci.* *110*, 2935–2941.
- Jolly, C., Mongelard, F., Robert-Nicoud, M., and Vourc'h, C. (1997b). Optimization of nuclear transcript detection by FISH and combination with fluorescence immunocytochemical detection of transcription factors. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* *45*, 1585–1592.
- Jolly, C., Vourc'h, C., Robert-Nicoud, M., and Morimoto, R.I. (1999a). Intron-independent association of splicing factors with active genes. *J. Cell Biol.* *145*, 1133–1143.
- Jolly, C., Usson, Y., and Morimoto, R.I. (1999b). Rapid and reversible relocalization of heat shock factor 1 within seconds to nuclear stress granules. *Proc. Natl. Acad. Sci.* *96*, 6769–6774.
- Jolly, C., Konecny, L., Grady, D.L., Kutsikova, Y.A., Cotto, J.J., Morimoto, R.I., and Vourc'h, C. (2002). In vivo binding of active heat shock transcription factor 1 to human chromosome 9 heterochromatin during stress. *J. Cell Biol.* *156*, 775–781.

- Jolly, C., Metz, A., Govin, J., Vigneron, M., Turner, B.M., Khochbin, S., and Vourc'h, C. (2004). Stress-induced transcription of satellite III repeats. *J. Cell Biol.* *164*, 25–33.
- Jurica, M.S., and Moore, M.J. (2003). Pre-mRNA splicing: awash in a sea of proteins. *Mol. Cell* *12*, 5–14.
- Jurica, M.S., Sousa, D., Moore, M.J., and Grigorieff, N. (2004). Three-dimensional structure of C complex spliceosomes by electron microscopy. *Nat. Struct. Mol. Biol.* *11*, 265–269.
- Kabakov, A.E., Budagova, K.R., Latchman, D.S., and Kampinga, H.H. (2002). Stressful preconditioning and HSP70 overexpression attenuate proteotoxicity of cellular ATP depletion. *Am. J. Physiol. - Cell Physiol.* *283*, C521–C534.
- Kafasla, P., Mickleburgh, I., Llorian, M., Coelho, M., Gooding, C., Cherny, D., Joshi, A., Kotik-Kogan, O., Curry, S., Eperon, I.C., et al. (2012). Defining the roles and interactions of PTB. *Biochem. Soc. Trans.* *40*, 815–820.
- Kamma, H., Portman, D.S., and Dreyfuss, G. (1995). Cell type-specific expression of hnRNP proteins. *Exp. Cell Res.* *221*, 187–196.
- Kampinga, H.H., Hageman, J., Vos, M.J., Kubota, H., Tanguay, R.M., Bruford, E.A., Cheetham, M.E., Chen, B., and Hightower, L.E. (2009). Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress Chaperones* *14*, 105–111.
- Kanopka, A., Mühlemann, O., Petersen-Mahrt, S., Estmer, C., Ohrmalm, C., and Akusjärvi, G. (1998). Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins. *Nature* *393*, 185–187.
- Kapur, K., Jiang, H., Xing, Y., and Wong, W.H. (2008). Cross-hybridization modeling on Affymetrix exon arrays. *Bioinforma. Oxf. Engl.* *24*, 2887–2893.
- Kataoka, N., Bachorik, J.L., and Dreyfuss, G. (1999). Transportin-SR, a nuclear import receptor for SR proteins. *J. Cell Biol.* *145*, 1145–1152.
- Katz, Y., Wang, E.T., Airoidi, E.M., and Burge, C.B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods* *7*, 1009–1015.
- Kedersha, N., and Anderson, P. (2007). Mammalian stress granules and processing bodies. *Methods Enzymol.* *431*, 61–81.
- Kedersha, N., Chen, S., Gilks, N., Li, W., Miller, I.J., Stahl, J., and Anderson, P. (2002). Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol. Biol. Cell* *13*, 195–210.
- Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol.* *169*, 871–884.
- Kelley, J.B., and Paschal, B.M. (2007). Hyperosmotic stress signaling to the nucleus disrupts the Ran gradient and the production of RanGTP. *Mol. Biol. Cell* *18*, 4365–4376.
- Kelley, J.B., Datta, S., Snow, C.J., Chatterjee, M., Ni, L., Spencer, A., Yang, C.-S., Cubenas-Potts, C., Matunis, M.J., and Paschal, B.M. (2011). The Defective Nuclear Lamina in Hutchinson-Gilford Progeria Syndrome Disrupts the Nucleocytoplasmic Ran Gradient and Inhibits Nuclear Localization of Ubc9. *Mol. Cell Biol.* *31*, 3378–3395.

- Khodor, Y.L., Rodriguez, J., Abruzzi, K.C., Tang, C.-H.A., Marr, M.T., 2nd, and Rosbash, M. (2011). Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in *Drosophila*. *Genes Dev.* *25*, 2502–2512.
- Khodor, Y.L., Menet, J.S., Tolan, M., and Rosbash, M. (2012). Cotranscriptional splicing efficiency differs dramatically between *Drosophila* and mouse. *RNA N. Y. N* *18*, 2174–2186.
- Kiang, J.G., and Tsokos, G.C. (1998). Heat Shock Protein 70 kDa: Molecular Biology, Biochemistry, and Physiology. *Pharmacol. Ther.* *80*, 183–201.
- Kiledjian, M., and Dreyfuss, G. (1992). Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J.* *11*, 2655–2664.
- Kim, Y.K., and Jang, S.K. (2002). Continuous heat shock enhances translational initiation directed by internal ribosomal entry site. *Biochem. Biophys. Res. Commun.* *297*, 224–231.
- Kim, G., Meriin, A.B., Gabai, V.L., Christians, E., Benjamin, I., Wilson, A., Wolozin, B., and Sherman, M.Y. (2012). The heat shock transcription factor Hsf1 is downregulated in DNA damage-associated senescence, contributing to the maintenance of senescence phenotype. *Aging Cell* *11*, 617–627.
- Kim, H.-J., Joo, H.J., Kim, Y.H., Ahn, S., Chang, J., Hwang, K.-B., Lee, D.-H., and Lee, K.-J. (2011). Systemic Analysis of Heat Shock Response Induced by Heat Shock and a Proteasome Inhibitor MG132. *PLoS ONE* *6*, e20252.
- King, C.R., Lemmer, J., Campbell, J.R., and Atkins, A.R. (1978). Osteosarcoma in a patient with Hutchinson-Gilford progeria. *J. Med. Genet.* *15*, 481–484.
- Kline, M.P., and Morimoto, R.I. (1997). Repression of the heat shock factor 1 transcriptional activation domain is modulated by constitutive phosphorylation. *Mol. Cell. Biol.* *17*, 2107–2115.
- Konishi, H., Matsuzaki, H., Tanaka, M., Takemura, Y., Kuroda, S., Ono, Y., and Kikkawa, U. (1997). Activation of protein kinase B (Akt/RAC-protein kinase) by cellular stress and its association with heat shock protein Hsp27. *FEBS Lett.* *410*, 493–498.
- Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J., and Nogues, G. (2004). Multiple links between transcription and splicing. *RNA N. Y. N* *10*, 1489–1498.
- Krecic, A.M., and Swanson, M.S. (1999). hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* *11*, 363–371.
- Kroeger, P.E., Sarge, K.D., and Morimoto, R.I. (1993). Mouse heat shock transcription factors 1 and 2 prefer a trimeric binding site but interact differently with the HSP70 heat shock element. *Mol. Cell. Biol.* *13*, 3370–3383.
- Krol, A., Branlant, C., Lazar, E., Gallinaro, H., and Jacob, M. (1981). Primary and secondary structures of chicken, rat and man nuclear U4 RNAs. Homologies with U1 and U5 RNAs. *Nucleic Acids Res.* *9*, 2699–2716.
- Kudlow, B.A., Stanfel, M.N., Burtner, C.R., Johnston, E.D., and Kennedy, B.K. (2008). Suppression of proliferative defects associated with processing-defective lamin A mutants by hTERT or inactivation of p53. *Mol. Biol. Cell* *19*, 5238–5248.
- Kukalev, A.S., Lobov, I.B., Percipalle, P., and Podgornaya, O.I. (2009). SAF-A/hnRNP-U localization in interphase and metaphase. *Cytogenet. Genome Res.* *124*, 288–297.

- Kukreja, R.C., Kontos, M.C., Loesser, K.E., Batra, S.K., Qian, Y.Z., Gbur, C.J., Jr, Naseem, S.A., Jesse, R.L., and Hess, M.L. (1994). Oxidant stress increases heat shock protein 70 mRNA in isolated perfused rat heart. *Am. J. Physiol.* *267*, H2213–2219.
- Kültz, D. (2003). Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function. *J. Exp. Biol.* *206*, 3119–3124.
- Kültz, D. (2005). Molecular and evolutionary basis of the cellular stress response. *Annu Rev Physiol* *67*, 225–257.
- Kutay, U., Bischoff, F.R., Kostka, S., Kraft, R., and Görlich, D. (1997). Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell* *90*, 1061–1071.
- Lacadie, S.A., and Rosbash, M. (2005). Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5' splice base pairing in yeast. *Mol. Cell* *19*, 65–75.
- Ladd, A.N., Charlet, N., and Cooper, T.A. (2001). The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol. Cell Biol.* *21*, 1285–1296.
- Lamond, A.I., and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. *Nat. Rev. Mol. Cell Biol.* *4*, 605–612.
- Landis, G.N., and Tower, J. (2005). Superoxide dismutase evolution and life span regulation. *Mech. Ageing Dev.* *126*, 365–379.
- Lange, T.S., and Gerbi, S.A. (2000). Transient nucleolar localization Of U6 small nuclear RNA in *Xenopus Laevis* oocytes. *Mol. Biol. Cell* *11*, 2419–2428.
- Lanneau, D., de Thonel, A., Maurel, S., Didelot, C., and Garrido, C. (2007). Apoptosis versus cell differentiation: role of heat shock proteins HSP90, HSP70 and HSP27. *Prion* *1*, 53–60.
- Lapointe, J., and Hekimi, S. (2010). When a theory of aging ages badly. *Cell. Mol. Life Sci. CMLS* *67*, 1–8.
- Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C., and Brenner, S.E. (2007). Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* *446*, 926–929.
- Larkindale, J., and Vierling, E. (2008). Core genome responses involved in acclimation to high temperature. *Plant Physiol.* *146*, 748–761.
- Lavigueur, A., La Branche, H., Kornblihtt, A.R., and Chabot, B. (1993). A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes Dev.* *7*, 2405–2417.
- Lazzaretti, D., Tournier, I., and Izaurralde, E. (2009). The C-terminal domains of human TNRC6A, TNRC6B, and TNRC6C silence bound transcripts independently of Argonaute proteins. *Rna* *15*, 1059–1066.
- Lee, B.S., Chen, J., Angelidis, C., Jurivich, D.A., and Morimoto, R.I. (1995). Pharmacological modulation of heat shock factor 1 by antiinflammatory drugs results in protection against stress-induced cellular damage. *Proc. Natl. Acad. Sci. U. S. A.* *92*, 7207.
- Lee, R., Chang, S.Y., Trinh, H., Tu, Y., White, A.C., Davies, B.S.J., Bergo, M.O., Fong, L.G., Lowry, W.E., and Young, S.G. (2010). Genetic studies on the functional relevance of the protein prenyltransferases in skin keratinocytes. *Hum. Mol. Genet.* *19*, 1603–1617.

- Lei, H., Dias, A.P., and Reed, R. (2011). Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 17985–17990.
- Lenz-Böhme, B., Wismar, J., Fuchs, S., Reifegerste, R., Buchner, E., Betz, H., and Schmitt, B. (1997). Insertional mutation of the *Drosophila* nuclear lamin Dm0 gene results in defective nuclear envelopes, clustering of nuclear pore complexes, and accumulation of annulate lamellae. *J. Cell Biol.* *137*, 1001–1016.
- Levinson, W., Oppermann, H., and Jackson, J. (1980). Transition series metals and sulfhydryl reagents induce the synthesis of four proteins in eukaryotic cells. *Biochim. Biophys. Acta* *606*, 170–180.
- Lewis, J.D., and Izaurralde, E. (1997). The role of the cap structure in RNA processing and nuclear export. *Eur. J. Biochem. FEBS* *247*, 461–469.
- Lewis, J.D., Izaurralde, E., Jarmolowski, A., McGuigan, C., and Mattaj, I.W. (1996). A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev.* *10*, 1683–1698.
- Li, X., Martin, T.A., and Jiang, W.G. (2012). COM-1/p8 acts as a tumour growth enhancer in colorectal cancer cell lines. *Anticancer Res.* *32*, 1229–1237.
- Libri, D., Piseri, A., and Fiszman, M.Y. (1991). Tissue-specific splicing in vivo of the beta-tropomyosin gene: dependence on an RNA secondary structure. *Science* *252*, 1842–1845.
- Lim, L.P., and Burge, C.B. (2001). A computational analysis of sequence features involved in recognition of short introns. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 11193–11198.
- Lin, R.Z., Hu, Z.W., Chin, J.H., and Hoffman, B.B. (1997). Heat shock activates c-Src tyrosine kinases and phosphatidylinositol 3-kinase in NIH3T3 fibroblasts. *J. Biol. Chem.* *272*, 31196–31202.
- Lindquist, S. (1980). Translational efficiency of heat-induced messages in *Drosophila melanogaster* cells. *J. Mol. Biol.* *137*, 151–158.
- Listerman, I., Sapra, A.K., and Neugebauer, K.M. (2006). Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat. Struct. Mol. Biol.* *13*, 815–822.
- Liu, A.Y., Lin, Z., Choi, H.S., Sorhage, F., and Li, B. (1989). Attenuated induction of heat shock gene expression in aging diploid fibroblasts. *J. Biol. Chem.* *264*, 12037–12045.
- Liu, B., Wang, J., Chan, K.M., Tjia, W.M., Deng, W., Guan, X., Huang, J., Li, K.M., Chau, P.Y., Chen, D.J., et al. (2005). Genomic instability in laminopathy-based premature aging. *Nat. Med.* *11*, 780–785.
- Liu, B., Ghosh, S., Yang, X., Zheng, H., Liu, X., Wang, Z., Jin, G., Zheng, B., Kennedy, B.K., Suh, Y., et al. (2012). Resveratrol rescues SIRT1-dependent adult stem cell decline and alleviates progeroid features in laminopathy-based progeria. *Cell Metab.* *16*, 738–750.
- Liu, B., Wang, Z., Zhang, L., Ghosh, S., Zheng, H., and Zhou, Z. (2013). Depleting the methyltransferase Suv39h1 improves DNA repair and extends lifespan in a progeria mouse model. *Nat. Commun.* *4*, 1868.
- Liu, J., Rolef Ben-Shahar, T., Riemer, D., Treinin, M., Spann, P., Weber, K., Fire, A., and Gruenbaum, Y. (2000). Essential roles for *Caenorhabditis elegans* lamin gene in nuclear organization, cell cycle progression, and spatial organization of nuclear pore complexes. *Mol. Biol. Cell* *11*, 3937–3947.

- Liu, Y., Rusinol, A., Sinensky, M., Wang, Y., and Zou, Y. (2006). DNA damage responses in progeroid syndromes arise from defective maturation of prelamin A. *J. Cell Sci.* *119*, 4644–4649.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* *25*, 402–408.
- Lobov, I.B., Tsutsui, K., Mitchell, A.R., and Podgornaya, O.I. (2000). Specific interaction of mouse major satellite with MAR-binding protein SAF-A. *Eur. J. Cell Biol.* *79*, 839–849.
- Lobov, I.B., Tsutsui, K., Mitchell, A.R., and Podgornaya, O.I. (2001). Specificity of SAF-A and lamin B binding in vitro correlates with the satellite DNA bending state. *J. Cell. Biochem.* *83*, 218–229.
- Lombard, D.B., Chua, K.F., Mostoslavsky, R., Franco, S., Gostissa, M., and Alt, F.W. (2005). DNA repair, genome stability, and aging. *Cell* *120*, 497–512.
- Long, J.C., and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.* *417*, 15–27.
- Louhichi, A., Fourati, A., and Rebaï, A. (2011). IGD: A resource for intronless genes in the human genome. *Gene* *488*, 35–40.
- Lourim, D., and Lin, J.J. (1992). Expression of wild-type and nuclear localization-deficient human lamin A in chick myogenic cells. *J. Cell Sci.* *103*, 863–874.
- Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science* *327*, 996–1000.
- Luco, R.F., Allo, M., Schor, I.E., Kornblihtt, A.R., and Misteli, T. (2011). Epigenetics in alternative pre-mRNA splicing. *Cell* *144*, 16–26.
- Ludérus, M.E., de Graaf, A., Mattia, E., den Blaauwen, J.L., Grande, M.A., de Jong, L., and van Driel, R. (1992). Binding of matrix attachment regions to lamin B1. *Cell* *70*, 949–959.
- Ludérus, M.E., den Blaauwen, J.L., de Smit, O.J., Compton, D.A., and van Driel, R. (1994). Binding of matrix attachment regions to lamin polymers involves single-stranded regions and the minor groove. *Mol. Cell. Biol.* *14*, 6297–6305.
- Lutz, R.J., Trujillo, M.A., Denham, K.S., Wenger, L., and Sinensky, M. (1992). Nucleoplasmic localization of prelamin A: implications for prenylation-dependent lamin A assembly into the nuclear lamina. *Proc. Natl. Acad. Sci. U. S. A.* *89*, 3000–3004.
- Lutzny, G., Kocher, T., Schmidt-Supprian, M., Rudelius, M., Klein-Hitpass, L., Finch, A.J., Dürig, J., Wagner, M., Haferlach, C., Kohlmann, A., et al. (2013). Protein kinase c- β -dependent activation of NF- κ B in stromal cells is indispensable for the survival of chronic lymphocytic leukemia B cells in vivo. *Cancer Cell* *23*, 77–92.
- Ly, D.H., Lockhart, D.J., Lerner, R.A., and Schultz, P.G. (2000). Mitotic misregulation and human aging. *Science* *287*, 2486–2492.
- Ma, N., Matsunaga, S., Morimoto, A., Sakashita, G., Urano, T., Uchiyama, S., and Fukui, K. (2011). The nuclear scaffold protein SAF-A is required for kinetochore-microtubule attachment and contributes to the targeting of Aurora-A to mitotic spindles. *J. Cell Sci.* *124*, 394–404.
- Ma, S., Bhattacharjee, R.B., and Bag, J. (2009). Expression of poly(A)-binding protein is upregulated during recovery from heat shock in HeLa cells. *FEBS J.* *276*, 552–570.

- Mahé, D., Mähl, P., Gattoni, R., Fischer, N., Mattei, M.-G., Stévenin, J., and Fuchs, J.-P. (1997). Cloning of Human 2H9 Heterogeneous Nuclear Ribonucleoproteins. *J. Biol. Chem.* *272*, 1827–1836.
- Mähl, P., Lutz, Y., Puvion, E., and Fuchs, J.-P. (1989). Rapid effect of heat shock on two heterogeneous nuclear ribonucleoprotein-associated antigens in HeLa cells. *J. Cell Biol.* *109*, 1921–1935.
- De Maio, A. (2011). Extracellular heat shock proteins, cellular export vesicles, and the Stress Observation System: a form of communication during injury, infection, and cell damage. It is never known how far a controversial finding will go! Dedicated to Ferruccio Ritossa. *Cell Stress Chaperones* *16*, 235–249.
- Malhas, A.N., Lee, C.F., and Vaux, D.J. (2009). Lamin B1 controls oxidative stress responses via Oct-1. *J. Cell Biol.* *184*, 45–55.
- Mallampalli, M.P., Huyer, G., Bendale, P., Gelb, M.H., and Michaelis, S. (2005). Inhibiting farnesylation reverses the nuclear morphology defect in a HeLa cell model for Hutchinson-Gilford progeria syndrome. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 14416–14421.
- Malmendal, A., Overgaard, J., Bundy, J.G., Sørensen, J.G., Nielsen, N.C., Loeschcke, V., and Holmstrup, M. (2006). Metabolomic profiling of heat stress: hardening and recovery of homeostasis in *Drosophila*. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* *291*, R205–R212.
- Mancini, M.A., Shan, B., Nickerson, J.A., Penman, S., and Lee, W.H. (1994). The retinoblastoma gene product is a cell cycle-dependent, nuclear matrix-associated protein. *Proc. Natl. Acad. Sci. U. S. A.* *91*, 418–422.
- Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature* *416*, 499–506.
- Manju, K., Muralikrishna, B., and Parnaik, V.K. (2006). Expression of disease-causing lamin A mutants impairs the formation of DNA repair foci. *J. Cell Sci.* *119*, 2704–2714.
- Manley, J.L., and Krainer, A.R. (2010). A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes Dev.* *24*, 1073–1074.
- Mao, L., and Yang, Y. (2013). Targeting the nuclear transport machinery by rational drug design. *Curr. Pharm. Des.* *19*, 2318–2325.
- Marchand, V., Méreau, A., Jacquenet, S., Thomas, D., Mougin, A., Gattoni, R., Stévenin, J., and Branlant, C. (2002). A Janus splicing regulatory element modulates HIV-1 tat and rev mRNA production by coordination of hnRNP A1 cooperative binding. *J. Mol. Biol.* *323*, 629–652.
- Marchand, V., Santerre, M., Aigueperse, C., Fouillen, L., Saliou, J.-M., Van Dorselaer, A., Sanglier-Cianfèrari, S., Branlant, C., and Motorin, Y. (2011). Identification of protein partners of the human immunodeficiency virus 1 tat/rev exon 3 leads to the discovery of a new HIV-1 splicing regulator, protein hnRNP K. *RNA Biol.* *8*, 325–342.
- Marin-Vinader, L., Shin, C., Onnekink, C., Manley, J.L., and Lubsen, N.H. (2006). Hsp27 enhances recovery of splicing as well as rephosphorylation of SRp38 after heat shock. *Mol. Biol. Cell* *17*, 886–894.
- Marji, J., O'Donoghue, S.I., McClintock, D., Satagopam, V.P., Schneider, R., Ratner, D., J. Worman, H., Gordon, L.B., and Djabali, K. (2010). Defective Lamin A-Rb Signaling in Hutchinson-Gilford Progeria Syndrome and Reversal by Farnesyltransferase Inhibition. *PLoS ONE* *5*, e11132.

- Marko, M., Leichter, M., Patrino-Georgoula, M., and Guialis, A. (2010). hnRNP M interacts with PSF and p54nrb and co-localizes within defined nuclear structures. *Exp. Cell Res.* *316*, 390–400.
- Marshall, W.F., and Sedat, J.W. (1999). Nuclear architecture. *Results Probl. Cell Differ.* *25*, 283–301.
- Martinez-Contreras, R., Fisette, J.-F., Nasim, F.H., Madden, R., Cordeau, M., and Chabot, B. (2006). Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS Biol.* *4*, e21.
- Martins, S.B., Eide, T., Steen, R.L., Jahnsen, T., Skålhegg B, S., and Collas, P. (2000). HA95 is a protein of the chromatin and nuclear matrix regulating nuclear envelope dynamics. *J. Cell Sci.* *113 Pt 21*, 3703–3713.
- Matsuura, H., Ishibashi, Y., Shinmyo, A., Kanaya, S., and Kato, K. (2010). Genome-wide analyses of early translational responses to elevated temperature and high salinity in *Arabidopsis thaliana*. *Plant Cell Physiol.* *51*, 448–462.
- Matsuzaki, H., Yamamoto, T., and Kikkawa, U. (2004). Distinct activation mechanisms of protein kinase B by growth-factor stimulation and heat-shock treatment. *Biochemistry (Mosc.)* *43*, 4284–4293.
- Mayeda, A., and Krainer, A.R. (1992). Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* *68*, 365–375.
- Mayeda, A., Sreaton, G.R., Chandler, S.D., Fu, X.D., and Krainer, A.R. (1999). Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. *Mol. Cell. Biol.* *19*, 1853–1863.
- Mayes, A.E., Verdone, L., Legrain, P., and Beggs, J.D. (1999). Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J.* *18*, 4321–4331.
- McClintock, D., Ratner, D., Lokuge, M., Owens, D.M., Gordon, L.B., Collins, F.S., and Djabali, K. (2007). The Mutant Form of Lamin A that Causes Hutchinson-Gilford Progeria Is a Biomarker of Cellular Aging in Human Skin. *PLoS ONE* *2*, e1269.
- McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* *385*, 357–361.
- McKeon, F.D., Kirschner, M.W., and Caput, D. (1986). Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. *Nature* *319*, 463–468.
- Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum. Mol. Genet.* *10*, 99–105.
- Meier, J., Campbell, K.H., Ford, C.C., Stick, R., and Hutchison, C.J. (1991). The role of lamin LIII in nuclear assembly and DNA replication, in cell-free extracts of *Xenopus* eggs. *J. Cell Sci.* *98 (Pt 3)*, 271–279.
- Mercier, P.A., Winegarden, N.A., and Westwood, J.T. (1999). Human heat shock factor 1 is predominantly a nuclear protein before and after heat stress. *J. Cell Sci.* *112 (Pt 16)*, 2765–2774.
- Metz, A., Soret, J., Vourc'h, C., Tazi, J., and Jolly, C. (2004). A key role for stress-induced satellite III transcripts in the relocalization of splicing factors into nuclear stress granules. *J. Cell Sci.* *117*, 4551–4558.

- Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.-B., Pagie, L., Kellis, M., Reinders, M., Wessels, L., and van Steensel, B. (2013). Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res.* *23*, 270–280.
- Michlewski, G., Sanford, J.R., and Cáceres, J.F. (2008). The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1. *Mol. Cell* *30*, 179–189.
- Miguet, L., Bogumil, R., Decloquement, P., Herbrecht, R., Potier, N., Mauvieux, L., and Van Dorselaer, A. (2006). Discovery and identification of potential biomarkers in a prospective study of chronic lymphoid malignancies using SELDI-TOF-MS. *J. Proteome Res.* *5*, 2258–2269.
- Milarski, K.L., and Morimoto, R.I. (1986). Expression of human HSP70 during the synthetic phase of the cell cycle. *Proc. Natl. Acad. Sci. U. S. A.* *83*, 9517.
- Miller, M.L., Andringa, A., Dixon, K., and Carty, M.P. (2002). Insights into UV-induced apoptosis: ultrastructure, trichrome stain and spectral imaging. *Micron Oxf. Engl.* *1993* *33*, 157–166.
- Millevoi, S., Loulergue, C., Dettwiler, S., Karaa, S.Z., Keller, W., Antoniou, M., and Vagner, S. (2006). An interaction between U2AF 65 and CF I(m) links the splicing and 3' end processing machineries. *EMBO J.* *25*, 4854–4864.
- Mirzayans, R., Andrais, B., Scott, A., Wang, Y.W., and Murray, D. (2013). Ionizing Radiation-Induced Responses in Human Cells with Differing TP53 Status. *Int. J. Mol. Sci.* *14*, 22409–22435.
- Misteli, T., and Spector, D.L. (1998). The cellular organization of gene expression. *Curr. Opin. Cell Biol.* *10*, 323–331.
- Misteli, T., and Spector, D.L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell* *3*, 697–705.
- Misteli, T., Cáceres, J.F., and Spector, D.L. (1997). The dynamics of a pre-mRNA splicing factor in living cells. *Nature* *387*, 523–527.
- Misteli, T., Cáceres, J.F., Clement, J.Q., Krainer, A.R., Wilkinson, M.F., and Spector, D.L. (1998). Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. *J. Cell Biol.* *143*, 297–307.
- Miyamoto, Y., Saiwaki, T., Yamashita, J., Yasuda, Y., Kotera, I., Shibata, S., Shigeta, M., Hiraoka, Y., Haraguchi, T., and Yoneda, Y. (2004). Cellular stresses induce the nuclear accumulation of importin alpha and cause a conventional nuclear import block. *J. Cell Biol.* *165*, 617–623.
- Modrek, B., and Lee, C. (2002). A genomic view of alternative splicing. *Nat. Genet.* *30*, 13–19.
- Moir, R.D., Spann, T.P., and Goldman, R.D. (1995). The dynamic properties and possible functions of nuclear lamins. *Int. Rev. Cytol.* *162B*, 141–182.
- Moir, R.D., Spann, T.P., Herrmann, H., and Goldman, R.D. (2000a). Disruption of nuclear lamin organization blocks the elongation phase of DNA replication. *J. Cell Biol.* *149*, 1179–1192.
- Moir, R.D., Yoon, M., Khuon, S., and Goldman, R.D. (2000b). Nuclear Lamins a and B1 Different Pathways of Assembly during Nuclear Envelope Formation in Living Cells. *J. Cell Biol.* *151*, 1155–1168.
- Mollet, S., Cougot, N., Wilczynska, A., Dautry, F., Kress, M., Bertrand, E., and Weil, D. (2008). Translationally Repressed mRNA Transiently Cycles through Stress Granules during Stress. *Mol. Biol. Cell* *19*, 4469–4479.

- Moore, M.J., and Sharp, P.A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* 365, 364–368.
- Morimoto, R.I., and Cuervo, A.M. (2009). Protein homeostasis and aging: taking care of proteins from the cradle to the grave. *J. Gerontol. A. Biol. Sci. Med. Sci.* 64, 167–170.
- Morimoto, R.I., and Santoro, M.G. (1998). Stress-inducible responses and heat shock proteins: new pharmacologic targets for cytoprotection. *Nat. Biotechnol.* 16, 833–838.
- Morimoto, R.I., Sarge, K.D., and Abravaya, K. (1992). Transcriptional regulation of heat shock genes. A paradigm for inducible genomic responses. *J. Biol. Chem.* 267, 21987–21990.
- Morita, M.T., Tanaka, Y., Kodama, T.S., Kyogoku, Y., Yanagi, H., and Yura, T. (1999). Translational induction of heat shock transcription factor β : evidence for a built-in RNA thermosensor. *Genes Dev.* 13, 655–665.
- Morley, J.F., Brignull, H.R., Weyers, J.J., and Morimoto, R.I. (2002). The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 99, 10417–10422.
- Moyzis, R.K., Albright, K.L., Bartholdi, M.F., Cram, L.S., Deaven, L.L., Hildebrand, C.E., Joste, N.E., Longmire, J.L., Meyne, J., and Schwarzacher-Robinson, T. (1987). Human chromosome-specific repetitive DNA sequences: novel markers for genetic analysis. *Chromosoma* 95, 375–386.
- Muh, S.J., Hovhannisyanyan, R.H., and Carstens, R.P. (2002). A Non-sequence-specific double-stranded RNA structural element regulates splicing of two mutually exclusive exons of fibroblast growth factor receptor 2 (FGFR2). *J. Biol. Chem.* 277, 50143–50154.
- Müller, B., Blackburn, J., Feijoo, C., Zhao, X., and Smythe, C. (2007). DNA-activated protein kinase functions in a newly observed S phase checkpoint that links histone mRNA abundance with DNA replication. *J. Cell Biol.* 179, 1385–1398.
- Munoz, M.J., Pérez Santangelo, M.S., Paronetto, M.P., de la Mata, M., Pelisch, F., Boireau, S., Glover-Cutter, K., Ben-Dov, C., Blaustein, M., Lozano, J.J., et al. (2009). DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation. *Cell* 137, 708–720.
- Muro, A.F., Caputi, M., Pariyarath, R., Pagani, F., Buratti, E., and Baralle, F.E. (1999). Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. *Mol. Cell. Biol.* 19, 2657–2671.
- Murray, J.I., Whitfield, M.L., Trinklein, N.D., Myers, R.M., Brown, P.O., and Botstein, D. (2004). Diverse and specific gene expression responses to stresses in cultured human cells. *Mol. Biol. Cell* 15, 2361–2374.
- Muto, Y., Pomeranz Krummel, D., Oubridge, C., Hernandez, H., Robinson, C.V., Neuhaus, D., and Nagai, K. (2004). The structure and biochemical properties of the human spliceosomal protein U1C. *J. Mol. Biol.* 341, 185–198.
- Nadeau, S.I., and Landry, J. (2007). Mechanisms of activation and regulation of the heat shock-sensitive signaling pathways. *Adv. Exp. Med. Biol.* 594, 100–113.
- Nagai, K., Oubridge, C., Jessen, T.H., Li, J., and Evans, P.R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* 348, 515–520.

- Nakahata, K., Miyakoda, M., Suzuki, K., Kodama, S., and Watanabe, M. (2002). Heat shock induces centrosomal dysfunction, and causes non-apoptotic mitotic catastrophe in human tumour cells. *Int. J. Hyperth. Off. J. Eur. Soc. Hyperthermic Oncol. North Am. Hyperth. Group* 18, 332–343.
- Nakai, A., and Ishikawa, T. (2001). Cell cycle transition under stress conditions controlled by vertebrate heat shock factors. *EMBO J.* 20, 2885–2895.
- Nakai, A., and Morimoto, R.I. (1993). Characterization of a novel chicken heat shock transcription factor, heat shock factor 3, suggests a new regulatory pathway. *Mol. Cell. Biol.* 13, 1983.
- Nakai, A., Tanabe, M., Kawazoe, Y., Inazawa, J., Morimoto, R.I., and Nagata, K. (1997). HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator. *Mol. Cell. Biol.* 17, 469–481.
- Narberhaus, F. (2010). Translational control of bacterial heat shock and virulence genes by temperature-sensing mRNAs. *RNA Biol.* 7, 84–89.
- Navarro, C.L., De Sandre-Giovannoli, A., Bernard, R., Boccaccio, I., Boyer, A., Geneviève, D., Hadj-Rabia, S., Gaudy-Marqueste, C., Smitt, H.S., Vabres, P., et al. (2004). Lamin A and ZMPSTE24 (FACE-1) defects cause nuclear disorganization and identify restrictive dermopathy as a lethal neonatal laminopathy. *Hum. Mol. Genet.* 13, 2493–2503.
- Nayler, O., Strätling, W., Bourquin, J.-P., Stagljar, I., Lindemann, L., Jasper, H., Hartmann, A.M., Fackelmayer, F.O., Ullrich, A., and Stamm, S. (1998). SAF-B protein couples transcription and pre-mRNA splicing to SAR/MAR elements. *Nucleic Acids Res.* 26, 3542–3549.
- Nikolova, V., Leimena, C., McMahon, A.C., Tan, J.C., Chandar, S., Jogia, D., Kesteven, S.H., Michalicek, J., Otway, R., Verheyen, F., et al. (2004). Defects in nuclear structure and function promote dilated cardiomyopathy in lamin A/C-deficient mice. *J. Clin. Invest.* 113, 357–369.
- Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457–463.
- Ninomiya, K., Kataoka, N., and Hagiwara, M. (2011). Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor. *J. Cell Biol.* 195, 27–40.
- Niwa, M., and Walter, P. (2000). Pausing to decide. *Proc. Natl. Acad. Sci. U. S. A.* 97, 12396–12397.
- Nover, L., Scharf, K.D., and Neumann, D. (1983). Formation of cytoplasmic heat shock granules in tomato cell cultures and leaves. *Mol. Cell. Biol.* 3, 1648–1655.
- Nover, L., Scharf, K.D., and Neumann, D. (1989). Cytoplasmic heat shock granules are formed from precursor particles and are associated with a specific set of mRNAs. *Mol. Cell. Biol.* 9, 1298–1308.
- Oesterreich, S., Lee, A.V., Sullivan, T.M., Samuel, S.K., Davie, J.R., and Fuqua, S.A. (1997). Novel nuclear matrix protein HET binds to and influences activity of the HSP27 promoter in human breast cancer cells. *J. Cell. Biochem.* 67, 275–286.
- Osorio, F.G., Navarro, C.L., Cadiñanos, J., López-Mejía, I.C., Quirós, P.M., Bartoli, C., Rivera, J., Tazi, J., Guzmán, G., Varela, I., et al. (2011). Splicing-directed therapy in a new mouse model of human accelerated aging. *Sci. Transl. Med.* 3, 106ra107.
- Ostareck-Lederer, A., and Ostareck, D.H. (2012). Precision mechanics with multifunctional tools: how hnRNP K and hnRNPs E1/E2 contribute to post-transcriptional control of gene expression in hematopoiesis. *Curr. Protein Pept. Sci.* 13, 391–400.

- Ostberg, J.R., Kaplan, K.C., and Repasky, E.A. (2002). Induction of stress proteins in a panel of mouse tissues by fever-range whole body hyperthermia. *Int. J. Hyperth. Off. J. Eur. Soc. Hyperthermic Oncol. North Am. Hyperth. Group* 18, 552–562.
- Östling, P., Björk, J.K., Roos-Mattjus, P., Mezger, V., and Sistonen, L. (2007). Heat Shock Factor 2 (HSF2) Contributes to Inducible Expression of hsp Genes through Interplay with HSF1. *J. Biol. Chem.* 282, 7077–7086.
- Ozaki, T., Saijo, M., Murakami, K., Enomoto, H., Taya, Y., and Sakiyama, S. (1994). Complex formation between lamin A and the retinoblastoma gene product: identification of the domain on lamin A required for its interaction. *Oncogene* 9, 2649–2653.
- Dhe-Paganon, S. (2002). Structure of the Globular Tail of Nuclear Lamin. *J. Biol. Chem.* 277, 17381–17384.
- Pahlich, S., Quero, L., Roschitzki, B., Leemann-Zakaryan, R.P., and Gehring, H. (2009). Analysis of Ewing sarcoma (EWS)-binding proteins: interaction with hnRNP M, U, and RNA-helicases p68/72 within protein-RNA complexes. *J. Proteome Res.* 8, 4455–4465.
- Palacios, C., Collins, M.K., and Perkins, G.R. (2001). The JNK phosphatase M3/6 is inhibited by protein-damaging stress. *Curr. Biol. CB* 11, 1439–1443.
- Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40, 1413–1415.
- Pandey, P., Farber, R., Nakazawa, A., Kumar, S., Bharti, A., Nalin, C., Weichselbaum, R., Kufe, D., and Kharbanda, S. (2000). Hsp27 functions as a negative regulator of cytochrome c-dependent activation of procaspase-3. *Oncogene* 19, 1975–1981.
- Paradisi, M., McClintock, D., Boguslavsky, R.L., Pedicelli, C., Worman, H.J., and Djabali, K. (2005). Dermal fibroblasts in Hutchinson-Gilford progeria syndrome with the lamin A G608G mutation have dysmorphic nuclei and are hypersensitive to heat stress. *BMC Cell Biol.* 6, 27.
- Park, E., Iaccarino, C., Lee, J., Kwon, I., Baik, S.M., Kim, M., Seong, J.Y., Son, G.H., Borrelli, E., and Kim, K. (2011). Regulatory roles of heterogeneous nuclear ribonucleoprotein M and Nova-1 protein in alternative splicing of dopamine D2 receptor pre-mRNA. *J. Biol. Chem.* 286, 25301–25308.
- Park, H.S., Lee, J.S., Huh, S.H., Seo, J.S., and Choi, E.J. (2001). Hsp72 functions as a natural inhibitory protein of c-Jun N-terminal kinase. *EMBO J.* 20, 446–456.
- Parker, R., and Sheth, U. (2007a). P bodies and the control of mRNA translation and degradation. *Mol. Cell* 25, 635–646.
- Parker, R., and Sheth, U. (2007b). P bodies and the control of mRNA translation and degradation. *Mol. Cell* 25, 635–646.
- Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat. Struct. Mol. Biol.* 11, 121–127.
- Patel, B., Khaliq, A., Jarvis-Evans, J., Boulton, M., Arrol, S., Mackness, M., and McLeod, D. (1995). Hypoxia induces HSP 70 gene expression in human hepatoma (HEP G2) cells. *Biochem. Mol. Biol. Int.* 36, 907–912.
- Patel, N.A., Kaneko, S., Apostolatos, H.S., Bae, S.S., Watson, J.E., Davidowitz, K., Chappell, D.S., Birnbaum, M.J., Cheng, J.Q., and Cooper, D.R. (2005). Molecular and genetic studies imply Akt-

- mediated signaling promotes protein kinase CbetaII alternative splicing via phosphorylation of serine/arginine-rich splicing factor SRp40. *J. Biol. Chem.* *280*, 14302–14309.
- Patriarca, E.J., and Maresca, B. (1990). Acquired thermotolerance following heat shock protein synthesis prevents impairment of mitochondrial ATPase activity at elevated temperatures in *Saccharomyces cerevisiae*. *Exp. Cell Res.* *190*, 57–64.
- Pawlik, A., Nowak, J.M., Grzanka, D., Gackowska, L., Michalkiewicz, J., and Grzanka, A. (2013). Hyperthermia induces cytoskeletal alterations and mitotic catastrophe in p53-deficient H1299 lung cancer cells. *Acta Histochem.* *115*, 8–15.
- Pendas, A.M., Zhou, Z., Cadiñanos, J., Freije, J.M.P., Wang, J., Hultenby, K., Astudillo, A., Wernerson, A., Rodríguez, F., Tryggvason, K., et al. (2002). Defective prelamin A processing and muscular and adipocyte alterations in Zmpste24 metalloproteinase-deficient mice. *Nat. Genet.* *31*, 94–99.
- Perales, R., and Bentley, D. (2009). “Cotranscriptionality”: the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell* *36*, 178–191.
- Pérez, V.I., Bokov, A., Van Remmen, H., Mele, J., Ran, Q., Ikeno, Y., and Richardson, A. (2009). Is the oxidative stress theory of aging dead? *Biochim. Biophys. Acta* *1790*, 1005–1014.
- Peter, M., and Nigg, E.A. (1991). Ectopic expression of an A-type lamin does not interfere with differentiation of lamin A-negative embryonal carcinoma cells. *J. Cell Sci.* *100 (Pt 3)*, 589–598.
- Peter, M., Nakagawa, J., Dorée, M., Labbé, J.C., and Nigg, E.A. (1990). In vitro disassembly of the nuclear lamina and M phase-specific phosphorylation of lamins by cdc2 kinase. *Cell* *61*, 591–602.
- Peteranderl, R., and Nelson, H.C. (1992). Trimerization of the heat shock transcription factor by a triple-stranded alpha-helical coiled-coil. *Biochemistry (Mosc.)* *31*, 12272–12276.
- Peters, A.H.F.M., Kubicek, S., Mechtler, K., O’Sullivan, R.J., Derijck, A.A.H.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., et al. (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell* *12*, 1577–1589.
- Pinol-Roma, S., Choi, Y.D., Matunis, M.J., and Dreyfuss, G. (1988). Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes Dev.* *2*, 215–227.
- Piva, F., Giulietti, M., Nocchi, L., and Principato, G. (2009). SpliceAid: a database of experimental RNA target motifs bound by splicing proteins in humans. *Bioinforma. Oxf. Engl.* *25*, 1211–1213.
- Plasilova, M., Chattopadhyay, C., Pal, P., Schaub, N.A., Buechner, S.A., Mueller, H., Miny, P., Ghosh, A., and Heinemann, K. (2004). Homozygous missense mutation in the lamin A/C gene causes autosomal recessive Hutchinson-Gilford progeria syndrome. *J. Med. Genet.* *41*, 609–614.
- Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C. (2007). Rapid, transcript-specific changes in splicing in response to environmental stress. *Mol. Cell* *27*, 928–937.
- Plesset, J., Palm, C., and McLaughlin, C.S. (1982). Induction of heat shock proteins and thermotolerance by ethanol in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* *108*, 1340–1345.
- Plocik, A.M., and Guthrie, C. (2012). Diverse forms of RPS9 splicing are part of an evolving autoregulatory circuit. *PLoS Genet.* *8*, e1002620.

- Polioudaki, H., Kourmouli, N., Drosou, V., Bakou, A., Theodoropoulos, P.A., Singh, P.B., Giannakouros, T., and Georgatos, S.D. (2001). Histones H3/H4 form a tight complex with the inner nuclear membrane protein LBR and heterochromatin protein 1. *EMBO Rep.* 2, 920–925.
- Pollard, V.W., Michael, W.M., Nakielny, S., Siomi, M.C., Wang, F., and Dreyfuss, G. (1996). A novel receptor-mediated nuclear protein import pathway. *Cell* 86, 985–994.
- Pollex, R.L., and Hegele, R.A. (2004). Hutchinson-Gilford progeria syndrome. *Clin Genet* 66, 375–381.
- Ponthier, J.L., Schluepen, C., Chen, W., Lersch, R.A., Gee, S.L., Hou, V.C., Lo, A.J., Short, S.A., Chasis, J.A., Winkelmann, J.C., et al. (2006). Fox-2 splicing factor binds to a conserved intron motif to promote inclusion of protein 4.1R alternative exon 16. *J. Biol. Chem.* 281, 12468–12474.
- Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R., and Bickmore, W.A. (2012). Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet.* 8, e1002717.
- Prasad, J., Colwill, K., Pawson, T., and Manley, J.L. (1999). The protein kinase Clk/Sty directly modulates SR protein activity: both hyper- and hypophosphorylation inhibit splicing. *Mol. Cell. Biol.* 19, 6991–7000.
- Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. *Cell* 108, 501–512.
- Rabindran, S.K., Haroun, R.I., Clos, J., Wisniewski, J., and Wu, C. (1993). Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. *Science* 259, 230–234.
- Rallu, M., Loones, M., Lallemand, Y., Morimoto, R., Morange, M., and Mezger, V. (1997). Function and regulation of heat shock factor 2 during mouse embryogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 94, 2392–2397.
- Raponi, M., and Baralle, D. (2010). Alternative splicing: good and bad effects of translationally silent substitutions. *FEBS J.* 277, 836–840.
- Rattan, S.I.S. (1998). The Nature of Gerontogenes and Vitagenes: Antiaging Effects of Repeated Heat Shock on Human Fibroblasts. *Ann. N. Y. Acad. Sci.* 854, 54–60.
- Rebbapragada, I., and Lykke-Andersen, J. (2009). Execution of nonsense-mediated mRNA decay: what defines a substrate? *Curr. Opin. Cell Biol.* 21, 394–402.
- Reddy, A.S.N., and Shad Ali, G. (2011). Plant serine/arginine-rich proteins: roles in precursor messenger RNA splicing, plant development, and stress responses. *Wiley Interdiscip. Rev. RNA* 2, 875–889.
- Reddy, T.R., Tang, H., Xu, W., and Wong-Staal, F. (2000). Sam68, RNA helicase A and Tap cooperate in the post-transcriptional regulation of human immunodeficiency virus and type D retroviral mRNA. *Oncogene* 19, 3570–3575.
- Al Refaii, A., and Alix, J.-H. (2009). Ribosome biogenesis is temperature-dependent and delayed in *Escherichia coli* lacking the chaperones DnaK or DnaJ. *Mol. Microbiol.* 71, 748–762.
- Reichert, V.L., Le Hir, H., Jurica, M.S., and Moore, M.J. (2002). 5' exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly. *Genes Dev.* 16, 2778–2791.
- Reijns, M.A.M., Alexander, R.D., Spiller, M.P., and Beggs, J.D. (2008). A role for Q/N-rich aggregation-prone regions in P-body localization. *J. Cell Sci.* 121, 2463–2472.

- Renault, L., Kuhlmann, J., Henkel, A., and Wittinghofer, A. (2001). Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1). *Cell* 105, 245–255.
- Renz, A., and Fackelmayer, F.O. (1996). Purification and molecular cloning of the scaffold attachment factor B (SAF-B), a novel human nuclear protein that specifically binds to S/MAR-DNA. *Nucleic Acids Res.* 24, 843–849.
- Resnick, R.J., Taylor, S.J., Lin, Q., and Shalloway, D. (1997). Phosphorylation of the Src substrate Sam68 by Cdc2 during mitosis. *Oncogene* 15, 1247–1253.
- Richard, V., Kaeffer, N., and Thuillez, C. (1996). Delayed protection of the ischemic heart—from pathophysiology to therapeutic applications. *Fundam. Clin. Pharmacol.* 10, 409–415.
- Richards, S.A., Muter, J., Ritchie, P., Lattanzi, G., and Hutchison, C.J. (2011). The accumulation of un-repairable DNA damage in laminopathy progeria fibroblasts is caused by ROS generation and is prevented by treatment with N-acetyl cysteine. *Hum. Mol. Genet.* 20, 3997–4004.
- Richmond, C.S., Glasner, J.D., Mau, R., Jin, H., and Blattner, F.R. (1999). Genome-wide expression profiling in *Escherichia coli* K-12. *Nucleic Acids Res.* 27, 3821–3835.
- Rivera-Torres, J., Acín-Perez, R., Cabezas-Sánchez, P., Osorio, F.G., Gonzalez-Gómez, C., Megias, D., Cámara, C., López-Otín, C., Enríquez, J.A., Luque-García, J.L., et al. (2013). Identification of mitochondrial dysfunction in Hutchinson-Gilford progeria syndrome through use of stable isotope labeling with amino acids in cell culture. *J. Proteomics*.
- Rizzi, N., Denegri, M., Chiodi, I., Corioni, M., Valgardsdottir, R., Cobianchi, F., Riva, S., and Biamonti, G. (2004). Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. *Mol. Biol. Cell* 15, 543–551.
- Rohlin, L., Trent, J.D., Salmon, K., Kim, U., Gunsalus, R.P., and Liao, J.C. (2005). Heat shock response of *Archaeoglobus fulgidus*. *J. Bacteriol.* 187, 6046–6057.
- Ropers, D., Ayadi, L., Gattoni, R., Jacquenet, S., Damier, L., Branlant, C., and Stévenin, J. (2004). Differential effects of the SR proteins 9G8, SC35, ASF/SF2, and SRp40 on the utilization of the A1 to A5 splicing sites of HIV-1 RNA. *J. Biol. Chem.* 279, 29963–29973.
- Roscigno, R.F., and Garcia-Blanco, M.A. (1995). SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA N. Y. N* 1, 692–706.
- Rosette, C., and Karin, M. (1996). Ultraviolet light and osmotic stress: activation of the JNK cascade through multiple growth factor and cytokine receptors. *Science* 274, 1194–1197.
- Rosonina, E., Ip, J.Y.Y., Calarco, J.A., Bakowski, M.A., Emili, A., McCracken, S., Tucker, P., Ingles, C.J., and Blencowe, B.J. (2005). Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo. *Mol. Cell. Biol.* 25, 6734–6746.
- Rossi, F., Labourier, E., Forné, T., Divita, G., Derancourt, J., Riou, J.F., Antoine, E., Cathala, G., Brunel, C., and Tazi, J. (1996). Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature* 381, 80–82.
- Ryan, K.J., and Cooper, T.A. (1996). Muscle-specific splicing enhancers regulate inclusion of the cardiac troponin T alternative exon in embryonic skeletal muscle. *Mol. Cell. Biol.* 16, 4014–4023.
- Sachsenmaier, C., Radler-Pohl, A., Zinck, R., Nordheim, A., Herrlich, P., and Rahmsdorf, H.J. (1994). Involvement of growth factor receptors in the mammalian UVC response. *Cell* 78, 963–972.

- Saito, Y., Yamagishi, N., and Hatayama, T. (2009a). Nuclear Localization Mechanism of Hsp105 β and its Possible Function in Mammalian Cells. *J. Biochem. (Tokyo)* *145*, 185–191.
- Saito, Y., Yamagishi, N., and Hatayama, T. (2009b). Nuclear localization mechanism of Hsp105 β and its possible function in mammalian cells. *J. Biochem. (Tokyo)* *145*, 185–191.
- Saitoh, N., Sakamoto, C., Hagiwara, M., Agredano-Moreno, L.T., Jiménez-García, L.F., and Nakao, M. (2012). The distribution of phosphorylated SR proteins and alternative splicing are regulated by RANBP2. *Mol. Biol. Cell* *23*, 1115–1128.
- Salgado-Garrido, J., Bragado-Nilsson, E., Kandels-Lewis, S., and Séraphin, B. (1999). Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J.* *18*, 3451–3462.
- Sandi, M.J., Hamidi, T., Malicet, C., Cano, C., Loncle, C., Pierres, A., Dagorn, J.C., and Iovanna, J.L. (2011). p8 expression controls pancreatic cancer cell migration, invasion, adhesion, and tumorigenesis. *J. Cell. Physiol.* *226*, 3442–3451.
- Sandqvist, A., Bjork, J.K., Akerfelt, M., Chitikova, Z., Grichine, A., Vourc'h, C., Jolly, C., Salminen, T.A., Nymalm, Y., and Sistonen, L. (2009). Heterotrimerization of Heat-Shock Factors 1 and 2 Provides a Transcriptional Switch in Response to Distinct Stimuli. *Mol. Biol. Cell* *20*, 1340–1347.
- De Sandre-Giovannoli, A. (2003). Lamin A Truncation in Hutchinson-Gilford Progeria. *Science* *300*, 2055–2055.
- Sanford, J.R., and Bruzik, J.P. (2001). Regulation of SR protein localization during development. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 10184–10189.
- Sanford, J.R., Gray, N.K., Beckmann, K., and Cáceres, J.F. (2004). A novel role for shuttling SR proteins in mRNA translation. *Genes Dev.* *18*, 755–768.
- Sanford, J.R., Ellis, J.D., Cazalla, D., and Cáceres, J.F. (2005). Reversible phosphorylation differentially affects nuclear and cytoplasmic functions of splicing factor 2/alternative splicing factor. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 15042–15047.
- Santoro, M.G. (2000). Heat shock factors and the control of the stress response. *Biochem. Pharmacol.* *59*, 55–63.
- Sapra, A.K., Ankö, M.-L., Grishina, I., Lorenz, M., Pabis, M., Poser, I., Rollins, J., Weiland, E.-M., and Neugebauer, K.M. (2009). SR protein family members display diverse activities in the formation of nascent and mature mRNPs in vivo. *Mol. Cell* *34*, 179–190.
- Sarge, K.D., Murphy, S.P., and Morimoto, R.I. (1993). Activation of heat shock gene transcription by heat shock factor 1 involves oligomerization, acquisition of DNA-binding activity, and nuclear localization and can occur in the absence of stress. *Mol. Cell. Biol.* *13*, 1392–1407.
- Sato, S., Fujita, N., and Tsuruo, T. (2000). Modulation of Akt kinase activity by binding to Hsp90. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 10832–10837.
- Saulière, J., Sureau, A., Expert-Bezançon, A., and Marie, J. (2006). The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the beta-tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit. *Mol. Cell. Biol.* *26*, 8755–8769.
- Savkur, R.S., Philips, A.V., and Cooper, T.A. (2001). Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet* *29*, 40–47.

- Scaffidi, P., and Misteli, T. (2005). Reversal of the cellular phenotype in the premature aging disease Hutchinson-Gilford progeria syndrome. *Nat. Med.* *11*, 440–445.
- Scaffidi, P., and Misteli, T. (2006). Lamin A-dependent nuclear defects in human aging. *Science* *312*, 1059–1063.
- Scaffidi, P., and Misteli, T. (2008). Lamin A-dependent misregulation of adult stem cells associated with accelerated ageing. *Nat. Cell Biol.* *10*, 452–459.
- Scherer, W.F., Syverton, J.T., and Gey, G.O. (1953). Studies on the Propagation in Vitro of Poliomyelitis Viruses Iv. Viral Multiplication in a Stable Strain of Human Malignant Epithelial Cells (strain Hela) Derived from an Epidermoid Carcinoma of the Cervix. *J. Exp. Med.* *97*, 695–710.
- Schmitt, E., Gehrman, M., Brunet, M., Multhoff, G., and Garrido, C. (2007). Intracellular and extracellular functions of heat shock proteins: repercussions in cancer therapy. *J. Leukoc. Biol.* *81*, 15–27.
- Schoenberg, D.R., and Maquat, L.E. (2012). Regulation of cytoplasmic mRNA decay. *Nat. Rev. Genet.* *13*, 246–259.
- Schor, I.E., Llères, D., Risso, G.J., Pawellek, A., Ule, J., Lamond, A.I., and Kornblihtt, A.R. (2012). Perturbation of chromatin structure globally affects localization and recruitment of splicing factors. *PLoS One* *7*, e48084.
- Schullery, D.S., Ostrowski, J., Denisenko, O.N., Stempka, L., Shnyreva, M., Suzuki, H., Gschwendt, M., and Bomsztyk, K. (1999). Regulated interaction of protein kinase Cdelta with the heterogeneous nuclear ribonucleoprotein K protein. *J. Biol. Chem.* *274*, 15101–15109.
- Schwoebel, E.D., Ho, T.H., and Moore, M.S. (2002). The mechanism of inhibition of Ran-dependent nuclear transport by cellular ATP depletion. *J. Cell Biol.* *157*, 963–974.
- Screaton, G.R., Cáceres, J.F., Mayeda, A., Bell, M.V., Plebanski, M., Jackson, D.G., Bell, J.I., and Krainer, A.R. (1995). Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *EMBO J.* *14*, 4336–4349.
- Seewald, M.J., Körner, C., Wittinghofer, A., and Vetter, I.R. (2002). RanGAP mediates GTP hydrolysis without an arginine finger. *Nature* *415*, 662–666.
- Sena, L.A., and Chandel, N.S. (2012). Physiological Roles of Mitochondrial Reactive Oxygen Species. *Mol. Cell* *48*, 158–167.
- Sengupta, S., Parihar, R., and Ganesh, S. (2009). Satellite III non-coding RNAs show distinct and stress-specific patterns of induction. *Biochem. Biophys. Res. Commun.* *382*, 102–107.
- Seno, J.D., and Dynlacht, J.R. (2004). Intracellular redistribution and modification of proteins of the Mre11/Rad50/Nbs1 DNA repair complex following irradiation and heat-shock. *J. Cell. Physiol.* *199*, 157–170.
- Shalev, S.A., De Sandre-Giovannoli, A., Shani, A.A., and Levy, N. (2007). An association of Hutchinson-Gilford progeria and malignancy. *Am. J. Med. Genet. A.* *143*, 1821–1826.
- Shen, H., and Green, M.R. (2004). A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. *Mol. Cell* *16*, 363–373.

- Shen, M., Bellaousov, S., Hiller, M., de La Grange, P., Creamer, T.P., Malina, O., Sperling, R., Mathews, D.H., Stoilov, P., and Stamm, S. (2013). Pyrvinium pamoate changes alternative splicing of the serotonin receptor 2C by influencing its RNA structure. *Nucleic Acids Res.* *41*, 3819–3832.
- Shen, S., Warzecha, C.C., Carstens, R.P., and Xing, Y. (2010). MADS+: discovery of differential splicing events from Affymetrix exon junction array data. *Bioinforma. Oxf. Engl.* *26*, 268–269.
- Shi, Y., and Manley, J.L. (2007). A complex signaling pathway regulates SRp38 phosphorylation and pre-mRNA splicing in response to heat shock. *Mol. Cell* *28*, 79–90.
- Shin, C., Feng, Y., and Manley, J.L. (2004). Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. *Nature* *427*, 553–558.
- Shoeman, R.L., and Traub, P. (1990). The in vitro DNA-binding properties of purified nuclear lamin proteins and vimentin. *J. Biol. Chem.* *265*, 9055–9061.
- Shukla, R.R., Dominski, Z., Zwierzynski, T., and Kole, R. (1990). Inactivation of splicing factors in HeLa cells subjected to heat shock. *J. Biol. Chem.* *265*, 20377–20383.
- Shumaker, D.K., Dechat, T., Kohlmaier, A., Adam, S.A., Bozovsky, M.R., Erdos, M.R., Eriksson, M., Goldman, A.E., Khuon, S., and Collins, F.S. (2006). Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc. Natl. Acad. Sci.* *103*, 8703–8708.
- Siatecka, M., Reyes, J.L., and Konarska, M.M. (1999). Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center. *Genes Dev.* *13*, 1983–1993.
- Sims, R.J., 3rd, Millhouse, S., Chen, C.-F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., and Reinberg, D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol. Cell* *28*, 665–676.
- Singh, R.K., and Cooper, T.A. (2012). Pre-mRNA splicing in disease and therapeutics. *Trends Mol. Med.* *18*, 472–482.
- Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z., and Moore, M.J. (2012). The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* *151*, 750–764.
- Singh, R., Mortazavi, A., Telu, K.H., Nagarajan, P., Lucas, D.M., Thomas-Ahner, J.M., Clinton, S.K., Byrd, J.C., Freitas, M.A., and Parthun, M.R. (2013). Increasing the complexity of chromatin: functionally distinct roles for replication-dependent histone H2A isoforms in cell proliferation and carcinogenesis. *Nucleic Acids Res.* gkt736.
- Sinnamon, J.R., and Czaplinski, K. (2011). mRNA trafficking and local translation: the Yin and Yang of regulating mRNA localization in neurons. *Acta Biochim. Biophys. Sin.* *43*, 663–670.
- Siomi, H., and Dreyfuss, G. (1995). A nuclear localization domain in the hnRNP A1 protein. *J. Cell Biol.* *129*, 551–560.
- Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.* *21*, 1193–1198.
- Sirand-Pugnet, P., Durosay, P., Clouet d'Orval, B.C., Brody, E., and Marie, J. (1995). beta-Tropomyosin pre-mRNA folding around a muscle-specific exon interferes with several steps of spliceosome assembly. *J. Mol. Biol.* *251*, 591–602.

- Sistonen, L., Sarge, K.D., Phillips, B., Abravaya, K., and Morimoto, R.I. (1992). Activation of heat shock factor 2 during hemin-induced differentiation of human erythroleukemia cells. *Mol. Cell. Biol.* *12*, 4104–4111.
- Smith, C.W., and Valcarcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.* *25*, 381–388.
- Smith, P.J., Zhang, C., Wang, J., Chew, S.L., Zhang, M.Q., and Krainer, A.R. (2006). An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum. Mol. Genet.* *15*, 2490–2508.
- Smolle, M., and Workman, J.L. (2013). Transcription-associated histone modifications and cryptic transcription. *Biochim. Biophys. Acta* *1829*, 84–97.
- Solovei, I., Wang, A.S., Thanisch, K., Schmidt, C.S., Krebs, S., Zwerger, M., Cohen, T.V., Devys, D., Foisner, R., Peichl, L., et al. (2013). LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* *152*, 584–598.
- Soret, J., Gattoni, R., Guyon, C., Sureau, A., Popielarz, M., Le Rouzic, E., Dumon, S., Apiou, F., Dutrillaux, B., Voss, H., et al. (1998). Characterization of SRp46, a novel human SR splicing factor encoded by a PR264/SC35 retropseudogene. *Mol. Cell. Biol.* *18*, 4924–4934.
- Spann, T.P., Moir, R.D., Goldman, A.E., Stick, R., and Goldman, R.D. (1997). Disruption of nuclear lamin organization alters the distribution of replication factors and inhibits DNA synthesis. *J. Cell Biol.* *136*, 1201–1212.
- Spann, T.P., Goldman, A.E., Wang, C., Huang, S., and Goldman, R.D. (2002). Alteration of nuclear lamin organization inhibits RNA polymerase II-dependent transcription. *J. Cell Biol.* *156*, 603–608.
- Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* *3*.
- Speit, G., and Schütz, P. (2013). Hyperthermia-induced genotoxic effects in human A549 cells. *Mutat. Res.*
- Spies, N., Nielsen, C.B., Padgett, R.A., and Burge, C.B. (2009). Biased chromatin signatures around polyadenylation sites and exons. *Mol. Cell* *36*, 245–254.
- Spriggs, K.A., Stoneley, M., Bushell, M., and Willis, A.E. (2008). Re-programming of translation following cell stress allows IRES-mediated translation to predominate. *Biol. Cell Auspices Eur. Cell Biol. Organ.* *100*, 27–38.
- Srivastava, S.P., Kumar, K.U., and Kaufman, R.J. (1998). Phosphorylation of eukaryotic translation initiation factor 2 mediates apoptosis in response to activation of the double-stranded RNA-dependent protein kinase. *J. Biol. Chem.* *273*, 2416–2423.
- Stepanova, L., Leng, X., Parker, S.B., and Harper, J.W. (1996). Mammalian p50Cdc37 is a protein kinase-targeting subunit of Hsp90 that binds and stabilizes Cdk4. *Genes Dev.* *10*, 1491–1502.
- Stewart, M. (2007). Molecular mechanism of the nuclear protein import cycle. *Nat. Rev. Mol. Cell Biol.* *8*, 195–208.
- Stickeler, E., Fraser, S.D., Honig, A., Chen, A.L., Berget, S.M., and Cooper, T.A. (2001). The RNA binding protein YB-1 binds A/C-rich exon enhancers and stimulates splicing of the CD44 alternative exon v4. *EMBO J.* *20*, 3821–3830.

- Stoneley, M., and Willis, A.E. (2004). Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. *Oncogene* 23, 3200–3207.
- Stoss, O., Olbrich, M., Hartmann, A.M., Konig, H., Memmott, J., Andreadis, A., and Stamm, S. (2001). The STAR/GSG family protein rSLM-2 regulates the selection of alternative splice sites. *J. Biol. Chem.* 276, 8665–8673.
- Strelkov, S.V., Schumacher, J., Burkhard, P., Aebi, U., and Herrmann, H. (2004). Crystal Structure of the Human Lamin A Coil 2B Dimer: Implications for the Head-to-tail Association of Nuclear Lamins. *J. Mol. Biol.* 343, 1067–1080.
- Stuurman, N., Sasse, B., and Fisher, P.A. (1996). Intermediate filament protein polymerization: molecular analysis of *Drosophila* nuclear lamin head-to-tail binding. *J. Struct. Biol.* 117, 1–15.
- Stuurman, N., Heins, S., and Aebi, U. (1998). Nuclear lamins: their structure, assembly, and interactions. *J. Struct. Biol.* 122, 42–66.
- Su, C., Gao, G., Schneider, S., Helt, C., Weiss, C., O'Reilly, M.A., Bohmann, D., and Zhao, J. (2004). DNA damage induces downregulation of histone gene expression through the G1 checkpoint pathway. *EMBO J.* 23, 1133–1143.
- Su, X., Lucas, D.M., Zhang, L., Xu, H., Zabrouskov, V., Davis, M.E., Knapp, A.R., Young, D.C., Payne, P.R.O., Parthun, M.R., et al. (2009). Validation of an LC-MS based approach for profiling histones in chronic lymphocytic leukemia. *Proteomics* 9, 1197–1206.
- Sullivan, T., Escalante-Alcalde, D., Bhatt, H., Anver, M., Bhat, N., Nagashima, K., Stewart, C.L., and Burke, B. (1999). Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. *J. Cell Biol.* 147, 913–920.
- Suzuki, Y., Minami, M., Suzuki, M., Abe, K., Zenno, S., Tsujimoto, M., Matsumoto, K., and Minami, Y. (2009). The Hsp90 Inhibitor Geldanamycin Abrogates Colocalization of eIF4E and eIF4E-Transporter into Stress Granules and Association of eIF4E with eIF4G. *J. Biol. Chem.* 284, 35597–35604.
- Swanson, C.M., Sherer, N.M., and Malim, M.H. (2010). SRp40 and SRp55 promote the translation of unspliced human immunodeficiency virus type 1 RNA. *J. Virol.* 84, 6748–6759.
- Tabuchi, Y., Takasaki, I., Wada, S., Zhao, Q.-L., Hori, T., Nomura, T., Ohtsuka, K., and Kondo, T. (2008). Genes and genetic networks responsive to mild hyperthermia in human lymphoma U937 cells. *Int. J. Hyperth. Off. J. Eur. Soc. Hyperthermic Oncol. North Am. Hyperth. Group* 24, 613–622.
- Takechi, H., Hosokawa, N., Hirayoshi, K., and Nagata, K. (1994). Alternative 5' splice site selection induced by heat shock. *Mol. Cell. Biol.* 14, 567–575.
- Takekawa, M., Kubota, Y., Nakamura, T., and Ichikawa, K. (2011). Regulation of stress-activated MAP kinase pathways during cell fate decisions. *Nagoya J. Med. Sci.* 73, 1–14.
- Tanabe, M., Kawazoe, Y., Takeda, S., Morimoto, R.I., Nagata, K., and Nakai, A. (1998). Disruption of the HSF3 gene results in the severe reduction of heat shock gene expression and loss of thermotolerance. *EMBO J.* 17, 1750–1758.
- Tang, D., Wu, D., Hirao, A., Lahti, J.M., Liu, L., Mazza, B., Kidd, V.J., Mak, T.W., and Ingram, A.J. (2002). ERK activation mediates cell cycle arrest and apoptosis after DNA damage independently of p53. *J. Biol. Chem.* 277, 12710–12717.
- Tange, T.Ø., Shibuya, T., Jurica, M.S., and Moore, M.J. (2005). Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA N. Y. N* 11, 1869–1883.

- Taniura, H., Glass, C., and Gerace, L. (1995). A chromatin binding site in the tail domain of nuclear lamins that interacts with core histones. *J. Cell Biol.* *131*, 33–44.
- Tarn, W.Y., and Steitz, J.A. (1994). SR proteins can compensate for the loss of U1 snRNP functions in vitro. *Genes Dev.* *8*, 2704–2717.
- Tazi, J., Bakkour, N., and Stamm, S. (2009). Alternative splicing and disease. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* *1792*, 14–26.
- Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J., and Guigó, R. (2009). Nucleosome positioning as a determinant of exon recognition. *Nat. Struct. Mol. Biol.* *16*, 996–1001.
- Toivola, D.M., Strnad, P., Habtezion, A., and Omary, M.B. (2010). Intermediate filaments take the heat as stress proteins. *Trends Cell Biol.* *20*, 79–91.
- Tolstorukov, M.Y., Goldman, J.A., Gilbert, C., Ogryzko, V., Kingston, R.E., and Park, P.J. (2012). Histone variant H2A.Bbd is associated with active transcription and mRNA processing in human cells. *Mol. Cell* *47*, 596–607.
- Tomita, M. (2010). Involvement of DNA-PK and ATM in Radiation- and Heat-induced DNA Damage Recognition and Apoptotic Cell Death. *J. Radiat. Res. (Tokyo)* *51*, 493–501.
- Toth, J.I., Yang, S.H., Qiao, X., Beigneux, A.P., Gelb, M.H., Moulson, C.L., Miner, J.H., Young, S.G., and Fong, L.G. (2005). Blocking protein farnesyltransferase improves nuclear shape in fibroblasts from humans with progeroid syndromes. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 12873–12878.
- Tournier, C., Hess, P., Yang, D.D., Xu, J., Turner, T.K., Nimnual, A., Bar-Sagi, D., Jones, S.N., Flavell, R.A., and Davis, R.J. (2000). Requirement of JNK for stress-induced activation of the cytochrome c-mediated death pathway. *Science* *288*, 870–874.
- Tsutsui, K.M., Sano, K., and Tsutsui, K. (2005). Dynamic view of the nuclear matrix. *Acta Med. Okayama* *59*, 113–120.
- Turner, J.G., Dawson, J., and Sullivan, D.M. (2012). Nuclear export of proteins and drug resistance in cancer. *Biochem. Pharmacol.* *83*, 1021–1032.
- Tycowski, K.T., You, Z.H., Graham, P.J., and Steitz, J.A. (1998). Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol. Cell* *2*, 629–638.
- Uffenbeck, S.R., and Krebs, J.E. (2006). The role of chromatin structure in regulating stress-induced transcription in *Saccharomyces cerevisiae*. *Biochem. Cell Biol. Biochim. Biol. Cell.* *84*, 477–489.
- Ulitzur, N., and Gruenbaum, Y. (1989). Nuclear envelope assembly around sperm chromatin in cell-free preparations from *Drosophila* embryos. *FEBS Lett.* *259*, 113–116.
- Ullrich, N.J., Kieran, M.W., Miller, D.T., Gordon, L.B., Cho, Y.-J., Silvera, V.M., Giobbie-Hurder, A., Neuberger, D., and Kleinman, M.E. (2013). Neurologic features of Hutchinson-Gilford progeria syndrome after lonafarnib treatment. *Neurology* *81*, 427–430.
- Utans, U., Behrens, S.E., Lüthmann, R., Kole, R., and Krämer, A. (1992). A splicing factor that is inactivated during in vivo heat shock is functionally equivalent to the [U4/U6.U5] triple snRNP-specific proteins. *Genes Dev.* *6*, 631–641.
- Vagner, S., Vagner, C., and Mattaj, I.W. (2000). The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. *Genes Dev.* *14*, 403–413.

- Valgardsdottir, R., Chiodi, I., Giordano, M., Cobianchi, F., Riva, S., and Biamonti, G. (2005a). Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. *Mol. Biol. Cell* *16*, 2597–2604.
- Valgardsdottir, R., Chiodi, I., Giordano, M., Cobianchi, F., Riva, S., and Biamonti, G. (2005b). Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. *Mol. Biol. Cell* *16*, 2597–2604.
- Valgardsdottir, R., Chiodi, I., Giordano, M., Rossi, A., Bazzini, S., Ghigna, C., Riva, S., and Biamonti, G. (2007a). Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res.* *36*, 423–434.
- Valgardsdottir, R., Chiodi, I., Giordano, M., Rossi, A., Bazzini, S., Ghigna, C., Riva, S., and Biamonti, G. (2007b). Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res.* *36*, 423–434.
- Vanhaesebroeck, B., and Alessi, D.R. (2000). The PI3K-PDK1 connection: more than just a road to PKB. *Biochem. J.* *346 Pt 3*, 561–576.
- Varela, I., Cadiñanos, J., Pendás, A.M., Gutiérrez-Fernández, A., Folgueras, A.R., Sánchez, L.M., Zhou, Z., Rodríguez, F.J., Stewart, C.L., Vega, J.A., et al. (2005). Accelerated ageing in mice deficient in Zmpste24 protease is linked to p53 signalling activation. *Nature* *437*, 564–568.
- Varela, I., Pereira, S., Ugalde, A.P., Navarro, C.L., Suárez, M.F., Cau, P., Cadiñanos, J., Osorio, F.G., Foray, N., Cobo, J., et al. (2008). Combined treatment with statins and aminobisphosphonates extends longevity in a mouse model of human premature aging. *Nat. Med.* *14*, 767–772.
- Vargas, D.Y., Shah, K., Batish, M., Levandoski, M., Sinha, S., Marras, S.A.E., Schedl, P., and Tyagi, S. (2011). Single-molecule imaging of transcriptionally coupled and uncoupled splicing. *Cell* *147*, 1054–1065.
- Vassileva, M.T., and Matunis, M.J. (2004). SUMO modification of heterogeneous nuclear ribonucleoproteins. *Mol. Cell. Biol.* *24*, 3623–3632.
- Vázquez Nin, G.H., Abolhassani-Dadras, S., Echeverría, O.M., Rouelle-Rossier, V.B., and Fakan, S. (1996). Phosphorus distribution in perichromatin granules and surrounding nucleoplasm as visualized by electron spectroscopic imaging. *Biol. Cell Auspices Eur. Cell Biol. Organ.* *87*, 171–177.
- Vázquez Nin, G.H., Echeverría, O.M., Ortiz, R., Ubaldo, E., and Fakan, S. (1997). Effects of hypophyseal hormones on transcription and RNA export to the cytoplasm. *Exp. Cell Res.* *236*, 519–526.
- Verheggen, C., and Bertrand, E. (2012). CRM1 plays a nuclear role in transporting snoRNPs to nucleoli in higher eukaryotes. *Nucl. Austin Tex* *3*, 132–137.
- Vernet, C., and Artzt, K. (1997). STAR, a gene family involved in signal transduction and activation of RNA. *Trends Genet. TIG* *13*, 479–484.
- Vigh, L., Horváth, I., Maresca, B., and Harwood, J.L. (2007). Can the stress protein response be controlled by “membrane-lipid therapy”? *Trends Biochem. Sci.* *32*, 357–363.
- Villa-Bellosta, R., Rivera-Torres, J., Osorio, F.G., Acín-Pérez, R., Enriquez, J.A., López-Otín, C., and Andrés, V. (2013). Defective extracellular pyrophosphate metabolism promotes vascular calcification in a mouse model of Hutchinson-Gilford progeria syndrome that is ameliorated on pyrophosphate treatment. *Circulation* *127*, 2442–2451.

- Vincent, A.J., Ren, S., Harris, L.G., Devine, D.J., Samant, R.S., Fodstad, O., and Shevde, L.A. (2012). Cytoplasmic translocation of p21 mediates NUPR1-induced chemoresistance: NUPR1 and p21 in chemoresistance. *FEBS Lett.* *586*, 3429–3434.
- Viteri, G., Chung, Y.W., and Stadtman, E.R. (2010). Effect of progerin on the accumulation of oxidized proteins in fibroblasts from Hutchinson Gilford progeria patients. *Mech. Ageing Dev.* *131*, 2–8.
- Vizlin-Hodzic, D., Runnberg, R., Ryme, J., Simonsson, S., and Simonsson, T. (2011). SAF-A forms a complex with BRG1 and both components are required for RNA polymerase II mediated transcription. *PLoS One* *6*, e28049.
- Voelker, R.B., and Berglund, J.A. (2007). A comprehensive computational characterization of conserved mammalian intronic sequences reveals conserved motifs associated with constitutive and alternative splicing. *Genome Res.* *17*, 1023–1033.
- Vorburger, K., Kitten, G.T., and Nigg, E.A. (1989). Modification of nuclear lamin proteins by a mevalonic acid derivative occurs in reticulocyte lysates and requires the cysteine residue of the C-terminal CXXM motif. *EMBO J.* *8*, 4007.
- Voss, A.K., Thomas, T., and Gruss, P. (2000). Mice lacking HSP90beta fail to develop a placental labyrinth. *Dev. Camb. Engl.* *127*, 1–11.
- Vytopil, M., Benedetti, S., Ricci, E., Galluzzi, G., Dello Russo, A., Merlini, L., Boriani, G., Gallina, M., Morandi, L., Politano, L., et al. (2003). Mutation analysis of the lamin A/C gene (LMNA) among patients with different cardiomyopathy phenotypes. *J. Med. Genet.* *40*, e132.
- Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* *136*, 701–718.
- Waldholm, J., Wang, Z., Brodin, D., Tyagi, A., Yu, S., Theopold, U., Farrants, A.K.Ö., and Visa, N. (2011). SWI/SNF regulates the alternative processing of a specific subset of pre-mRNAs in *Drosophila melanogaster*. *BMC Mol. Biol.* *12*, 46.
- Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008a). Alternative isoform regulation in human tissue transcriptomes. *Nature* *456*, 470–476.
- Wang, L.L., Richard, S., and Shaw, A.S. (1995). P62 association with RNA is regulated by tyrosine phosphorylation. *J. Biol. Chem.* *270*, 2010–2013.
- Wang, R.-H., Sengupta, K., Li, C., Kim, H.-S., Cao, L., Xiao, C., Kim, S., Xu, X., Zheng, Y., Chilton, B., et al. (2008b). Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice. *Cancer Cell* *14*, 312–323.
- Wang, X., Grammatikakis, N., Sigano, A., and Calderwood, S.K. (2003). Regulation of molecular chaperone gene transcription involves the serine phosphorylation, 14-3-3 epsilon binding, and cytoplasmic sequestration of heat shock factor 1. *Mol. Cell. Biol.* *23*, 6013–6026.
- Wang, Z., Rolish, M.E., Yeo, G., Tung, V., Mawson, M., and Burge, C.B. (2004). Systematic identification and analysis of exonic splicing silencers. *Cell* *119*, 831–845.
- Watson, M. (1962). Observations on a granule associated with chromatin in the nuclei of cells of rat and mouse. *J. Cell Biol.* *13*, 162–167.
- Weighardt, F., Biamonti, G., and Riva, S. (1995). Nucleo-cytoplasmic distribution of human hnRNP proteins: a search for the targeting domains in hnRNP A1. *J. Cell Sci.* *108 (Pt 2)*, 545–555.

- Weighardt, F., Cobianchi, F., Cartegni, L., Chiodi, I., Villa, A., Riva, S., and Biamonti, G. (1999). A novel hnRNP protein (HAP/SAF-B) enters a subset of hnRNP complexes and relocates in nuclear granules in response to heat shock. *J. Cell Sci.* *112*, 1465–1476.
- Welch, W.J., and Suhan, J.P. (1985). Morphological study of the mammalian stress response: characterization of changes in cytoplasmic organelles, cytoskeleton, and nucleoli, and appearance of intranuclear actin filaments in rat fibroblasts after heat-shock treatment. *J. Cell Biol.* *101*, 1198–1211.
- Welch, W.J., and Suhan, J.P. (1986). Cellular and biochemical events in mammalian cells during and after recovery from physiological stress. *J. Cell Biol.* *103*, 2035–2052.
- Westerheide, S.D., Anckar, J., Stevens, S.M., Jr, Sistonen, L., and Morimoto, R.I. (2009). Stress-inducible regulation of heat shock factor 1 by the deacetylase SIRT1. *Science* *323*, 1063–1066.
- Wilhelm, B.T., Marguerat, S., Aligianni, S., Codlin, S., Watt, S., and Bähler, J. (2011). Differential patterns of intronic and exonic DNA regions with respect to RNA polymerase II occupancy, nucleosome density and H3K36me3 marking in fission yeast. *Genome Biol.* *12*, R82.
- Will, C.L., and Luhrmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.* *13*, 290–301.
- Wong, G., Müller, O., Clark, R., Conroy, L., Moran, M.F., Polakis, P., and McCormick, F. (1992). Molecular cloning and nucleic acid binding properties of the GAP-associated tyrosine phosphoprotein p62. *Cell* *69*, 551–558.
- Wong, R.S., Dynlacht, J.R., Cedervall, B., and Dewey, W.C. (1995). Analysis by pulsed-field gel electrophoresis of DNA double-strand breaks induced by heat and/or X-irradiation in bulk and replicating DNA of CHO cells. *Int. J. Radiat. Biol.* *68*, 141–152.
- Workman, E., Kolb, S.J., and Battle, D.J. (2012). Spliceosomal small nuclear ribonucleoprotein biogenesis defects and motor neuron selectivity in spinal muscular atrophy. *Brain Res.* *1462*, 93–99.
- Worman, H.J., and Bonne, G. (2007). “Laminopathies”: A wide spectrum of human diseases. *Exp. Cell Res.* *313*, 2121–2133.
- Wu, B.J., and Morimoto, R.I. (1985). Transcription of the human hsp70 gene is induced by serum stimulation. *Proc. Natl. Acad. Sci. U. S. A.* *82*, 6070.
- Wu, J.Y., and Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* *75*, 1061–1070.
- Xiao, S.H., and Manley, J.L. (1997). Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev.* *11*, 334–344.
- Xiao, H., Perisic, O., and Lis, J.T. (1991). Cooperative binding of *Drosophila* heat shock factor to arrays of a conserved 5 bp unit. *Cell* *64*, 585–593.
- Xiao, R., Tang, P., Yang, B., Huang, J., Zhou, Y., Shao, C., Li, H., Sun, H., Zhang, Y., and Fu, X.-D. (2012). Nuclear matrix factor hnRNP U/SAF-A exerts a global control of alternative splicing by regulating U2 snRNP maturation. *Mol. Cell* *45*, 656–668.
- Yaglom, J., O’Callaghan-Sunol, C., Gabai, V., and Sherman, M.Y. (2003). Inactivation of dual-specificity phosphatases is involved in the regulation of extracellular signal-regulated kinases by heat shock and hsp72. *Mol. Cell. Biol.* *23*, 3813–3824.

- Yamagishi, N., Fujii, H., Saito, Y., and Hatayama, T. (2009a). Hsp105beta upregulates hsp70 gene expression through signal transducer and activator of transcription-3. *FEBS J.* 276, 5870–5880.
- Yamagishi, N., Fujii, H., Saito, Y., and Hatayama, T. (2009b). Hsp105 β upregulates hsp70 gene expression through signal transducer and activator of transcription-3: Mechanism of Hsp105 β -induced Hsp70 expression. *FEBS J.* 276, 5870–5880.
- Yang, S.H., Meta, M., Qiao, X., Frost, D., Bauch, J., Coffinier, C., Majumdar, S., Bergo, M.O., Young, S.G., and Fong, L.G. (2006). A farnesyltransferase inhibitor improves disease phenotypes in mice with a Hutchinson-Gilford progeria syndrome mutation. *J. Clin. Invest.* 116, 2115–2121.
- Yang, Z., Jakymiw, A., Wood, M.R., Eystathioy, T., Rubin, R.L., Fritzler, M.J., and Chan, E.K.L. (2004). GW182 is critical for the stability of GW bodies expressed during the cell cycle and cell proliferation. *J. Cell Sci.* 117, 5567–5578.
- Yasuda, Y., Miyamoto, Y., Saiwaki, T., and Yoneda, Y. (2006a). Mechanism of the stress-induced collapse of the Ran distribution. *Exp. Cell Res.* 312, 512–520.
- Yasuda, Y., Miyamoto, Y., Saiwaki, T., and Yoneda, Y. (2006b). Mechanism of the stress-induced collapse of the Ran distribution. *Exp. Cell Res.* 312, 512–520.
- Ye, Q., Callebaut, I., Pezhman, A., Courvalin, J.C., and Worman, H.J. (1997). Domain-specific interactions of human HP1-type chromodomain proteins and inner nuclear membrane protein LBR. *J. Biol. Chem.* 272, 14983–14989.
- Yeo, G., and Burge, C.B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol. J. Comput. Mol. Cell Biol.* 11, 377–394.
- Yost, H.J., and Lindquist, S. (1986). RNA splicing is interrupted by heat shock and is rescued by heat shock protein synthesis. *Cell* 45, 185–193.
- Yuan, J., Simos, G., Blobel, G., and Georgatos, S.D. (1991). Binding of lamin A to polynucleosomes. *J. Biol. Chem.* 266, 9211–9215.
- Zahler, A.M., Lane, W.S., Stolk, J.A., and Roth, M.B. (1992). SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* 6, 837–847.
- Zahler, A.M., Neugebauer, K.M., Stolk, J.A., and Roth, M.B. (1993). Human SR proteins and isolation of a cDNA encoding SRp75. *Mol. Cell. Biol.* 13, 4023–4028.
- Zeng, C., and Berget, S.M. (2000). Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. *Mol. Cell. Biol.* 20, 8290–8301.
- Zhang, X.H.-F., and Chasin, L.A. (2006). Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13427–13432.
- Zhang, C., Jenkins, H., Goldberg, M.W., Allen, T.D., and Hutchison, C.J. (1996). Nuclear lamina and nuclear matrix organization in sperm pronuclei assembled in *Xenopus* egg extract. *J. Cell Sci.* 109 (Pt 9), 2275–2286.
- Zhang, Q., Bethmann, C., Worth, N.F., Davies, J.D., Wasner, C., Feuer, A., Ragnauth, C.D., Yi, Q., Mellad, J.A., Warren, D.T., et al. (2007). Nesprin-1 and -2 are involved in the pathogenesis of Emery Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. *Hum. Mol. Genet.* 16, 2816–2833.

- Zhao, K., Harel, A., Stuurman, N., Guedalia, D., and Gruenbaum, Y. (1996). Binding of matrix attachment regions to nuclear lamin is mediated by the rod domain and depends on the lamin polymerization state. *FEBS Lett.* *380*, 161–164.
- Zhong, X.Y., Ding, J.H., Adams, J.A., Ghosh, G., and Fu, X.D. (2009). Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones. *Genes Dev.* *23*, 482–495.
- Zhou, H.-L., Hinman, M.N., Barron, V.A., Geng, C., Zhou, G., Luo, G., Siegel, R.E., and Lou, H. (2011). Hu proteins regulate alternative splicing by inducing localized histone hyperacetylation in an RNA-dependent manner. *Proc. Natl. Acad. Sci. U. S. A.* *108*, E627–635.
- Zhou, Z., Luo, M.J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000). The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* *407*, 401–405.
- Zou, J., Guo, Y., Guettouche, T., Smith, D.F., and Voellmy, R. (1998). Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. *Cell* *94*, 471–480.
- Zraly, C.B., and Dingwall, A.K. (2012). The chromatin remodeling and mRNA splicing functions of the Brahma (SWI/SNF) complex are mediated by the SNR1/SNF5 regulatory subunit. *Nucleic Acids Res.* *40*, 5975–5987.

Le syndrome de Hutchinson-Gilford, également appelé progéria, est une pathologie génétique orpheline qui se caractérise par l'apparition de symptômes assimilés à un vieillissement prématuré. Les enfants atteints de cette maladie ont une apparence similaire à celle d'une personne âgée, et leur espérance de vie excède rarement 13 à 14 ans. Ils décèdent la plupart du temps des complications liées à l'âge comme l'athérosclérose. Les mutations à l'origine de la progéria affectent le gène *LMNA*, codant la lamine A, une protéine d'une superstructure appelée *lamina* qui tapisse l'enveloppe nucléaire. Cette structure joue un rôle majeur dans la formation, la structuration et la résistance du noyau. Les recherches récentes ont pu mettre en évidence plusieurs autres rôles de la *lamina* dans des processus importants comme la réplication et la transcription de l'ADN nucléaire. Les mutations identifiées activent toutes l'utilisation de sites 5' alternatif ou cryptique d'épissage présents au sein de l'exon 11 du pré-ARNm LMNA en amont du site normalement utilisé. Il en résulte une délétion d'un fragment de la lamine A, qui contient un site de clivage essentiel à sa maturation. La délétion de ce site empêche la lamine A d'être séparée de l'enveloppe nucléaire, notamment au cours de la mitose, ce qui a de graves conséquences au niveau de la physiologie de la cellule. La première partie de ma thèse a porté sur l'étude des mécanismes par lesquels les mutations à l'origine de la progéria favorisent l'utilisation des sites d'épissage alternatif et cryptique aux dépens du site canonique de l'exon 11. Ces études ont révélé un effet des mutations sur la structure secondaire, qui permet l'augmentation de l'utilisation des sites d'épissage mutants. De plus, nous avons également montré l'implication de plusieurs protéines SR, SRSF1, SRSF5 et SRSF6, dans la régulation de l'utilisation des différents sites d'épissage.

D'autre part, la formation de granules de stress au sein des noyaux des cellules de patients atteints de progéria avait été observée. Ces granules ou nuclear stress bodies (nSB) ont la particularité de contenir des protéines régulatrices de l'épissage. Elles sont relocalisées après la synthèse à partir de régions péricentromériques des chromosomes d'une classe d'ARN contenant des séquences répétées, les ARN satellite III. Des granules très similaires sont formés dans les cellules saines soumises à divers stress, comme le stress thermique, le stress oxydant, aux UV, aux métaux lourds ou d'autres stress génotoxiques. Une hypothèse de travail est que ces nSB séquestrent ces facteurs d'épissage afin de réguler le profil d'épissage alternatif global des cellules pendant la régénération après un stress. Nous avons initié des travaux visant à terme à comprendre comment les nSB peuvent affecter l'épissage alternatif. Pour cela, nous avons étudié leur formation au travers du modèle du stress thermique. Nous avons entrepris la purification des protéines associées aux ARN satellite III *in vitro* afin de trouver de nouveaux composants des nSB. Nous avons également analysé, par emploi de puces jonction-exon, le transcriptome de cellules soumises à un choc thermique afin de caractériser les effets globaux du stress thermique sur l'épissage.

Mots clés : syndrome de Hutchinson-Gilford, progéria, lamine A, épissage alternatif, granules de stress nucléaires, choc thermique, micropuces "exon-jonction"

The Hutchinson-Gilford syndrome, also called progeria, is a rare genetic disease, characterized by symptoms that can be assimilated to accelerated natural ageing. Children affected by this disease have the appearance of elderly, and their lifespan rarely overcome 13 or 14 years. Their death is mostly caused by age-linked complications like atherosclerosis. Mutations that cause progeria affect the *LMNA* gene, which codes the lamin A, a protein included in a superstructure called *lamina* that covers the inner face of the nuclear envelope. This structure plays a major role in the shaping, maintenance and resistance of the nucleus. Recent research shed light on several other roles of *lamina* in important metabolic processes like DNA transcription and replication. These mutations lead to the activation of alternative or cryptic 5' splice sites located within the exon 11 of LMNA pre-mRNA upstream from the normal 5' splice site. The subsequent deletion of a lamin A fragment containing a cleavage site essential for lamin A processing prevents lamin A from being separated from the nuclear envelope, especially during mitosis, and has dramatic consequences regarding cell physiology. The first part of my thesis concerned the study of the mechanisms by which the mutations leading to progeria favor the use of the alternative and cryptic 5' splice sites at the expense of the exon 11 canonical splice site. Our work revealed an effect of the mutations on the 2D RNA structure of the splice sites, which contributes to the increased use of the mutant sites. On top of it, we showed the impact of several SR proteins, SRSF1, SRSF5 and SRSF6, on the regulation of the use of the exon 11 5' splice sites.

On the other hand, it was shown in the literature that nuclear stress bodies (nSB) are formed in the nucleus of cells from progeria patients. These bodies contain several splicing regulatory proteins. Those proteins are relocalized in nSBs after the synthesis, from pericentromeric chromosomal regions, of a class of RNA containing repeated sequences, the satellite III RNAs. Similar granules are formed in healthy cells submitted to various stresses like heat shock, oxidative stress, UV, heavy metals or other genotoxic stresses. A work hypothesis is that those nSBs sequester splicing factors in order to regulate the global alternative splicing profile in cells during the recovery period after stress. We initiated a project aiming to a better understanding of how the nSBs can affect alternative splicing. For this purpose, we studied their formation thanks to the heat shock model. We undertook the purification of proteins associated with sat III RNA *in vitro*, to find new components of the nSB. We also analyzed using exon junction microarrays, the transcriptome of cells subjected to heat shock in order to characterize the global effect of heat stress on alternative splicing.

Keywords : Hutchinson-Gilford Progeria Syndrome, lamin A, alternative splicing, nuclear stress bodies, heat shock, exon junction microarray