



HAL
open science

Décryptage des mécanismes de régulation de l'épissage de l'exon 5 du pré-ARNm de la troponine T cardiaque : étude du rôle de l'épissage alternatif des pré-ARNm dans la réponse des cellules de vertébrés au stress oxydant

Jean-Vincent Philippe

► **To cite this version:**

Jean-Vincent Philippe. Décryptage des mécanismes de régulation de l'épissage de l'exon 5 du pré-ARNm de la troponine T cardiaque : étude du rôle de l'épissage alternatif des pré-ARNm dans la réponse des cellules de vertébrés au stress oxydant. Biochimie, Biologie Moléculaire. Université de Lorraine, 2015. Français. NNT : 2015LORR0298 . tel-01754594

HAL Id: tel-01754594

<https://hal.univ-lorraine.fr/tel-01754594>

Submitted on 30 Mar 2018

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 Jean-Vincent PHILIPPE

**Décryptage des mécanismes de régulation de l'épissage de l'exon
5 du pré-ARNm de la troponine T cardiaque.**

**Etude du rôle de l'épissage alternatif des pré-ARNm dans la
réponse des cellules de vertébrés au stress oxydant**

Le 16 décembre 2015

Membres du jury :

Rapporteurs :	Mme D.WEIL	Directeur de recherche, CNRS, Paris
	M M.DUTERTRE	Chargé de recherche, INSERM, Paris
Examineurs :	Mme M-L.CAILLET-BOUDIN	Chargée de recherche, INSERM, Lille
	M A.VISVIKIS	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
Membre invité :	M L.VALLAR	Docteur, CRP-santé, Luxembourg

**UMR 7365 CNRS-Université de Lorraine
Laboratoire d'ingénierie Moléculaire et Physiopathologie Articulaire (IMoPA)
Bâtiment Biopôle, Faculté de médecine de Nancy, Vandœuvre-lès-Nancy, France.**

REMERCIEMENTS

Je tiens à exprimer ma plus profonde gratitude à Madame Isabelle Behm-Ansmant pour sa direction, son implication et pour m'avoir toujours soutenu au cours de ce travail. Elle est de ces personnes qui possèdent de belles et rares qualités : enthousiasme, dynamisme, volonté d'aller toujours plus loin avec des compétences immenses dont elle m'a fait bénéficier généreusement ; à son contact je n'ai cessé d'apprendre. Elle a toujours su accepter mes erreurs (et quelques chocolats), a pris le temps de les corriger et de m'aider ainsi à en tirer profit et à progresser. Travailler sous sa direction fut un honneur, pour tout ce qu'elle m'a donné, humainement et scientifiquement, je lui rends hommage et la remercie du fond du cœur.

Je tiens à remercier sincèrement Madame Christiane Branlant pour m'avoir accepté au sein de son laboratoire, m'offrant ainsi les meilleures conditions pour réaliser ce travail. Directrice de recherche, elle m'a permis de profiter de son expérience, je la remercie particulièrement de sa disponibilité, de sa rigueur scientifique et de son soutien qui m'ont encouragé à faire plus, à faire mieux.

De même je suis particulièrement reconnaissant envers Monsieur Martin Dutertre pour avoir accepté de faire partie de mon comité de suivi de thèse et pour ses conseils avisés. Je le remercie ainsi que Madame Dominique Weil, pour avoir accepté d'être rapporteurs de mon manuscrit de thèse et d'évaluer mon travail. Je tiens à remercier Madame Marie-Laure Caillet-Boudin, Monsieur Laurent Vallar et Monsieur Athanase Visvikis pour avoir accepté de juger mon travail et d'être membres de mon jury de thèse.

J'exprime toute ma reconnaissance à Athanase pour qui j'ai un profond respect ; il a toujours pris le temps, tout au long de mon parcours universitaire, de répondre à mes questions et m'a donné de précieux conseils lors de discussions scientifiques enrichissantes.

Je remercie Laurent Vallar pour m'avoir accueilli au CRP-Santé du Luxembourg et pour le travail en collaboration que nous avons réalisé. A lui et son équipe : Nathalie Nicot, Tony Kaoma et Arnaud Muller, j'adresse mes remerciements pour leur accueil chaleureux, pour m'avoir transmis leurs connaissances ; leur disponibilité et leurs compétences ont permis la réalisation de mon travail.

Je remercie tout particulièrement Christelle Aigueperse qui fut la première à m'accueillir lors de mon premier jour au laboratoire et qui, tout au long de ma thèse, m'a supporté avec patience et m'a apporté une aide précieuse et efficace. Je remercie Florence Oillo-Blanloeil pour toute son aide dans le projet DM1.

Je remercie également tous les membres anciens et actuels du box2/Labo4 pour leur gentillesse et leur bienveillance, notamment Valentin, le précieux ami sans lequel mes années auraient été bien monotones. Il a toujours été là pour me soutenir, m'écouter et partager moments difficiles et fous rires. Un grand merci à Audrey Vautrin dont j'ai poursuivi les travaux sur la DM1 et qui m'a aidé et soutenu lors de mon arrivée au laboratoire. Je remercie également Mouna, Camille et Géraud qui ont participé à mon travail.

Je remercie tous les membres du Laboratoire IMoPA (tous étages confondus), notamment, Guillaume, Antoine et Yoann.

Je finirai ces remerciements par ma famille qui m'a toujours soutenu et qui a toujours été là pour moi (je check M-A). Je remercie tout particulièrement ma maman pour son soutien inconditionnel et sa confiance ainsi que mon père sans lequel je ne serais pas le même, tu n'es plus là, je ne t'oublie pas.

**ABREVIATIONS
ET
ACRONYMES**

ABREVIATIONS ET ACRONYMES

°C : degré Celsius
A : adénine
aa : acide(s) amine(s)
A/mA : ampère/milliampère
ADN : Acide DésoxyriboNucléique
ADNc : ADN complémentaire
ARN : Acide RiboNucléique
ARNm : ARN messenger
ARNr : ARN ribosomique
ARNt : ARN de transfert
ATP : adenosine 5' triphosphate
BSA : "Bovine Serum Albumin" (Albumine de sérum bovin)
C : cytosine
CBC : Cap Binding Complex
CBP : Creb-Binding Protein
CELF : CUG-BP and ETR-3 Like Factors
CpG : Cytosine-phosphate-Guanosine
cpm : coups par minute
cps : coups par seconde
cSG : "cytoplasmic Stress Granules" (granules de stress cytoplasmiques)
CTD : Carboxy Terminal Domain
C-ter : C-terminal(e)
CUG-BP : CUG Binding Protein
Da : Dalton
DAVID : "Database for Annotation, Visualization and Integrated Discovery"
DM1 : Dystrophie Myotonique de type 1
DMEM : Dulbecco's Modified Eagle's Medium
DMPK : Dystrophie Myotonique protéine kinase
DNase : désoxyribonucléase
dNTP : désoxyribonucléotide
DSB : "Double Strand Break" (cassures double brin)
DTT : dithiothréitol
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
G : guanine
GO : ("Gene Ontology")

GPX : Glutathion Peroxidase(s)
GSH : Glutathion
GST : Glutathion-S-Transferase
HAT : Histone AcétylTransférase
hcTNT : human cardiac Troponin T
HDAC : Histone Déacétylase
HEK : Human Embryonic Kidney
HJAY : "Human Junction Array"
hnRNP : "heteronuclear RiboNucleic Particle"
IF : "Intermediate Filament"
IR : Insulin Receptor
IRES : "Internal Ribosome Entry Site"
ISE : "Intronic Splicing Enhancer(s)"
ISS : "Intronic Splicing Silencer(s)"
Kd : constante de dissociation
kb : kilobase
kDa : kilo Dalton
KEGG : "Kyoto Encyclopedia of Genes and Genomes"
LB : Luria-Bertani
LNA : "Locked Nucleic Acid" ("acide nucléique verrouillé")
M : A ou C
MBNL : Muscleblind-Like
MBP : "Maltose Binding Protein"
MCS : "Multi Cloning Site" (site de clonage)
MEC : Matrice Extracellulaire
miR, miRNA : microRNA
MMS : méthylméthanesulfonate
N : nucléotide A, C, T ou U
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)
PAGE : PolyAcrylamide Gel Electropheris
PCR : "Polymerase Chain Reaction"
PDGF : "Platelet Derived Growth Factor"
PG : "Perichromatin Granules"
pol : polymérase
pré-ARNm : Acide Ribo Nucléique pré-messenger
PRX : Peroxiredoxine(s)
PTB : Polyprymidine Tract Binding
R : purine

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
S : G ou C
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"
SRPK : SR protein kinase
SRSF : "Serin-Arginin Rich" Splicing Factor
T : Thymin
Tm : "melting Température" (température de fusion)
TNF- α : "Tumor Necrosis Factor- α "
TNNT3 : troponine squelettique
TPR : "Tetratricopeptide Repeats" (domaine protéique)
TRX : Thioredoxine(s)
U : uracile
UPR : "Unfolded Proteins Response"
UsnRNA : "Uridine-rich Small Nuclear Ribonucleic Acids"
UsnRNP : "Uridine-rich Small Nuclear Ribonucleic Particle"
UV : Ultra-Violet
v/v : volume à volume
WB : "Western Blot" ou "Western Blotting"
WT : "Wild Type" (type sauvage)

Liste des figures et des tableaux

N° de Figure	Titre	Page
Chapitre I : L'ÉPISSAGE DES PRE-ARNm CHEZ L'HOMME.		
Figure I.1	Représentation schématique de l'organisation d'un intron situé entre deux exons	9
Figure I.2	Représentation schématique de la réaction d'épissage.	10
Figure I.3	Biogenèse des U snRNA transcrits par l'ARN polymérase II : rôle du complexe SMN	13
Figure I.4	Assemblage séquentiel des spliceosomes majeur et mineur.	15
Figure I.5	Les différents modes d'épissage alternatif des pré-ARNm.	19
Figure I.6	Schéma d'un domaine RRM.	24
Figure I.7	Exemples de protéines à domaines RS autres que les protéines SR.	26
Figure I.8	Rôle des protéines SR dans l'épissage des ARNm	28
Figure I.9	Modes de régulation de l'épissage par les protéines hnRNP.	36
Figure I.10	Antagonisme entre protéines hnRNP et SR dans la régulation de l'épissage.	38
Figure I.11	Modèles de couplage entre la transcription et l'épissage des pré-ARNm.	42
Figure I.12	Représentation schématique des mutations ponctuelles à l'origine de maladies génétiques affectant l'épissage en cis	50
Chapitre II : LA DYSTROPHIE MYOTONIQUE DE TYPE I.		
Figure II.1	Structure des régions ARN présentant des répétitions CUG.	57
Figure II.2	Les différentes isoformes protéiques produites à partir du gène MBNL1.	62
Figure II.3	Structure cristallographique de la paire à doigts de Zn ²⁺ 3 et 4 (TZF34) de MBNL1 en complexe avec un oligonucléotide r(CGCUGU).	62
Figure II.4	Schéma des différents éléments NLS présents au sein de l'ARNm codant la protéine MBNL1.	63
Figure II.5	Schématisation des différents modèles proposés pour la formation d'ARN circulaires.	66
Figure II.6	Schéma de l'organisation structurale des protéines CELF.	70
Figure II.7	Schéma de la localisation de deux sites de fixation de la protéine MBNL1 sur l'ARNm hcTNT.	76
Figure II.8	Schéma du minigène IR-hcTNT.	78
Figure II.9	Détermination de la taille des régions introniques en amont et en aval de l'exon 5 du pré-ARNm hcTNT requises pour récapituler l'ensemble des régulations de l'inclusion de l'exon 5.	78
Figure II.10	Schéma représentant les différents sites identifiés pour la fixation de la protéine MBNL1.	79
Figure II.11	Modèle de structure 2D de la région du pré-ARNm hcTNT requise pour récapituler l'ensemble des régulations et dérégulations de l'inclusion de l'exon 5, établi à partir des données d'hydrolyse enzymatiques.	80
Figure II.12	Analyse par RT-PCR de l'effet des mutations des différents groupes de sites sur l'inclusion de l'exon 5 dans les ARN produits à partir du minigène IR-hcTNT 150-E5-150.	82
Figure II.13	Rôle des différentes régions du minigène IR-hcTNT dans la régulation de l'inclusion de l'exon 5 du pré-ARNm hcTNT.	83

N° de Figure	Titre	Page
Chapitre III : DECRYPTAGE DES MECANISMES DE REGULATION DE L'INCLUSION DANS L'ARNM DE L'EXON 5 DU PRE-ARNM DE LA TROPONINE T CARDIAQUE HUMAINE. DECOUVERTE D'UN ROLE DE LA PROTEINE HNRNP H.		
Figure III.1	Présentation des différents mutants du minigène hcTNT 150-E5-150 utilisé lors de cette étude.	91
Figure III.2	Analyse par RT-PCR de l'effet des mutations des sites 1, 2 et 3 du groupe 2 sur l'intégration de l'exon 5 dans l'ARNm produit à partir du minigène IR-hcTNT 150-E5-150.	95
Figure III.3	Analyse par RT-PCR de l'effet des mutations des sites 1, 2 et 3 du groupe 2 sur l'intégration de l'exon 5 dans l'ARNm produit à partir du minigène IR-hcTNT 150-E5-150 ΔG.	96
Figure III.4	Analyse par RT-PCR de l'effet des mutations des sites 4, 5 et 6 du groupe 3 sur l'intégration de l'exon 5 dans l'ARNm produit à partir du minigène IR-hcTNT 150-E5-150 ΔG.	97
Figure III.5	Description des différents mutants produits dans le minigène RTB300.	99
Figure III.6	Analyse par RT-PCR de l'effet des mutations des sites MBNL1 des groupes 1, 2 et 3 sur l'intégration de l'exon 5 dans les ARNm produits à partir du minigène RTB300.	100
Figure III.7	Description des différents mutants produits dans les minigènes RTB300 150-E5-150 et RTB 150-E5-372 utilisés lors de cette étude.	102
Figure III.8	Analyse par RT-PCR de l'effet des mutations des sites MBNL1 des sites 1, 2 et 3 sur l'intégration de l'exon 5 dans les ARNm produits à partir des minigènes RTB300 150-150 et RTB300 150-372.	102
Figure III.9	Schéma du principe de la purification des complexes ribonucléoprotéiques par chromatographie d'affinité.	105
Figure III.10	Schéma des constructions réalisées pour les expériences de purification par chromatographie d'affinité.	107
Figure III.11	Purifications des RNP et des protéines par chromatographie d'affinité sur billes d'amylose.	109
Figure III.12	Analyse des fractions d'éluion obtenues avec chacun des ARN de fusion par Western blot.	110
Figure III.13	Etude de l'interaction de la protéine hnRNP H-Nter avec l'ARN 150-E5-150 et ses sous fragments.	111
Figure III.14	Analyse des protections générées par la fixation de la protéine hnRNP H-Nter sur l'ARN hcTNT 150-E-150.	113
Figure III.15	Etude de l'effet de la sur-expression de la protéine hnRNP H sur l'intégration de l'exon 5 hcTNT dans le transcrit du minigène IR-hcTNT.	116
Figure III.16	Etude de l'effet de la sur-expression de la protéine hnRNP H sur l'inclusion de l'exon 5 hcTNT dans le transcrit du minigène RTB300.	117
Figure III.17	Etude de l'effet de la déplétion de la protéine hnRNP H sur l'inclusion de l'exon 5 hcTNT dans le transcrit du minigène IR-hcTNT.	118
Figure III.18	Etude de l'importance fonctionnelle des séquences introniques -150 à -100 et +50 à +150 dans la régulation de l'inclusion de l'exon 5 du pré-ARNm hcTNT par la protéine hnRNP H.	120
Figure III.19	Etude de l'importance fonctionnelle de la séquence exonique dans la régulation de l'exon 5 du pré-ARNm hcTNT par la protéine hnRNP H.	122
Figure III.20	Effet de la variation de la quantité de hnRNP H-Nter sur le taux d'inclusion de l'exon 5 du pré-ARNm hcTNT <i>in vitro</i> .	124
Figure III.21	Effet de la déplétion de la protéine hnRNP H sur la l'expression de la protéine MBNL1.	126
Figure III.22	Etude de l'effet de la compétition in cellulo entre les protéines hnRNP H et MBNL1 sur l'inclusion de l'exon 5 hcTNT dans le transcrit du minigène IR-hcTNT.	127
Figure III.23	Effet de la modulation de la quantité de protéine hnRNP H intracellulaire sur la régulation de l'épissage de pré-ARNm dont l'épissage est régulé par MBNL1 et dérégulé dans le cadre de la DM1.	129
Figure III.24	Nouveau modèle proposé pour la régulation de l'inclusion de l'exon 5 du pré-ARNm hcTNT.	133

N° de Figure	Titre	Page
Chapitre IV : ETUDE DU ROLE DE L'ÉPISSAGE ALTERNATIF DES PRE-ARNM DANS LA REPONSE DES CELLULES DE VERTEBRES AU STRESS OXYDANT.		
Figure IV.1	Schéma de la chaîne respiratoire	142
Figure IV.2	Réactions de Fenton et de Haber-Weiss	143
Figure IV.3	Schéma du processus de peroxydation des lipides.	148
Figure IV.4	Représentation de l'impact d'un stress oxydant selon le niveau de ROS sur la peroxydation lipidique et ses conséquences cellulaires.	150
Figure IV.5	Représentation des différentes bases de l'ADN pouvant être produites lors d'un stress oxydant	151
Figure IV.6	Principales modifications des acides aminés cystéines, méthionine et tyrosine par des espèces réactives de l'oxygène.	153
Figure IV.7	Schéma de l'organisation des domaines fonctionnels de la protéine Nrf2	155
Figure IV.8	Schéma du mécanisme d'activation de la voie Nrf2 suite à un stress oxydant et de l'activation de l'expression des gènes ARE-dépendant.	156
Figure IV.9	Schématisation de la formation des PML	161
Figure IV.10	Organisation des séquences d'ARN Sat III	163
Figure IV.11	Effets de différents stress sur la production d'ARN sat III et la cinétique des nSB	164
Figure IV.12	Modèle proposé actuellement pour la régulation globale de l'épissage alternatif dans le noyau par formation des granules de stress nucléaires.	166
Chapitre V : RESULTATS DE L'ETUDE DES CONSEQUENCES D'UN STRESS OXYDANT SUR L'ÉPISSAGE ALTERNATIF.		
Figure V.1	Analyse de la génération de ROS intracellulaires suite à un traitement de 24h.	173
Figure V.2	Analyse par RT-PCR de l'effet sur l'épissage alternatif du pré-ARNm POLB du peroxyde d'hydrogène, du paraquat et du DNP dans des cellules HeLa.	175
Figure V.3	Analyse cinétique de la cytotoxicité et de l'induction de ROS par l'H ₂ O ₂ dans des cellules HeLa	177
Figure V.4	Détermination par RT-qPCR du taux d'induction des ARN Sat III au cours de la réponse à un stress oxydant	179
Figure V.5	Suivi par RNA-FISH de la localisation des ARN Sat III au cours de la réponse de cellules HeLa S3 à un stress oxydant généré par l'ajout de 500 µM d'H ₂ O ₂ .	180
Figure V.6	Visualisation de la co-localisation entre les ARN sat III et la protéine hnRNP HAP dans des cellules HeLa S3 soumises depuis 24 h à un stress oxydant par ajout de 500 µM d'H ₂ O ₂ .	181
Figure V.7	Suivi par immunofluorescence de la localisation des protéines HSF1 et hnRNP HAP au cours de la réponse de cellules HeLa S3 à un stress oxydant généré par l'ajout de 500 µM d'H ₂ O ₂ .	184
Figure V.8	Effet de la déplétion du facteur de transcription HSF1 sur l'induction de l'expression des ARN Sat III en réponse à un stress oxydant	185
Figure V.9	Principe des puces à exons utilisées.	188
Figure V.10	Représentation schématique des différentes isoformes d'ARNm du gène Tra2b et de la protéine tra2β de taille entière	210
Figure V.11	Analyse du taux d'augmentation de l'exclusion de l'exon 3 au sein des ARNm tra2β au cours de la réponse à un stress oxydant.	212
Figure V.12	Représentation schématique des séquences exoniques et introniques alternatives de la séquence 3'-UTR du gène SERCA2 et des différentes isoformes connues de l'ARNm SERCA2	214
Figure V.13	Analyse du taux d'inclusion de l'exon 22 au sein de l'ARNm SERCA2 au cours de la réponse à un stress oxydant.	214
Figure V.14	Analyse du taux d'inclusion des exons 7 et 10 au sein de l'ARNm ECM2 24 h après l'induction d'un stress oxydant.	215
Figure V.15	Analyse du taux d'inclusion de l'exon 1 au sein de l'ARNm FBXW10 24 h après l'induction d'un stress oxydant.	217
Figure V.16	Représentation schématique des isoformes NR4A3α, NR4A3β et NR4A3α+exon3 produites à partir du gène NR4A3.	218

N° de Figure	Titre	Page
Chapitre V (suite) : RESULTATS DE L'ETUDE DES CONSEQUENCES D'UN STRESS OXYDANT SUR L'EPISSAGE ALTERNATIF.		
Figure V.17	Analyse du taux d'augmentation de l'exclusion de l'exon 3 au sein de l'ARNm NR4A3 au cours de la réponse à un stress oxydant.	218
Figure V.18	Analyse de la quantité des UsnRNA des spliceosomes majeur et mineur au cours de la réponse au stress oxydant.	222
Figure V.19	Analyse de la rétention d'introns U12 –dépendants au sein d'ARNm dont l'expression diminue lors de la phase précoce de la réponse au stress oxydant.	226
Figure V.20	Vérification de la qualité du fractionnement des compartiments cytoplasmiques et nucléaires.	228
Figure V.21	Analyse préliminaire de l'effet du stress oxydant sur l'expression et la localisation des protéines SART1, U1A et SNRNP40 au temps 1h, 2h et 4h.	230

N° de Tableau	Titre	Page
Chapitre I : L'EPISSAGE DES PRE-ARNm CHEZ L'HOMME.		
Tableau I.1	Liste des protéines SR exprimées chez l'homme.	23
Tableau I.2	Séquences reconnues par les protéines SR.	25
Tableau I.3	Structure générale et domaines des protéines hnRNP.	32
Tableau I.4	Séquences reconnues par les différentes protéines hnRNP.	33
Tableau I.5	Exemples de maladies associées à la présence de mutations affectant en cis l'épissage des pré-ARNm les portants.	49
Tableau I.6	Exemples de pathologies associées à la présence de mutations affectant en trans l'épissage des pré-ARNm.	51
Chapitre II : LA DYSTROPHIE MYOTONIQUE DE TYPE I.		
Tableau II.1	Tableau récapitulatif des caractéristiques des différents phénotypes de la DM1.	55
Tableau II.2	Principaux symptômes des formes adultes de la dystrophie myotonique de type 1.	55
Tableau II.3	Principales anomalies d'épissage détectées dans différents tissus de patients DM1.	59
Chapitre V : RESULTATS DE L'ETUDE DES CONSEQUENCES D'UN STRESS OXYDANT SUR L'EPISSAGE ALTERNATIF.		
Tableau V.1	Nombre de gènes dont l'expression varie de manière significative (p -value $\leq 10^{-3}$) 1 h, 2 h, 4 h et 24 h après le début du stress oxydant.	190
Tableau V.2	Groupes fonctionnels de gènes pour lesquels des variations significatives d'expression sont détectées avec les meilleurs scores statistiques par le logiciel DAVID au temps 4 h.	195
Tableau V.3	Groupes fonctionnels de gènes pour lesquels des variations significatives d'expression sont détectées avec les meilleurs scores statistiques par le logiciel DAVID au temps 24 h.	197
Tableau V.4	Voies métaboliques identifiées par le logiciel Ingenuity pour lesquelles l'expression d'un nombre significatif de gènes est modulée aux différents temps	199
Tableau V.5	Principaux résultats des analyses et prédictions en amont des données d'expression des puces par le logiciel Ingenuity au temps 4 h.	203
Tableau V.6	Principaux résultats des analyses et prédictions en amont des données d'expression des puces par le logiciel Ingenuity au temps 24 h.	204
Tableau V.7	Nombre d'évènements d'épissage modulé de manière significative (p -value $\leq 10^{-3}$) 1 h, 2 h, 4 h et 24 h après le début du stress oxydant.	205
Tableau V.8	Groupes fonctionnels établis par le logiciel DAVID à partir des listes de gènes pour lesquels des variations d'épissage significatives ont été détectées 2 h, 4 h ou 24 h après induction du stress oxydant.	207
Tableau V.9	Gènes dont l'expression diminue sur les puces, qui contiennent un ou plusieurs introns de type U12 et qui respectent les critères minima pour une analyse par RT-PCR.	224

N° de Tableau	Titre	Page
Chapitre VI : Matériel et Méthodes		
Tableau VI.1	Caractéristiques et conditions d'utilisation des anticorps.	301
Tableau VI.2	Désignations, séquences et usages des oligodésoxynucléotides employés lors de ce travail.	302
Tableau VI.3	Caractéristiques des siRNA utilisés.	306
Tableau VI.4	Mélange réactionnel pour la réalisation d'une mutagenèse dirigée.	311
Tableau VI.5	Les conditions expérimentales des cycles de la PCR.	311
Tableau VI.6	Réaction de PCR pour la synthèse de matrice de transcription.	312
Tableau VI.7	Réaction de la transcription in vitro.	313
Tableau VI.8	Protocole pour la déphosphorylation en 5' des ARN.	316
Tableau VI.9	Protocole pour le marquage en 5' des ARN déphosphorylés.	316
Tableau VI.10	Composition du mélange réactionnel pour la formation de complexes ARN-protéine avant ajout de protéine purifiée.	317
Tableau VI.11	Mélange réactionnel pour la transcription d'ARN pré-messagers.	318
Tableau VI.12	Mélange réactionnel pour la réalisation d'un épissage in vitro d'ARN pré-messagers.	319
Tableau VI.13	Mélange réactionnel de PCR pour la détection de produits spécifiques à partir d'ARN totaux rétrotranscrits.	321
Tableau VI.14	Conditions expérimentales des cycles de PCR pour la détection de produits spécifiques à partir d'ARN totaux rétrotranscrits.	321
Tableau VI.15	Mélange réactionnel pour la réalisation d'une réaction de qPCR.	322
Tableau VI.16	Conditions expérimentales des cycles pour la réaction de qPCR pour la détection de produits spécifique à partie d'ARN totaux rétrotranscrits.	322

TABLE DES MATIERES

Table des matières

AVANT-PROPOS	1
CHAPITRE I : L'ÉPISSAGE DES pré-ARNm CHEZ L'HOMME.....	8
I.1. LE PROCESSUS D'ÉPISSAGE.....	8
I.1.1. Les séquences permettant la définition des introns.....	8
I.1.2. La réaction d'épissage.....	9
I.2. FORMATION DES COMPLEXES SPLICEOSOMAX D'ÉPISSAGE.....	10
I.2.1. Les particules UsnRNP : biogénèse et maturation.....	10
I.2.2. Assemblage des complexes spliceosomaux : un processus dynamique.....	13
I.3. L'ÉPISSAGE ALTERNATIF.....	17
I.3.1. Les différents types d'évènements d'épissage alternatif.....	18
I.3.2. Les séquences cis-régulatrices de l'épissage.....	20
I.4. LES COMPLEXES SPLICEOSOMAX : UNE MACHINERIE COMPLEXE, DYNAMIQUE ET METTANT EN JEU UN TRES GRAND NOMBRE DE PROTEINES.....	21
I.4.1. Les protéines SR.....	22
I.4.2. Les protéines hnRNP.....	30
I.4.2.4. Antagonisme entre protéines hnRNP et protéines SR dans l'épissage.....	37
I.4.3. Les autres protéines régulatrices de l'épissage.....	39
I.5. L'ÉPISSAGE ALTERNATIF : UN PROCESSUS COUPLE AVEC DE NOMBREUX MECANISMES CELLULAIRES.....	41
I.5.1. Couplage entre l'épissage et la transcription des pré-ARNm.....	41
I.5.2. Liens entre épissage et état de la chromatine au travers des modifications épigénétiques.....	44
I.5.3. Couplage de l'épissage avec le transport et la stabilité des ARNm.....	46
I.6. ÉPISSAGE ALTERNATIF ET PATHOLOGIES.....	47
I.6.1. Pathologies associées à des mutations en cis des pré-ARNm.....	48
I.6.2. Pathologies associées à des mutations affectant <i>en trans</i> l'épissage des pré-ARNm.....	50
CHAPITRE II : LA DYSTROPHIE MYOTONIQUE DE TYPE I.....	53
II.1. LA DYSTROPHIE MYOTONIQUE DE TYPE I.....	53
II.1.1. Les aspects génétiques à l'origine de la maladie.....	53
II.1.2. Les aspects cliniques associés à la maladie.....	54
II.2. CONSÉQUENCES DE LA PRÉSENCE DES RÉPÉTITIONS CTG DANS LA DM1.....	56
II.2.1. Impact des répétitions CTG sur le gène <i>DMPK</i> et sur l'expression des gènes voisins.....	56
II.2.2. Impact des répétitions CUG sur l'ARNm <i>DMPK</i> et sur la DM1.....	56
II.3. LA PROTÉINE MBNL1 ET SON IMPLICATION DANS LA DM1.....	58
II.3.1. Structure de la protéine MBNL1 et reconnaissance des séquences ARN.....	60

II.3.2. Localisation et fonction de la protéine MBNL1.....	63
II.3.3. Implication de la protéine MBNL1 dans la dystrophie myotonique de type I.....	67
II.4. LES PROTEINES DE LA FAMILLE CELF.....	69
II.4.1. Structure des protéines CELF et interaction avec l'ARN.....	70
II.4.2. Fonction des protéines CELF.....	71
II.4.3. Implication des protéines CELF dans la DM1.....	73
II.5. LA DEREGULATION DE L'ÉPISSAGE DE L'EXON 5 DE LA TROPONINE T CARDIAQUE DANS LA DM1.....	74
II.5.1. La troponine T et son rôle dans la contraction musculaire.....	74
II.5.2. Régulation et dérégulation de l'épissage alternatif de l'exon 5 de hCTNT dans la DM1.....	75
II.6. OBJECTIFS.....	84
CHAPITRE III : DECRYPTAGE DES MECANISMES DE REGULATION DE L'INCLUSION DANS L'ARNM DE L'EXON 5 DU pré-ARNm DE LA TROPONINE T CARDIAQUE HUMAINE. DECOUVERTE D'UN ROLE DE LA PROTEINE hnRNP H.....	89
III.1. ETUDE DE L'IMPORTANCE DES DIFFERENTS SITES DE FIXATION DE MBNL1 DANS LA REGULATION DE L'INCLUSION DE L'EXON 5.....	89
III.1.1. Construction des minigènes nécessaire à l'étude.....	90
III.1.2. Mise en évidence de l'importance fonctionnelle individuelle des sites MBNL1 des groupes 2 et 3 dans la régulation de l'épissage alternatif de l'exon 5 du pré-ARNm hCTNT....	92
III.1.3. Les résultats obtenus <i>in vivo</i> par l'utilisation du minigène RTB300 différent de ceux obtenus avec le minigène IR-hCTNT.....	98
III.2. RECHERCHE DE NOUVEAUX REGULATEURS DE L'ÉPISSAGE ALTERNATIF DE L'EXON 5 DU pré-ARNm hCTNT.....	104
III.2.1. Principe de la méthode de chromatographie d'affinité basé sur le système MS2/MBP.....	104
III.2.2. Construction des plasmides puc18-hCTNT-MS2 permettant de produire par transcription <i>in vitro</i> des fragments de 50 nucléotides du pré-ARNm hCTNT fusionnés à 3 motifs MS2.....	106
III.2.3. Purification des complexes ribonucléoprotéiques formés avec les ARN hCTNT (A, B, C, D, E et F)-MS2 et identification de la protéine hnRNP H.....	108
III.3. ETUDE <i>IN CELLULO</i> DE L'EFFET DE LA VARIATION DE LA QUANTITE DE hnRNP H SUR L'INTEGRATION DE L'EXON 5 DU pré-ARNm hCTNT.....	115
III.3.1. La sur-expression de la protéine hnRNP h dans des cellules HEK293T entraîne une augmentation du taux d'inclusion de l'exon 5 du pré-ARNm hCTNT.....	115
III.3.2. L'absence de la protéine hnRNP H dans des cellules hek293 entraîne une diminution du taux d'inclusion de l'exon 5 du pré-ARNm hCTNT.....	117
III.3.3. Etude de l'importance des séquences introniques et exoniques pour l'activité de la protéine hnRNP H dans la régulation de l'épissage de l'exon 5 du pré-ARNm hCTNT.....	118
III.4. ETUDE <i>IN VITRO</i> DE L'EFFET DE LA VARIATION DE LA QUANTITE DE hnRNP H SUR L'INCLUSION DE L'EXON 5 DU pré-ARNm hCTNT.....	123

III.5. ETUDE DE LA RELATION ENTRE LES PROTEINES MBNL1 ET hnRNP H.....	125
III.6. CONCLUSION ET PERSPECTIVES.....	130
III.6.1. Mise en évidence d'un grand nombre de sites de régulation de l'épissage de l'exon 5 de hcTNT par MBNL1 dans le contexte génétique du minigène IR-150-E5-150.....	130
III.6.2. Proposition d'un modèle de régulation de l'inclusion de l'exon 5 de hcTNT dans le contexte du minigène IR-150-E5-150.....	131
III.6.3. Identification de la protéine hnRNP H comme régulateur de l'inclusion de l'exon 5... 134	134
CHAPITRE IV : ETUDE DU ROLE DE L'EPISSAGE ALTERNATIF DES pré-ARNm DANS LA REPONSE DES CELLULES DE VERTEBRES AU STRESS OXYDANT.....	140
IV.1. DEFINITION ET ORIGINE D'UN STRESS OXYDANT.....	140
IV.1.1. Définition d'un stress cellulaire.....	140
IV.1.2. Le stress oxydant et ses principaux acteurs.....	141
IV.2. DOMMAGES DES ROS ET DES RNS DANS LA CELLULE.....	146
IV.2.1. La peroxydation des lipides.....	147
IV.2.2. Impact du stress oxydant sur l'ADN.....	151
IV.2.3. Effet des ROS et des RNS sur les protéines.....	152
IV.3. LES PRINCIPAUX MECANISMES DE REPONSE AU STRESS OXYDANT.....	154
IV.3.1. La voie Nrf2 : la voie antioxydante.....	155
IV.3.2. Adaptation métabolique au stress oxydant : les granules de stress cytoplasmiques.....	157
IV.3.3. Adaptation métabolique au stress oxydant : les granules nucléaires.....	160
IV.3.4. Régulation de l'épissage et de l'épissage alternatif suite à différents stress.....	167
IV.4. Objectifs.....	169
CHAPITRE V : RESULTATS DE L'ETUDE DES CONSEQUENCES D'UN STRESS OXYDANT SUR L'EPISSAGE ALTERNATIF.....	171
V.1. MISE AU POINT DES DIFFERENTES CONDITIONS EXPERIMENTALES POUR LA GENERATION D'UN STRESS OXYDANT.....	171
V.1.1. Importance du lot de cellules HeLa utilisé et du nombre de passage des cellules pour la réalisation de l'étude.....	172
V.1.2. Vérification des conditions expérimentales sur la modulation de l'épissage alternatif liée au stress oxydant.....	174
V.2. CARACTERISATION DE LA REPONSE DES CELLULES HELA A UN STRESS OXYDANT GENERE PAR H ₂ O ₂	176
V.2.1. Impact du stress oxydant sur la viabilité cellulaire et le taux de ROS intracellulaires.....	176
V.2.2. Impact du stress oxydant sur l'expression des ARN sat III.....	179
V.2.3. Identification du facteur de transcription responsable de l'expression des ARN sat III.....	182
V.3. ANALYSE GLOBALE DES EFFETS D'UN STRESS OXYDANT SUR L'EXPRESSION ET L'EPISSAGE ALTERNATIF DES PRE-ARNM DE CELLULES HELA PAR UTILISATION DE PUCES A ADN DETECTANT LES EXONS.....	187

V.3.1. Description des puces à ADN utilisées lors de cette étude.	187
V.3.2. Caractéristiques des échantillons analysés sur les puces à exons.	189
V.3.3. Données concernant les variations d'expression des gènes.	189
V.3.4. Données concernant les variations d'épissage alternatif.	205
V.4. CONFIRMATION PAR RT-PCR OU RT-QPCR DES EVENEMENTS D'EPISSAGE ALTERNATIF DETECTES APRES UN STRESS OXYDANT.	208
V.4.1. Validation de la modulation de l'épissage de l'exon 3 du pré-ARNm codant la protéine TRA2B, un facteur de régulation de l'épissage alternatif.	209
V.4.2. Validation de la modulation de l'épissage de l'exon 22 du pré-ARNm codant la protéine SERCA2, un transporteur d'ions Ca ²⁺	212
V.4.3. Validation de la modulation de l'épissage des exons 7 et 10 du pré-ARNm codant la protéine ECM2.	215
V.4.4. Validation de la modulation de l'épissage de l'exon 1 du pré-ARNm codant la protéine FBXW10.	216
V.4.5. Validation de la modulation de l'épissage de l'exon 3 du pré-ARNm codant la protéine NR4A3.	217
V.5. ETUDE DES MECANISMES CELLULAIRE A L'ORIGINE DES VARIATIONS D'EPISSAGE OBSERVEES APRES UN STRESS OXYDANT.	220
V.5.1. Etudes de la variation de la quantité de snRNA spliceosomaux au cours de la réponse à un stress oxydant.	220
V.5.2. Analyse d'évènement d'épissage U12-dépendant par RT-PCR.	223
V.5.3. Etude de la relocalisation des UsnRNP et UsnRNA spliceosomaux dans le cytoplasme.	227
V.6. CONCLUSION ET PERSPECTIVES.	231
ANNEXES 239	239
ANNEXE 1 : RESULTATS DE L'ANALYSE FONCTIONNELLE PAR LE LOGICIEL DAVID DES TERMES GO ASSOCIES AUX GENES DONT L'EXPRESSION EST MODIFIEE LORS D'UN STRESS OXYDANT AU TEMPS 4 H.	239
ANNEXE 2 : RESULTATS DE L'ANALYSE FONCTIONNELLE PAR LE LOGICIEL DAVID DES TERMES GO ASSOCIES AUX GENES DONT L'EXPRESSION EST MODIFIEE LORS D'UN STRESS OXYDANT AU TEMPS 24 H.	252
ANNEXE 3 : RESULTATS DES ANALYSES ET PREDICTIONS EN AMONT DES DONNEES D'EXPRESSION DES PUCES PAR LE LOGICIEL INGENUITY AU TEMPS 2 H, 4 H ET 24 H.	260
ANNEXE 4 : RESULTATS DE L'ANALYSE FONCTIONNELLE PAR LE LOGICIEL DAVID DES TERMES GO ASSOCIES AUX GENES DONT L'EPISSAGE EST MODULE LORS D'UN STRESS OXYDANT AU TEMPS 2H, 4 H ET 24 H.	277
Matériel et Méthodes.....	299
Matériel.....	299
I. Souches bactériennes, cellules humaines et plasmides.....	299
II. Anticorps.....	301

III. Oligodésoxyribonucléotides et siRNA.	301
IV. Milieux de culture et tampons.....	306
Méthodes	308
V. Méthodes utilisées pour la culture cellulaire.....	308
VI. Méthodes d'analyse in vitro.	311
VII. Méthode d'analyse RT-PCR et RT-qPCR.	320
VIII. Microscopie de fluorescence.	323
PUBLICATIONS	325
Tau exon 2 responsive elements deregulated in myotonic dystrophy type I are proximal to exon 2 and synergistically regulated by MBNL1 and MBNL2.....	327
Probing Small Non-Coding RNAs Structures	329
Bibliographie	2

AVANT-PROPOS

AVANT-PROPOS

L'expression des gènes nucléaires au sein des cellules eucaryotes nécessite de nombreuses étapes réactionnelles afin d'obtenir des protéines fonctionnelles. La majorité des gènes codant des protéines sont transcrits par l'ARN polymérase II sous forme de transcrits primaires non fonctionnels, les pré-ARN messagers (pré-ARNm). Cette non-fonctionnalité est essentiellement due à la présence de séquences excédentaires, les introns, qui sont localisés entre les segments correspondants à l'ARN mature, les exons. L'obtention d'un ARN messager mature (ARNm) fonctionnel nécessite donc l'élimination des introns tout en maintenant les exons. Ce processus permettant d'exciser les introns au sein du pré-ARNm dans le noyau est appelé épissage. Ainsi, au cours de la réaction d'épissage, les introns des pré-ARNm sont éliminés et les séquences exoniques sont jointes les unes aux autres pour former un pré-ARNm épissé. Tôt lors de la transcription, les pré-ARNm sont ensuite coiffés en 5'. Ils sont plus tard polyadénylés en 3', ce qui permet la production des ARNm matures et fonctionnels qui sont ensuite traduits en protéine dans le cytoplasme par les ribosomes.

Le processus d'épissage est un mécanisme complexe qui concerne plus de 95% des gènes nucléaires, chez les vertébrés. Il met en jeu plus de 200 000 introns avec une taille moyenne de 1800 nucléotides (la taille de certains pouvant dépasser 700 000 nucléotides) et environ autant d'exons différents (pour revue Hoskins and Moore, 2012). La reconnaissance des introns et des exons est au centre du mécanisme d'épissage, mécanisme qui fait intervenir près de 300 protéines. Le génome humain contient une vingtaine de milliers de gènes de protéines or près de 80 000 protéines peuvent être synthétisées dans la cellule (Gencode 22, release oct 2014). Cet écart, entre le nombre de gènes et le nombre de protéines, est expliqué par le phénomène d'épissage alternatif, qui permet l'exclusion ou inclusion d'un ou plusieurs exons de façon différentielle au sein de la séquence d'un ARNm. L'épissage alternatif permet aussi d'augmenter le nombre de protéines distinctes exprimées à partir d'un seul gène (Nilsen and Graveley, 2010). L'inclusion ou l'exclusion différentielle des exons par épissage alternatif peut entraîner des modulations des propriétés de la protéine, telle que sa localisation dans la cellule, mais également modifier son activité, voir même conduire à la production d'une protéine différente.

Le nombre important d'exons soumis à épissage alternatif fait que ce processus dynamique est essentiel pour le développement des organismes eucaryotes et la différenciation cellulaire (pour revue Pedrotti and Cooper, 2014).

L'épissage d'une manière générale, et tout particulièrement l'épissage alternatif, sont extrêmement régulés et ces régulations peuvent être perturbées à la fois par des mutations dans les séquences nécessaires au mécanisme de base de reconnaissance des introns et des exons, ainsi que par des mutations qui affectent les signaux reconnus par des facteurs impliqués dans la régulation de l'épissage. Ces mutations qui affectent l'épissage peuvent être la cause directe de pathologies ou contribuer à la gravité de certaines maladies (Ward and Cooper, 2010). Il est estimé qu'au moins 20 % des maladies génétiques sont associées à des mutations affectant directement l'épissage et dont la modulation serait à l'origine des principaux symptômes de ces maladies (pour revue Singh and Cooper, 2012a). La plupart des mutations affectant l'épissage qui ont été répertoriées agissent en *cis* et conduisent à une altération de l'ARNm portant la mutation, cependant certaines mutations peuvent affecter l'épissage alternatif d'autres pré-ARNm (mutations agissant en *trans*). C'est le cas en particulier des expansions de triplets CTG à l'origine de la dystrophie myotonique de type I (DM1) ou dystrophie de Steinert, qui est la plus fréquente des dystrophies musculaires chez l'adulte. C'est une maladie génétique à évolution progressive, caractérisée par une dégénérescence des muscles squelettiques, ainsi qu'une myotonie. Les expansions instables de trinuécléotides CTG au sein de la région 3' non traduite (3'-UTR) du gène DMPK (Dystrophia Myotonica Protein Kinase) séquestrent le facteur d'épissage MBNL1 (Muscleblind-Like 1) au sein de *foci* nucléaires et conduisent à une augmentation de la concentration en facteur d'épissage CUG-BP1, un antagoniste de MBNL1. Ces modifications de la disponibilité des protéines MBNL1 et CUG-BP1 conduisent à l'épissage alternatif aberrant d'un grand nombre de pré-ARNm et sont directement responsables de la majorité des symptômes présentés par les patients. Lors de mon stage de master et au début de ma thèse, l'objectif de mon travail était de mieux comprendre les bases moléculaires de la régulation de l'épissage alternatif de l'exon 5 du pré-ARNm codant la protéine hcTNT (human cardiac Troponin T), exon dont l'épissage est altéré dans la DM1. C'est pourquoi la première partie de mon manuscrit de thèse sera dédiée à la présentation des processus d'épissage et d'épissage alternatif et à leurs régulations, puis je ferai l'état de connaissances

concernant la dystrophie myotonique de type 1 et je présenterai mes objectifs et les résultats concernant cette partie.

La deuxième partie de ce manuscrit sera consacrée à la présentation de mon second projet de thèse que j'ai initié après avoir complété ce travail sur les régulations de l'épissage alternatif de la troponine T cardiaque. L'objectif de ce second projet était de mieux comprendre le rôle de l'épissage alternatif et de ses modulations dans la réponse au stress oxydant des cellules humaines. Le stress oxydant est provoqué par une augmentation des dérivés actifs de l'oxygène, appelés également radicaux libres oxygénés (RLO ou ROS pour Reactive Oxygen Species), et représente une des plus importantes conditions de stress auxquelles les cellules de vertébrés doivent s'adapter. Il est reconnu que le stress oxydant est un des facteurs clés participant au déclenchement de nombreuses maladies dont les cancers, et certaines maladies neurodégénératives telles que maladie d'Alzheimer, la maladie de Parkinson idiopathique et la sclérose latérale amyotrophique (Desport et Couratier, 2002). De plus, il est établi que la réponse cellulaire à un stress oxydant implique une reprogrammation globale de l'expression des gènes et des altérations spécifiques de l'épissage alternatif de certains pré-ARNm avaient pu être mises en évidence (Disher and Skandalis, 2007; Skandalis and Uribe, 2004; Takeo et al., 2009). Cependant, lorsque j'ai débuté ma thèse il n'existait aucune étude globale des effets d'un stress oxydant sur l'épissage alternatif et rien n'était connu quant aux mécanismes responsables de ces variations. C'est pourquoi la deuxième partie de mon travail de thèse a consisté en une étude exhaustive des effets du stress oxydant sur l'épissage alternatif pour ensuite développer une étude des bases moléculaires à l'origine des variations d'épissage observées. De ce fait, dans la deuxième partie de mon manuscrit de thèse, j'introduirai la définition du stress oxydant et je ferai l'état des connaissances des effets de ce stress chez les vertébrés, puis je présenterai mes objectifs de mon travail et les résultats obtenus lors de cette étude.

PARTIE 1 :

**DECRYPTAGE DES MECANISMES DE
REGULATION DE L'EPISSAGE DE L'EXON 5
DU PRE-ARNm DE LA
TROPONINE T CARDIAQUE.**

INTRODUCTION

CHAPITRE I : L'ÉPISSAGE DES pré-ARNm CHEZ L'HOMME.

I.1. LE PROCESSUS D'ÉPISSAGE.

I.1.1. Les séquences permettant la définition des introns.

L'épissage des pré-ARNm nucléaires est un processus crucial pour l'expression des gènes. La reconnaissance des introns par le spliceosome est par conséquent un événement clé pour la synthèse d'un ARNm fonctionnel. Cette reconnaissance des séquences introniques repose sur la présence de séquences fortement conservées au cours de l'évolution qui sont situées aux extrémités des introns. Ces séquences sont appelées site 5' d'épissage (ou site donneur) et site 3' d'épissage (ou site accepteur) (Figure I.1). Des séquences consensus de ces sites chez les vertébrés ont pu être déterminées par comparaison de séquences. La séquence consensus au niveau du site 5' donneur d'épissage est AG/GURAGU, où "/" représente la jonction exon/intron et R, un résidu 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. Le site 3' d'épissage est constitué de 3 éléments, dont chacun est essentiel aux processus de reconnaissance de l'intron et d'épissage. Il comprend, tout comme le site 5', une séquence située à la jonction de l'intron et de l'exon (site d'épissage *stricto sensu*) dont la séquence consensus est YAG/GU, où "/" représente la jonction exon/intron et Y, un résidu pyrimidine. En amont du site *stricto sensu* (18 à 40 nt) se trouve la boîte de branchement de séquence YNYURAC qui contient un résidu adénosine conservé appelé point de branchement. Enfin, le site accepteur comporte aussi une séquence polypyrimidine, localisée entre le site 3' *stricto sensu* et la boîte de branchement, qui renferme de long enchainements de pyrimidine (10 à 20 résidus C ou U). Il est à noter l'existence d'introns qualifiés de mineurs en référence à leur faible représentation dans le génome (882 introns mineurs sur environ 1 million introns référencés). Ces introns se caractérisent également par la présence de sites 5' et 3' d'épissage, mais les séquences consensus de ces sites sont différentes. En effet, le site 5' est caractérisé par une séquence consensus /RUAUCCU, le site 3' *stricto sensu* par (YA(C ou G)/N) et la boîte de branchement par UCCUUAACU (où le point de branchement correspond au résidu adénosine indiqué en gras).

L'ensemble de ces séquences (site 5' et 3' d'épissage) permet la reconnaissance des introns par le spliceosome lors du processus d'épissage. L'efficacité de cette reconnaissance et, par conséquent, de la réaction d'épissage diminue en fonction de la dégénérescence des séquences, et est à la base de l'épissage alternatif, une reconnaissance moins efficace d'un exon alternatif, laissant place à une régulation plus fine pour favoriser l'inclusion ou l'exclusion de l'exon (Burge et al., 1998; Hoskins and Moore, 2012; Moore and Sharp, 1993).

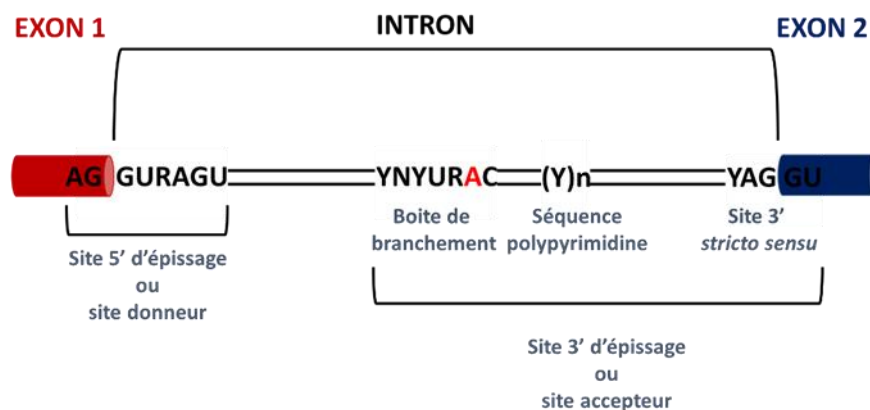


Figure I.1 : Représentation schématique de l'organisation d'un intron situé entre deux exons. Les séquences conservées jouant un rôle essentiel dans la réaction d'épissage sont représentées en gras (Y pour pyrimidine, R pour purine, N pour un des 4 résidus A, C, G ou T).

I.1.2. La réaction d'épissage.

L'épissage des pré-ARNm nucléaires implique deux étapes successives de trans-estérification qui permettent d'éliminer les introns (Figure I.2). La première étape débute par l'attaque nucléophile du phosphate en 5' de l'intron (site 5' donneur) par l'hydroxyle en 2' du résidu adénosine de la boîte de branchement (point de branchement). Cette réaction génère deux intermédiaires réactionnels : l'exon avec un hydroxyle libre en 3' et un intermédiaire intron-exon dans lequel le nucléotide du point de branchement est impliqué dans trois liaisons phosphodiester, formant ainsi une structure ayant la forme d'un lasso. Lors de la deuxième étape de trans-estérification, l'hydroxyle en 3' de l'exon, libéré lors de la première étape, réalise une attaque nucléophile sur le phosphate en 3' de l'intron (site 3' *stricto sensu*). Cette seconde étape conduit à la ligation des deux exons et à la libération de l'intron sous forme d'un lasso (pour revue Hoskins and Moore, 2012; Moore and Sharp, 1993).

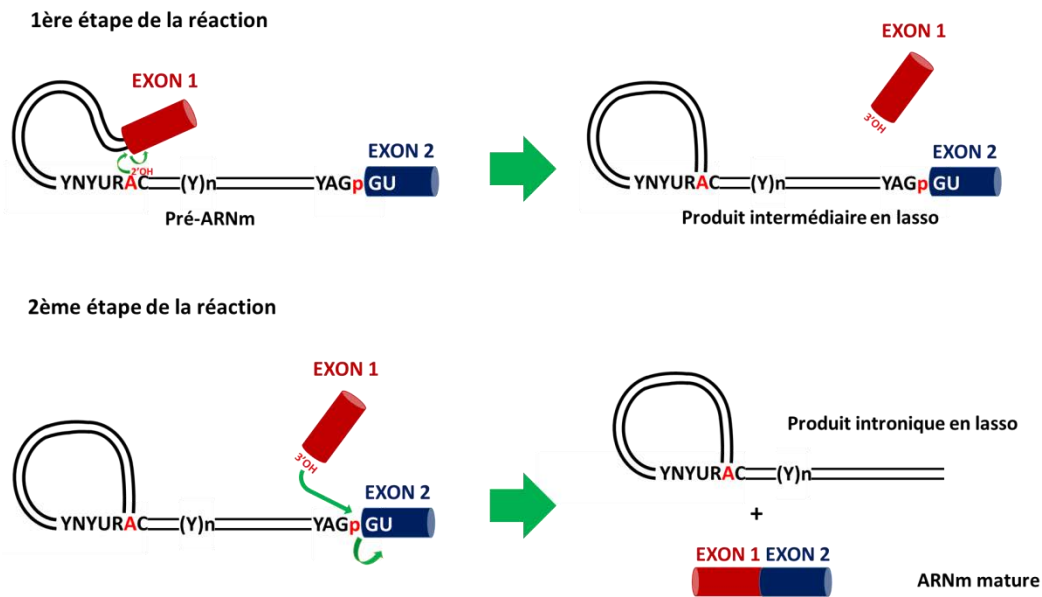


Figure 1.2 : Représentation schématique de la réaction d'épissage. Cette réaction se déroule en deux étapes correspondant chacune à une étape de trans-estérification. Dans la première étape, l'attaque par le 2' OH du résidu A de la boîte de branchement sur le phosphate du résidu à l'extrémité 5' de l'intron provoque la formation d'un intermédiaire en lasso. Dans la seconde étape, l'attaque par le 3' OH de l'exon 1 du phosphate à l'extrémité 3' de l'intron permet la libération de l'intron et de l'ARNm épissé.

I.2. FORMATION DES COMPLEXES SPLICEOSOMAux D'ÉPISSAGE.

I.2.1. Les particules UsnRNP : biogénèse et maturation.

Bien que la réaction d'épissage soit simple sur le plan chimique, la réalisation des deux étapes enzymatiques implique une dynamique très importante de mise en place et de rupture d'interactions ARN-ARN et protéines-protéines incluant, l'association au pré-ARNm de particules ribonucléoprotéiques ou snRNP ou UsnRNP (pour Uridine-rich Small Nuclear Ribonucleic Particle). L'association ces snRNP et d'un grand nombre de protéines au pré-ARNm conduit à la formation d'un macrocomplexe, le spliceosome, dont la composition et la conformation évoluent au cours de la réaction d'épissage. Le spliceosome est ainsi composé des snRNP U1, U2, U5 et U4/U6, les UsnRNP U4 et U6 étant associées en une seule di-UsnRNP. De plus, il existe, pour l'épissage des introns mineurs, un spliceosome spécifique dit mineur, composé des snRNP U11/U12, U5, U4atac/U6atac. Seule la snRNP U5 est commune aux spliceosomes mineurs et majeurs. Chaque UsnRNP est composée d'un nombre variable

de protéines, ainsi que d'un ou deux petits ARN nucléaires appelés UsnRNA (pour "Uridine rich small nuclear ribonucleic acids").

Les UsnRNA sont transcrits par l'ARN polymérase II, à l'exception des snRNA U6 et U6atac qui sont transcrits par l'ARN polymérase III. La transcription par l'ARN polymérase II et la maturation des UsnRNA sont couplées et ce processus de maturation est distinct de celui qui génère les ARNm. (Egloff et al., 2010; Henry et al., 1998; Hernandez and Weiner, 1986; pour revue Matera and Wang, 2014). Suite à la transcription par l'ARN polymérase II, les UsnRNA portent une coiffe monométhylée (m7G) en 5' et comportent des nucléotides additionnels à leur extrémité 3'. Ils sont exportés vers le cytoplasme afin d'initier l'assemblage des UsnRNP selon un processus séquentiel (Figure I.3).

La première étape d'assemblage des UsnRNP est l'association de sept protéines, appelées protéines Sm (SM B / B', D1, D2, D3, E, F et G) qui forment un anneau heptamérique autour d'une séquence spécifique, riche en résidus uridine et appelée site Sm, présente au sein de chacun des UsnRNA (Branlant et al., 1980, 1982, 1983). L'assemblage des protéines Sm est un processus séquentiel et ordonné qui fait intervenir, dans un premier temps, l'association concomitante des complexes D1-D2 et E-F-G de protéines Sm, puis le recrutement du complexe B-D3 ou B'-D3 (B et B' étant codées par le même gène et ne différant que par leurs 11 acides aminés C-terminaux) (Raker et al., 1996). La formation de l'anneau de protéines Sm est nécessaire pour la stabilité *in vivo*, l'import nucléaire et l'activité d'épissage des snRNP (Achsel et al., 2001; Kambach et al., 1999; Stark et al., 2001). Cette étape est réalisée grâce au complexe SMN (Survival Motor Neuron) composé de la protéine SMN, mais également de nombreuses autres protéines telles que les protéines Gemin (Gemin 2 à 8), qui se lient directement aux snRNA et aux protéines Sm (pour revue Battle et al., 2006).

Après la transcription, les snRNA sont exportés vers le cytoplasme où ils sont liés par Gemin5 et dirigés ensuite vers le complexe SMN pour l'assemblage de l'anneau de protéines Sm. La protéine SMN est une protéine multi-fonctionnelle indispensable à l'assemblage des complexes protéines-ARN (Eggert et al., 2006; Fischer et al., 2011). L'assemblage des snRNP par le complexe SMN est accompli de façon modulaire, regroupant plusieurs protéines de liaison de l'ARN avec plusieurs ARN " substrats ", ce qui facilite l'assemblage de protéines spécifiques sur les ARN cibles (Workman et al., 2012).

L'assemblage des protéines Sm sur le snRNA s'accompagne de la conversion de la coiffe monométhylée m7G en une coiffe triméthylée m2,2,7G (TMG) et d'une maturation de l'extrémité 3' (Paushkin et al., 2002) (Figure I.3). La particule est ensuite réimportée dans le noyau au sein de corpuscules nucléaires appelés corps de Cajal (Cajal Bodies). C'est au niveau des corps de Cajal que les snRNP vont achever leur maturation en subissant des modifications post-transcriptionnelles au niveau des snRNA (2'-O-méthylation et pseudouridylation de certains résidus), mais également en s'associant à des protéines spécifiques qui sont essentielles à leur activité.

Cependant, dans le cas des snRNA U6 et U6atac, transcrits par l'ARN polymérase III, le processus de maturation des ARN est différent. En effet, la maturation de ces snRNA est un processus exclusivement nucléaire. Les snRNA U6 et U6atac ne possèdent pas de site Sm et la formation des snRNP U6 et U6atac implique l'association d'un anneau hétéro-heptamérique de protéines similaires aux protéines Sm, les protéines LSm (Sm-Like), au niveau d'une séquence poly-U située à l'extrémité 3' de ces snRNA (Vidal et al., 1999 ; Achsel et al., 1999). Les protéines Lsm2 à 8 ont des propriétés identiques aux protéines Sm à l'exception qu'elles sont capables de former un anneau hétéro-heptamérique en absence de l'ARN (Achsel et al., 1999). Ainsi, l'assemblage de l'anneau LSm sur les snRNA U6 et U6atac pourrait être un processus en une seule étape. Cette anneau LSm sert par ailleurs de signal de rétention nucléaire pour les snRNA U6 et U6atac (Achsel et al., 1999; Spiller et al., 2007a, 2007b). De plus, ces snRNA sont modifiés post-transcriptionnellement dans le nucléole, et non dans les corps de Cajal, qu'ils rejoignent ensuite pour s'associer aux snRNP U4 et U4atac.

Les snRNP, une fois maturées, sont ensuite relocalisées avec leurs cofacteurs associés dans les speckles qui sont des structures nucléaires dynamiques de taille et de forme variables (Pour revue Spector and Lamond, 2011). Les études menées sur les speckles montrent que ces structures contiennent peu ou pas d'ADN et que l'activité d'épissage se localise essentiellement au niveau de la frontière entre les speckles et les domaines chromatiques, ce qui laisse à penser que les speckles auraient un rôle de compartiment de stockage des snRNP et de facteurs d'épissage. Le processus de relocalisation des snRNP recrutées au niveau des speckles jusqu'aux sites de transcription n'est pas à l'heure actuelle élucidé (Girard et al., 2012; Hall et al., 2006).

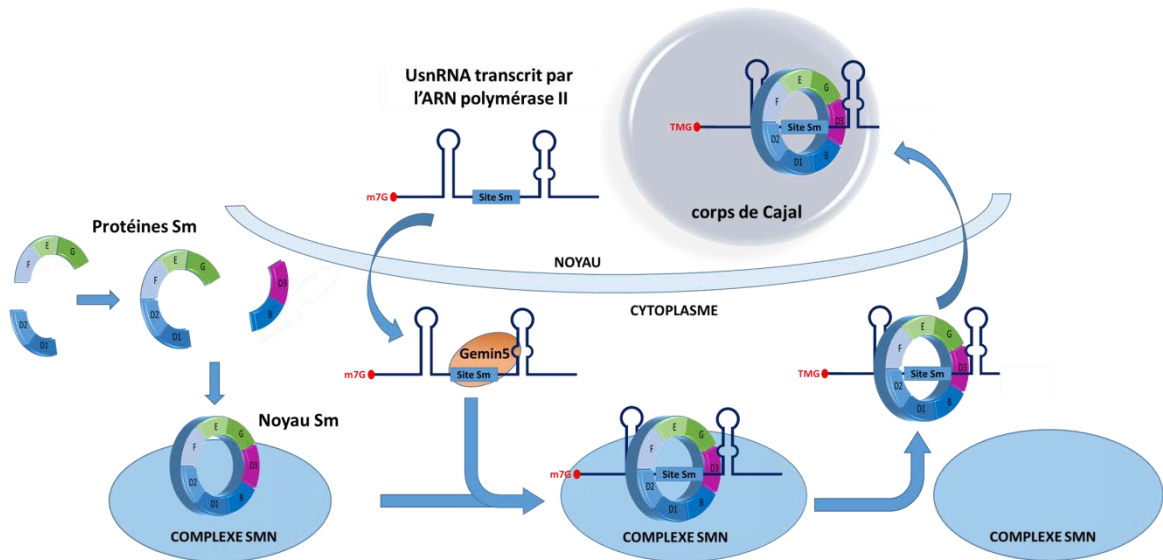


Figure I.3 : Biogenèse des UsnRNA transcrits par l'ARN polymérase II : rôle du complexe SMN. Après leur transcription dans le noyau, les U snRNA sont exportés dans le cytoplasme. Associés à la protéine Gemin 5 les UsnRNA sont redirigés ensuite vers le complexe SMN où les protéines Sm sous forme d'anneau heptamérique reconnaissent le site Sm. L'assemblage des protéines Sm sur le snRNA s'accompagne de la conversion de la coiffe monométhylée m7G en une coiffe triméthylée m2,2,7G (TMG) et d'une maturation de l'extrémité 3'. La snRNP formée est enfin importée dans le noyau, grâce à l'anneau de protéines Sm, au sein de corpuscules nucléaires appelés corps de Cajal (Cajal Bodies). C'est au niveau des corps de Cajal que les snRNP vont achever leur maturation en subissant des modifications post-transcriptionnelles au niveau des snRNA et en s'associant à de nouvelles protéines essentielles à leur activité.

1.2.2. Assemblage des complexes spliceosomaux : un processus dynamique.

1.2.2.1. Assemblage du spliceosome majeur.

L'assemblage du spliceosome est un processus complexe réalisé séquentiellement, qui nécessite la présence des snRNP, mais également l'intervention de nombreuses autres protéines. La formation du spliceosome est un phénomène dynamique qui implique de nombreuses interactions ARN-ARN et protéines-ARN ainsi que des transitions structurales des complexes spliceosomaux lors de l'activation du spliceosome et des étapes catalytiques d'épissage. Les appariements intermoléculaires mis en place dans le spliceosome sont bien caractérisés. Néanmoins, on est encore loin de comprendre les mécanismes à la base de la dynamique de cette architecture.

L'assemblage du spliceosome nécessite la formation de plusieurs complexes intermédiaires transitoires qui sont nécessaires pour l'épissage d'un intron. Dès leur synthèse, et avant même l'initiation de l'assemblage du spliceosome, les transcrits Pol II sont associés à un ensemble de protéines, dont des facteurs de régulation de l'épissage, formant

un complexe ribonucléoprotéique appelé complexe H (Figure I.4A). L'assemblage du spliceosome est ensuite initié par l'association de la snRNP U1 qui reconnaît, par appariement de bases, le site 5' donneur du pré-ARNm formant ainsi le premier complexe dit d'engagement (complexe E) (Figure I.4A). Cette reconnaissance est facilitée par la protéine U1-C de la snRNP U1 (Muto et al., 2004). L'interaction entre le site 5' donneur et la snRNP U1 pour la formation du complexe E est facilitée par d'autres facteurs, tels que les protéines SR (Serine/Arginine) qui, comme nous le verrons plus en détails par la suite, jouent un rôle essentiel dans l'assemblage du spliceosome et dans la régulation de l'épissage (Cho et al., 2011). En effet, la plupart des interactions ARN-ARN formées à l'intérieur du spliceosome sont faibles chez les vertébrés et nécessitent généralement l'aide de protéines qui renforcent leur stabilité. L'établissement du complexe E comprend également la reconnaissance du point de branchement par la protéine SF1/mBBP (Splicing Factor 1/ mammalian Branchpoint Binding Protein) qui interagit en coopération avec le facteur U2AF. La protéine U2AF est un hétérodimère composé d'une sous unité de 65 kDa (U2AF65) responsable de l'interaction avec SF1/BBP, mais également de la reconnaissance de la séquence polypyrimidine du pré-ARNm et d'une sous unité de 35 kDa (U2AF35) qui interagit avec le dinucléotide AG du site 3' accepteur d'épissage (Figure I.4A). L'ensemble de ces interactions moléculaires conduit à l'établissement du complexe spliceosomal E et joue un rôle crucial dans la reconnaissance des extrémités d'un intron.

La formation du complexe E permet ensuite le recrutement de la snRNP U2 qui se fixe sur le pré-ARNm au niveau de la boîte de branchement et interagit avec la snRNP U1 pour former le pré-spliceosome (complexe A) entraînant le déplacement de SF1/mBBP, mais également un rapprochement des exons situés à chaque extrémité de l'intron à épisser (Figure I.4A). L'établissement de ce nouveau complexe nécessite de nombreuses protéines, en particulier les protéines appartenant aux familles SF3a et SF3b, ainsi que des hélicases qui permettent et contrôlent la transition structurale des ARN. Au sein du complexe A, le pré-ARNm va subir une transition structurale au niveau du point de branchement, afin de repositionner l'adénosine portant le groupement 2'OH nécessaire à la réalisation de la première étape catalytique de l'épissage.

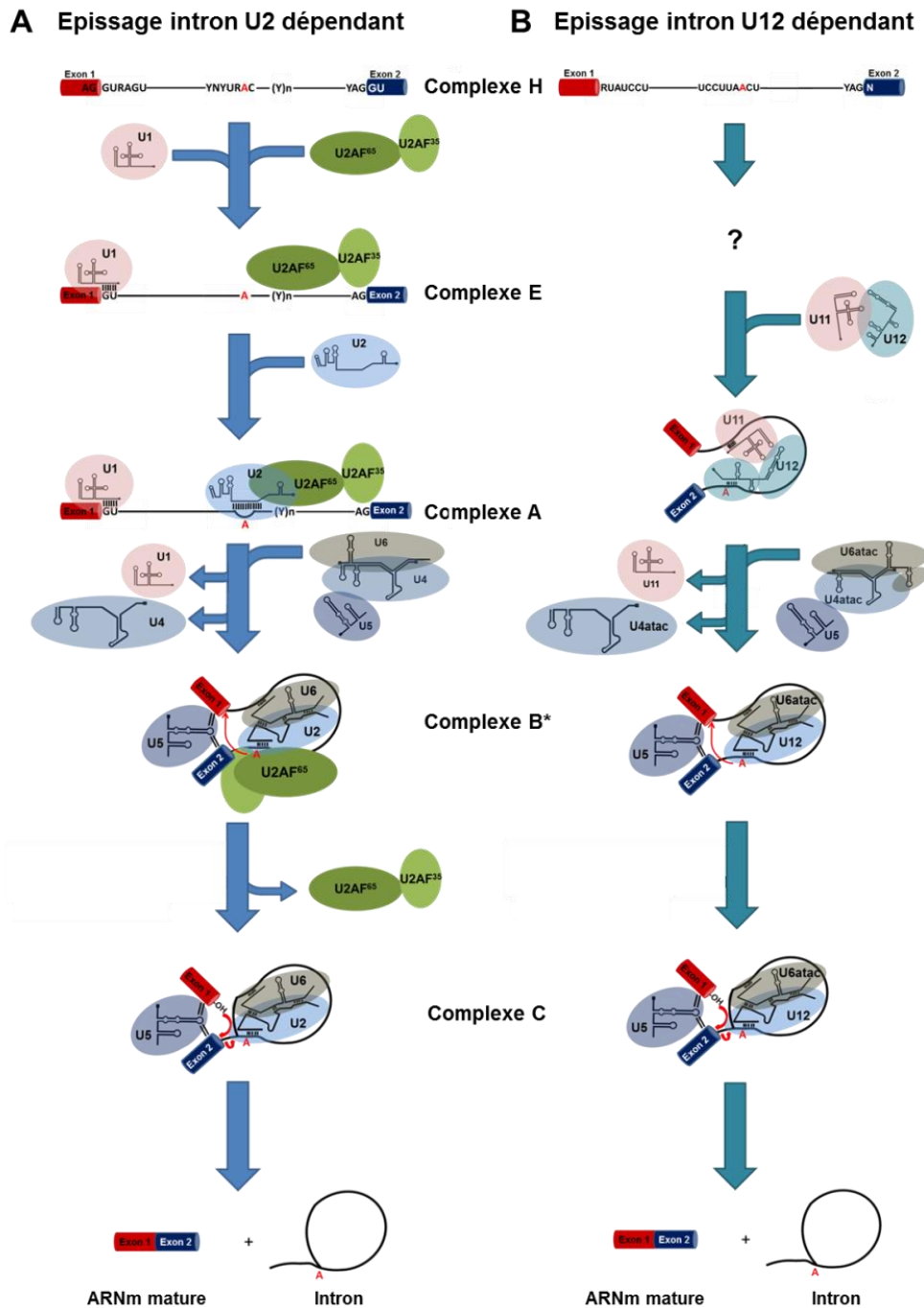


Figure I.4 : Assemblage séquentiel des spliceosomes majeur et mineur (d'après Turunen et al., 2013). A. L'assemblage du spliceosome majeur pour l'épissage des introns U2-dépendant commence par l'association de la snRNP U1 avec le site 5' d'épissage et du facteur U2AF avec la séquence polypyrimidine située au site 3' (formation du complexe E). Suite à la fixation de la snRNP U2 et la formation du complexe A, la triparticule U4/U6.U5 s'associe au complexe entraînant la dissociation de la snRNP U1. Les snRNA U2 et U6 interagissent entre eux et avec l'intron (formation du complexe B* actif, qui réalise la première étape de la réaction d'épissage). Après la première réaction d'épissage, des réarrangements structuraux ont lieu conduisant à la formation du complexe C dans lequel la snRNA U5 aligne les deux extrémités des deux exons pour faciliter leur ligation au cours de la seconde étape de la réaction d'épissage. L'ARNm formé et l'intron sous forme de lasso sont libérés. B. L'assemblage du spliceosome mineur, utilisé pour l'épissage des introns U12-dépendant, ainsi que les différentes étapes enzymatiques sont très similaires à ceux du spliceosome majeur. Cependant, il existe des différences entre les deux processus, en particulier les premières étapes d'assemblage qui ne sont pas encore totalement élucidées et qui mettent en jeu une diparticule U11/U12.

Par la suite, les snRNP U4, U5 et U6 qui sont préassemblées en tri-snRNP U4/U6.U5 sont recrutées pour la formation du complexe B (Figure I.4A). Bien que toutes les snRNP soient présentes dans ce complexe, ce dernier est encore inactif et va subir de nombreux réarrangements de conformation et de composition afin de permettre la réalisation de la première réaction catalytique. Ainsi, des interactions entre le snRNA U5 et les deux exons vont conduire à un rapprochement entre les deux exons et perturber l'interaction des snRNP U4/U6 entre elles et permettre à l'extrémité 5' du snRNA U6 d'interagir avec les paires de bases du site 5' donneur d'épissage. Ceci provoque le déplacement du snRNA U1, mais également la libération de la snRNP U4. De plus, un réseau d'appariement de bases étendu est formé entre les snRNA U6 et U2, permettant la juxtaposition du site 5' donneur d'épissage et du point de branchement avec le résidu adénosine positionné favorablement. L'établissement de cette configuration (complexe B*) entraîne l'activation du spliceosome qui réalise la première réaction de trans-estérification générant ainsi le complexe C (Figure I.4A). Avant la deuxième étape catalytique, des réarrangements supplémentaires se produisent afin de repositionner les intermédiaires d'épissage de la seconde étape de l'épissage et permettre une attaque nucléophile de l'exon 5' au site 3' d'épissage (Konarska et al., 2006). Cette deuxième réaction de trans-estérification entraîne la dissociation du spliceosome libérant les différentes snRNP U2, U5 et U6 qui sont ensuite recyclés pour la réalisation d'un nouveau processus d'épissage.

I.2.2.2 ASSEMBLAGE DU SPLICEOSOME MINEUR.

La voie d'assemblage du spliceosome mineur est très similaire à celle du spliceosome majeur dans sa globalité, mais il subsiste des différences (Figure I.4B). La principale étant l'établissement d'une di-snRNP U11/12 préformée qui reconnaît le site 5' donneur et le point de branchement de manière coopérative, sans la formation d'un complexe d'engagement et forme un pré-spliceosome (Frilander and Steitz, 1999). De plus, l'appariement de bases entre la snRNP U11 et le site 5' donneur n'est pas identique à celui réalisé avec la snRNP U1. En effet, l'interaction entre la snRNP U11 et le site 5' donneur ne passe pas directement par le snRNA, mais se fait par l'intermédiaire des domaines en doigt de zinc de la protéine U11-48K de la snRNP U11 (Turunen et al., 2008). L'interaction du snRNA U12 avec le point de branchement est similaire à celle formée par le snRNA U2, et il

est possible que le facteur SF3b participe également au processus de reconnaissance du spliceosome mineur afin de stabiliser le complexe formé (Schellenberg et al., 2011). Les introns de type U12 ne présentent pas de séquence polypyrimidine et ne nécessitent donc pas U2AF dans le processus, mais une autre protéine, appelée Urp (U2AF35 related protein), est nécessaire pour la reconnaissance et la formation du complexe (Shen et al., 2010). Suite à la reconnaissance de la di-snRNP U11/U12, la tri-snRNP U4atac/U6atac/U5 s'associe sur le pré-spliceosome pour former un complexe B catalytiquement actif (Figure I.4B). Les étapes suivantes sont ensuite très similaires à celles décrites pour le spliceosome majeur (Pour revue Turunen et al., 2013).

I.3. L'ÉPISSAGE ALTERNATIF.

Il existe chez l'homme de nombreux introns dont les sites 5' et 3' d'épissage sont dégénérés par rapport aux séquences consensus. Par conséquent, la seule reconnaissance de ces sites n'est pas suffisante pour la définition et l'épissage des introns, ce qui nécessite l'intervention de séquences régulatrices supplémentaires. Ces séquences régulatrices permettent des modulations du taux d'utilisation des sites d'épissage entraînant un épissage dit alternatif des exons (Lim and Burge, 2001). L'épissage alternatif est un phénomène majeur chez l'homme, qui au cours de l'évolution a permis d'augmenter significativement la capacité codante de notre génome en exprimant plusieurs protéines avec des fonctions ou des spécificités distinctes voir même opposées, à partir d'un seul gène. A l'heure actuelle, il est estimé que près de 93% des gènes humains subissent au niveau de leurs pré-ARNm un épissage alternatif qui peut varier d'un tissu à l'autre (Wang et al., 2008). L'épissage alternatif joue un rôle crucial dans la plupart des processus cellulaires dont la différenciation et la croissance cellulaire, mais également dans la capacité de répondre rapidement et de manière adaptée aux changements environnementaux et aux différents stress cellulaires (Barbosa-Morais et al., 2012; Pleiss et al., 2007; pour revue Chen, 2015). Le spliceosome réalise à la fois les épissages constitutif et alternatif. De ce fait, la régulation de son assemblage est un point de contrôle clé dans le processus d'épissage alternatif. Des modulations aberrantes de l'épissage alternatif peuvent avoir des conséquences très négatives pour la cellule voir pour l'organisme et être ainsi à l'origine de nombreuses pathologies (pour revue Cooper et al., 2009).

I.3.1. Les différents types d'évènements d'épissage alternatif.

Huit grands types d'évènements d'épissage alternatif différents ont pu être mis en évidence et sont illustrés dans la figure I.5 :

- Les exons cassettes : c'est l'évènement d'épissage alternatif le plus fréquent. L'exon alternatif peut être soit inclus, soit exclu de l'ARNm de manière tissu spécifique ou en fonction des conditions physiologiques.

- Les exons mutuellement exclusifs : bien que deux exons mutuellement exclusifs aient chacun des sites 5' et 3' fonctionnels, la reconnaissance par le spliceosome d'un des couples de sites abolit la reconnaissance de l'autre. Par conséquent, ces exons ne sont jamais présents ensemble au sein du même ARNm.

- Les sites 5' ou 3' alternatifs d'épissage: il peut exister au niveau de certains exons plusieurs sites 5' ou 3' d'épissage fonctionnels qui entrent alors en compétition. Selon le site utilisé lors de l'épissage, les conséquences de la sélection d'un site par rapport à un autre sont la formation d'isoformes d'ARNm plus ou moins long.

- La rétention d'intron : c'est le cas le plus rare, mais dans certains cas, les sites 5' et 3' d'épissage ne sont pas reconnus et l'intron associé est alors maintenu au sein de l'ARNm. Ce phénomène est généralement lié à un défaut lors de la reconnaissance de l'intron par la machinerie d'épissage.

- Les exons alternatifs en 5' du pré-ARNm : il peut exister au niveau de la séquence génomique plusieurs sites d'initiation alternatifs de la transcription, qui conduisent à la synthèse de différents transcrits. Ces différents sites d'initiation peuvent moduler soit l'inclusion d'exons en 5' mutuellement exclusifs, soit l'inclusion d'un exon ou d'un groupe d'exons localisés en aval du site d'initiation entraînant la formation d'une isoforme de taille différente.

- Les exons en 3' du pré-ARNm alternatifs : différents exons d'un même gène peuvent contenir un site de polyadénylation. Par conséquent, la reconnaissance d'un premier site de polyadénylation peut empêcher l'inclusion d'un ou plusieurs exons en aval du site de polyadénylation, conduisant à la formation de différentes isoformes avec des exons terminaux différents.

- Sites de polyadénylation alternatifs : il est également possible qu'au sein d'un même exon il existe plusieurs sites de polyadénylation alternatifs dont l'utilisation peut conduire à la synthèse de différentes isoformes d'ARNm.

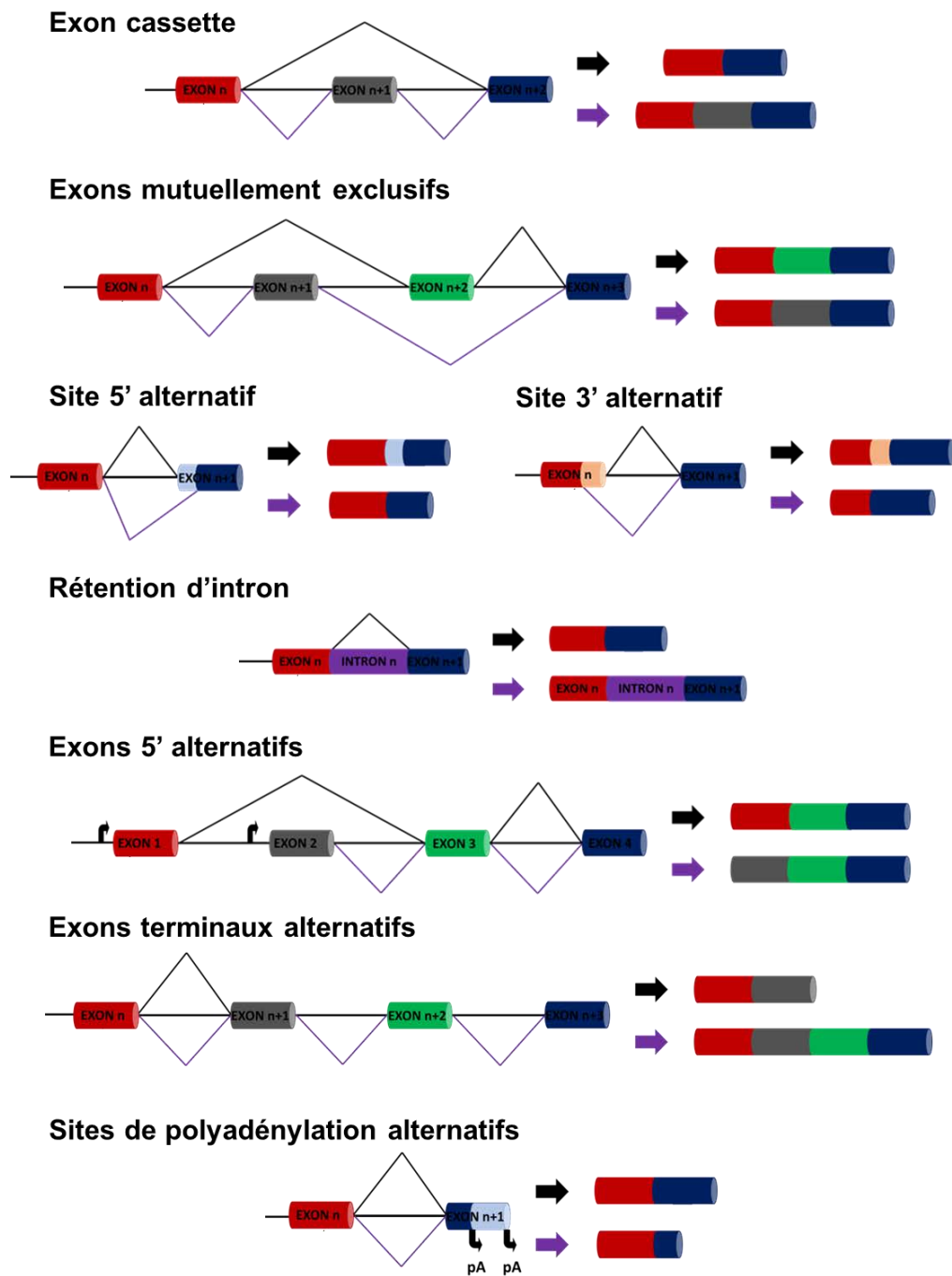


Figure I.5 : Les différents modes d'épissage alternatif des pré-ARNm. Les flèches désignent les promoteurs alternatifs et les pA, les signaux de polyadénylation.

Les différents types d'évènements d'épissage alternatif peuvent être combinés au sein d'un même pré-ARNm et conduire à différentes isoformes d'ARNm issues d'un même gène. La production d'une isoforme au dépend d'une autre est régulée par la nature des sites d'épissage, la présence de séquences cis-régulatrices, ainsi que la concentration des facteurs protéiques de régulation de l'épissage. De plus, le type cellulaire, les conditions environnementales, le patrimoine génétique, ainsi que l'état global physiologique de la cellule influencent le taux de production d'une isoforme par rapport aux autres (Lu et al., 2012; pour revue Wang and Burge, 2008).

I.3.2. Les séquences cis-régulatrices de l'épissage.

Comme mentionné précédemment, la dégénérescence des séquences des sites 5' et 3' d'épissage conduit à l'existence de sites d'épissage dits sous-optimaux, moins efficacement reconnus par le spliceosome. L'utilisation de ces sites nécessite la présence de facteurs protéiques qui interagissent soit au niveau de l'exon, soit au niveau de l'intron en s'associant à des séquences appelées cis-régulatrices afin de favoriser ou inhiber le recrutement du spliceosome.

Les séquences cis-régulatrices activatrices sont appelées ESE (pour "Exonic Splicing Enhancer"), lorsqu'elles sont localisées au sein d'un exon, ou ISE (pour "Intronic Splicing Enhancer"), lorsqu'elles sont localisées dans un intron. De même, les séquences cis-régulatrices inhibitrices sont appelées ESS (pour "Exonic Splicing Silencer") et ISS (pour "Intronic Splicing Silencer") (Wang et al., 2006; pour revue Blencowe, 2006).

Au niveau d'un même site d'épissage, il est possible de trouver simultanément des séquences cis-régulatrices inhibitrices et activatrices qui entrent alors en compétition. Les séquences cis-régulatrices agissent toujours par l'intermédiaire de protéines partenaires, dont les protéines SR qui sont majoritairement associées aux éléments activateurs, et les protéines hnRNP qui sont, le plus souvent, associées aux séquences inhibitrices, mais également d'autres protéines telles que FOX1, NOVA, TDP43 et MBNL1 (Chen et al., 2015b; Matlin et al., 2005). De ce fait, la régulation de l'épissage alternatif est souvent dépendante de la concentration relative des facteurs activateurs et inhibiteurs de l'inclusion des exons.

I.4. LES COMPLEXES SPLICEOSOMAX : UNE MACHINERIE COMPLEXE, DYNAMIQUE ET METTANT EN JEU UN TRES GRAND NOMBRE DE PROTEINES.

Les complexes spliceosomaux sont des complexes ribonucléoprotéiques renfermant un très grand nombre de protéines, qui peuvent représenter jusqu'à plus de deux tiers de leur masse. Les complexes spliceosaux humains peuvent nécessiter au total plus de 170 protéines distinctes qui permettent d'assurer le mécanisme de base de l'épissage et qui sont associées pour certaines directement aux snRNP, ou aux constitutants des sous-complexes protéiques nécessaires à la formation des corps spliceosomaux, à leur régulation ou à leur activité. En fait, les complexes spliceosomaux évoluent énormément tant dans la nature des interactions ARN-ARN, ARN-protéine que protéine-protéine, depuis l'assemblage des pré-spliceosomes jusqu'aux complexes portant l'activité pour la 1^{er} étape de la réaction d'épissage, puis à ceux catalysant la 2^{ème} étape et ceux permettant la libération d'un ARNm après élimination des introns. D'où la notion de complexes spliceosomaux et non plus d'une et unique machinerie d'épissage : le spliceosome. Cette complexité permet d'assurer l'efficacité et la précision des épissages constitutif et alternatif. Par ailleurs, de nombreuses protéines assurant la régulation de l'épissage (environ 300) sont retrouvées aux complexes spliceosomaux, dont la plupart sont préassemblées en complexes hétéromères stables afin d'assurer leur association aux pré-complexes spliceosomaux (Rappsilber et al., 2002; pour revue Wahl et al., 2009). La famille des protéines SR et SR-like joue un rôle très particulier dans les mécanismes d'épissage, dans la mesure où d'un côté elles sont requises pour le mécanisme de base et de l'autre elles interviennent dans la régulation de l'épissage en contrecarrant souvent l'effet négatif des protéines hnRNP.

I.4.1. Les protéines SR.

L'épissage des pré-ARNm chez les métazoaires, et en particulier les vertébrés, est dépendant des protéines SR, une famille de facteurs d'épissage essentiels à la fois pour la mise en place d'interactions protéine-protéine permettant l'assemblage de la machinerie de base de l'épissage, mais aussi la définition des exons, et la régulation de leur utilisation. Les protéines SR comportent un domaine C-terminal qui est riche en répétitions de dipeptides arginine-sérine (appelé domaines RS). Les domaines RS ont pour fonction de permettre des interactions avec d'autres protéines possédant le même type de domaines (Zahler et al., 1992). Ces interactions protéine-protéine permettent de faciliter l'assemblage du spliceosome dans la mesure où un nombre important de protéines associées directement aux snRNP possèdent des domaines RS. Les domaines RS peuvent également renforcer la stabilité des interactions avec les ARN. Les protéines SR sont également caractérisées par le fait qu'elles possèdent un ou deux domaines d'interactions avec les l'ARN de type RRM (pour RNA Recognition Motif). Les protéines SR sont subdivisées en deux sous familles sur la base du nombre de domaines RRM qu'elles possèdent. On distingue d'une part la sous famille SRSF1, qui possède deux domaines RRM, et d'autre part, la sous famille SRSF2 qui n'en possède qu'un. C'est sur la base de ces caractéristiques de domaines et de séquences, qu'en 2010, une nouvelle classification des protéines SR a été proposée en les définissant comme « toutes protéines ayant un ou deux domaines RRM suivie d'un domaine SR d'au moins 50 acides aminés avec une teneur en arginine et sérine de 40% » (Manley and Krainer, 2010). Cette nouvelle classification, beaucoup plus stricte que les précédentes, a permis d'établir l'existence de 12 protéines SR chez l'homme dont les noms sont donnés dans le tableau I.1 et qui sont maintenant dénommées "SRSF" (Serine/arginine-rich splicing factor), suivi d'un numéro issu de la nomenclature de leurs gènes et reflétant l'ordre chronologique de leur découverte (Tableau I.1).




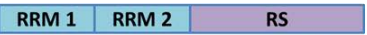
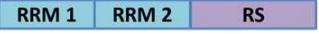
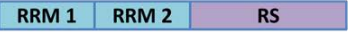



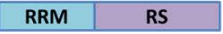


Symbole du gène	Noms précédents de la protéines	Nouvelles nomenclature	Domaines protéiques
SFSR1	ASF/SF2; ASF;SRp30a	SFSR1	
SFSR2	SC35; PR264; SRp30b	SFSR2	
SFSR3	SRp20	SFSR3	
SFSR4	SRp75	SFSR4	
SFSR5	SRp40; HRS	SFSR5	
SFSR6	SRp55; B52	SFSR6	
SFSR7	9G8	SFSR7	
SFSR2B	SRp46	SFSR8	
SFSR9	SRp30c	SFSR9	
SFSR13A	SRp38; TASR1; Srrp40	SFSR10	
SFSR11	SRp54; p54	SFSR11	
SFSR13B	Srrp35	SFSR12	

Tableau I.1 : Liste des protéines SR exprimées chez l'homme. (Adapté de (Manley and Krainer, 2010).

Le symbole du gène, l'ancienne nomenclature, la nouvelle nomenclature proposée en 2010 ainsi que l'organisation générale des différentes protéines SR humaines sont indiqués. RS : domaine riche en arginines et en sérines à l'origine d'interactions protéine-protéine. RRM : pour RNA Recognition Motif, domaine d'interaction avec l'ARN. Zn : motif doigt à zinc.

I.4.1.1. Structure des protéines SR et leur interaction avec les ARN.

La structure des domaines RRM de la plupart des protéines SR a pu être établie, en particulier par RMN et par cristallographie (Daubner et al., 2013). Un domaine RRM (RRM1) est localisé en N-ter de toutes les protéines SR (RRM1). Il est constitué de 4 feuillets β et deux hélices α qui adoptent une topologie de type $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$ (Figure I.6) (Maris et al., 2005). Les feuillets $\beta 1$ et 3 constituent une surface d'interaction avec l'ARN caractérisé par la présence de deux séquences conservées, appelées RNP1 et RNP2, respectivement localisées au niveau des feuillets $\beta 3$ et $\beta 1$ (Figure I.6). Ces séquences RNP participent à la spécificité de reconnaissance des ARN et sont en général essentielles à l'interaction des domaines RRM avec l'ARN. Néanmoins, les études réalisées par l'équipe de F. Allain mettent en évidence une grande variété de modèles d'interaction des ARN avec les domaines RRM et en particulier l'intervention d'interactions avec les hélices.



Figure I.6 : Schéma d'un domaine RRM. Les feuillets β sont représentés par une flèche verte et les hélices α par un rectangle rose.

Le second domaine RRM (RRM2), aussi appelé domaine pseudo RRM, qui caractérise les protéines de la sous famille SRSF1, est structuralement proche du domaine RRM1 (Birney et al., 1993). Cependant les séquences RNP retrouvées dans les domaines RRM2 présentent généralement des dégénérescences de séquences par rapport à celles des domaines RRM1 (Cáceres et al., 1997). Grâce à des expériences de SELEX (pour Systematic Evolution of Ligands by EXponential enrichment), les séquences ARN de haute affinité pour les protéines SR ont été déterminées, ce qui a permis de réaliser des analyses structurales poussées des interactions par RMN montrant la complexité des mécanismes qui assurent la spécificité de reconnaissance des ARN (Tableau I.2). (Nagai et al., 1990; Daubner et al., 2013).

Symbole du gène	Noms précédents de la protéines	Nouvelles nomenclature	Site de fixation à l'ARN
SFSR1	ASF/SF2; ASF;SRp30a	SFSR1	RGAAGAAC (SELEX) AGGACRRAGC (SELEX) SRSASGA (S100) (SELEX fonctionnel) CRSMGW (S100) (SELEX fonctionnel) UGRWG (CLIP)
SFSR2	SC35; PR264; SRp30b	SFSR2	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)
SFSR3	SRp20	SFSR3	WCAKCUUUA (SELEX) AUCCNNA (SELEX) WCAUCGAY (SELEX) YWCUUCAU (SELEX) CUWCAAC (SELEX) GGUCCUCUUC(EMSA) WCWWC (Test épissage) CUCKUCY (EN) (SELEX fonctionnel) CCUCGUCC (EN) (SELEX fonctionnel)
SFSR4	SRp75	SFSR4	GAAGGA (Pontage UV)
SFSR5	SRp40; HRS	SFSR5	GAGCAGUCGGCUC (SELEX) UGGGAGCRGUYRGCUCGY (SELEX) YRCRKM (S100) (SELEX fonctionnel) ACDGS (S100) (SELEX fonctionnel)
SFSR6	SRp55; B52	SFSR6	GRUCAACNNGGCGACMG (SELEX) USCGKM (S100) (SELEX fonctionnel) YYWCWSG (S100) (SELEX fonctionnel)
SFSR7	9G8	SFSR7	UCAACA (Pontage UV) AGACKACGAY (SELEX) ACGAGAGAY (SELEX) GGACGACGAG (EN) (SELEX fonctionnel)
SFSR9	SRp30c	SFSR9	GACGAC (SELEX fonctionnel) AAAGAGCUCGG (SELEX fonctionnel) CUGGAUU (EMSA)
SFSR13A	SRp38; TASR1; Srrp40	SFSR10	ACAAAGACAAA (SELEX)
SFSR11	SRp54; p54	SFSR11	Riche en C (Pontage UV)

Tableau I.2 : Séquences reconnues par les protéines SR. Pour chaque séquence présentée, nous avons précisé la méthode d'identification. N= tout nucléotides ; R=purine ; Y=pyrimidine ; D= A, G ou U ; K=U ou G ; M= A ou C ; S= G ou C ; W=A ou U. (d'après la revue de Long et Caceres, 2009).

I.4.1.2. Les protéines "SR-like".

Il existe un grand nombre de protéines qui ne sont pas conformes à l'ensemble des critères nécessaires pour appartenir à la famille des protéines SR, bien qu'elles possèdent également un ou plusieurs domaines RS. Ces protéines appelées SR-like ou SRrp (pour SR protein-related) participent à de nombreux processus dont l'épissage, certaines interviennent dans le couplage entre transcription et épissage, d'autres facilitent la libération des ARNm du spliceosome ainsi que leur export (Ohno and Shimura, 1996; pour revue Long and Caceres, 2009). Ces protéines se distinguent par le fait qu'elles ne possèdent pas nécessairement de domaines de liaison à l'ARN de type RRM, et/ou que leurs domaines RS ne soient pas localisés à l'extrémité N-terminale. En plus, elles peuvent posséder d'autres domaines qui leur confèrent des activités enzymatiques. C'est notamment le cas des protéines SR-like HRH1 et hPRP16, qui possèdent en plus d'un domaine RS situé dans la région C-terminale, une boîte DEXD/H qui confère à ces protéines leur activité hélicase. Les protéines SR-like présentent une vaste diversité d'activité et interviennent à toutes les étapes de l'épissage, de la formation des snRNP pour les protéines U1-70K, U4/U6.U5-27K et U5-100K, ainsi qu'au recrutement du spliceosome avec les protéines U2AF35 et U2AF65, jusqu'à la régulation de l'épissage par hTRA2 α et hTRA2 β (Figure I.7) (Tacke et al., 1998; pour revue Graveley, 2000).

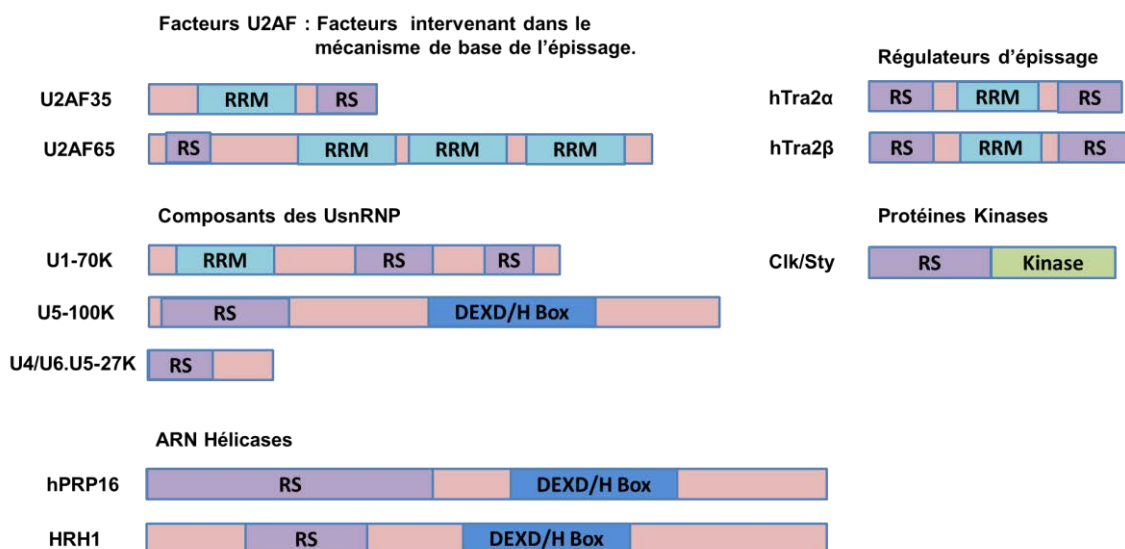


Figure I.7 : Exemples de protéines à domaines RS autres que les protéines SR (Graveley, 2000 ; Long and Caceres, 2009).

I.4.1.3. Rôles des protéines SR dans le mécanisme de base de l'épissage constitutif et ses régulateurs.

Les protéines SR, de par leur structure, établissent de nombreuses interactions avec les pré-ARNm ainsi qu'avec différentes protéines tout au long du processus d'épissage. En effet, les protéines SR interviennent dans la reconnaissance de l'exon en 5' ceci en interagissant avec les protéines CBP20 et CBP80 du complexe CBC fixé sur la coiffe afin de permettre le recrutement de la snRNP U1 (Lewis et al., 1996). Dès les premières étapes d'assemblage du spliceosome, les protéines SR contribuent également à la reconnaissance des introns et exons en interagissant avec les éléments ESE ou ISE. Cette reconnaissance des séquences cis-activatrices favorise le recrutement de la snRNP U1, par l'intermédiaire de la protéine U1-70K, au site 5' d'épissage, ainsi que du complexe U2AF sur la séquence polypyrimidine (Wu and Maniatis, 1993; Tarn and Steitz, 1994; Cho et al., 2011; pour revue Bradley et al., 2015). Cet ensemble d'interactions participe au rapprochement des sites 5' et 3' lors de l'assemblage du spliceosome. De manière générale les protéines SR favorisent et stabilisent l'assemblage du spliceosome, et semblent également participer au respect du bon ordre de ligation des exons au cours de l'épissage d'un transcrit (Ibrahim et al., 2005).

Les protéines SR interviennent aussi dans la régulation de l'épissage alternatif. En effet, elles permettent de réguler l'inclusion d'exons alternatifs avec des sites spécifiques de reconnaissance sur le pré-ARNm grâce à l'interaction et à la coordination de protéines SR entre elles (Figure I.8). Ainsi, une variation d'expression d'une protéine SR peut entraîner des modulations de l'épissage alternatif et favoriser l'inclusion ou l'exclusion d'un grand nombre d'exons (Pandit et al., 2013 ; Han et al., 2011; pour revue Bradley et al., 2015). Enfin les protéines SR sont associées à la reconnaissance du dernier exon d'un pré-ARNm en interagissant avec la machinerie de polyadénylation et la protéine PAP (pour poly A Polymérase) qui favorise le recrutement du facteur U2AF65 au niveau du site 3' d'épissage. Elles interviennent également dans la régulation de l'utilisation des sites de polyadénylation alternatifs (Blechingberg et al., 2007; Kyburz et al., 2006; Vagner et al., 2000).

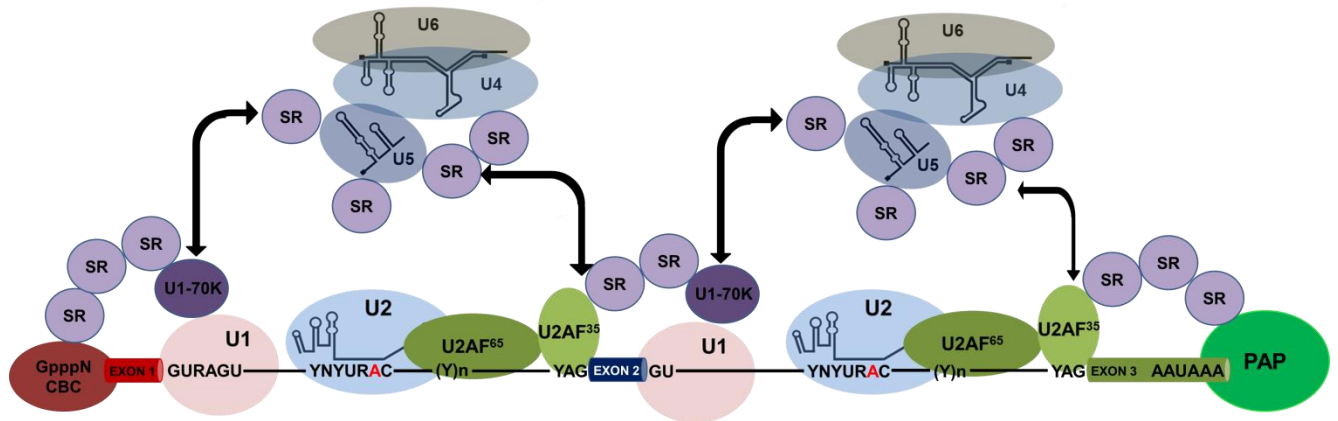


Figure I.8 : Rôle des protéines SR dans l'épissage des ARNm. Les protéines SR, par l'intermédiaire de leur domaine SR, favorisent la reconnaissance des sites 5' et 3' d'épissage par la snRNP U1 et le complexe U2AF lors de la définition des exons. Les protéines SR participent également à la reconnaissance des exons en 5' et en 3' des pré-ARNm en interagissant avec le CBC (Cap Binding Complex) et en favorisant le recrutement de la protéine polyA Polymérase (PAP).

1.4.1.4. Modification post-traductionnelle et régulation de l'activité des protéines SR.

La régulation de l'activité des protéines SR, ainsi que de toutes les protéines ayant des domaines SR, est largement dépendante de la présence de modifications post-traductionnelles. La majorité de ces modifications post-traductionnelles sont des phosphorylations réalisées principalement par deux familles de protéines kinases : les protéines kinases SR spécifiques (SRPK) et les protéines CLK/Sty (pour cdc2-like kinase/serine, thréonine et tyrosine). Bien que, ces deux familles de kinases aient des propriétés communes pour permettre la phosphorylation des protéines SR au niveau de leur domaine SR, elles diffèrent en termes de spécificité et de mécanisme d'action. En effet, les protéines SRPK phosphorylent les protéines SR au niveau de plusieurs sites sur une longue répétition Arg-Ser à l'extrémité N-terminale du domaine RS, tandis que les protéines CLK/Sty reconnaissent des répétitions plus courtes présentant plusieurs dipeptides Ser-Pro au niveau de l'extrémité C-terminale du domaine SR (Aubol et al., 2013; Colwill et al., 1996a, 1996b). Le degré de phosphorylation du domaine RS est par ailleurs régulé par de nombreuses phosphatases cellulaires. Cette régulation de l'état de phosphorylations des protéines SR permet de moduler leur affinité pour d'autres protéines à domaine RS, leur localisation cellulaire ainsi que leur activité. Ainsi, une même protéine SR peut présenter différents états de phosphorylation au cours du temps et selon les processus cellulaires dans lesquels elle est impliquée. De récentes études montrent que la régulation des différentes formes

phosphorylées des protéines SR peut en partie être sous le contrôle de longs ARN non codant (Cooper et al., 2014; Tripathi et al., 2010). Ce cycle dynamique de phosphorylation et de déphosphorylation des protéines SR est essentiel pour de nombreux processus cellulaires, en particulier l'initiation et l'assemblage des complexes spliceosomaux, ainsi que la régulation de l'épissage alternatif (Cao et al., 1997; Xiao and Manley, 1997; pour revue Howard and Sanford, 2015).

I.4.1.5 LOCALISATION DES PROTEINES SR DANS LA CELLULE.

Comme dit précédemment, la localisation des protéines SR est fortement associée au degré de phosphorylation de leur domaine SR. Les domaines RS contiennent généralement aussi un signal de localisation nucléaire qui peut parfois s'étendre jusqu'au domaine de fixation à l'ARN, ce qui favorise leur concentration dans le noyau et plus particulièrement dans les speckles. La localisation des protéines SR dans les speckles n'est pas définitive et ces protéines selon leur degré de phosphorylation peuvent se déplacer d'un speckle à un autre, mais également rejoindre des sites de transcription environnant (Misteli and Spector, 1999; Misteli et al., 1997). Par ailleurs, une hyperphosphorylation du domaine RS des protéines SR entraîne leur relocalisation dans le cytoplasme, où elles ont des activités différentes dans de nombreux processus cellulaires autre que l'épissage.

I.4.1.6 LES PROTEINES SR : DES PROTEINES MULTIFONCTIONNELLES.

Le degré de phosphorylation des protéines SR affecte, non seulement leur localisation, mais également leur activité dans l'épissage, ainsi que leurs activités dans d'autres processus biologiques. En effet, au cours de la dernière décennie, le rôle des protéines SR a été étendu à tout un ensemble de processus cellulaires autres que l'épissage. Par exemple, les protéines SRSF1 et SRSF2 peuvent réguler le recrutement de l'ARN polymérase II et moduler l'efficacité de transcription et donc d'élongation du pré-ARNm naissant (Ji et al., 2013; Lin et al., 2008). Les protéines SR peuvent également intervenir dans le transport des ARNm du noyau vers le cytoplasme. En effet, les protéines SR servent, grâce à leurs domaines RRM et RS, d'adaptateurs entre les ARNm et des facteurs protéiques d'exports nucléaires (Hautbergue et al., 2008; Huang et al., 2003). De plus, les protéines SR

sont impliquées dans la régulation de la traduction. Ainsi, SRSF1 est associée à près de 1500 ARNm au sein des polysomes, en particulier des ARNm codant des protéines impliquées dans le cycle cellulaire. Elle est également nécessaire à leur traduction. Il a été également rapporté que les protéines SR sont des acteurs importants dans les mécanismes de surveillance de la qualité des ARNm en favorisant l'activité du NMD (Nonsense Mediated mRNA Decay) (Maslon et al., 2014; Zhang and Krainer, 2004; Sanford et al., 2004). Les protéines SR peuvent également intervenir dans la stabilité des ARNm, le maintien de la stabilité génomique et la transformation oncogénique (Zhong et al., 2009, 2005; pour revue Long and Caceres, 2009; Huang and Steitz, 2005). Tous ces processus sont donc liés par les protéines SR, ce qui suggère que ces phénomènes sont couplés et non distincts ou segmentés comme proposé à l'origine (Pour revue Braunschweig et al., 2013a; Das and Krainer, 2014).

I.4.2. Les protéines hnRNP.

Les protéines hnRNP (pour heterogeneous nuclear ribonucleoprotein particule) ont été initialement décrites pour leur capacité à interagir avec les ARN nouvellement synthétisés par l'ARN polymérase II, entraînant la formation de particules hétérogènes dont elles tirent leur nom. Cette famille protéique est composée d'une vingtaine de membres dont certains font partie des protéines les plus abondantes au sein de la cellule, telle la protéine hnRNP A1 qui est présente en plus grande quantité que la snRNP U1 (Dreyfuss et al., 1993; pour revue Han et al., 2010). Les protéines hnRNP sont classées par ordre alphabétiques de A (34kDa) à U (120 kDa) en fonction de leurs poids moléculaires. Les protéines hnRNP forment des complexes ARN-protéines avec les ARN et interviennent dans leur maturation, en particulier l'épissage. Elles interviennent aussi dans de nombreux processus cellulaires nucléaires, tels que l'activation de la transcription, le remodelage de la chromatine et également dans l'export des ARN dans le cytoplasme, leur stabilité et leur traduction. Par conséquent, ces protéines jouent plusieurs rôles essentiels dans la cellule au même titre que les protéines SR.

I.4.2.1. Structure des protéines hnRNP et interaction avec l'ARN.

Toutes les protéines hnRNP contiennent un ou plusieurs domaines d'interaction avec l'ARN : domaines RRM, RRMH (pour RRM homology) ou KH (pour K homology) (Tableau I.3). Les domaines RRM sont similaires à ceux retrouvés au sein des protéines SR. Les domaines RRMH se distinguent des domaines RRM canoniques par la présence d'une boucle β_3 supplémentaire et l'absence de séquences RNP consensus. Ils ont un mode d'interaction avec l'ARN différent de celui des domaines RRM classiques, pouvant varier selon la protéine hnRNP (Dominguez and Allain, 2005, 2006; Vitali et al., 2006). Les domaines KH ont été nommés ainsi en raison de leur découverte au sein de la protéine hnRNP K. Le domaine KH de type I est classiquement retrouvé chez les eucaryotes et présente une organisation structurale de type β_1 - α_1 - α_2 - β_2 - β' - α' , distincte de celle des domaines RRM (pour revue Valverde et al., 2008). Les protéines hnRNP présentent donc une vaste diversité de domaines pour interagir avec l'ARN.

Cette diversité de domaines d'interaction à l'ARN permet la reconnaissance d'une multitude de séquences ARN et les spécificités d'interactions de certaines des protéines hnRNP avec l'ARN ont été déterminées par des expériences de SELEX (Tableau I.4). Parmi ces séquences, on notera qu'un certain nombre de protéines hnRNP présente une forte affinité pour les séquences riches en résidu G (Pour revue Han et al., 2010). En plus des domaines de fixation à l'ARN, les protéines hnRNP contiennent un ou plusieurs domaines auxiliaires riches en glycine (domaine Gly) qui sont caractérisés par la présence de motifs tri-peptidiques Arg-Gly-Gly (motif RGG) intercalés de résidus phénylalanine et tyrosine. Les domaines Gly ont un rôle essentiel pour les protéines hnRNP car ils stabilisent l'interaction des protéines hnRNP avec l'ARN et permettent également d'établir des interactions protéine-protéine. Ainsi, le domaine Gly de la protéine 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). Cette propriété de la protéine hnRNP A1 est importante pour la régulation de l'épissage alternatif par cette protéine. En effet, la fixation de hnRNP A1 à un site de haute affinité grâce aux domaines RRM1 et RRM2 peut s'accompagner de la fixation d'autres molécules de protéine hnRNP A1 sur des sites adjacents de plus faible affinité par oligomérisation, mais également d'interactions avec d'autres protéines, telle la protéine

hnRNP H (Fisette et al., 2010; Hallay et al., 2006). Les motifs RGG sont d'autant plus essentiels pour la protéine hnRNP U, qui ne possède pas de domaine classique de fixation à l'ARN, et interagit avec ses partenaires ARN uniquement via son domaine Gly (Kiledjian and Dreyfuss, 1992). Cette diversité structurale des protéines hnRNP permet une modularité et une grande variabilité en termes de combinaisons et d'interactions protéiques, qui sont à la base d'une diversité fonctionnelle de ces protéines au sein de la cellule.

Symbole du gène (hnRNP)	Domaines protéiques
A/B	RRM RRM Gly
C	RRM Asp-Glu
D	RRM RRM Gly
E	KH KH KH
F	RRMH RRMH RRMH
G	RRM Gly
H	RRMH RRMH Gly RRMH Gly
2H9	RRMH RRMH
I	RRMH RRMH RRMH RRMH
K	KH KH RGG KH
L	Gly RRMH RRMH RRMH RRMH
M	RRM RRM RRM
P	Gly RRM Gly
Q	RRM RRM RRM RGG
R	Asp-Glu RRM RRM RRM Gly RGG
U	Asp-Glu Gln RGG Gly

Tableau I.3 : Structure générale et domaines des protéines hnRNP. hnRNP A/B représente les hnRNP (A0, A1, A2/B1 and A3) qui ont des structures similaires, bien qu'elles soient exprimées par des gènes différents. hnRNP M et Q représentent respectivement les hnRNP M1 à 2 et les hnRNP Q1 à 5. Les différents domaines ou motifs des hnRNP sont représentés par des rectangles : RRM (RNA Recognition Motif) ; RRMH (RNA Recognition Motif Homologue) ; KH (K Homology) ; Gly (riche en glycine) ; motif RGG (Arg-Gly-Gly) ; Asp-Glu/Gln (domaine riche en résidus Asp-Glu ou Gln). D'après Han et al., 2010 ; Krecic and swanson., 1999 ; Dreyfuss et al., 2002.

Symbole du gène (hnRNP)	Séquence d'interaction ARN	
A/B	G₄(AU₃)₄A (A0) GCCAAGGAGCC (A3)	UAGGG, (Py)UAG (A1,A1B) GCCAAGGAGCC (A2,B1)
C	Riche en U	AGUA(U)SGUGGA
D	Riche en A/U	UUAG/G
E	Poly (C)	Poly (G)
F		Poly (G)
G		AAGU
H	GGGA/C	Poly (G)
2H9		GGGA
I		Riche en C/U
K	Poly (C)	AU(C)_{3/4}(A/U)₂
L		Répétition de C/A
M		GAGGAGG
P		
Q		
R		
U	Poly (G)	Riche en A/U

Tableau I.4 : Séquences reconnues par les différentes protéines hnRNP (méthode SELEX). HnRNP A/B représente les hnRNP (A0, A1, A2/B1 and A3), pour chaque séquence d'interaction il est indiqué en marron la protéine hnRNP qui est spécifiquement associée à cette séquence. HnRNP M et Q représentent respectivement les hnRNP M1 à 2 et les hnRNP Q1 à 5. Les résidus G sont représentés en rouge. D'après Han et al., 2010 ; Krecic and Swanson., 1999 ; Dreyfuss et al., 2002.

I.4.2.2. Localisation des protéines hnRNP dans la cellule.

Les protéines hnRNP ont globalement une localisation diffuse dans le nucléoplasme, mais sont exclues des nucléoles. Cependant, les protéines hnRNP prises individuellement peuvent avoir des spécificités de localisation dans le noyau comme c'est le cas pour la protéine hnRNP L qui est localisée dans des *foci* non-nucléolaires associés avec des chromosomes (Piñol-Roma et al., 1988). La localisation nucléaire des protéines hnRNP est essentiellement due à la présence d'un signal de localisation nucléaire appelé NLS (pour "Nuclear Localization Signal"). Toutefois certaines protéines, comme hnRNP A1, n'ont pas de NLS, mais un motif riche en glycines et en résidus aromatique appelé M9, permettant d'interagir avec une protéine d'import nucléaire, la transportine. Ainsi la localisation de ces protéines est régulée par leur degré de phosphorylation, en particulier la phosphorylation

d'une séquence en aval du motif M9, qui une fois phosphorylée, diminue l'interaction avec la transportine et maintient ces protéines dans le cytoplasme. Certaines protéines hnRNP (A/B, D, E, I et K), au même titre que les protéines SR, ont la capacité de transiter entre le noyau et le cytoplasme. Cette capacité peut être ou non liée à la transcription selon que la protéine hnRNP possède ou non un domaine NLS. Ainsi, les protéines hnRNP A1 et I, qui n'ont pas de domaine NLS classique, transitent depuis le noyau vers le cytoplasme en couplage avec la transcription, alors que le mouvement de la protéine hnRNP K, qui possède un domaine NLS classique, est indépendant de la transcription (Gama-Carvalho and Carmo-Fonseca, 2001). Ces propriétés de mouvement au sein de la cellule confèrent aux hnRNP un rôle dans le transport des ARNm vers le cytoplasme. C'est notamment le cas des protéines hnRNP A2 et B1 qui sont impliquées dans le transport des ARNm long (Sinnamon and Czaplinski, 2011). Bien qu'à l'origine, il ait été montré que les protéines hnRNP C et U ne sortent pas du noyau, ces protéines ont ensuite été détectées dans certaines conditions au niveau de granules de stress cytoplasmique avec des ARN. Ainsi, suite à une infection virale, la protéine hnRNP C peut être relocalisée dans ces granules cytoplasmiques avec également les protéines hnRNP A1 et K (Kanai et al., 2004; Piñol-Roma and Dreyfuss, 1992). La localisation des protéines hnRNP semble être essentielle dans la réponse cellulaire à une variation environnementale et leur relocalisation au sein de granules de stress cytoplasmiques suggère un rôle physiologique important. C'est notamment le cas de la protéine hnRNP A1 qui, suite à un stress oxydant, est hyperphosphorylée et s'accumule au sein des granules de stress cytoplasmiques, affectant ainsi la traduction et la régulation de l'épissage alternatif (Guil et al., 2006; Cammas et al., 2007)

1.4.2.3. Rôles des protéines hnRNP dans l'épissage constitutif et alternatif.

Par leurs propriétés d'interaction avec l'ARN et avec des protéines partenaires, les protéines hnRNP sont impliquées dans la régulation de l'épissage en modulant la reconnaissance des introns/exons par le spliceosome. Les protéines hnRNP sont essentielles à la régulation du taux d'épissage des introns et de l'épissage alternatif et une diminution du taux de ces protéines peut affecter près de 6555 événements d'épissage alternatif dans 4638 gènes chez l'homme (Huelga et al., 2012). Lors de l'épissage, l'influence des protéines hnRNP peut se faire au travers de la reconnaissance des séquences cis-régulatrices

exoniques ou introniques ou bien directement au niveau des sites donneurs ou accepteurs d'épissage. La fixation des protéines hnRNP entraîne la plupart du temps un encombrement stérique empêchant la fixation des composants du spliceosome au niveau des sites d'épissage du pré-ARNm (Figure I.9B et C). C'est notamment le cas pour la protéine hnRNP I (également appelée PTB, pour Polypyrimidine Tract-Binding protein), qui interagit préférentiellement avec la séquence polypyrimidine du site 3' accepteur d'épissage (Singh et al., 1995). En effet, la spécificité de liaison de hnRNP I est similaire à celle de la protéine U2AF65 qui, comme nous l'avons vu, est impliquée dans le recrutement de la snRNP U2. Ainsi la protéine hnRNP I peut interférer avec la reconnaissance du point de branchement et donc du site 3' accepteur en empêchant l'assemblage du spliceosome (Izquierdo et al., 2005; Sharma et al., 2008). Cet encombrement stérique peut être renforcé par la capacité de certaines protéines hnRNP à s'oligomériser le long du pré-ARNm (Figure I.9D) (Zhu et al., 2001; Okunola and Krainer, 2009; pour revue Busch and Hertel, 2012). Cependant, ce type de multimérisation pourrait aussi dans certains cas être utilisé pour favoriser le recrutement du spliceosome et l'inclusion d'exons (Ortuño-Pineda et al., 2012; Xue et al., 2009).

Les protéines hnRNP peuvent également intervenir dans la régulation de l'épissage en modulant la reconnaissance des sites d'épissage par la formation de grandes boucles d'ARN (Figure I.9E). La formation de ces boucles est réalisée par la fixation des protéines hnRNP au niveau de séquences introniques situées de part et d'autre d'un exon. Suite à cette fixation, les protéines hnRNP interagissent entre elles, ce qui a pour conséquences un rapprochement des séquences introniques et la formation d'une boucle d'ARN contenant l'exon. Cette boucle dans la structure ARN conduit à un isolement de l'exon qui n'est plus reconnu par le spliceosome. Ce mode de régulation a notamment été décrit pour les protéines hnRNP H, I (PTB), L ainsi que pour la protéine hnRNP A1 qui régule ainsi l'épissage alternatif de l'exon 7B de son propre pré-ARNm conduisant à la formation de plusieurs isoformes d'hnRNP A1 (Martinez-Contreras et al., 2006; pour revue He and Smith, 2009).

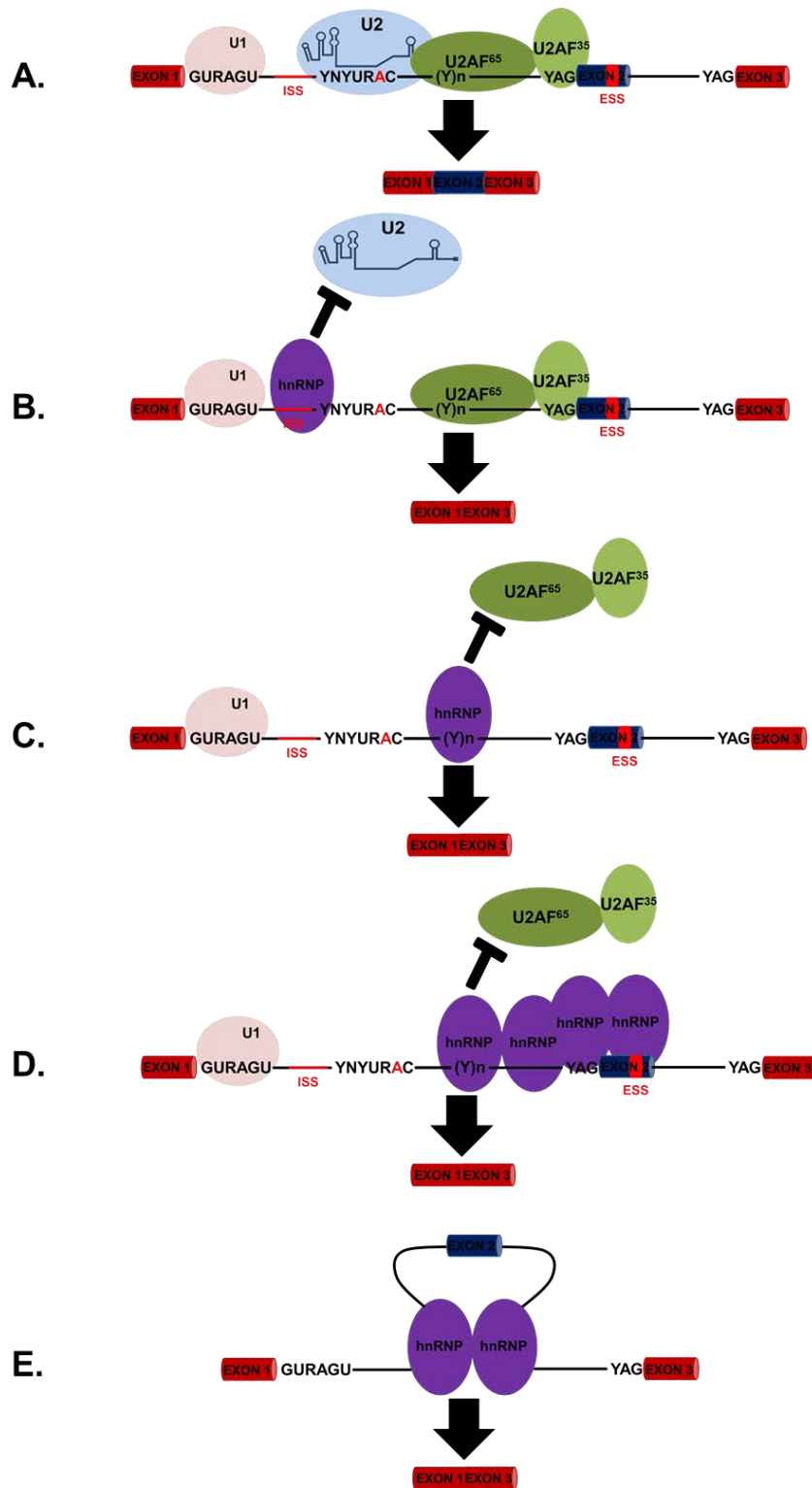


Figure I.9 : Modes de régulation de l'épissage par les protéines hnRNP. A. Activation de l'épissage en absence d'interaction du pré-ARNm avec des protéines hnRNP. Les séquences régulatrices inhibitrices de l'épissage intronique (ISS) et exonique (ESS) sont représentées en rouge. B. Inhibition de l'épissage par une protéine hnRNP en limitant le recrutement de la snRNP U2. C. Inhibition de l'épissage par fixation d'une hnRNP au niveau de la séquence polypyrimidine, inhibant ainsi par encombrement stérique la fixation du complexe U2AF. D. Inhibition de l'épissage par oligomérisation de hnRNP le long du pré-ARNm. E. Inhibition de l'épissage par rapprochement d'introns suite à l'interaction de protéines hnRNP entre elles entraînant la formation d'une boucle d'ARN contenant l'exon alternatif et inhibant la reconnaissance de l'exon par le spliceosome (d'après He et Smith, 2009).

I.4.2.4. Antagonisme entre protéines hnRNP et protéines SR dans l'épissage alternatif.

L'inhibition de l'inclusion d'un exon par les protéines hnRNP ne se limite pas à empêcher stériquement la reconnaissance des sites d'épissage par les composants des complexes spliceosomaux, mais peut également être liée à leur capacité à interférer avec la fixation des protéines SR activatrices. Cet antagonisme entre les protéines hnRNP et SR est un mécanisme général de régulation de l'épissage qui intervient lorsque des séquences cis-régulatrices activatrices et inhibitrices, respectivement reconnues par des protéines SR et hnRNP, se chevauchent ou se juxtaposent. Les protéines sont alors en compétition pour la fixation sur ces séquences, la fixation de l'une empêchant la fixation de l'autre par encombrement stérique. De ce fait, l'utilisation d'un site d'épissage par le spliceosome est déterminée d'une part par le taux relatif de chaque protéine régulatrice au sein du noyau et ceci de manière tissu-spécifique et d'autre part par leur affinité pour la séquence d'intérêt (Figure I.10 A et B). Ainsi, la régulation de l'épissage alternatif de l'exon 2 du pré-ARNm codant l'alpha-tropomyosine met en compétition les protéines hnRNP H et SRSF7 dont les séquences cis-régulatrices sont juxtaposées (Crawford and Patton, 2006). A forte concentration en protéines hnRNP, leur capacité à contrecarrer l'interaction des protéines SR sur l'ARN est exacerbée par leurs propriétés d'oligomérisation (Figure I.10 C) (Eperon et al., 2000). Il a ainsi pu être établi au laboratoire que la régulation de l'utilisation du site accepteur A7 du VIH (Virus de l'Immunodéficience Humaine), qui est essentielle pour la production de la protéine virale Tat, met en jeu une fixation coopérative de la protéine hnRNP A1 sur trois sites de régulation de forte affinité. La fixation de la protéine hnRNP A1 sur ces sites permet à la protéine de s'oligomériser et de recouvrir une partie des séquences régulatrices de l'ARN, dont le site A7, qui n'est plus accessible à la protéine SRSF1, entraînant une inhibition de la reconnaissance du site A7 (Marchand et al., 2002). Plus récemment un degré supplémentaire de complexité dans ce mécanisme de régulation a été démontré puisqu'il a pu être montré que la régulation du site A7 peut également faire intervenir la protéine hnRNP K qui, en favorisant la fixation de la protéine hnRNP A1, inhibe l'utilisation de ce site (Marchand et al., 2011) .

De manière générale, la régulation de l'utilisation d'un site d'épissage peut être dépendante de plusieurs mécanismes de régulation qui sont associés à un panel de protéines. Bien que les protéines SR et hnRNP soient des protéines essentielles à la régulation de l'épissage, il existe d'autres familles de protéines qui interviennent dans la régulation de l'épissage et qui jouent un rôle majeur dans la détermination de l'inclusion ou l'exclusion d'un exon.

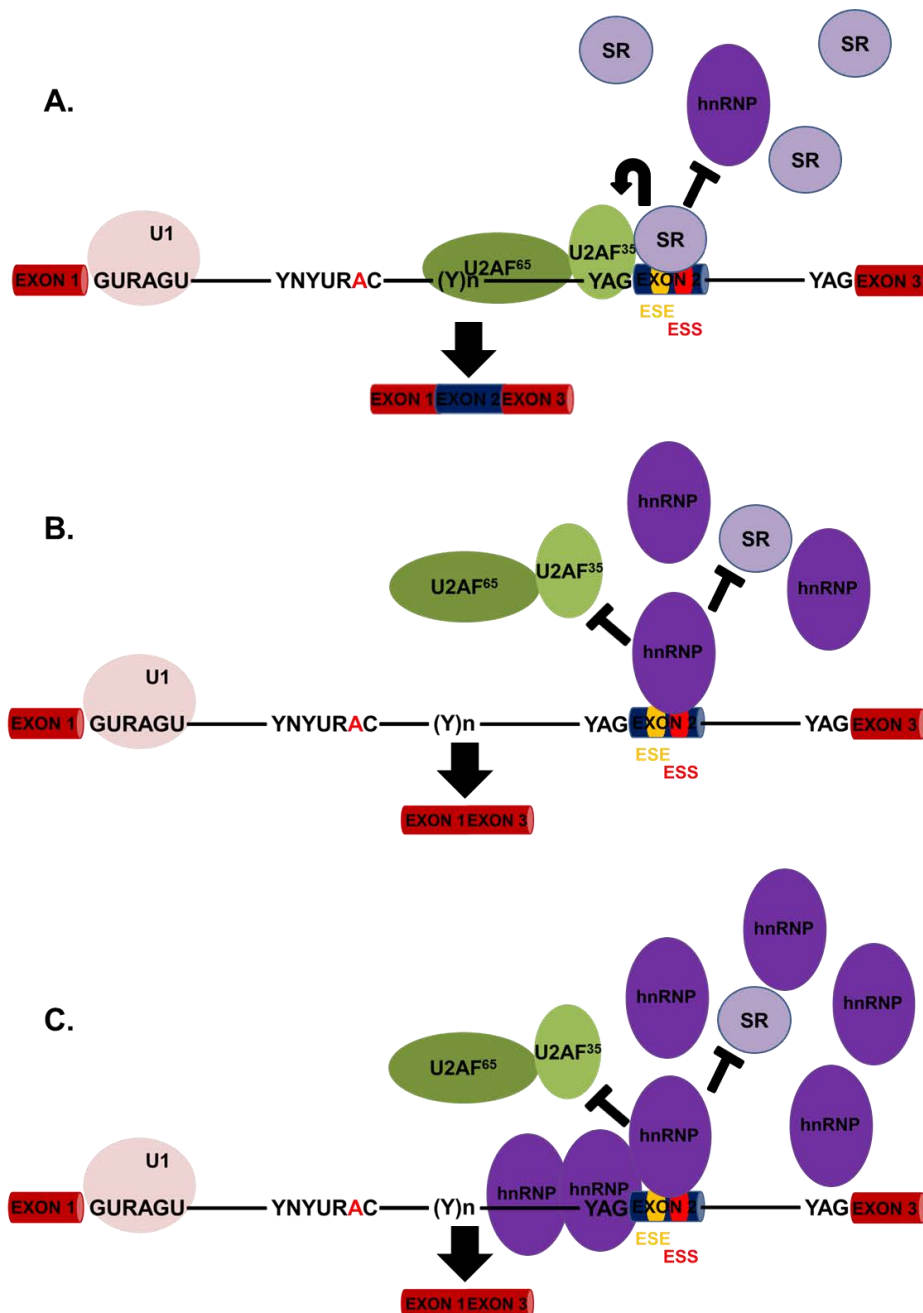


Figure I.10 : Antagonisme entre protéines hnRNP et SR dans la régulation de l'épissage. A. Activation de l'épissage et inclusion d'un exon par fixation d'une protéine SR qui favorise le recrutement du complexe U2AF et inhibe la fixation d'une protéine hnRNP. **B.** Inhibition de l'épissage d'un exon par fixation d'une protéine hnRNP qui limite la fixation du complexe U2AF et la fixation d'une protéine SR. **C.** Inhibition de l'épissage par oligomérisation des protéines hnRNP présentes en forte concentration.

I.4.3. Les autres protéines régulatrices de l'épissage.

Parmi ces facteurs, il existe des protéines de régulation de l'épissage dont l'activité est tissu-spécifique et qui sont regroupées en familles protéiques. Deux exemples de ces familles vont être décrits ci-dessous, néanmoins il en existe plusieurs autres dont les familles MBNL et CELF qui seront décrites ultérieurement dans la partie de l'introduction bibliographique consacrée à la présentation de la dystrophie myotonique.

I.4.3.1. Les protéines Nova.

Les protéines Nova, représentées par les protéines Nova 1 et 2, ont été les premiers facteurs d'épissage tissu-spécifiques identifiés. Elles sont exclusivement exprimées dans les neurones du système nerveux central et sont caractérisées par la présence de 3 domaines KH de fixation à l'ARN, qui reconnaissent les répétitions de séquence YCAY (où Y est un résidu pyrimidine) (Buckanovich et al., 1993; Jensen et al., 2000). Une étude à haut débit, basée sur l'utilisation de puces "exon-jonction", a permis d'établir que les protéines Nova sont impliquées dans la régulation de l'épissage alternatif d'environ 50 pré-ARNm, dont la majorité codent des protéines synaptiques (Ule et al., 2005). De façon intéressante, il a pu être établi que la position relative des séquences YCAY sur les pré-ARNm détermine l'issue de la régulation de l'épissage de l'exon. Ainsi, la fixation de la protéine Nova en amont ou sur l'exon favorise l'exclusion de ce dernier, tandis qu'une interaction au niveau de l'intron en aval favorise l'inclusion de l'exon (Licatalosi et al., 2008; Ule et al., 2006). Les protéines Nova peuvent donc activer ou inhiber l'inclusion d'un exon donné selon la localisation de leur site de fixation au sein des pré-ARNm, mais également selon leur taux d'expression dans la cellule dans la mesure où ces protéines peuvent également entrer en compétition avec d'autres protéines de régulation de l'épissage. De manière générale, l'effet inhibiteur des protéines Nova est associé à une inhibition du recrutement de la snRNP U1 et leur effet activateur à un antagonisme avec un autre facteur de régulation comme c'est le cas dans la régulation de l'épissage de l'exon 6 du pré-ARNm codant la dopamine où Nova 1 est en compétition avec la protéine hnRNP M. Dans ce cas, la fixation de la protéine Nova 1 empêche la fixation de la protéine hnRNP M, un inhibiteur de l'inclusion de l'exon 6 (Darnell, 2013; Park et al., 2011).

1.4.3.2. Les protéines Fox.

La famille des protéines Fox, ou Rbfox (pour RNA-binding protein Fox), est constituée des protéines Rbfox 1, 2 et 3. Ces protéines sont d'importants régulateurs tissus-spécifiques de l'épissage alternatif. Contrairement à ses paralogues, Rbfox1 et Rbfox3, qui sont principalement exprimés dans les tissus musculaires et neuronaux, Rbfox2 est le seul membre de la famille Rbfox exprimé dans les cellules souches embryonnaires ainsi que dans un certain nombre de lignées cellulaires, dont les cellules cancéreuses épithéliales mésenchymateuses (Lee et al., 2009; Venables et al., 2009; Warzecha et al., 2009). Leurs activités sont associées à la présence d'un domaine d'interaction à l'ARN de type RRM qui reconnaît spécifiquement la séquence UGCAUG (pour revue Kuroyanagi, 2009). Récemment, des études par HITS-CLIP (high-throughput sequencing) ont permis de montrer que 59 à 65% des interactions des protéines Rbfox sur les ARN se situent au niveau de séquences introniques. En parallèle, l'emploi de l'approche RNA-Seq dans des cellules où Rbfox2 était réprimée a permis d'identifier plus d'une centaine d'exons régulés par cette protéine (Weyn-Vanhentenryck et al., 2014). De plus, un enrichissement en séquences UGCAUG a été observé en amont des exons exclus en présence des protéines Fox et en aval des exons dont l'inclusion est favorisée par ces protéines (Auweter et al., 2006; Wang et al., 2008; Yeo et al., 2009; Zhang et al., 2008a). De manière générale, la fixation des protéines Rbfox en 5' de l'intron situé en amont de l'exon à épisser empêche la reconnaissance du point de branchement par la protéine SF1, ainsi que la formation du pré-complexe E du spliceosome, ce qui entraîne l'exclusion de l'exon. De la même manière, les protéines Rbfox peuvent inhiber l'utilisation d'ESE et favoriser l'exclusion d'exons (Fukumura et al., 2007; Zhou and Lou, 2008). Il est à noter que les protéines Rbfox 1 et 2 ont des spécificités similaires de reconnaissance de l'ARN et peuvent par conséquent réguler l'épissage alternatif d'un même exon. C'est notamment le cas de l'épissage du transcrite codant la protéine Rbfox1. En effet, l'un des exons codant en partie le domaine RRM de la protéine est soumis à un épissage alternatif et son exclusion entraîne la formation d'un dominant négatif qui ne peut plus interagir avec l'ARN (Damianov and Black, 2010). L'exclusion de cet exon est non seulement positivement auto-régulée par la protéine Rbfox1, mais également par Rbfox2 et 3 (Weyn-Vanhentenryck et al., 2014).

Une étude récente, chez la souris, montre que la protéine Rbfox 2 régule de nombreux événements d'épissage alternatif entraînant l'activation du NMD de pré-ARNm codant des protéines qui interagissent avec l'ARN, ceci afin de moduler leur expression (Jangi et al., 2014).

I.5. L'ÉPISSAGE ALTERNATIF : UN PROCESSUS COUPLE AVEC DE NOMBREUX MECANISMES CELLULAIRES.

Bien que la réaction d'épissage puisse être récapitulée *in vitro* dans un extrait nucléaire en l'absence du génome complet et en absence des autres processus nucléaires, l'épissage *in vivo* est couplé avec plusieurs processus nucléaires, dont la transcription, et ces couplages sont nécessaires pour l'efficacité du processus d'épissage. Ainsi, alors que dans les cellules l'épissage des introns de grande taille (>1500 nts) est très rapide, l'épissage *in vitro* des introns de taille supérieure à 500 nt est peu efficace et nécessite plus de temps. L'ajout d'ARN pol II dans un extrait nucléaire permettant l'épissage renforce l'efficacité de cet extrait. In *cellulo*, la reconnaissance des introns et des exons dépend aussi d'autres paramètres, comme l'état de la chromatine et de ces modifications épigénétiques qui peut moduler l'efficacité d'épissage d'un exon. L'épissage alternatif est un processus qui permet de réguler le taux de production de certains ARNm et qui influence par l'inclusion ou l'exclusion d'exons le devenir des ARNm, ainsi que l'activité des protéines produites à partir de ces ARNm. Par conséquent, l'épissage alternatif joue un rôle central dans la régulation du métabolisme cellulaire et est couplé avec de nombreux autres processus cellulaires.

I.5.1. Couplage entre l'épissage et la transcription des pré-ARNm.

Aujourd'hui le couplage entre les processus de transcription et d'épissage est parfaitement établi dans la littérature. Deux modèles non mutuellement exclusifs ont été proposés pour expliquer ce couplage : le modèle de couplage par recrutement et le modèle de couplage cinétique (Figure I.11).

Le modèle de couplage par recrutement implique l'association physique de facteurs d'épissage avec la machinerie transcriptionnelle, en particulier par l'intermédiaire du domaine CTD (pour Carboxy Terminal domain) de la grande sous-unité de l'ARN

polymérase II (Das et al., 2007), qui *in vitro* peut interagir avec des protéines possédant des domaines RS, facilitant l'assemblage des protéines SR et donc des complexes spliceosomaux (Millhouse and Manley, 2005). L'influence de l'ARN polymérase II se retrouve dès les phases précoces de l'épissage mais également tout le long du processus. En effet, le domaine CTD est impliqué, tôt au cours de la transcription dans la mise en place de la coiffe m7G des transcrits Pol II qui permet ensuite la fixation du complexe CBC ("Cap Binding Complex ") nécessaire à la définition de l'exon en 5'. Plus généralement, le domaine CTD en liaison avec les protéines SR peut promouvoir et coordonner l'initiation de l'assemblage du spliceosome en interagissant avec la snRNP U1 dans un premier temps, puis avec le complexe U2AF65/PRP19 lors des phases plus tardives du processus d'épissage, au moment de l'activation de la réaction catalytique (David et al., 2011; pour revue Lee and Tarn, 2013). L'implication du CTD dans l'épissage des pré-ARNm est d'autant plus importante dans ce modèle que ce domaine est nécessaire à la localisation des complexes de transcription à l'abord des speckles nucléaires qui sont le lieu de stockage d'un grand nombre de facteurs requis pour l'épissage (Misteli et al., 1997).

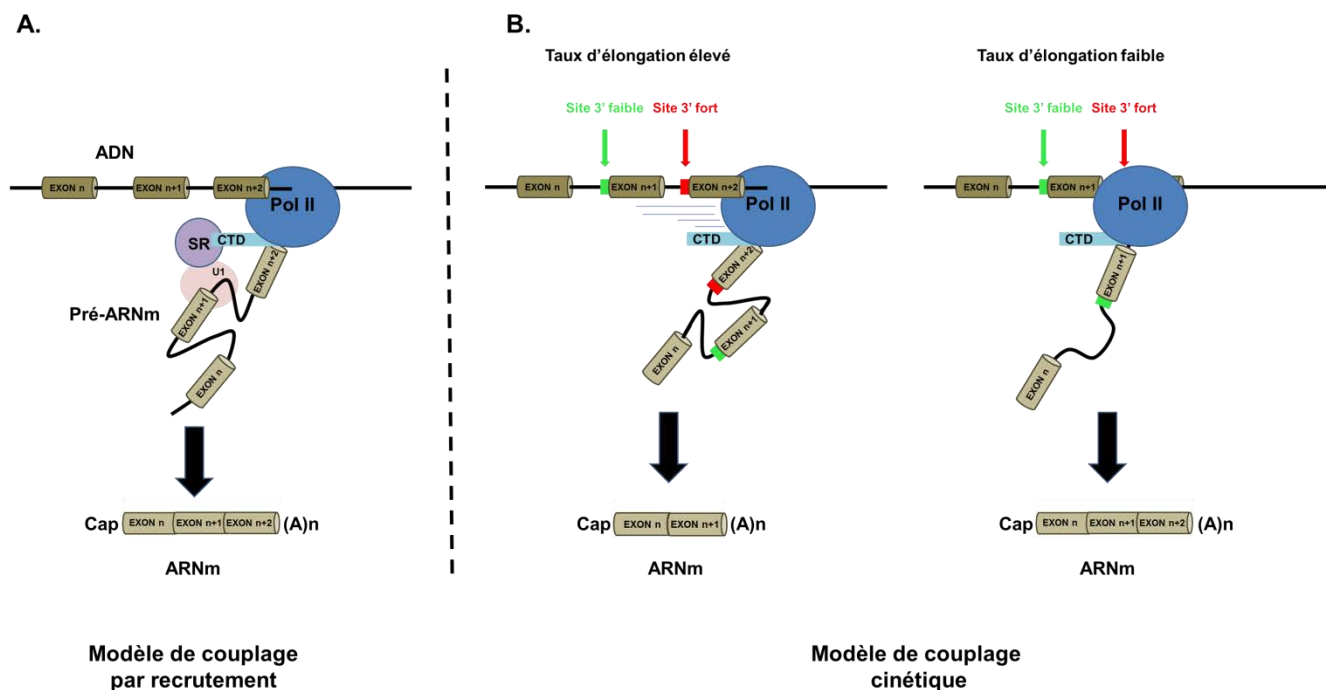


Figure I.11 : Modèles de couplage entre la transcription et l'épissage des pré-ARNm. A. Modèle de couplage par recrutement. Le domaine CTD peut moduler l'épissage d'un exon en favorisant le recrutement de la snRNP U1 par l'intermédiaire d'interactions avec des protéines SR. B. Modèle de couplage cinétique. Une vitesse d'élongation élevée de la polymérase II peut favoriser l'exclusion d'un exon associé à un site 3' faible (vert) précédant un exon avec un site 3' plus fort (rouge). Une faible vitesse d'élongation de la polymérase II permet la reconnaissance du site 3' faible par le spliceosome avant l'utilisation du site 3' fort en aval.

Le modèle de couplage cinétique repose sur la coordination spatio-temporelle de la transcription et de l'épissage ainsi que le fait que la plupart des événements d'épissage se produisent de manière co-transcriptionnelle. Ce modèle de couplage a été clairement mis en évidence chez les métazoaires, où la vitesse d'élongation des transcrits Pol II affecte l'efficacité de l'épissage. En effet, plus la vitesse de synthèse des transcrits Pol II est lente, plus la reconnaissance des jonctions introns/exons est efficace, même en présence de séquences dégénérées. Ce résultat est expliqué par le fait que le processus de transcription par l'ARN Pol II n'est pas cinétiquement linéaire et qu'au cours de l'élongation des nouveaux transcrits, l'enzyme peut réaliser des pauses dans son activité (Kornblihtt et al., 2004). Étant donné que l'ordre des sites d'épissage des transcrit émergents est défini par le sens de transcription de 5' vers 3', un ralentissement de l'activité de la Pol II permet une plus grande fenêtre de temps pour la reconnaissance de sites d'épissage de faible affinité en amont d'un exon au dépend des sites en compétition en aval (Braberg et al., 2013; de la Mata et al., 2003; Moehle et al., 2014). À l'inverse, en absence de pause ou de ralentissement de la Pol II, le spliceosome n'a pas le temps de reconnaître efficacement des sites d'épissage de faible affinité et l'utilisation des sites de plus forte affinité situés en aval est favorisée. De la même manière une cinétique de transcription lente facilite la fixation de facteurs d'activation de faible affinité requis pour l'épissage de l'intron (Braunschweig et al., 2013b; Ip et al., 2011). Il en est d'ailleurs de même pour la fixation de facteurs inhibant l'épissage, ainsi un ralentissement de la vitesse d'élongation de l'ARN polymérase II peut favoriser la fixation de la protéine ETR-3 au niveau du site 3' d'épissage de l'exon 9 du pré-ARNm CFTR entraînant l'exclusion de l'exon par encombrement stérique limitant la reconnaissance de la protéine U2AF65 (Dujardin et al., 2014). Des études sur la drosophile et la souris ont montré que le couplage cinétique entre la transcription et l'épissage dépend de la longueur de l'intron et de sa position par rapport à l'extrémité 5' de l'ARN. Ces études montrent également que la polyadénylation et le clivage en 3' des pré-ARNm sont réalisés, la plupart du temps, avant que tous les introns soient épissés, les transcrits restant associés à la chromatine jusqu'à que l'épissage soit achevé (Bhatt et al., 2012; Khodor et al., 2012). L'état de la chromatine qui influence la cinétique de transcription, a, de ce fait, un effet indirect sur l'épissage, mais peut également l'influencer directement comme nous allons le voir dans le prochain paragraphe.

I.5.2. Liens entre épissage et état de la chromatine au travers des modifications épigénétiques.

I.5.2.1. Influence de l'état de la chromatine sur la transcription.

Dans le paragraphe précédent, nous avons vu que le réseau d'interactions contractées par le domaine CTD et la vitesse d'élongation de l'ARN Pol II influencent l'épissage. Cependant, la processivité de l'ARN Pol II dépend de nombreux facteurs dont l'état de la chromatine. En effet, au cours de la transcription, la chromatine subit un cycle dynamique d'acétylation et de désacétylation des histones et ce cycle de modifications des histones est nécessaire pour faciliter l'accès de l'ARN Pol II au génome. La vitesse d'élongation de l'ARN polymérase dépend alors de l'addition et de l'élimination des groupements acétyles au niveau des histones qui sont respectivement ajoutés et éliminés par les enzymes de la famille HAT (pour Histone Acétyltransférases) et de la famille HDAC (pour Histone Déacétylase) (pour revue Li et al., 2007; Shahbazian and Grunstein, 2007). Une inhibition de l'activité de la protéine HDAC1 conduit à une augmentation de l'acétylation de l'histone H4 entraînant la déstabilisation de l'interaction entre l'ADN et l'histone, et favorisant l'élongation du nouveau transcrit par l'ARN Pol II. Dans ce cas, l'augmentation de la processivité de l'ARN polymérase par inhibition de la protéine HDAC1 réduit l'association co-transcriptionnelle de facteurs de régulation de l'épissage entraînant une modulation de l'épissage alternatif (Hnilicová et al., 2011; Zhou et al., 2014).

Il existe d'autres modifications des histones telles que les méthylations H3K9me3 et H3K9me2 qui peuvent influencer l'épissage alternatif en modulant le taux d'élongation de l'ARN polymérase II. Il a été mis en évidence, dans le cas de l'épissage alternatif du pré-ARNm codant la protéine CD44, que la région chromosomique englobant les exons variables de CD44 est fortement enrichie en H3K9me3 et que cette marque épigénétique est en corrélation avec un fort taux d'inclusion des exons concernés dans l'ARNm. Ce résultat est expliqué par la reconnaissance des marquages H3K9me3 par la protéine CBX3 qui interagit par la suite avec un complexe de remodelage de la chromatine, le complexe SWI/SNF, qui facilite l'inclusion des exons en réduisant localement la vitesse d'élongation lors de la transcription (Batsché et al., 2006; Saint-André et al., 2011).

De récentes études mettent en évidence un rôle essentiel des marquages épigénétiques dans l'épissage alternatif aussi bien par son influence sur le taux de transcription que par son influence sur le recrutement de facteurs d'épissage assurant la reconnaissance des exons.

I.5.2.2. Influence de l'état de la chromatine sur la reconnaissance des exons.

Chez les mammifères deux classes d'architecture de la chromatine ont été décrites. La première classe est caractérisée par des exons présentant une proportion de résidus G/C plus importante par rapport aux introns bordant et une seconde classe où la proportion de résidus G/ C de l'exon est similaire à celle des introns bordants. Ces deux classes se distinguent également par des taux différents d'occupation des nucléosomes, au niveau des exons et des introns. La formation de ces nucléosomes est favorisée par des hautes teneurs en G/C. De plus, les marquages épigénétiques portés par les exons et les introns entre les deux classes sont différents. Dans les deux cas, les exons présentent un signal beaucoup plus fort de méthylation de l'ADN par rapport à celui détecté pour les introns (Gelfman et al., 2013; Lev Maor et al., 2015). Ainsi, des données récentes ont permis de montrer qu'une variation du taux de G/C au niveau des exons/introns peut entraîner une modulation du niveau d'inclusion d'un exon. Ces données suggèrent que la méthylation de l'ADN, ainsi que les taux de G/C exonique et intronique, affectent la reconnaissance des exons lors de l'épissage (Amit et al., 2012).

I.5.2.3. Influence de l'état de la chromatine sur le recrutement de facteurs d'épissage.

De nombreuses expériences ont mis en évidence un effet direct de l'état de la chromatine sur le recrutement des facteurs d'épissage. Ainsi, des expériences d'immunoprécipitation de la chromatine (ChIP) ont permis de montrer, chez l'homme, le recrutement de nombreux facteurs d'épissage au niveau des sites d'épissage lors de la transcription (Listerman et al., 2006; Sapro et al., 2009). Ce recrutement est associé à des modifications épigénétiques de la chromatine, telles que l'acétylation et la méthylation H3K36me3 ou encore H3K9me3. La méthylation H3K9me3 permet le recrutement des protéines CBX1, CBX3 et CBX5. Ces protéines appartiennent à la famille protéique HP1 dont les membres peuvent s'associer à différents facteurs d'épissage. Ainsi, dans des cellules S3,

la protéine CBX3 est retrouvée en complexe avec de nombreuses protéines SR et hnRNP, ainsi qu'avec les protéines argonautes (AGO), qui sont impliquées dans l'inhibition de l'expression des gènes (Ameyar-Zazoua et al., 2012). Il a été proposé que les modifications H3K9me3 de la chromatine au niveau d'exons permettent le recrutement des protéines HP1 qui favorisent ensuite le recrutement de différents facteurs d'épissage et qu'une modulation de cette méthylation entraîne une altération globale de l'épissage alternatif associée aux protéines HP1 (Bannister et al., 2001; Salton et al., 2014; Yearim et al., 2015). De la même manière, 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). De façon plus générale, les modifications épigénétiques permettent le recrutement de protéines adaptatrices, comme les protéines HP1 ou Psip1, qui vont ensuite favoriser le recrutement de différentes protéines dont les facteurs d'épissage tels que la snRNP U2, U2AF65 ou les protéines SR et hnRNP (pour revue Luco et al., 2011 ; Lev Maor et al., 2015).

I.5.3. Couplage de l'épissage avec le transport et la stabilité des ARNm.

Le transport et la stabilité des ARNm impliquent un contrôle qualité des nouveaux transcrits par un complexe appelé EJC (pour "Exon Junction Complex"). Le complexe EJC est initialement recruté au niveau des introns avant la deuxième étape enzymatique d'épissage, puis suite à un changement conformationnel, il se fixe sur les ARNm 20 à 24 nucléotides en amont de chaque jonction exon-exon (Ideue et al., 2007; Le Hir et al., 2000). Ce complexe est constitué de quatre protéines "cœur", conservées au cours de l'évolution, les protéines MAGO, Y14, eIF4AIII et BTZ, qui forment un complexe protéique et qui ne sont dissociées de l'ARNm qu'au moment du passage du ribosome lors des premiers cycles de la traduction (Dostie and Dreyfuss, 2002; Isken et al., 2008; pour revue Bono and Gehring, 2011). Ce complexe protéique constitue une plateforme qui permet le recrutement de nombreuses protéines, dont des protéines à domaine RS qui, une fois associées au complexe EJC, forment un macrocomplexe favorisant la structuration et le compactage de l'ARNm pour son transport. Elles protègent également certaines séquences de l'ARNm contre la dégradation par des nucléases (Singh et al., 2012). Cependant, le complexe EJC est surtout associé au système de surveillance des ARNm, le système NMD (pour "Nonsense Mediated mRNA Decay"). Ce système est constitué des protéines cœur Upf1, Upf2, Upf3 ainsi que Upf3X

(pour "Upstream Frameshift") et de protéines de régulation, les protéines SMG (SMG1, 5, 6 et 7), qui contrôlent les cycles de phosphorylation/déphosphorylation de la protéine Upf1. Le NMD a pour fonction la détection et la dégradation des ARNm comportant un codon stop prématuré, de façon à empêcher la synthèse de protéines tronquées, potentiellement néfastes pour la cellule. Chez les vertébrés, le NMD repose principalement sur le fait qu'un « vrai » codon stop est toujours situé dans le dernier exon ou à proximité directe de la dernière jonction exonique de l'ARN pré-messager. Ainsi ce mécanisme ne fonctionne pas pour les gènes qui ne contiennent pas d'introns (Histone H4, Hsp70, récepteur à la mélanocortine 4...) (Zhang et al., 1998 MCB; Maquat, 2004 NMCB). Si le codon stop est bien situé sur le dernier exon, il n'y a pas de jonction exonique en aval de ce codon stop et il ne reste normalement pas d'EJC sur l'ARNm lorsque le ribosome l'atteint. Si un codon stop est situé à plus de 50-55 nucléotides en amont d'un complexe EJC, alors les facteurs de traduction eRF1 et eRF3 interagissent avec le ou les complexes EJC encore présents sur l'ARNm et les protéines Upf du NMD. Ceci conduit à la terminaison de la traduction et au relargage de l'ARNm qui est ensuite dégradé par les systèmes classiques de dégradation des ARNm impliquant l'élimination de la coiffe, la déadénylation et l'action d'exoribonucléases (pour revue Braunschweig et al., 2013). Ce système permet donc d'éviter la production de protéines tronquées suite à des mutations génétiques, ainsi que d'éliminer des transcrits aberrants non épissés ou partiellement épissés. De plus, le système NMD couplé à l'épissage alternatif permet à la cellule de réguler le niveau d'expression des gènes (Ge and Porse, 2014). Ainsi, il a été estimé qu'environ 10 à 20 % des événements d'épissage alternatif ont le potentiel d'introduire des codons stop prématurés, ce qui permet de réguler le taux d'expression de certains ARNm, notamment en favorisant la rétention d'introns contenant un codons stop prématuré (Lareau et al., 2007; Pan et al., 2006; pour revue Pan et al., 2015).

I.6. EPISSAGE ALTERNATIF ET PATHOLOGIES.

On estime aujourd'hui que 20 à 30 % des maladies génétiques, dont certains cancers, sont associés à des mutations affectant l'épissage alternatif. Pendant de nombreuses années seules les mutations modifiant la séquence codante ont été prises en compte comme cause potentielle de pathologies. Cependant, il est clairement établi à l'heure actuelle que de nombreuses mutations silencieuses, localisées au niveau des introns, impactent l'épissage

alternatif des pré-ARNm (Raponi and Baralle, 2010). Les mutations à l'origine de pathologies génétiques liées à l'épissage peuvent être soit directement portées par le pré-ARNm dont l'épissage est altéré (mutations *en cis*), soit affecter des acteurs, ARN ou protéines, intervenant *en trans* dans le processus d'épissage ou dans la maturation du spliceosome (Gerstberger et al., 2014; Singh and Cooper, 2012b; Tazi et al., 2009).

I.6.1. Pathologies associées à des mutations *en cis* des pré-ARNm.

Les mutations dites « *en cis* » peuvent affecter, soit les séquences des sites 5' ou 3' d'épissage soit des séquences impliquées dans les régulations de l'épissage. Ces mutations sont pour la plupart ponctuelles et peuvent entraîner la formation d'isoformes d'ARNm aberrantes conduisant à la formation de protéines non fonctionnelles, ou avec une activité modifiée du fait de la perte ou de l'ajout d'un exon, ce qui peut altérer la structure et/ou l'activité de la protéine produite. Ces mutations peuvent également affecter la séquence des régions 5'- ou 3'-UTR de l'ARNm induisant des modifications de l'efficacité de traduction, de la stabilité ou de la localisation de l'ARNm voir même être à l'origine de la formation d'aggrégats nucléaires ou cytoplasmiques. La rétention d'introns, comme nous venons de le voir dans le paragraphe précédent, est souvent associé à l'introduction d'un codon stop prématuré et à la prise en charge de l'ARNm par le système NMD, d'où une diminution du taux global d'ARNm et de la protéine associée dans la cellule (Ge and Porse, 2014). A titre d'exemple quelques-unes des pathologies associées à des mutations *en cis* sont listées dans le Tableau I.5. Les mutations *en cis* peuvent être subdivisées en plusieurs catégories, selon qu'elles affectent la fixation des complexes spliceosomaux, la fixation des facteurs de régulations ou encore qu'elles entraînent la formation de nouveaux sites d'épissage au sein du pré-ARNm. Les mutations affectant directement les sites 5' ou 3' d'épissage, conduisent généralement à l'abolition de la reconnaissance de ces sites par le spliceosome et à l'exclusion des exons concernés par ces mutations (phénomène d'exon skipping). Il arrive cependant que la présence d'une mutation dans un site d'épissage favorise sa reconnaissance et augmente ainsi le taux d'inclusion d'un exon. De façon moins fréquente, la mutation d'un site d'épissage peut conduire à un phénomène de rétention d'intron.

Pathologie	Mutation en Cis (gène impliqué)	Fonction de la protéine sauvage
Anémie de Fanconi	<i>FANCG</i>	Réparation de l'ADN
Cancer du sein et de l'ovaire	<i>BRCA1</i>	Prolifération cellulaire (composant de l'ARN Pol II)
Encéphalomyélopathie de Leigh	<i>Pyruvate déshydrogénase $\epsilon 1\alpha$</i>	Métabolisme du pyruvate
Fibrose kystique	<i>CFTR</i>	Canal chlore
Hyperammoniémie	<i>OCT</i>	Cycle de l'urée
Immunodéficience Combinée sévère	<i>Adénosine désaminase</i>	Métabolisme de l'adénosine
Maladie de Sandhoff	<i>Hexoaminidase</i>	Métabolisme des gangliosides
Neurofibromatose De type 1	<i>NF1</i>	Inhibiteur de la voie Ras
Porphyrie intermittente aiguë	<i>PBGD</i>	Synthèse de l'hème
Porphyrie cutanée	<i>UPC</i>	Métabolisme des porphyrines
Syndrome de Ehlers-Danlos du collagène	<i>Lysyl-hydroxylase</i>	Formation d'hydroxylysines
Syndrome de jaeken	<i>PMM2</i>	Synthèse de GDP-mannose
Syndrome Lesh-Nyhan	<i>HPRT</i>	Métabolisme de l'ADN
Syndrome de Marfan	<i>Fibrilline-1</i>	Constituant des microfibrilles
Tyrosinémie de type 1	<i>FAH</i>	Métabolisme de la tyrosine

Tableau I.5 : Exemples de maladies associées à la présence de mutations affectant en *cis* l'épissage des pré-ARNm les portants. (d'après Corcos and Solier, 2005).

Les mutations localisées dans les séquences cis-régulatrices, exoniques ou introniques, peuvent selon leur localisation au sein d'une séquence inhibitrice ou activatrice, favoriser ou inhiber l'inclusion d'un exon (Figure I.12). Enfin, certaines mutations peuvent conduire à la création *de novo* d'un site alternatif d'épissage ou augmenter l'utilisation d'un site sous optimal que l'on qualifie de « cryptique » (Figure I.12). Enfin, il existe également des pathologies génétiques qui ne sont pas associées à des mutations ponctuelles, mais à des éléments de séquences répétés (Tazi et al., 2009). C'est par exemple le cas de répétitions UG qui sont présentes en amont de l'exon 9 du pré-ARNm associé au gène *CFTR* et qui permettent le recrutement d'une protéine de régulation de l'épissage, le facteur TDP-43. Il a été montré que des variations du nombre de répétitions UG sont à l'origine d'un épissage aberrant de l'exon 9, conduisant à des fibroses kystiques (Buratti et al., 2001; Cuppens et al., 1998).

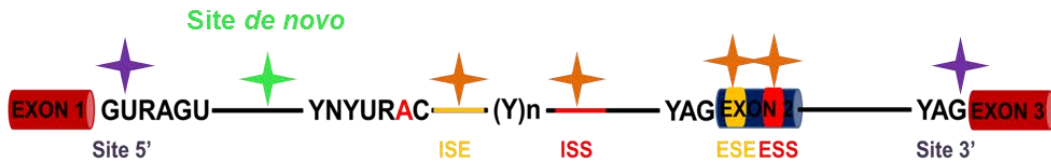


Figure I.12 : Représentation schématique des mutations ponctuelles à l'origine de maladies génétiques affectant l'épissage en cis. Les étoiles indiquent les différents types de mutations possibles, mutations de sites d'épissage préexistants (violet), mutations de séquences de régulations de l'épissage (orange). L'étoile verte symbolise la possibilité qu'une mutation entraîne la formation d'un nouveau site.

I.6.2. Pathologies associées à des mutations affectant *en trans* l'épissage des pré-ARNm.

Il existe plusieurs pathologies d'origine génétique dans lesquelles une mutation dite « *en trans* » affecte des éléments constitutifs du spliceosome ou des facteurs de régulation de l'épissage entraînant une perturbation globale de l'épissage dans la cellule (Tableau I.6). Dans le cas des mutations affectant les composants du spliceosome, les mutations sont le plus souvent situées au niveau de la région codante des protéines et peuvent avoir des répercussions catastrophiques voire létales chez les patients, en fonction du rôle de la protéine impliquée. Néanmoins, certaines pathologies sont associées à des mutations des composants ARN du spliceosome, les snRNA. Ainsi, le nanisme microcéphalique ostéodysplasique primordial de type 1 (MOPD1) ou syndrome de Taybi-Linder (TALS) a pour origine une mutation dans le gène codant le snRNA U4atac qui est transcrit en antisens de l'intron 2 du gène CLASP1. Huit mutations différentes du snRNA U4atac ont été identifiées chez les patients MOPD1 (Edery et al., 2011; He et al., 2011a). Ces mutations sont, pour la plupart, présentes au sein d'une structure tige-boucle conservée au sein des snRNA U4atac et U4. Cette structure tige-boucle correspond au site de fixation de la protéine 15,5K et la présence de ces mutations est prédite pour empêcher l'association des protéines 15,5K et Prp31 (61K) spécifiques de la tri-snRNP U4atac/U6atac.U5 et nécessaires à l'assemblage et la stabilité de cette tri-snRNP (Schultz et al., 2006; pour revue Padgett, 2012; Turunen, 2013). Ainsi, il a été observé chez les patients MOPD1 une diminution de l'activité du spliceosome mineur qui est à l'origine des différents symptômes de la pathologie (He et al., 2011b; Singh and Cooper, 2012a).

D'autres pathologies ont pour origine des mutations qui ne touchent pas les composants essentiels de la machinerie d'épissage, mais affectent des facteurs de régulation de l'épissage tant au point de vue de leur intégrité, en terme de domaines fonctionnels, que de leur niveau d'expression ou de leur distribution cellulaire. C'est notamment le cas de la dystrophie myotonique de type I qui a fait l'objet de la première partie de mes travaux de thèse et dont je parlerai plus en détails dans le prochaine chapitre.

Pathologie	Protéine impliquée dans l'épissage
Dystrophie myotonique	CUGBP1 et 2
Leucémie, sarcomes	FUSIP1
Alzheimer	GSK-3 β
Alzheimer	HMGA1 α
Dystrophie myotonique	MBNL1,-2,-3
Syndrome paranéoplasique	NOVA1
Rétinite pigmentaire	PrPF3,-8,-31
Azoospermie	RBM11
Carcinome hépatocellulaire	HCC1
Carcinome papillaire rénal	SFPQ
Atrophie musculaire spinal	SMN1 et 2
Syndrome de Hay-Wells	PMM2

Tableau I.6 : Exemples de pathologies associées à la présence de mutations affectant en trans l'épissage des pré-ARNm (d'après Corcos and Solier, 2005). Le nom de la pathologie ainsi que celui du facteur d'épissage dont l'activité est affectée dans la pathologie sont indiqués.

CHAPITRE II : LA DYSTROPHIE MYOTONIQUE DE TYPE I.

II.1. LA DYSTROPHIE MYOTONIQUE DE TYPE I.

La dystrophie myotonique de type 1 (DM1) ou maladie de Steinert est la dystrophie musculaire la plus fréquente chez l'adulte avec une prévalence estimée à 1 pour 10000. C'est une maladie génétique à évolution progressive et multisystémique qui est caractérisée principalement par une myotonie et des faiblesses musculaires issues d'une dégénérescence des muscles squelettiques. De plus, la DM1 est une maladie héréditaire et la transmission de la maladie dans les générations suivantes se manifeste par une aggravation toujours plus importante des symptômes. Cependant, la DM1 se distingue des autres maladies et myotonies, telles que la DM2, par certaines caractéristiques électrophysiologiques et la manifestation de symptômes non musculaires comme des cataractes, une résistance à l'insuline et une arythmie cardiaque.

II.1.1. Les aspects génétiques à l'origine de la maladie.

La DM1 est causée par une expansion instable de répétitions CTG dans la région 3'UTR du gène *DMPK*. Ces répétitions existent naturellement chez les individus sains et le nombre de répétitions est variable selon les individus avec une gamme allant de 4 à 37 répétitions. Les patients présentant 38 à 49 répétitions ne présentent pas de symptômes, mais sont à l'état de pré-mutation et risquent de transmettre un nombre de répétitions potentiellement pathologique à leur descendance. C'est à partir de 50 répétitions que les patients sont cliniquement reconnus comme atteints de DM1. A ce stade, la pathologie se développe tardivement et les patients présentent peu de symptômes. La forme principale de la maladie, majoritairement retrouvée, est caractérisée par la présence de 100 à 1000 répétitions avec l'apparition des symptômes entre 10 et 30 ans chez les patients. Il existe cependant une forme plus sévère de la maladie où les patients possèdent plusieurs milliers de répétitions. Cette forme sévère est congénitale et les symptômes apparaissent dès la naissance et provoquent une mortalité infantile élevée. L'âge d'apparition de la pathologie,

ainsi que sa sévérité, sont donc directement corrélés à la longueur des expansions de répétitions CTG (Brook et al., 1992; Kumar et al., 2013) (Tableau 7). De plus, le nombre de répétitions augmente chez les patients atteints de DM1 au cours de leur vie et conduit à la transmission d'expansions plus grandes que celles héritées à la naissance, on parle alors de transmission par anticipation. De ce fait les individus affectés dans les générations successives présentent des symptômes toujours plus sévères à un âge toujours plus précoce jusqu'à générer une forme congénitale (Höweler et al., 1989; Pratte et al., 2015). De manière générale, il a été proposé que les expansions instables de faibles tailles sont transmises par le père, tandis que les expansions associées aux formes congénitales sont transmises par la mère (Brunner et al., 1993). Cependant, une étude récente a mis en évidence que le nombre de répétitions CTG transmis à la descendance est plus important quand la pathologie est transmise par le père. Le mécanisme majeur pouvant expliquer l'expansion des répétitions CTG serait un défaut au moment de la réplication et de la réparation de l'ADN et concerne les cellules somatiques, mais également les cellules germinales au moment de la gamétogenèse expliquant ainsi le phénomène d'anticipation (De Temmerman et al., 2004; Liu et al., 2012; Reddy et al., 2014).

II.1.2. Les aspects cliniques associés à la maladie.

Comme mentionné précédemment, il existe plusieurs phénotypes de la DM1 qui se distinguent par l'âge d'apparition de la maladie, du nombre de répétitions et de la sévérité des symptômes. Quatre phénotypes ont ainsi être proposés : la forme asymptomatique (prémutation), la forme "moyenne", la forme "classique" et la forme congénitale. Les principaux signes cliniques associés à ces différents phénotypes sont répertoriés dans le Tableau II.1, ainsi que l'âge d'apparition de la pathologie et la durée de vie moyenne des patients (Kumar et al., 2013). De plus, la pathologie peut présenter des caractéristiques électrophysiologiques, des symptômes non musculaires et des troubles neurologiques qui caractérisent le phénotype multisystémique de la pathologie (Tableau II.2) (Udd and Krahe, 2012).

Phénotype	Signes cliniques	Nbr répétitions CTG	Age d'apparition	Durée de vie (moyenne)
Asymptomatique (pré-mutation)	aucun	35 à 49		
Moyen	Myotonie moyenne Cataractes	50 à ± 150	20 à 70 ans	60 ans avec une vie normale
Classique	Myotonie Faiblesses musculaires Cataractes Arythmie cardiaque Calvitie	± 100 à ± 1000	10 à 30 ans	48 à 55 ans
Congénital	Hypotonie infantile Détresse respiratoire Retard intellectuel	± 100 à ± 1000	Depuis la naissance	48 à 55 ans

Tableau II.1 : Tableau récapitulatif des caractéristiques des différents phénotypes de la DM1. Pour chaque phénotype sont présentés les principaux signes cliniques, le nombre de répétitions CTG, l'âge d'apparition des symptômes et la durée de vie moyenne des patients (d'après Kumar et al., 2013).

Système/Organe	Atteinte
Endocrinien	- Résistance à l'insuline - Atrophie testiculaire - Taux élevés de FSH
Yeux	- Cataracte - Ptose - Dégénérescence de la rétine
Cœur	- Défaut de conduction - Arythmie cardiaque
Respiratoire	- Faiblesse et myotonie du diaphragme et des muscles respiratoires - Anomalies de contraction du pharynx et de l'œsophage - Hypoventilation alvéolaire - Aspiration branchiale
Cerveau	- Retard mental dans les formes congénitales - Hypersomnie - Déclin cognitif variable - Troubles de la personnalité : apathie élevée, évitement et comportement d'agressivité passive
Peau	- Calvitie prématurée
Muscle lisse	- Constipation et diarrhée - Douleurs abdominales de type colique - Syndrome pseudo-occlusif - Contractions non-coordonnées de l'utérus durant l'accouchement
Muscle squelettique	- Myotonie - Faiblesse et perte musculaires - Faiblesse des muscles distaux et tardivement des muscles proximaux - Faiblesse des muscles de la face, de la mâchoire, de la langue et des muscles fléchisseurs du cou - Atrophie du sternocléidomastoidien

Tableau II.2 : Principaux symptômes des formes adultes de la dystrophie myotonique de type 1.

II.2. CONSEQUENCES DE LA PRESENCE DES REPETITIONS CTG DANS LA DM1.

II.2.1. Impact des répétitions CTG sur le gène *DMPK* et sur l'expression des gènes voisins.

Les études menées lors de la dernière décennie ont mis en évidence que les répétitions CTG forment au niveau de l'ADN une structure particulière en forme de tige boucle stable qui est à l'origine des défauts de réplication des répétitions et de leur expansion (Reddy et al., 2014). De plus, la formation de ces tiges boucles sur l'ADN peut modifier la structure de la chromatine et affecter localement l'expression des gènes voisins, notamment du gène *SIX5*. En effet, il a été mis en évidence que les répétitions CTG chez les patients DM1 peuvent entraîner une diminution de l'expression du gène *SIX5* en empêchant la fixation d'activateurs transcriptionnels (Inukai et al., 2000; Thornton et al., 1997). Ainsi, il est suggéré que la diminution d'expression du gène *SIX5* participe à l'établissement de la DM1 et favorise la formation de cataracte chez les patients DM1, mais le déficit seul de l'expression du gène *SIX5* n'explique pas l'ensemble des troubles observés dans la maladie (Klesert et al., 2000; Sarkar et al., 2000). En effet, après transcription, les répétitions CTG sont retrouvées au niveau de la région 3'UTR de l'ARNm sous forme d'expansions de triplets CUG, dont la présence a de nombreuses conséquences sur le métabolisme cellulaire et conduit au caractère multisystémique de la DM1.

II.2.2. Impact des répétitions CUG sur l'ARNm *DMPK* et sur la DM1.

Les répétitions CUG forment au niveau des transcrits de longues structures tiges-boucles très stables qui comportent de longues hélices β de type A comme celles formées par l'ADN (Figure II.I) (Kiliszek and Rypniewski, 2014; Kiliszek et al., 2009; Sobczak et al., 2003). La formation de ces structures sur l'ARNm entraîne d'une part une diminution de l'expression de la protéine *DMPK*, et d'autre part génère un gain de fonction de l'ARNm. En effet, la structuration des répétitions CUG en tige-boucle confère à l'ARNm de nouvelles propriétés, dont la capacité d'interagir avec de nouvelles protéines et de recruter la protéine *MBNL1* qui est un facteur d'épissage.

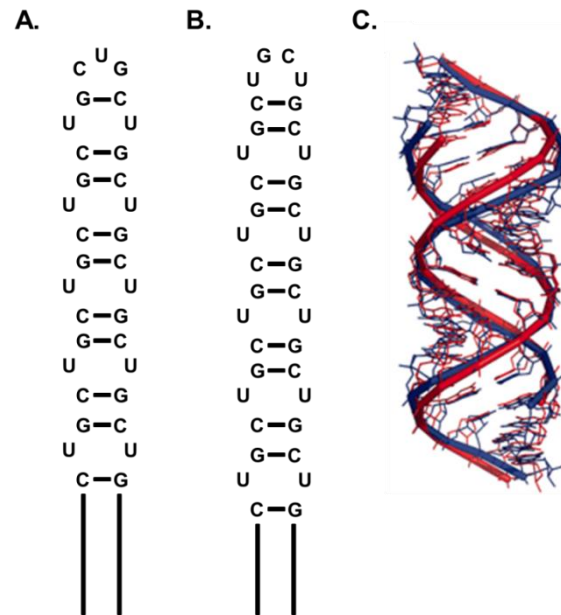


Figure II.1 : Structure des régions ARN présentant des répétitions CUG. A. Boucle terminale à 3 nucléotides possible dans les structures tige-boucles formées par un nombre impair de répétitions CUG. B. Boucle terminale à 4 nucléotides possible dans les structures tige-boucles formées par un nombre pair de répétitions CUG. C. Superposition de la structure établie pour r(CUG)₆ (bleu) et d'une hélice de type A (rouge) (d'après Mooers et al., 2005).

Les ARNm DMPK mutants associés à la protéine MBNL1 sont retenus dans le noyau des cellules et s'agrègent sous forme de *foci* nucléaires entraînant une réduction de 50 à 57% du taux de protéine DMPK dans les myoblastes DM1 (Furling et al., 2001a, 2001b). Les fonctions exactes de la protéine DMPK sont encore mal connues, on sait néanmoins, qu'il s'agit d'une kinase impliquée dans la régulation de l'excitabilité des cellules musculaires. Ainsi, une diminution de l'expression de la protéine DMPK pourrait favoriser les défauts de contractilités des muscles et contribuer ainsi à la pathologie (Mankodi et al., 2000). Cependant, c'est la séquestration de la protéine MBNL1 dans les *foci* qui est à l'origine du phénotype multisystémique de la maladie. En effet, la formation des *foci* conduit non seulement à une perte de fonction de la protéine MBNL1, mais s'accompagne également d'une augmentation de la stabilité de la protéine CUG-BP1, une protéine de régulation de l'épissage de la famille des protéines CELF qui est un antagoniste de MBNL1 (Fardaei et al., 2001; Mankodi et al., 2001; Michalowski et al., 1999; Savkur et al., 2001). La perte de fonction de la protéine MBNL1 associée au gain de fonction de la protéine CUG-BP1 entraîne une modulation de l'épissage alternatif de nombreux ARNm qui codent des protéines qui sont en corrélation directe avec chacun des symptômes de la pathologie hormis le

développement de cataracte (Tableau II.3). En effet, Au niveau du cerveau, les désordres neurodégénératifs observés chez les malades sont associés à la modulation de l'épissage du pré-ARNm *MAPT* qui code la protéine Tau avec une augmentation de l'exclusion des exons 2, 3 et 10 (Buée et al., 2000). De même, une inclusion anormale de l'exon 5 du gène codant la troponine T cardiaque humaine, qui correspond une forme exprimée lors du développement embryonnaire, participe aux troubles cardiaque et musculaires observés chez les patients (Philips et al., 1998). La dérégulation de l'épissage de l'exon 5 associée à la DM1 a fait l'objet d'une partie de mon travail de thèse et sera introduite plus en détails dans le paragraphe **II.5** de l'introduction.

II.3. LA PROTEINE MBNL1 ET SON IMPLICATION DANS LA DM1.

La protéine MBNL1 appartient à la famille de régulateurs de l'épissage alternatif Muscleblind-like (MBNL). Elle participe plus généralement au métabolisme des ARN. L'expression des protéines MBNL est tissu dépendante et ces protéines sont exprimées majoritairement dans les yeux, le système nerveux central et dans la plupart des types musculaires. Cette famille est composée de trois membres, MBNL1, MBNL2 et MBNL3 qui sont codés par trois gènes distincts. Les protéines MBNL1 et 2 sont exprimées de manière ubiquitaire dans la plupart des tissus, mais c'est MBNL1 qui est majoritairement exprimée, principalement dans le cœur et dans les muscles squelettiques (pour revue Konieczny et al., 2014). Il est à noter que les protéines MBNL1 et 2 présentent des activités partiellement redondantes et qu'une perte fonctionnelle de la protéine MBNL1 entraîne la surexpression de la protéine MBNL2 chez la souris (Lee et al., 2013). L'expression des protéines MBNL est régulée au cours du développement. Ainsi, l'expression de la protéine MBNL3 est plus importante chez les fœtus que dans les tissus adultes où elle est relativement restreinte, alors qu'à l'inverse, les protéines MBNL1 et 2 sont de plus en plus exprimées au cours du développement et de la différenciation cellulaire (Kanadia et al., 2003a). De même, la localisation des différentes isoformes des protéines MBNL varie au cours du développement. En effet, les pré-ARNm codant les protéines MBNL sont soumis à un épissage alternatif, ainsi les exons 5 et 7 sont majoritairement inclus dans les ARNm codant MBNL1 dans les muscles et le cerveau fœtaux avec une localisation cellulaire nucléaire.

Système/Organe	Pré-ARNm	Dérégulation exon/intron	Inclusion/exclusion	Référence	
Cœur	Troponine T cardiaque (TNNT2/cTNT)	Exon 5	Inclusion	Philips et al., 1998	
Muscle squelettique	Récepteur à l'insuline (INSR)	Exon 11	Exclusion	Savkur et al., 2001	
	Canal chlore (CLCN1)	Exon 7A/intron 2	Inclusion/inclusion	Charlet et al., 2002 Mankodi et al., 2002	
	Troponine T squelettique (TNNT3)	Exon fœtal	Inclusion	Kanadia et al., 2003	
	Amphyphisine 2 (BIN1)	Exon 11	Exclusion	Fugier et al., 2011	
	Récepteur à la ryanodine (RyR)	Exon 70	Exclusion	Kimura et al., 2005 Lin et al., 2006	
	Sarcoplasmic/endoplasmic Réticulum Ca ²⁺ ATPase 1 (SERCA1)	Exon 22	Exclusion		
	Sarcoplasmic/endoplasmic Réticulum Ca ²⁺ ATPase 2 (SERCA2)	Intron 19	Inclusion		
	LIM domain binding 3 (LB3,ZASP)	Exon 11	Inclusion	Lin et al., 2006	
	Titin (TTN)	Exon Zr4 et 5	Inclusion/ Inclusion		
	Nebulin-related anchoring Protein (NRAP)	Exon 12	Inclusion		
	Calpain 3 (CAPN3)	Exon 16	Exclusion		
	Attractin-like (ATRNL1,ALP)	Exon 5a et b	Inclusion		
	Forming homology 2 domain Contening 1 (FHOD1)	Exon 11a	Exclusion		
	Glutamine-fructose 6-phosphate transaminase 1 (GFPT1)	Exon 10	Exclusion		
	MBNL1	Exon 7	Inclusion		
	MBNL2	Exon 7	Inclusion		
	Facteur nucléaire 1	Exon 123	Inclusion		
	Cerveau	SET and MYND domain Contening (SMYD1)	Exon 39	Inclusion	Du et al., 2010
		Sperm associated antigen 9	Exon 39	Inclusion	
		Myotubularin-related Protein 1 (MTMR1)	Exon 2.1 et 2.3	Exclusion/ Exclusion	
Alpha-dystrobrevin (DTNA)		Exon 11a et 12	Inclusion / Inclusion	Nakamori et al., 2008	
Cerveau	Tau (MAPT)	Exon 2, 3, 6 et 10	Exclusion exon 2, 3 et 6c Inclusion exon 6d	Sergeant et al., 2001 Jiang et al., 2004 Leroy et al., 2004	
	Récepteur-N-methyl_D-Aspartate (NMDAR1)	Exon 7A intron 2	Exclusion	Jiang et al., 2004	
	Amyloid precursor prtein (APP)	Exon fœtal	Inclusion		
	MBNL1	Exon 11	Inclusion	Dhaenens et al., 2010	

Tableau II.3 : Principales anomalies d'épissage détectées dans différents tissus de patients DM1.

Ces exons sont ensuite exclus dans les tissus adultes où la protéine est retrouvée essentiellement dans le cytoplasme. Ces données suggèrent que la régulation fonctionnelle des protéines MBNL provoque une transition dans la régulation de l'épissage alternatif qui participe au développement et à la différenciation cellulaire (Kalsotra et al., 2008; Lin et al., 2006). Chez les patients DM1 une augmentation de l'inclusion des exons 5 et 7 dans les ARNm de MBNL1 et donc un retour vers une forme foetale est observé (Tran et al., 2011). Par ailleurs, chez la souris l'inhibition de l'expression de la protéine MBNL1 suffit à entraîner les mêmes symptômes musculaires et neurologiques que la DM1 sans la formation de *foci* mettant en évidence le rôle essentiel de la protéine MBNL1 dans la pathologie (Kanadia et al., 2003b).

II.3.1. Structure de la protéine MBNL1 et reconnaissance des séquences ARN.

Le gène MBNL1 est composé de 12 exons et la séquence codante est distribuée sur 10 exons (numérotés de 1 à 10) dont 4 sont alternatifs (exon 3, 5, 7 et 9), ce qui permet la production de 9 isoformes de la protéine MBNL1 (Figure II.2). Chacune des isoformes présente au niveau de sa région N-terminale 4 domaines en doigt de zinc (Zn1 à 4) qui se structurent en deux paires en tandem très similaires, appelés TZ12 (couple Zn1/Zn2) et TZ34 (couple Zn3/Zn4). Ces domaines confèrent à la protéine MBNL1 la capacité d'interagir avec l'ARN au niveau de séquence de type YGCY, le résidu G de ce motif étant essentiel à la fixation de MBNL1 (Goers et al., 2010). Des études cristallographiques ont montré que la région reliant les doigts à Zinc au sein du motif TZ34 est fixe, et structurée de façon à ce que les surfaces de fixation à l'ARN de chaque doigt de zinc soient opposées avec une orientation antiparallèle (Figure II.3). Ces données suggèrent ainsi une conformation en boucle de l'ARN lié, les domaines doigt à Zn étant situés entre les séquences complémentaires de l'ARN (Teplova and Patel, 2008). En effet, il a été montré que la protéine MBNL1 interagit avec une plus grande affinité avec des séquences qui présentent plusieurs sites YGCY. Il est proposé que ces sites puissent interagir entre eux et former une structure en tige-boucle reconnue préférentiellement par la protéine MBNL1. La structure ainsi formée serait similaire à celle générée par les répétitions CUG dans la DM1 du fait de la présence de mésappariements au niveau de l'hélice (Goers et al., 2010; Warf and Berglund, 2007).

La protéine MBNL1 peut également interagir avec un seul site YGCY, mais avec une affinité moins importante (Rau et al., 2011a). Il est à noter que la présence de la séquence de liaison qui sépare les deux motifs TZ12 et TZ34 et qui est codée par l'exon alternatif 3 module l'affinité de MBNL1 pour ses pré-ARNm cibles et pour les répétitions CUG *in vitro* (Tran et al., 2011; Carpentier et al., 2014). Cependant cette séquence, bien que nécessaire pour la modulation de l'épissage des ARNm cibles de MBNL1, n'est pas requise pour la séquestration de MBNL1 dans les foci formés par les répétitions CUG *in cellulo* et dans des fibroblastes ou myoblastes DM1 (Tran et al., 2011; Rehman et al., 2014; Carpentier et al., 2014). Le domaine C-terminal de MBNL1 est codé par les exons 5 à 10 et n'intervient pas dans l'interaction de la protéine avec l'ARN. Néanmoins le domaine C-terminal est essentiel pour l'activité de la protéine et sa fonction dans la régulation de l'épissage alternatif. En effet, il a été montré qu'en son absence la protéine perd sa capacité à réguler l'épissage alternatif, mais également à s'oligomériser (Kino et al., 2004; Yuan et al., 2007). En effet, les exons 5 et 7 qui codent une partie du domaine C-terminal sont soumis à un épissage alternatif et influencent respectivement la localisation de la protéine dans la cellule (exon 5), ainsi que sa capacité à s'oligomériser (exon 7). Ces modulations du domaine C-terminal confèrent à la protéine MBNL1 différentes localisations cellulaires et modifient sa fonction (Tran et al., 2011).

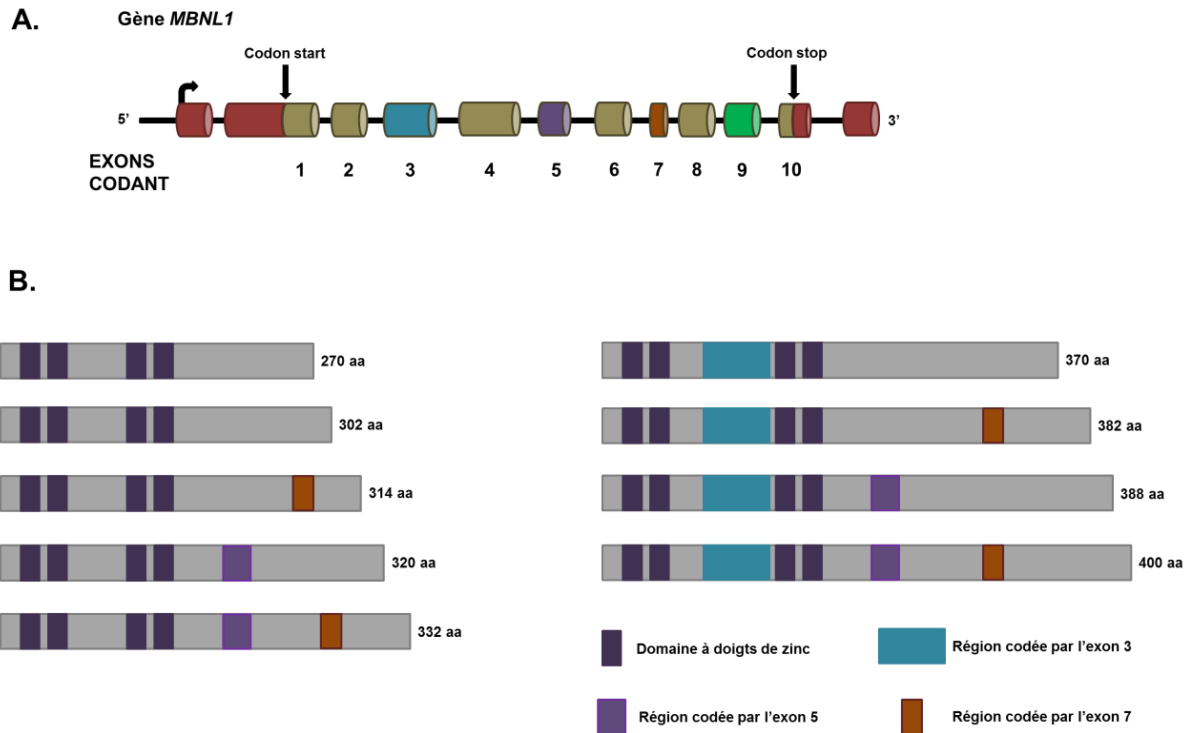


Figure II.2 : Les différentes isoformes protéiques produites à partir du gène *MBNL1*. **A.** Organisation génomique du gène humain *MBNL1*. Les exons non codants et codants sont respectivement indiqués en rouge et en marron. Les exons alternatifs 3, 5, 7 et 9 sont respectivement indiqués en bleu, violet, orange et vert. Les codons start et stop sont indiqués par une flèche noire. **B.** Représentation schématique des différentes isoformes protéiques de *MBNL1*. Les 2 premiers motifs en doigts à Zn^{2+} sont codés par l'exon 2 et les 2 derniers par l'exon 4. Le nombre d'acides aminés (aa) de chaque isoforme est indiqué (d'après Tran et al., 2011).

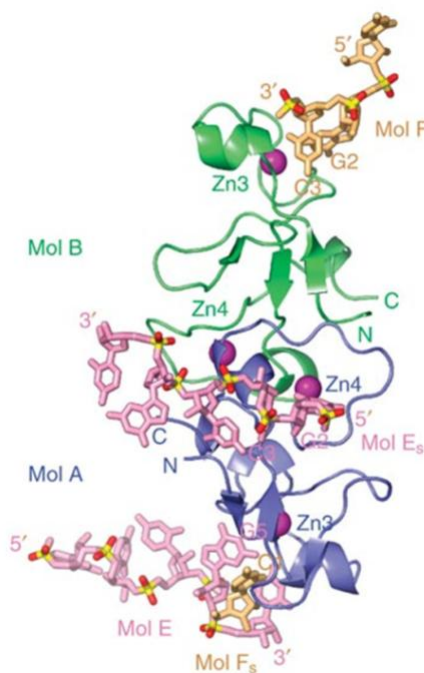


Figure II.3 : Structure cristallographique de la paire à doigts de Zn^{2+} 3 et 4 (TZF34) de *MBNL1* en complexe avec un oligonucléotide r(CGCGUGU). Représentation en ruban et en bâton du complexe contenant 4 molécules d'ARN interagissant avec 2 paires TZF34. Les molécules A (en bleu) et B (en vert), correspondent chacune à un motif TZF34, et interagissent entre elles via le motif ZF4. Les molécules d'ARN E et Es (en rose) et F et Fs (en orange) interagissent avec les molécules de protéines A et B. Chaque domaine ZF3 et ZF4 de *MBNL1* interagit avec un oligonucléotide, l'un au niveau de la séquence GC et l'autre de la séquence GCU, révélant ainsi un motif de reconnaissance GC commun. Les atomes coordonnés de zinc sont représentés par des sphères pourpres. (D'après Teplova et Patel, 2008).

II.3.2. Localisation et fonction de la protéine MBNL1.

II.3.2.1. Localisation de la protéine MBNL1.

La localisation de la protéine MBNL1 est un déterminant majeur pour ses fonctions cellulaires et dépend d'un signal NLS bipartite (biNLS) constitué de deux clusters de résidus KR au niveau des exons 5 et 6 (Tran et al., 2011). En absence de l'inclusion de l'exon 5 alternatif, la protéine MBNL1 perd sa localisation strictement nucléaire pour devenir nucléocytoplasmique, ce qui impacte son rôle de régulateur de l'épissage alternatif. Cependant, une étude récente a mis en évidence un second NLS dit conformationnel (conNLS), dont l'activité ne semble pas être associée à la voie des protéines importines (Figure II.4). En effet, il a été proposé que les domaines KRP(A)LE des exons 5 et 6 puissent chacun agir avec une région conservée localisée au niveau du domaine TZ12. Ainsi, l'inclusion de l'exon 5 pourrait agir comme un commutateur pour multiplier les voies d'import nucléaires, en assurant une localisation efficace de la protéine dans le noyau afin d'assurer certaines fonctions nucléaires comme la régulation de l'épissage alternatif. De plus, l'implication de régions situées au niveau des domaines doigt à zinc 1 et 2 pour former un conNLS, suggère que l'état de liaison de la protéine MBNL1 à l'ARN pourrait moduler l'activité des deux conNLS possibles (Kino et al., 2015; Tran et al., 2011).

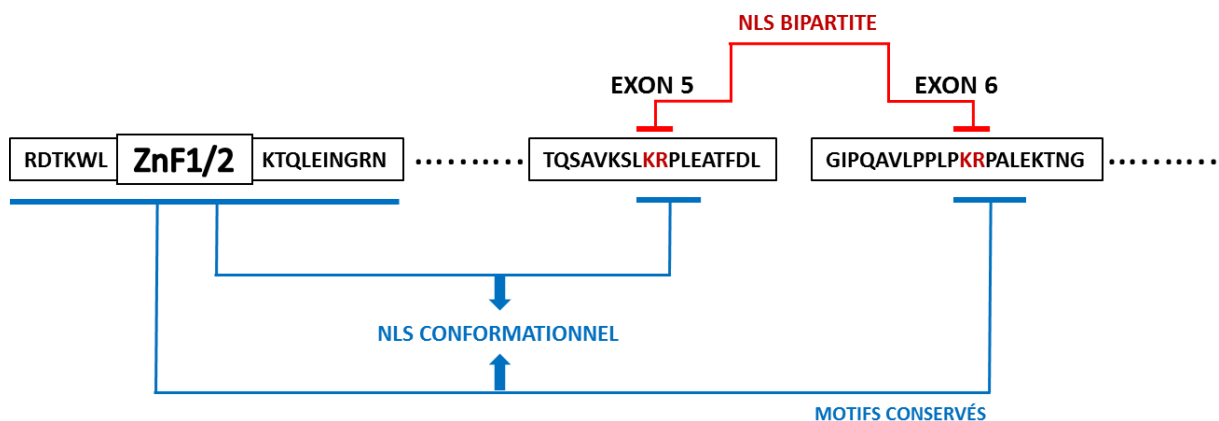


Figure II.4 : Schéma des différents éléments NLS présents au sein de l'ARNm codant la protéine MBNL1. Les séquences impliquées dans le NLS conformationnel sont indiquées en bleues et les séquences impliquées dans le NLS bipartite en rouge, ainsi que les résidus KR.

II.3.2.2. La protéine MBNL1 : un régulateur de l'épissage alternatif.

La protéine MBNL1 a été caractérisée en tant que régulateur de l'épissage alternatif et l'inactivation de la protéine chez la souris conduit à la dérégulation de l'épissage de plus de 200 pré-ARNm, dont le propre pré-ARNm de MBNL1 (Du et al., 2010; Gates et al., 2011; Ho et al., 2004a). Selon les pré-ARNm, la protéine MBNL1 peut soit favoriser, soit inhiber l'inclusion d'un exon et le niveau d'activité de la protéine est en corrélation avec le degré d'affinité de MBNL1 avec le pré-ARNm cible. En effet, différentes études ont mis en évidence que l'activité de régulation de l'épissage de la protéine MBNL1 nécessite la présence de l'exon 6, et de l'exon 3 qui code la région de la protéine située entre les deux couples de doigts à zinc TZ12 et TZ34 impliqués dans la fixation de la protéine sur les pré-ARNm (Edge et al., 2013; Grammatikakis et al., 2011; Purcell et al., 2012; Tran et al., 2011). Il a été montré que la localisation des sites de fixation de MBNL1 par rapport à l'exon peut avoir un impact sur les propriétés activatrices ou inhibitrices de la protéine MBNL1. En effet, il a été proposé que la fixation de MBNL1 sur les séquences introniques en amont de l'exon favorise l'exclusion de l'exon, tandis que la fixation de la protéine en aval de l'exon favorise l'inclusion de l'exon (Goers et al., 2010; Wang et al., 2012; Echeverria and Cooper, 2014). MBNL1 agirait sur l'épissage alternatif en bloquant des sites activateurs ou inhibiteurs de l'épissage par oligomérisation (Kino et al., 2009). De plus, des études ont montré que la protéine MBNL1 peut reconnaître et stabiliser des structures ARN en tige-boucle adjacentes à des sites accepteurs ou inhibiteurs de l'épissage, limitant ainsi leur utilisation lors de l'épissage alternatif. C'est notamment le cas pour l'intron situé en amont de l'exon F du pré-ARNm TNNT3 où MBNL1 reconnaît et stabilise une structure tige-boucle située à proximité du site 3' d'épissage, qui n'est alors plus utilisé, ce qui conduit à l'exclusion de l'exon. Fait intéressant, une étude récente a montré que le domaine C-terminal de MBNL1 est suffisant à lui seul pour interagir avec l'ARN et permettre le recrutement de la protéine PTB afin d'induire la répression de l'inclusion de l'exon 3 du pré-ARNm codant la protéine TPM1 (Yuan et al., 2007; Gooding et al., 2013; pour revue Konieczny et al., 2014). Enfin, une étude récente chez la souris a montré l'implication de la protéine MBNL1 dans la sélection alternative de sites de polyadénylation de centaines d'ARNm dans des fibroblastes (Batra et al., 2014).

II.3.2.3. Implication de la protéine MBNL1 dans la formation d'ARN circulaire.

L'existence d'ARN circulaires libres au cours de l'épissage avait été observée depuis de nombreuses années. Au cours de ces dernières années, de nombreuses études ont mis en évidence le fait que ce type d'ARN circulaires (circRNA) est exprimé de manière abondante chez les eucaryotes. Les mécanismes conduisant à la production de ces ARN sont encore mal compris (Salzman et al., 2012). Ces ARN circulaires sont constitués soit d'un ou de plusieurs exons, ou d'un intron. Les fonctions de ces ARN sont aussi largement incomprises. Cependant, il a été proposé qu'ils puissent agir en tant que régulateurs de la transcription ou comme compétiteurs de leurs homologues linéaires pour l'interaction avec des miRNA ou des RBP (RNA Binding Protein) (Jeck et al., 2013). En effet, certains circRNA ont un rôle « d'éponge » afin de moduler localement la concentration de miRNA et de RBP libres (Hansen et al., 2013; Memczak et al., 2013). De plus, il est proposé que les circRNA puissent être traduits par les ribosomes et puissent générer des protéines avec une fonction différente de la protéine issue de l'ARNm linéaire homologue (Chen and Sarnow, 1995). Récemment, des études ont établi l'implication directe de la protéine MBNL1 dans la formation d'un ARN circulaire (appelé circMBL) lors de l'épissage alternatif de son propre pré-ARNm (Petkovic and Müller, 2015). En effet, les études ont révélé que le nombre de sites YGCY de part et d'autre de l'exon 3 du pré-ARNm influence la production de l'ARN circMBL à partir de cet exon et cette production nécessite la fixation de la protéine MBNL1. Ainsi, la mutation ou la déplétion d'un ou plusieurs sites YCGY défavorise la production de l'ARN circMBL, tandis que la surexpression de la protéine favorise la production de circMBL. Ces résultats sont cohérents avec les derniers modèles proposés pour la formation des circRNA, notamment avec un modèle dans lequel les introns non épissés flanquant un exon, interagissent ensemble soit par appariement de bases, soit par l'intermédiaire de RBP. Cette interaction entre les introns provoquerait la formation d'une structure tige-boucle qui permettrait de juxtaposer le point de branchement de l'intron en aval avec le site 3' accepteur de l'intron en amont (site 3' donneur) qui réaliserait une attaque nucléophile. Ensuite, le site 3' donneur attaquerait le site 5' donneur de l'intron en aval, joignant ainsi les deux introns et libérant l'exon circularisé (Figure II.5A). Un second modèle basé sur le même schéma enzymatique, mais au sein d'un lasso classique au cours de l'épissage a également été proposé (Figure II.5B) (Petkovic and Müller, 2015). Le fait que la protéine MBNL1 puisse

interagir avec les ARN circMBL soulève la possibilité d'un mécanisme sophistiqué de régulation du niveau de MBNL1 libre. En effet, lorsque la protéine MBNL1 est en excès, la production de circMBL est favorisée et pourrait jouer un rôle d'éponge vis à vis de la protéine MBNL1 en excès en se liant à elle (Ashwal-Fluss et al., 2014).

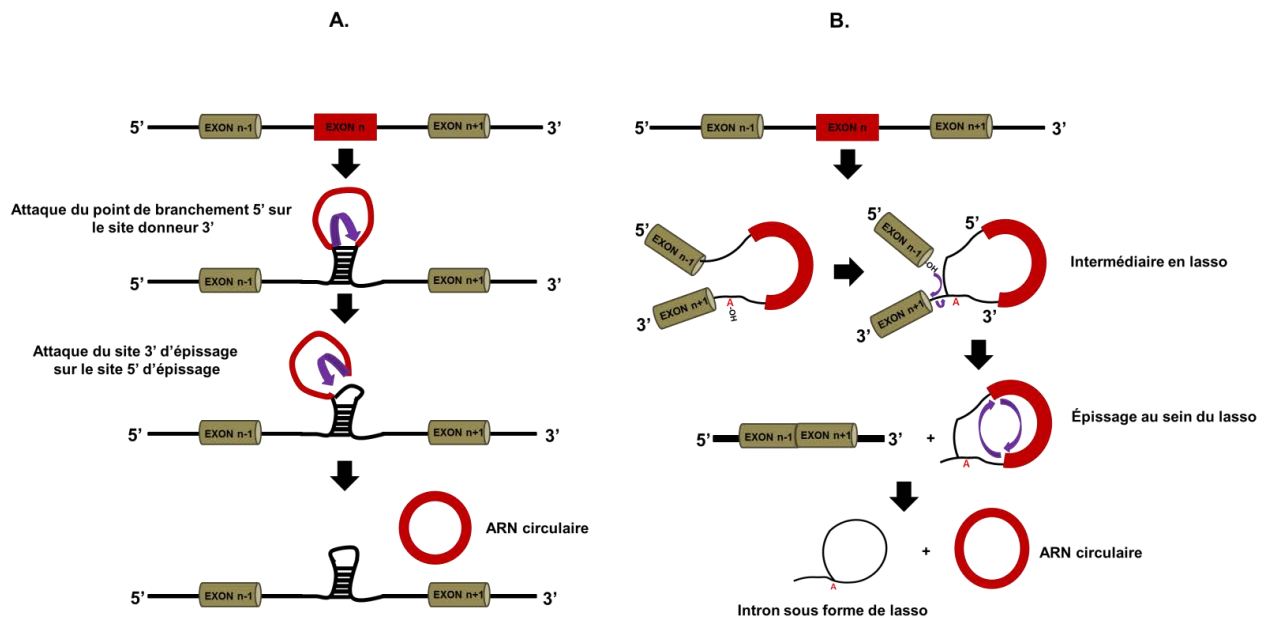


Figure II.5 : Schématisation des différents modèles proposés pour la formation d'ARN circulaires. A. Modèle de circularisation par formation d'une structure tige-boucle. **B.** Modèle de circularisation au sein d'un lasso classique au cours de l'épissage (d'après Petkovic and Müller, 2015).

II.3.2.4. Les autres fonctions de la protéine MBNL1.

Au cours de ces dernières années, de nombreuses études ont mis en évidence l'implication de la protéine MBNL1 dans plusieurs fonctions cellulaires. Comme mentionnée précédemment, la protéine MBNL1 peut interagir avec des motifs YGCY au niveau de la séquence 3'UTR d'un pré-ARNm afin d'inhiber l'utilisation d'un site de polyadénylation en faveur d'un autre. Cette implication dans la polyadénylation des pré-ARNm a un impact direct sur le devenir des ARNm, tant sur leur localisation que sur leur stabilité. Cette régulation a pour conséquence de moduler la présence de séquences régulatrices au sein de la région 3'UTR qui peuvent être reconnues par des micros ARN (miRNA ou miR) et/ou par des protéines de transport des ARNm (Batra et al., 2014; Wang et al., 2012). Par ailleurs, un rôle de MBNL1 dans la production des miRNA vient d'être proposé. En effet, des motifs YGCY ont été mis en évidence au sein des boucles de nombreux pré-miR, ainsi que la fixation de la

protéine MBNL1 sur les pré-miR-1 et 1307 (Rau et al., 2011b; Tang et al., 2015). L'interaction de la protéine MBNL1 sur ces pré-miR peut soit favoriser le processus de maturation du miRNA, soit l'inhiber en interférant avec l'activité de DICER1 sur le pré-miR. Ainsi, il a été proposé que la protéine MBNL1 puisse avoir comme certaines protéines hnRNP un rôle de régulateur de la maturation des miR ce qui a une influence sur la stabilité des ARNm cibles de ces miRNA (Rau et al., 2011b; Tang et al., 2015). De plus, Une étude récente a aussi montré que la protéine MBNL1 réprime l'expression de protéines qui seraient issues d'ARNm présentant des répétitions anormales de séquences CUG ou CAG qui conduisent à la production de protéines contenant des polyglutamines (polyQ) (Kino et al., 2015). Cette répression est associée à une relocalisation des ARNm mutants dans le noyau avec la protéine MBNL1. Ainsi, il est proposée qu'en fonction de sa localisation, la protéine MBNL1 reconnaisse les ARNm mutants avec des expansions de répétitions et maintienne ces ARN au sein du noyau afin de réprimer leur expression. Ce processus peut être considéré comme un mécanisme de contrôle qualité des ARNm qui limiterait la production de protéines aberrantes non physiologiques. Cette nouvelle fonction de la protéine MBNL1 suggère donc que, contrairement à l'idée établie à partir de l'étude de la DM1, la protéine MBNL1 ne serait pas simplement séquestrée par la formation des *foci* nucléaires, mais serait un acteur de leur formation, agissant comme un répresseur de la production de protéines aberrantes (Kino et al., 2015).

II.3.3. Implication de la protéine MBNL1 dans la dystrophie myotonique de type I.

Parmi les protéines impliquées dans la DM1, la protéine MBNL1 est la plus étudiée et semble avoir de nombreux rôles dans la pathologie. En effet, il est proposé aujourd'hui que la protéine MBNL1 soit à l'origine de la formation des *foci* nucléaires observés chez les patients, par sa fonction de répresseur d'expression d'ARNm renfermant des répétitions CUG (Kino et al., 2015). Cette hypothèse est d'autant plus plausible qu'une invalidation de la protéine MBNL1 dans des myoblastes de patients DM1 conduit à une diminution significative du nombre de *foci* (Dansithong et al., 2005). Par ailleurs, la formation de *foci* nucléaires dans le noyau des patients conduit à une diminution importante de la protéine MBNL1 libre, qui ne peut plus assurer correctement ses autres fonctions, en particulier son

rôle de régulateur de l'épissage alternatif. L'analyse comparative, par approche puces à exons, du transcriptome de souris modèles de la DM1 et de souris n'exprimant pas la protéine MBNL1 montre ainsi que plus de 80% des dérégulations de l'épissage alternatif sont communes entre les deux modèles. Néanmoins, chez l'homme, la perte de la protéine MBNL1 n'explique que 50% des modulations de l'épissage alternatif observées chez les patients DM1 (Kanadia et al., 2003b, 2006). La diminution de la quantité de protéine MBNL1 libre chez les patients DM1 impacte donc directement l'épissage alternatif de nombreux pré-ARNm, mais pourrait également l'impacter indirectement dans la mesure où MBNL1 peut interagir avec d'autres protéines de régulation de l'épissage, dont les protéines hnRNP et ceci de manière ARN indépendante (Paul et al., 2011). La stœchiométrie et l'activité globale des protéines intervenant dans le processus d'épissage sont alors modifiées (Paul et al., 2011). Ainsi, de nombreux travaux ont étudié l'implication de la protéine MBNL1 dans l'épissage d'un grand nombre de pré-ARNm. C'est notamment le cas de l'épissage du pré-ARNm MAPT qui code six isoformes de la protéine Tau dans le cerveau humain d'un adulte, du fait de l'inclusion alternative des exons 2, 3 et 10. Dans les cerveaux de patients DM1, une augmentation de l'exclusion de ces exons favorisant l'expression de la protéine Tau fœtale est observée. Elle contribue aux symptômes neuro-dégénératifs observés chez les malades. Une étude récente, réalisée par l'équipe de N. Sergeant et ML. Caillet-Boudin de Lille et à laquelle j'ai contribué et qui est présentée dans la partie Publications, a permis de mettre en évidence l'implication de la protéine MBNL1 en synergie avec la protéine MBNL2 dans la régulation de l'épissage alternatif de l'exon 2 du pré-ARNm codant la protéine Tau. En effet, une diminution du taux de protéine MBNL1 favorise l'exclusion de l'exon 2 du pré-ARNm contribuant ainsi à la formation de l'isoforme fœtale de la protéine (Carpentier et al., 2014; Sergeant et al., 2001). Afin de mieux comprendre l'altération du profil d'épissage chez les patients DM1, d'autres travaux ont été réalisés. Ces derniers révèlent l'importance d'une autre protéine de régulation de l'épissage alternatif, la protéine CUG-BP1.

II.4. LES PROTEINES DE LA FAMILLE CELF.

La famille des protéines CELF (pour CUG-BP and ETR3-Like Factor) est la seconde famille de régulateurs de l'épissage impliquée dans la DM1. Le premier membre de cette famille a été identifié dans la cadre d'une étude recherchant une protéine capable de se lier à une séquence ARN composée de 8 répétitions CUG. C'est ainsi que la protéine CELF1 a été identifiée et nommée également CUG-BP1 pour CUG Binding Protein. Puis par homologie de séquence, la protéine ETR-3 (pour Emryonically lethal abnormal vision-type RNA binding protéine 3 ou CELF2) a été identifiée établissant ainsi la famille des protéines CELF (Ladd et al., 2001a; Timchenko et al., 1996).

Chez l'homme, cette famille est composée de 6 membres, les protéines CELF1 à 6, également appelées BRUNOL1-6 (pour BRUNO-Like) en raison de leur similarité de séquences avec les protéines Bruno chez la drosophile. Les protéines CELF sont conservées chez les eucaryotes et peuvent être classées en deux sous familles sur la base de la phylogénie. La première sous-famille correspond aux protéines CUG-BP1 et ETR-3 (Barreau et al., 2006; Ladd and Cooper, 2004). Ces protéines sont exprimées de manière ubiquitaire avec un enrichissement dans le cœur, le cerveau et les muscles. La seconde sous-famille correspond aux protéines CELF3 à 6 dont l'expression est tissu-dépendante : les protéines CELF3, 4 et 5 étant exprimées au niveau du système nerveux central chez l'adulte alors que la protéine CELF6 est préférentiellement trouvée dans le cerveau et le rein (Loria et al., 2003; Meins et al., 2002). L'expression des protéines CELF joue un rôle essentiel dans la régulation de l'épissage au cours du développement et évolue en fonction des différents stades de développement (Ladd et al., 2001b). Ainsi, la protéine CELF6 est majoritairement exprimée dans le cerveau et le rein chez l'adulte, tandis que dans le fœtus, la protéine est beaucoup moins exprimée en particulier dans le rein fœtal (Ladd et al., 2004; pour revue Dasgupta and Ladd, 2012). Il est à noter qu'au cours du développement, les protéines CUG-BP1 et ETR-3 sont fortement exprimées dans les premiers stades embryonnaires, puis sous-exprimées plus de 10 fois dans le muscle squelettique et le cœur lors du développement post-natal, en restant à des niveaux faibles dans les tissus adultes (Kalsotra et al., 2008; Ladd et al., 2005).

II.4.1. Structure des protéines CELF et interaction avec l'ARN.

Les protéines CELF sont toutes constituées de trois domaines de reconnaissance à l'ARN (RRM) qui sont conservés entre les protéines. Les deux premiers domaines RRM (RRM1 et 2) sont localisés dans la partie N-Terminale de la protéine et le troisième dans la partie C-Terminale (RRM3) (Figure II.6). Le RRM3 est séparé des deux autres domaines par une région de 160 à 230 résidus dite divergente, non conservée et spécifique à chaque membre de la famille (Ladd et al., 2001b).

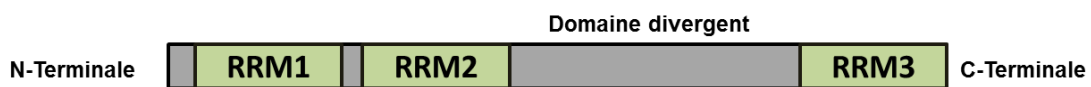


Figure II.6 : Schéma de l'organisation structurale des protéines CELF.

De nombreuses études montrent que les domaines RRM1 et 2 sont essentiels pour l'interaction des protéines CELF avec les ARN contenant des répétitions CUG. Cependant, d'autres études ont mis en évidence que le domaine RRM3 associé au domaine divergent était suffisant pour interagir avec des répétitions de GC. Ensemble, ces études suggèrent que les différents RRM peuvent se lier à des substrats distincts. Fait intéressant, des travaux réalisés avec l'approche SELEX, ont montré que les protéines CUG-BP1 et ETR-3 ont respectivement une meilleure affinité pour les séquences de type UGU (CUG-BP1) et UGUU (Faustino and Cooper, 2005; Marquis et al., 2006). Bien que sa fonction n'ait pas encore été clairement établie, le domaine divergent semble jouer un rôle dans l'interaction avec l'ARN et pourrait être impliqué dans la localisation sub-cellulaire de certaines protéines CELF (Dujardin et al., 2010; Han and Cooper, 2005). En effet, la déplétion de ce domaine dans la protéine CUG-BP1 affecte son exclusion du noyau, sa translocation au sein de granules de stress cytoplasmiques et sa localisation dans les compartiments périnucléaires (Fujimura et al., 2008).

II.4.2. Fonction des protéines CELF.

Les protéines CELF peuvent être localisées dans différents compartiments de la cellule, ce qui suggère que ces protéines pourraient avoir différents rôles selon leur localisation cellulaire. En effet, les protéines CELF ont, au sein du cytoplasme, plusieurs fonctions : elles interviennent dans la stabilité des ARNm, la déadénylation et la traduction (Paillard et al., 2003; Vlasova et al., 2008; Barreau et al., 2006). Au niveau du noyau, les protéines CELF jouent essentiellement un rôle dans la régulation de l'épissage alternatif. L'activité des protéines CELF est régulée par leur niveau de phosphorylation. Ainsi, CUG-BP1 peut être hyperphosphorylée par la protéine PKC (Protein Kinase C), mais aussi par de multiples autres kinases et ceci influence à la fois sa stabilité, sa localisation nucléaire et donne sa capacité à réguler l'épissage (pour revue Dasgupta and Ladd, 2012; Beisang et al., 2012).

II.4.2.1 Les protéines CELF dans l'épissage alternatif.

Les protéines CELF participent à la régulation de l'épissage alternatif en favorisant aussi bien l'inclusion que l'exclusion d'exons au sein des pré-ARNm. Les sites de fixation des protéines CELF au sein d'un pré-ARNm peuvent être situés en amont et/ou en aval des exons régulés. Différents mécanismes de régulation de l'épissage alternatif par les protéines CELF ont été proposés. Tout d'abord, les protéines CELF peuvent interagir avec la snRNP U2 et favoriser son recrutement au niveau du point de branchement. C'est notamment le cas dans l'épissage de l'exon 5 du pré-ARNm codant la troponine T cardiaque chez le poulet où ETR-3 induit l'inclusion de l'exon par interaction directe avec la snRNP U2 (Goo and Cooper, 2009). De plus, les protéines CELF peuvent agir indirectement en agissant comme compétiteur pour la fixation sur un site régulateur. Dans ce cas, les protéines CELF se fixent sur le site d'un activateur ou d'un répresseur pour empêcher son action. De même, il est proposé que la protéine ETR-3 puisse empêcher la reconnaissance du site polypyrimidine par U2AF et ainsi favoriser l'exclusion d'un exon (Dujardin et al., 2010). Bien que la régulation de l'épissage alternatif soit généralement considérée comme la fonction principale des protéines CELF dans le noyau, les protéines CELF sont également impliquées dans d'autres fonctions.

II.4.2.2. Les autres fonctions des protéines CELF.

Comme pour la plupart des protéines se liant aux ARNm, les protéines CELF sont impliquées dans divers processus cellulaires dont la régulation de la stabilité des ARNm. En effet, La protéine CUG-BP1 peut interagir avec des motifs conservés GRE (GU-Rich Element) qui sont enrichis dans la séquences 3'UTR de nombreux ARNm. Cette interaction permet de promouvoir la dégradation des ARNm (Lee et al., 2010; Vlasova-St Louis and Bohjanen, 2011). A l'inverse, la protéine ETR-3 stabilise les ARNm en reconnaissant des séquences riches en résidus A/U au sein de la région 3'UTR, mais empêche leur traduction par le ribosome (Mukhopadhyay et al., 2003). Les protéines CELF peuvent également fonctionner comme activateur ou répresseur de la traduction en interagissant avec la région 5'UTR des ARNm au niveau de séquences riches en G/C (Iakova et al., 2004). Enfin, en lien avec leur fonction de régulateur de stabilité des ARNm, les protéines CELF jouent également un rôle important dans la déadénylation des ARNm. Il a été clairement montré qu'en interagissant avec la 3'UTR d'ARNm cibles, CUG-BP1 pouvait promouvoir leur déadénylation en facilitant le recrutement ou la stabilisation de la déadénylase PARN (Poly(A)-specific RiboNuclease). Il a même été proposé que CUG-BP1 puisse induire elle-même la déadénylation (Kim and Richter, 2006; Moraes et al., 2006). Cette activité des protéines CELF est modulée par leur état de phosphorylation et la présence de cofacteurs, ainsi seules les protéines CELF déphosphorylées peuvent réguler la déadénylation des ARNm.

II.4.2.3. Les protéines CELF : des antagonistes des protéines MBNL.

Il est établi aujourd'hui que les protéines CELF et particulièrement les protéines CUG-BP1 et ETR-3 sont des antagonistes des protéines MBNL. Cet antagonisme a été mis en évidence pour la régulation de l'épissage alternatif de plus de 200 exons (Wang et al., 2012, 2015). Il ne reposerait pas sur un modèle de compétition pour la fixation sur un ARN cible, dans la mesure où les sites préférentiels de fixation des deux familles protéiques sont distincts et que la fixation d'une des protéines n'empêche par la fixation de l'autre (Dansithong et al., 2005; Ho et al., 2004a). De plus, une surexpression de la protéine CUG-BP1 dans des myoblastes de souris n'influence ni l'expression de la protéine MBNL1, ni sa localisation dans

la cellule. De même, le niveau d'inclusion d'exons régulés uniquement par MBNL1, et non par les protéines CELF, ne montre aucune variation (Lin et al., 2006; Wang et al., 2015). Ces observations suggèrent donc que l'antagonisme des protéines CELF envers les protéines MBNL résulte d'un effet direct sur l'épissage alternatif (Wang et al., 2015). L'antagonisme entre les protéines CELF et MBNL se retrouve également dans d'autres fonctions de régulation et plus particulièrement dans la régulation de l'expression des ARNm (Kalsotra et al., 2008; Masuda et al., 2012; Wang et al., 2015). Des analyses comparatives de la séquence de la région 3'UTR des ARNm ont permis de mettre en évidence plus de 1000 3'UTR d'ARNm permettant la fixation des deux familles de protéines. Ces analyses ont également permis d'observer la présence de sites pour la fixation des protéines CUG-BP1 et MBNL1 regroupés en cluster, au sein d'un nombre significatif de 3'UTR. Malgré la proximité des différents sites, les premières expérimentations suggèrent également que l'antagonisme entre CUG-BP1 et MBNL1 repose sur un modèle d'inhibition de fonction et non de fixation (Wang et al., 2015). Cet antagonisme entre les protéines CUG-BP1 et MBNL1 est également observé dans le cadre de la DM1 et la protéine CUG-BP1 contribue de manière importante à l'altération du profil d'épissage chez les patients (Ladd, 2013).

II.4.3. Implication des protéines CELF dans la DM1.

Les protéines CELF jouent un rôle essentiel dans le développement de maladies neuromusculaires et en particulier dans la DM1. En effet, bien que le niveau d'expression des protéines CELF diminue au cours du développement chez les patients DM1, une augmentation de la quantité des protéines CELF dans les muscles squelettiques et le cœur a été observée (Kalsotra et al., 2008; Timchenko et al., 2004). Cette augmentation s'accompagne d'une modulation de la localisation cellulaire des protéines CUG-BP1 et ETR-3 qui s'accumulent dans le noyau des cellules indépendamment de la présence des répétitions CUG dans la 3'UTR de l'ARNm *DMPK*. L'augmentation de la quantité de CUG-BP1 est due à son hyperphosphorylation par la protéine PKC qui confère à la protéine CUG-BP1 une plus grande stabilité. A l'heure actuelle, il est proposé que la dérégulation de l'activité de la protéine PKC soit associée à une réduction de l'expression de miRNA (Kalsotra et al., 2010; Kuyumcu-Martinez et al., 2007). De plus, il a été montré dans la souris, que la seule

surexpression de la protéine CUG-BP1 dans des cellules de muscles squelettiques permettait de reproduire certains des défauts d'épissage alternatif et des altérations physiologiques retrouvés chez les patients (Ho et al., 2005; Ward et al., 2010). Ainsi, l'augmentation de la stabilité de la protéine CUG-BP1 et de sa concentration nucléaire associée à la diminution des protéines MBNL libres au sein des noyaux contribue à exacerber l'antagonisme entre les deux familles protéiques et à favoriser un profil d'épissage de type fœtal des pré-ARNm chez les malades. Une implication de la protéine CUG-BP1 dans la DM1 dans la stabilité des ARNm et la formation de granules de stress cytoplasmiques entraînant une diminution de la traduction (Huichalaf et al., 2010).

II.5. LA DEREGULATION DE L'ÉPISSAGE DE L'EXON 5 DE LA TROPONINE T CARDIAQUE DANS LA DM1.

Comme mentionné précédemment, la dérégulation de l'épissage alternatif de l'exon 5 de hcTNT pourrait participer aux symptômes cardiaques observés chez les patients DM1, bien que la contribution directe de ce transcrit n'ait pas encore été démontrée. Néanmoins, l'atteinte cardiaque chez les patient DM1 est à l'origine de 20% des décès chez les malades, d'où l'importance d'étudier la régulation de l'épissage de l'exon 5 de hcTNT.

II.5.1. La troponine T et son rôle dans la contraction musculaire.

Chez les vertébrés, la contraction musculaire des muscles striés, met en jeu de nombreux complexes protéiques afin de permettre un mouvement ou le maintien d'une posture. La troponine T participe à ce processus et forme un complexe appelé troponine avec les troponines C et I (TNC et TNI). La troponine T joue un rôle essentiel dans la fixation du complexe troponine sur les filaments d'actines, par l'intermédiaire de la tropomyosine, aussi que dans l'organisation des différentes sous-unités dans le complexe. Elle participe à la régulation de la contraction musculaire. Dans le cœur, La troponine T cardiaque humaine est codée par le gène hcTNT (ou TNNT2) qui est composé de 17 exons et se situe sur le chromosome 1 en position 1q32 (pour revue Katrukha, 2013). Chez l'homme, il existe quatre isoformes de la protéine (TnT1 à 4) en fonction de l'inclusion ou l'exclusion des exons

4 et 5 dans l'ARNm qui sont les seuls exons alternatifs du pré-ARNm. Le taux relatif de chaque isoforme dans les cellules musculaires évolue au cours du temps. Ainsi, dans le cœur d'un fœtus, c'est l'isoforme TnT1, présentant l'inclusion des deux exons, qui est l'isoforme majoritaire. Tandis que chez l'adulte, c'est l'isoforme TnT3 qui est favorisée, suite à l'augmentation de l'exclusion de l'exon 5. Les deux autres isoformes TnT2 et 4 sont également des isoformes fœtales (Gomes et al., 2002). L'exon 5 du gène hcTNT code une séquence de 10 acides aminés très acide (EEEDWREDED) localisée dans la région N-terminale de la protéine. L'inclusion de l'exon 5 conduit à la synthèse d'une isoforme plus acide ce qui affecte le fonctionnement du complexe troponine. Il a été montré que l'expression de l'isoforme TnT3 (dépourvue de l'exon 5) chez l'adulte améliore la force maximale de contraction des myofibrilles ainsi que la sensibilité des myocytes au calcium en favorisant la dissociation du Ca²⁺ au niveau du complexe TNC. Par son impact sur la structure du complexe troponine et la dissociation de Ca²⁺, l'épissage alternatif de l'exon 5 représente donc un événement important dans la contraction musculaire (Gomes et al., 2004; pour revue Katrukha, 2013). En effet, une modulation de son taux d'inclusion peut entraîner des défauts dans la conduction et le rythme cardiaque qui peuvent à long terme s'avérer fatal. L'inclusion anormale de l'exon 5 dans l'ARNm hcTNT dans les cellules DM1 conduit donc à une mauvaise contraction musculaire et pourrait être impliquée dans les anomalies cardiaques de patients (Philips et al., 1998; Vinereanu et al., 2004).

II.5.2. Régulation et dérégulation de l'épissage alternatif de l'exon 5 de hcTNT dans la DM1.

II.5.2.1. Bases moléculaires de la régulation de l'épissage alternatif de l'exon 5 de hcTNT.

Au cours du développement, le profil d'épissage de l'exon 5 de hcTNT change pour passer de l'inclusion vers l'exclusion. Cette modulation du profil d'épissage est retrouvée chez de nombreuses espèces comme la souris, le rat et le poulet suggérant un mécanisme conservé chez les mammifères. Initialement, l'étude moléculaire de la régulation de l'épissage de l'exon 5 de hcTNT a été débutée chez le poulet, puis poursuivie chez l'homme. Deux sites de fixation de la protéine MBNL1 ont été identifiés entre les positions 41 et 21 en amont de l'exon 5 humain (Figure II.7) et il a été montré que la surexpression de la protéine MBNL1 entraînait une diminution de l'inclusion de l'exon 5 (Ho et al., 2004b).

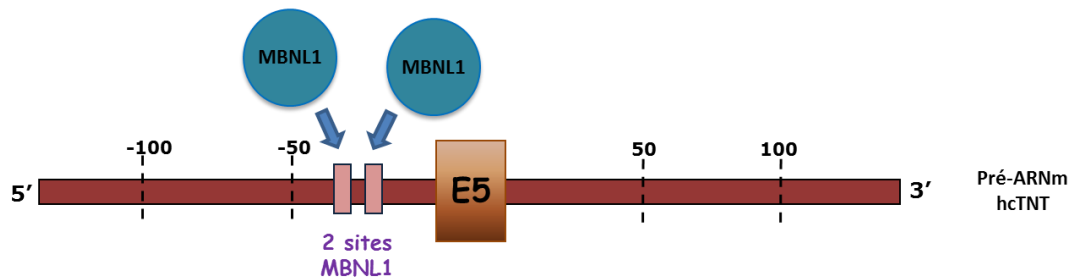


Figure II.7 : Schéma de la localisation de deux sites de fixation de la protéine MBNL1 sur l'ARNm hcTNT. Les sites (en rose) sont situés entre les positions 41 et 21 en amont de l'exon 5 (en marron)

Ainsi, il a été proposé que la fixation de la protéine MBNL1 stabilise localement la formation d'une structure tige-boucle formée par l'appariement de ces deux sites de fixation. La structure ainsi stabilisée empêcherait la reconnaissance par U2AF65 de la séquence polypyrimidine située entre les 2 sites MBNL1 limitant la reconnaissance de l'exon par le spliceosome (Warf et al., 2009). Des sites de fixations pour les protéines CELF ont été également identifiés au niveau des 50 premiers nucléotides de l'intron en aval de l'exon 5 (Ho et al., 2004b). Chez le poulet, une régulation positive de l'inclusion de l'exon 5 par les protéines CELF a été mise en évidence et deux mécanismes de régulation non exclusifs ont été proposés : dans le premier, la fixation des protéines CELF activerait directement le recrutement du spliceosome, alors que dans le second, les protéines CELF bloqueraient l'action d'un complexe répresseur composé des protéines MBNL1 et PTB. Ce mécanisme de compétition ne serait cependant pas basé sur une compétition directe des protéines pour la fixation sur le pré-ARNm cTNT, puisque les sites de fixation des protéines CELF, MBNL1 et CUGBP sont distincts (Ladd et al., 2005). Cependant le mécanisme d'action de ces protéines pour favoriser l'inclusion de l'exon 5 du pré-ARNm cTNT chez l'homme n'a pas été élucidé. Dans le cadre de la DM1, l'augmentation de l'inclusion de l'exon 5 observée chez les patients est corrélée avec l'augmentation de la stabilité de la protéine CUG-BP1 et la diminution de protéine MBNL1 libre, ce qui correspond à un retour à un profil fœtal d'épissage de l'exon.

II.5.2.2. Identification de nouveaux sites intervenants dans l'épissage alternatif du pré-ARNm hcTNT.

Dans le cadre du consortium français d'étude de la dystrophie myotonique financé par l'AFM, l'équipe de Christiane Branlant (IMoPA, ex-AREMS), en collaboration avec l'équipe de Nicolas Sergeant et Marie-Laure Caillet-Boudin à Lille, a mené des recherches visant à mieux comprendre les bases moléculaires des dérégulations de l'épissage alternatif observées chez les patients souffrant de DM1. Pour cela, avant mon arrivée au laboratoire, l'équipe avait entrepris de définir quelles étaient les séquences minimales requises en aval et en amont de l'exon 5 pour récapituler les régulations de son épissage, y compris l'effet *en trans* d'ARN contenant des répétitions de triplets CUG. Différents minigènes contenant l'exon 5 du pré-ARNm hcTNT, avec des régions introniques flanquant l'exon 5 entières ou variant de 50 à 250 nt avaient été produits. Dans ces minigènes l'exon 5 du pré-ARNm hcTNT et ses séquences introniques flanquantes avaient été insérés dans l'intron 2 du récepteur à l'insuline (IR), lui-même flanqué de ses exons bordants (Figure II.8). Ces minigènes avaient été cotransfectés avec une construction exprimant ou non des répétitions CTG insérées dans le contexte de séquence dans lequel elles sont présentes dans le gène DMPK. Les événements d'épissage avaient été analysés par RT-PCR et les données obtenues avaient révélé que les séquences impliquées dans les diverses régulations s'exerçant sur l'exon 5 s'étendent sur environ 150 nt en amont et/ou en aval de cet exon (Figure II.9).

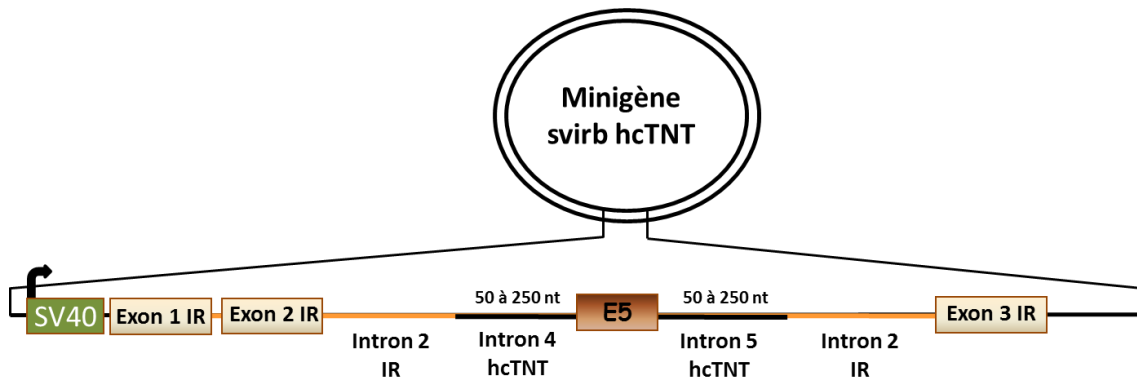


Figure II.8 : Schéma du minigène IR-hcTNT. Le minigène contient l'exon 5 du pré-ARNm hcTNT, avec ces régions introniques flanquantes entières ou variant de 50 à 250 nt. Dans ce minigène l'exon 5 du pré-ARNm hcTNT et ses séquences introniques flanquantes ont été insérés dans l'intron 2 du récepteur à l'insuline (IR) de rat, lui-même flanqué de ses exons bordants et sous contrôle du promoteur SV40.

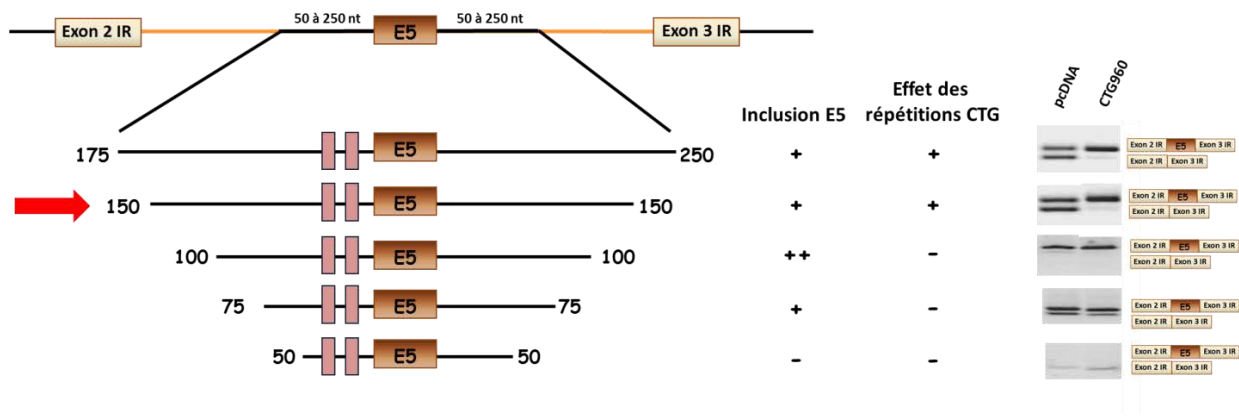


Figure II.9 : Détermination de la taille des régions introniques en amont et en aval de l'exon 5 du pré-ARNm hcTNT requises pour récapituler l'ensemble des régulations de l'inclusion de l'exon 5. La taille des différentes séquences introniques utilisées pour cette étude est indiquée. Chaque variant du minigène IR-hcTNT avait été cotransfecté dans des cellules T98 avec une construction exprimant ou non des répétitions CTG dans le contexte du gène *DMPK*. . Après extraction des ARN, l'épissage du minigène hcTNT avait été analysé par RT-PCR et les produits d'amplification avaient été séparés sur gel d'agarose. Les taux d'inclusion observés pour l'exon 5 sont indiqués : + correspond à une inclusion d'environ 50 % de l'exon 5 ; ++ correspond à une inclusion d'environ 90% de l'exon 5 ; - correspond à une forte exclusion de l'exon 5. L'effet des répétitions CUG sur l'inclusion de l'exon 5 est également indiqué. Le minigène indiqué par une flèche rouge présente la région minimale requise pour récapituler l'ensemble des régulations s'exerçant sur l'exon 5.

L'équipe de Christiane Branlant a alors recherché si ces séquences contenaient des sites potentiels de fixation de MBNL1 de consensus YGCU. C'est ainsi qu'au total huit sites potentiels ont été détectés. Parmi ces huit sites, seuls deux avaient été préalablement décrits et correspondent à ceux mentionnés dans le paragraphe précédent (appelés sites 1' et 2' et formant le groupe 1) (Ho et al., 2004b). Trois des six sites potentiels supplémentaires ainsi identifiés, sont situés entre les positions 92 et 56 en amont de l'exon 5 (appelés sites 1, 2 et 3 et formant le groupe 2), les trois autres sont localisés entre les positions 25 et 44 en aval de l'exon 5 (appelés sites 4, 5 et 6 et formant le groupe 3) (Figure II.10). Ayant identifié la région requise pour récapituler l'ensemble des régulations s'exerçant sur l'exon 5, il était important d'en établir sa structure 2D, afin de pouvoir ensuite définir les sites de fixation de ses partenaires protéiques par la technique d'empreinte. La structure 2D obtenue a révélé la présence d'une longue hélice centrale impliquant à la fois les séquences de 50 nt en 5' et 3' de l'exon (Figure II.11).

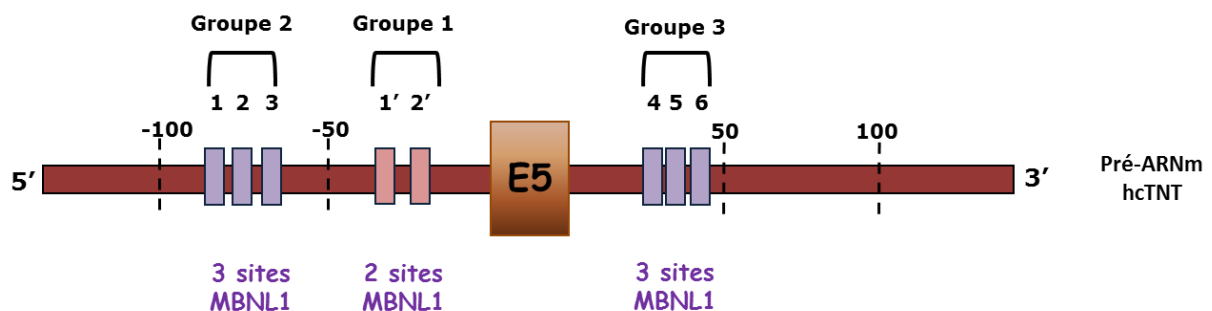


Figure II.10 : Schéma représentant les différents sites identifiés pour la fixation de la protéine MBNL1. Les sites 1' et 2' (rose) constituent les sites du groupe 1 identifié par Ho et al., 2004b. Les sites 1,2 et 3 constituent le groupe de sites 2 et les sites 4,5 et 6, le groupe de sites 3, groupes identifiés au laboratoire représentés en violet.

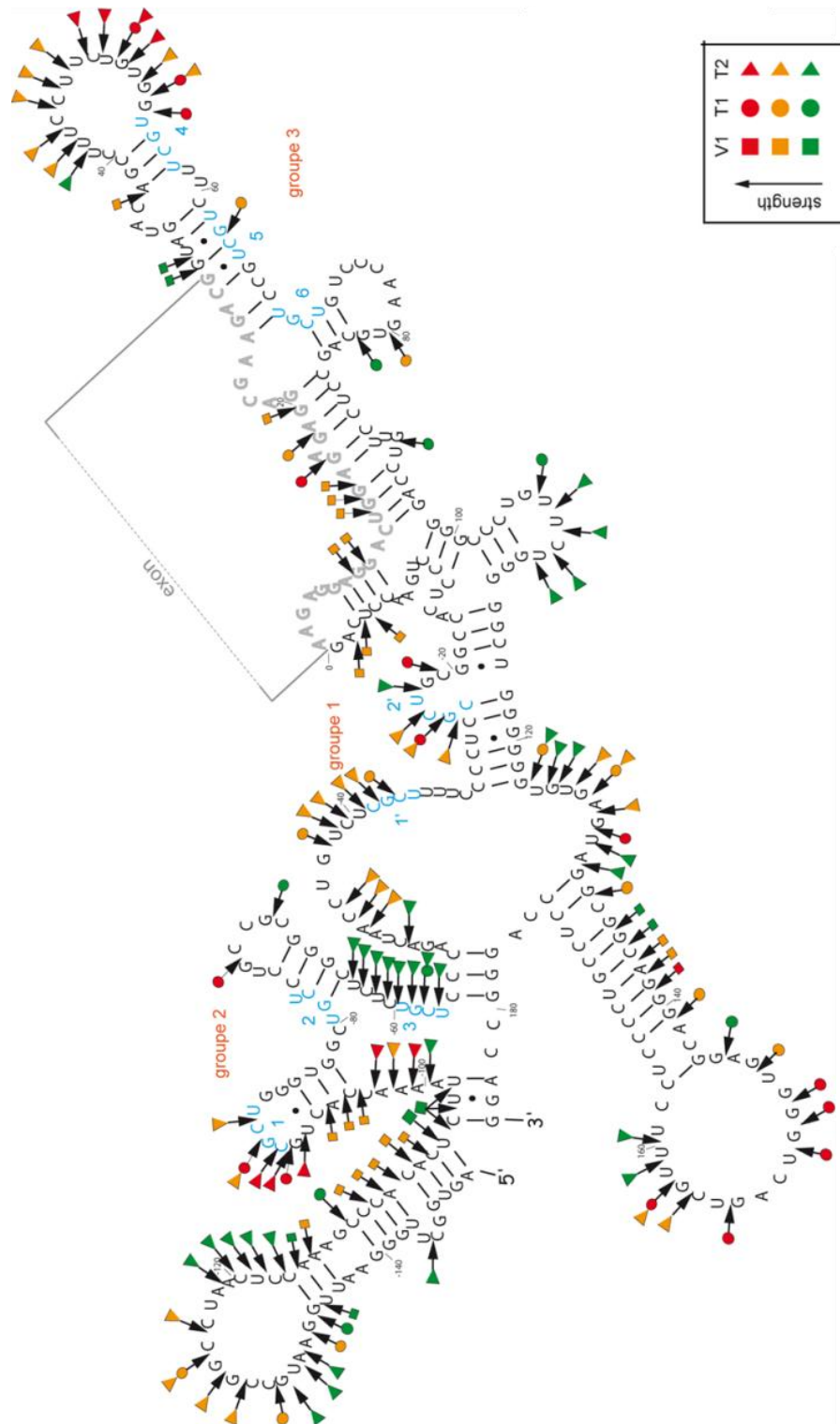


Figure II.11 : Modèle de structure 2D de la région du pré-ARNm hCNT requise pour récapituler l'ensemble des régulations et dérégulations de l'inclusion de l'exon 5, établi à partir des données d'hydrolyse enzymatiques. Les clivages par les RNases V1, T1 et T2 sont représentés par des carrés, des losanges et des ronds respectivement. La couleur des symboles représente l'intensité du clivage (faible en vert, moyenne en orange et forte en rouge). La séquence de l'exon 5 est indiquée en gris. Les huit sites MBNL1 sont indiqués en bleu et les groupes de sites (groupes 1, 2 et 3) sont indiqués.

Selon le modèle établi par Isabelle Behm-Ansmant, les 3 sites de fixation de MBNL1 situés plus en amont entre les positions 92 et 56 se structurent indépendamment du reste de la région. Par contre, les sites de fixation situés à la bordure 5' de l'exon 5 interagissent avec une séquence située en aval de l'exon 5, et les sites de fixation situés en aval de l'exon 5 interagissent avec la séquence exonique. Une analyse de la fixation de la protéine MBNL1 sur l'ARN 150-E5-150 a été réalisée par des expériences d'empreintes en solution. Des protections au niveau des huit sites potentiels de fixation de MBNL1 ont été observées, confirmant l'existence de 3 segments distincts de fixation de MBNL1 sur les régions régulatrices de l'exon 5. Afin de déterminer l'importance fonctionnelle des différents sites de fixation de MBNL1, les équipes de C. Branlant et de N. Sergeant ont alors muté individuellement les groupes de sites de fixation de MBNL1 (mutation YGCY en YACY) au sein du minigène hcTNT 150-E5-150. Dans le cas de la mutation du résidu G au sein du site 4, il s'est avéré que cette mutation entraînait l'apparition d'un site cryptique d'épissage. La délétion du résidu G24 (Δ G) permettait d'abolir l'utilisation de ce site sans impacter la régulation de l'épissage de l'exon 5. L'effet négatif de ces mutations sur la fixation de la protéine MBNL1 Δ 101 a été confirmé par des expériences de retard sur gel. Les expériences réalisées *in cellulo* indiquaient que, dans le contexte du minigène utilisé, les sites MBNL1 requis pour la régulation de l'épissage de l'exon 5 sont les 6 nouveaux sites de fixation de MBNL1 identifiés entre les positions 92 et 56 et en aval de l'exon 5, puisque la mutation des sites du groupe 2 ou du groupe 3 abolissait la régulation de l'inclusion de l'exon 5 par MBNL1 (Figure II.12). Les deux sites de fixation localisés en amont de l'exon 5 et initialement décrits comme importants ne semblaient pas jouer un rôle majeur dans la régulation de l'épissage malgré la capacité de MBNL1 à se fixer sur ces sites. L'importance individuelle de chacun des 6 nouveaux sites restait donc à déterminer.

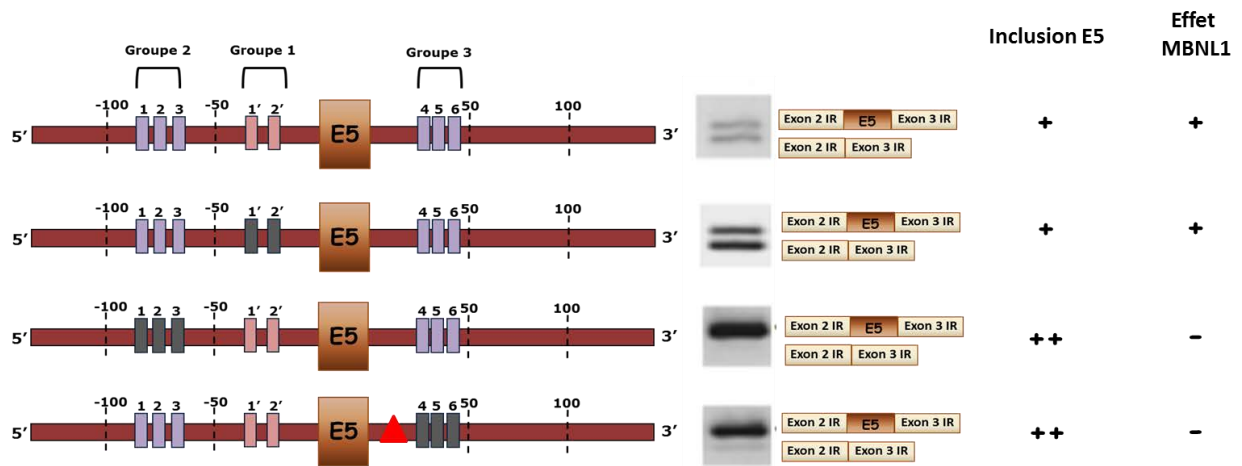


Figure II.12 : Analyse par RT-PCR de l'effet des mutations des différents groupes de sites sur l'inclusion de l'exon 5 dans les ARN produits à partir du minigène IR-hcTNT 150-E5-150. Les sites MBNL1 du groupe 1 sont indiqués en rose et les sites MBNL1 des groupes 2 et 3 sont indiqués en violet. La présence d'une mutation ΔG est indiquée par un triangle rouge. La mutation des sites MBNL1 (YGCY en YACY) est indiquée par un rectangle gris. Chaque minigène a été cotransfecté dans des cellules T98. Après extraction des ARN, l'épissage du minigène hcTNT a été analysé par RT-PCR et les produits d'amplification ont été séparés par électrophorèse en gel d'agarose. Le taux d'inclusion de l'exon 5 est indiqué : + correspond à une inclusion d'environ 50% de l'exon 5 ; ++ correspond à une inclusion d'environ 90% de l'exon 5 ; - correspond à une forte exclusion de l'exon 5.

Après avoir identifié les sites de fixation de MBNL1 jouant un rôle dans la régulation de l'utilisation de l'exon 5, les 2 équipes ont cherché à définir les autres régions nécessaires à la régulation normale de l'épissage de l'exon 5 et sa dérégulation par les répétitions CUG. Pour cela, différents variants asymétriques ont été produits à partir de la construction IR-hcTNT 150-E5-150, dans lesquels la taille des régions introniques situées soit en amont, soit en aval de l'exon 5 du pré-ARNm hcTNT variaient. Les résultats obtenus sont illustrés dans la figure 23 et indiquent que : i) les séquences nécessaires à la réponse aux répétitions CUG sont localisées dans les 150 nucléotides en amont de l'exon 5 et dans les 50 premiers nucléotides en aval, ii) des éléments activateurs sont localisés entre les positions +50 et +150 de la construction 150-E5-150 (en aval de l'exon 5) et, iii) des éléments inhibiteurs sont localisés entre les positions -150 à -100 (en amont de l'exon 5) (Figure II.13A). Ces données permettaient d'essayer d'assigner un rôle activateur ou inhibiteur à chacun des segments définis au sein de la région -150 à +150 encadrant l'exon 5 (Figure II.13B). Il restait à identifier les facteurs d'épissage, autres que MBNL1, impliqués dans ces propriétés activatrices ou inhibitrices, définir leurs sites de fixation, ainsi que leur mode d'action et leur éventuel antagonisme ou synergie avec MBNL1.

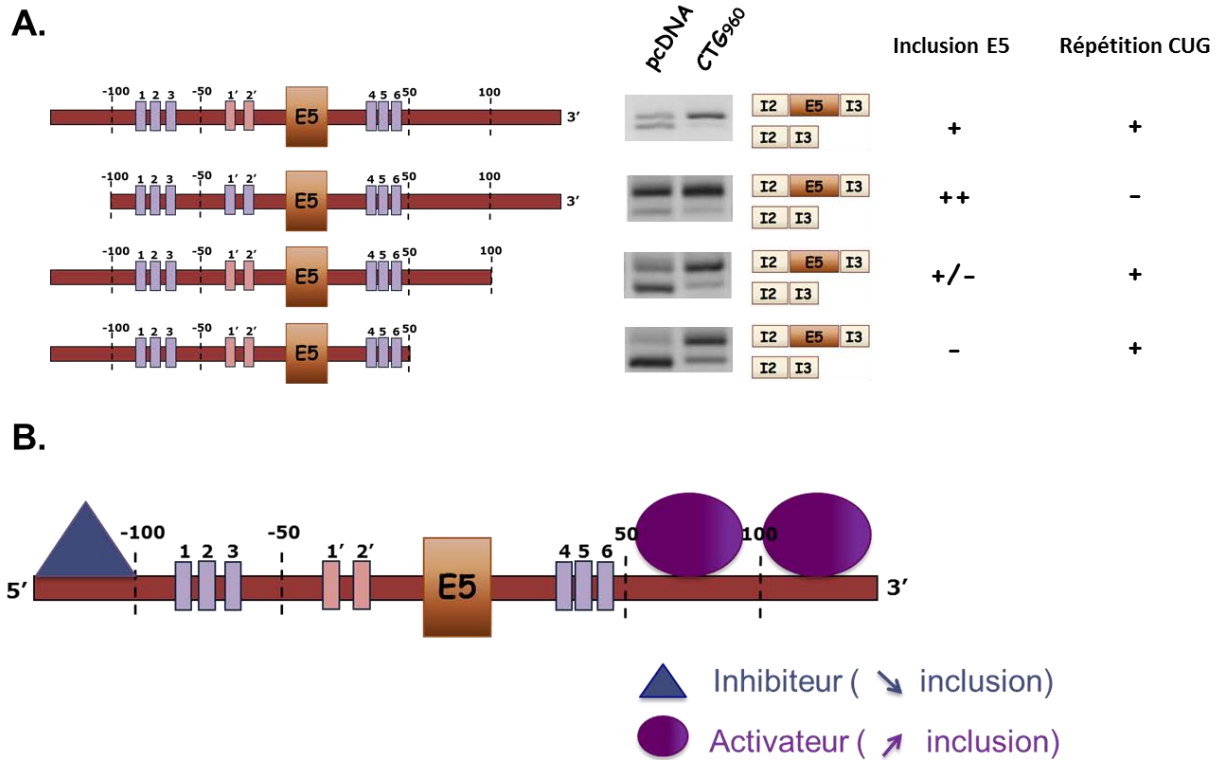


Figure II.13 : Rôle des différentes régions du minigène IR-hcTNT dans la régulation de l'inclusion de l'exon 5 du pré-ARNm hcTNT. A. Analyse par RT-PCR *in cellulo* des transcrits de minigènes contenant l'exon 5 de hcTNT et des fragments de tailles variables des introns flanquant l'exon 5. Les sites MBNL1 du groupe 1 sont indiqués en rose et les sites MBNL1 des groupes 2 et 3 sont indiqués en violet. Chaque minigène avait été cotransfecté dans des cellules T98. Après extraction des ARN, l'épissage du minigène hcTNT avait été analysé par RT-PCR et les produits d'amplification avaient été séparés sur gel d'agarose. Le taux d'inclusion de l'exon 5, ainsi que l'effet de la surexpression de MBNL1 ou de l'expression des répétitions CUG sont indiqués. **B.** Représentation schématique de la localisation des séquences impliquées dans la régulation de l'inclusion de l'exon 5. Les séquences inhibitrices sont représentées par un triangle bleu et les séquences activatrices par des sphères violettes.

II.6. OBJECTIFS.

La dystrophie myotonique de type 1 est une maladie multisystémique qui est causée par une expansion instable de répétitions CTG dans la région 3'UTR du gène *DMPK*. Aujourd'hui il est admis que ces répétitions entraînent la séquestration de la protéine MBNL1, une protéine de régulation de l'épissage alternatif, au sein de foci nucléaires, conduisant à une diminution de la protéine sous sa forme libre. Cette diminution de protéine MBNL1 libre participe à la majorité des altérations de l'épissage de pré-ARNm observées chez les patients, dont celui de la troponine T cardiaque (hcTNT) qui code une protéine importante dans la contraction cardiaque. En effet, il a été mis en évidence une dérégulation de l'inclusion de l'exon 5 de ce transcrit qui conduit à une mauvaise contraction musculaire et il est proposé que cette dérégulation pourrait être impliquée dans les anomalies cardiaques qui touchent 80% des patients atteints de DM1 (Philips et al., 1998). Lors de mon arrivée au laboratoire, les équipes de C. Branlant et de N. Sergeant avaient réalisé des mutations par groupes des sites de fixations de MBNL1 identifiés de part et d'autre de l'exon 5 (mutations simultanées des 3 sites du groupe 2, des 2 sites du groupe 1 et des 3 sites du groupe 3). Cependant aucune mutation n'avait encore été réalisée de façon individuelle. On pouvait en effet se demander si les différents sites appartenant à un même groupe avaient la même importance pour l'activité de la protéine MBNL1 dans la régulation de l'épissage alternatif de l'exon 5 du gène hcTNT. Pour définir l'importance individuelle des différents sites, j'ai donc entrepris la réalisation des mutations ponctuelles sur chacun des sites, ainsi que des combinaisons de doubles mutations affectant deux des trois sites des groupes 2 et 3. Mon premier objectif a donc été de produire les minigènes hcTNT 150-E5-150 portant ces mutations et de tester par transfection dans des cellules HEK293T, l'effet de ces mutations sur l'épissage alternatif de l'exon 5 du gène hcTNT dans la construction hcTNT 150-E5-150.

Par ailleurs, les expériences réalisées en collaboration avec l'équipe de N. Sergeant avaient permis d'établir l'existence d'autres éléments régulateurs introniques intervenant dans la régulation de l'épissage alternatif du gène hcTNT, en favorisant soit l'inclusion (séquences +50 à +100 +100 à +150), soit l'exclusion de l'exon 5 (séquence -150 à -100). Cependant les facteurs de régulation se fixant sur ces éléments n'avaient pas été identifiés. Mon second objectif a donc était de rechercher les facteurs régulateurs de l'épissage pouvant interagir

avec les séquences situées aux extrémités de l'insert hcTNT au sein du transcrit produit par la construction 150-E5-150 et de tester leur implication dans la régulation de l'épissage alternatif de l'exon 5 du gène hcTNT. Pour cela nous avons choisi de délimiter des segments de 50 nucléotides au sein de l'insert hcTNT de l'ARN hcTNT 150-E5-150 et nous avons recherché au sein d'un extrait nucléaire de cellules HeLa, compétent pour l'épissage, quelles protéines pouvaient se fixer sur chacun de ces fragments de 50 nucléotides en employant la méthode de chromatographie d'affinité basée sur l'emploi de la protéine de fusion MBP-MS2. Cela impliquait : i) la préparation de constructions permettant de produire par transcription *in vitro* les différents ARN hybrides cTNT-MS2, ii) la purification des complexes protéiques associés à ces ARN et l'identification par spectrométrie de masse avec confirmation par western-blot de l'identité des facteurs régulant l'épissage interagissant avec les ARN MS2, iii) la validation expérimentale de l'implication du ou des facteurs ainsi identifiés dans la régulation de l'épissage alternatif de l'exon 5 du pré-ARNm hcTNT et iv) l'étude des bases moléculaires de la régulation de l'épissage par ce ou ces facteurs.

RESULTATS DE LA PARTIE 1

CHAPITRE III : DECRYPTAGE DES MECANISMES DE REGULATION DE L'INCLUSION DANS L'ARNM DE L'EXON 5 DU pré-ARNm DE LA TROPONINE T CARDIAQUE HUMAINE. DECOUVERTE D'UN ROLE DE LA PROTEINE hnRNP H.

III.1. ETUDE DE L'IMPORTANCE DES DIFFERENTS SITES DE FIXATION DE MBNL1 DANS LA REGULATION DE L'INCLUSION DE L'EXON 5.

Comme mentionné dans l'introduction, il avait été proposé que les sites de fixation de la protéine MBNL1 dans les pré-ARNm s'associent deux à deux pour former des structures non canoniques du même type que celles formées par les répétitions CUG (Warf and Berglund, 200). L'idée étant que ces couples de sites forment de courtes structures tige-boucle particulières, pouvant être les éléments de reconnaissance de MBNL1 dans les pré-ARNm, en particulier dans le pré-ARNm hcTNT (Warf and Berglund, 2007). Dans ce cas, on pouvait s'attendre à ce que l'effet de MBNL1 nécessite la présence d'au moins deux séquences consensus YGCY pouvant former ce type de structure tige-boucle non canonique. En utilisant la construction IR-hcTNT 150-E5-150, il avait déjà été observé au laboratoire que, contrairement à ce qui avait été proposé par Warf et Berglund qui avaient utilisé une autre construction (construction RTB300), la mutation des deux séquences YGCY du groupe 1 en amont de l'exon 5 (site 1' et 2') n'abolit pas l'effet négatif de MBNL1 sur l'intégration de l'exon 5 de hcTNT (Figure II.12). Notre équipe avait montré, toujours avec la construction IR-hcTNT, que par contre la mutation simultanée des 3 sites de fixation de MBNL1 du groupe 2, ainsi que celle des 3 sites de fixation du groupe 3 limite très fortement l'action négative de MBNL1 sur l'inclusion de l'exon 5. On pouvait se demander si au sein de ces deux groupes de sites, chaque site possède la même capacité à participer à l'action de MBNL1 ou si ces sites agissent par paires en formant des structures tiges-boucles non canoniques. Pour répondre à cette question, nous avons donc décidé de muter chaque site au sein de chaque groupe de sites d'une part individuellement et d'autre part par paires, ceci avec ensuite l'étude de l'effet des mutations sur l'action négative de MBNL1 et sa suppression par surexpression d'enchaînements de CUG pouvant piéger MBNL1.

III.1.1. Construction des minigènes nécessaire à l'étude.

Pour cette étude nous avons utilisé le minigène IR-hcTNT qui avait été déjà bien caractérisé et utilisé au laboratoire. Il permet d'étudier les régulations de l'inclusion de l'exon 5 de hcTNT *in cellulo*. Il avait été montré que la substitution par un résidu A du résidu G de la séquence YGCY, correspondant aux consensus des séquences de fixation de MBNL1, abolissait la fixation de la protéine (Teplova and Patel, 2008). Nous avons donc choisi d'introduire cette substitution dans chacun des sites à muter, ceci de manière individuelle ou en combinaison par paires de sites. Les mutations simultanées des 3 sites des groupes 2 et 3 avaient déjà réalisées et étaient donc disponibles. Par ailleurs, l'équipe avait observé que la mutation G>A dans le site 4 entraîne l'utilisation d'un nouveau site 5' d'épissage cryptique (UGGUA) et perturbe l'analyse de l'intégration de l'exon 5. La délétion du résidu G24 (mutation Δ G) abolit l'utilisation de ce site supplémentaire sans modifier le taux d'inclusion de l'exon 5. C'est pourquoi, afin de pouvoir étudier l'importance fonctionnelle du site 4 et de la comparer à celle des autres sites, j'ai produit deux séries de constructions IR-hcTNT portant des mutations dans les sites MBNL1 avec ou sans la mutation Δ G (Figure III.1).

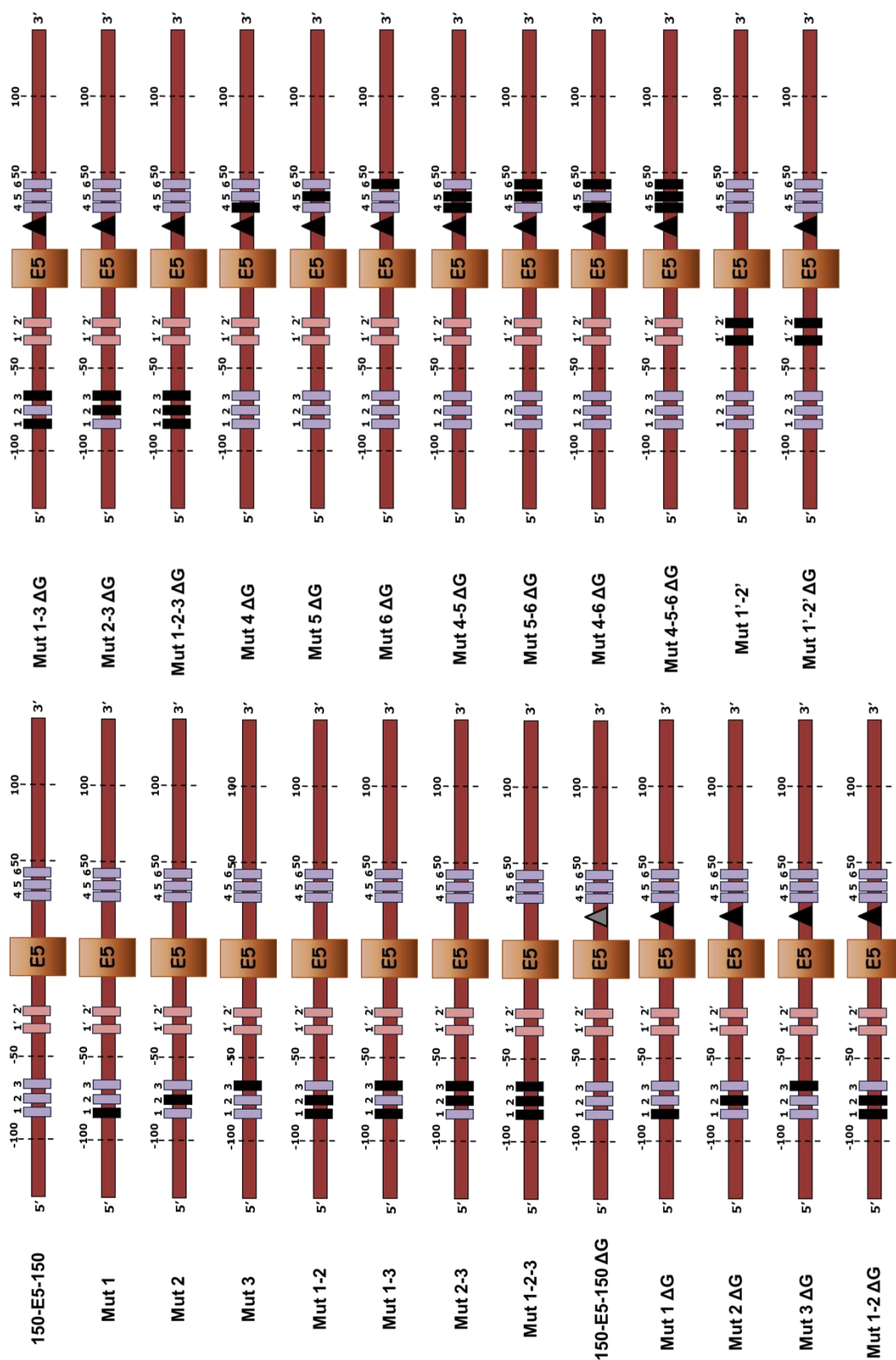


Figure III.1 : Présentation des différents mutants du minigène hcTNT 150-E5-150 utilisés lors de cette étude. Les sites MBNL1 du groupe de sites 1 sont indiqués par des rectangles roses, et ceux des groupes 2 ou 3 par des rectangles violets. Les mutations générées au sein d'un site MBNL1 sont indiquées par un rectangle noir. La présence de la mutation ΔG est indiquée par un triangle noir. Le nom de la construction est indiqué à gauche.

III.1.2. Mise en évidence de l'importance fonctionnelle individuelle des sites MBNL1 des groupes 2 et 3 dans la régulation de l'épissage alternatif de l'exon 5 du pré-ARNm hcTNT.

J'ai ensuite testé, dans des cellules humaines de la lignée HEK293T, l'effet des mutations sur l'épissage du transcrit du minigène IR-hcTNT en absence ou en présence d'une surexpression de la protéine MBNL1. Pour cela, les constructions IR-hcTNT sauvage ou mutées ont été co-transfectées avec le plasmide contrôle pcDNA3.1-flag n'exprimant que le peptide Flag ou avec le plasmide pcDNA3.1-MBNL1 surexprimant la protéine MBNL1. Les cellules ont été cultivées 48 h après la co-transfection. Les ARN totaux ont été extraits et le taux d'inclusion de l'exon 5 dans les produits d'épissage a été analysé par RT-PCR, en utilisant des oligonucléotides s'hybridant dans les exons IR situés en amont et en aval de l'exon 5 hcTNT selon le protocole décrit dans la partie Matériels et Méthodes. Les résultats ont été exprimés en pourcentage d'ARN contenant l'exon 5 par rapport à l'ensemble des ARN épissés. Ce pourcentage a été calculé à partir des quantifications des produits PCR contenant ou ne contenant pas l'exon 5 séparés par électrophorèse en gel d'agarose à bas point de fusion 4%. La construction IR-hcTNT contenant une mutation simultanée des sites 1' et 2' (groupe 1) a été introduite pour comparaison dans chacune des expériences. Les résultats obtenus avec les minigènes IR-hcTNT mutés au niveau des sites du groupe 2 (sites 1, 2 et 3) en absence de la mutation ΔG sont présentés dans la figure III.2, ceux obtenus avec les minigènes IR-hcTNT mutés au niveau du groupe 2 (sites 1, 2 et 3) et des sites du groupes 3 (sites 4, 5 et 6) en présence de la mutation ΔG sont respectivement présentés dans les figures III.3 et III.4.

Comme attendu, la présence de la mutation ΔG n'affectait pas le taux d'inclusion de l'exon 5 dans l'ARNm sauvage (69% sans mutation ΔG , 63% avec). Ce taux est réduit à 27 et 17%, respectivement, lorsque MBNL1 est surexprimée. Si l'on examine maintenant l'ensemble des résultats obtenus en absence de la mutation ΔG pour les sites du groupe 2 et sans sur-expression de MBNL1 (seulement MBNL1 endogène) dans les cellules HEK293 (Figure III.2), on constate une nette augmentation du taux d'intégration de l'exon 5 quelle que soit la mutation réalisée dans les sites de ce groupe, ce qui confirme le rôle de ces sites. Lorsque l'on compare l'effet individuel de la mutation de chacun des sites, on observe un effet légèrement plus important de la mutation du site 3 sur l'intégration de l'exon 5 par rapport aux deux autres sites, et cet effet positif est encore plus marqué en présence de la mutation ΔG (figure III.3). Aussi bien avec que sans la mutation ΔG , c'est la mutation du site 1 qui a le

moins grand effet renforçateur de l'inclusion de l'exon 5. Néanmoins, ces données suggèrent une importance fonctionnelle de chacun des sites du groupe 2. Nous n'avons pas observé d'effet très marqué de la disparition d'une paire de sites qui auraient pu former une structure tige-boucle non canonique telle que proposée par Warf et Berglund. En absence de la mutation ΔG , une intégration maximale de l'exon 5 est observée pour le triple mutant, mais la différence de taux d'inclusion pour ce mutant par rapport aux mutants simples et notamment par rapport à la mutation du site 3 est très faible. Par contre, en présence d'une surexpression de MBNL1, seul le triple mutant conduit à une intégration presque totale de l'exon 5 (Figure III.2). Nous concluons donc à partir de l'ensemble de ces données que chacun des 3 sites du groupe 2 est fonctionnel, le site 1 étant légèrement moins efficace que le site 3, lequel est légèrement plus efficace que les deux autres sites du groupe 2. A noter que la mutation simultanée des sites 1' et 2' du groupe 1 n'affectait que de façon très modérée le taux d'inclusion de l'exon 5 en absence de la mutation ΔG . Par contre, en présence de la mutation ΔG et en présence de MBNL1, nous observons un effet positif de ses deux sites qui est supérieur à celui observé pour la triple mutation 1, 2, et 3 sur l'inclusion de l'exon 5. Ces données suggèrent un lien fonctionnel possible entre les séquences situées en amont et en aval de l'exon.

Les résultats obtenus pour les mutations des sites du groupe 3 étaient moins faciles à interpréter probablement car tous les mutants comportaient en plus la mutation ΔG (Figure III.4). Nous faisons néanmoins un constat similaire à celui fait pour les sites du groupe 2 à partir de ces données : chacun des sites 4 et 5 est fonctionnel et participe à la limitation de l'intégration de l'exon 5. Le site 4 semble être le plus efficace et la seule mutation du site 6 ne semble pas affecter l'intégration de l'exon 5. La mutation des trois sites 4, 5 et 6 a un effet très marqué en présence d'une sur-expression de MBNL1. Au travers de ces données, nous démontrons donc que les 6 sites de fixation de MBNL1 qui ont été découverts au laboratoire participent, dans le cadre de la construction IR-hcTNT 150-E5-150, à l'action négative de MBNL1 sur l'intégration de l'exon 5 du pré-ARNm hcTNT.

De façon surprenante, nos données ne sont pas en parfait accord avec les résultats publiés par les équipes de T. Cooper et A. Berglund montrant l'implication des sites 1' et 2' dans la régulation de l'inclusion de l'exon 5 du pré-ARNm hcTNT par MBNL1 (Ho et al., 2004 ; Warf et al., 2007 ; Warf et al., 2009). Il faut noter à cet égard que ces deux équipes ont utilisé un

autre minigène (le même pour les deux équipes), le minigène RTB300. Les mutations qu'ils ont réalisées pour invalider les sites 1' et 2' étaient différentes de celles que nous avons produites, puisqu'ils ont remplacé le résidu G du motif YGCY par un résidu cytosine (YGCY→YCCY), alors que dans notre cas ce résidu G a été remplacé par un résidu adénosine (YGCY→YACY). La séquence polypyrimidine de l'intron 4 hcTNT étant située entre les sites 1' et 2' il est possible que les mutations (YGCY→YCCY) insérées par les équipes de T. Cooper et A. Berglund renforce cette séquence polypyrimidine optimisant ainsi le taux d'inclusion de l'exon 5 indépendamment de l'effet de MBNL1. Afin de comprendre les différences de résultats de notre équipe et de celles de T. Cooper et A. Berglund et démontrer définitivement le rôle dans la régulation de l'inclusion de l'exon 5 du pré-ARNm hcTNT des 6 nouveaux sites MBNL1 que nous avons identifiés, nous avons entrepris de vérifier si les effets observés de MBNL1 en fonction des mutations de ses sites de fixation étaient dépendants du minigène utilisé.

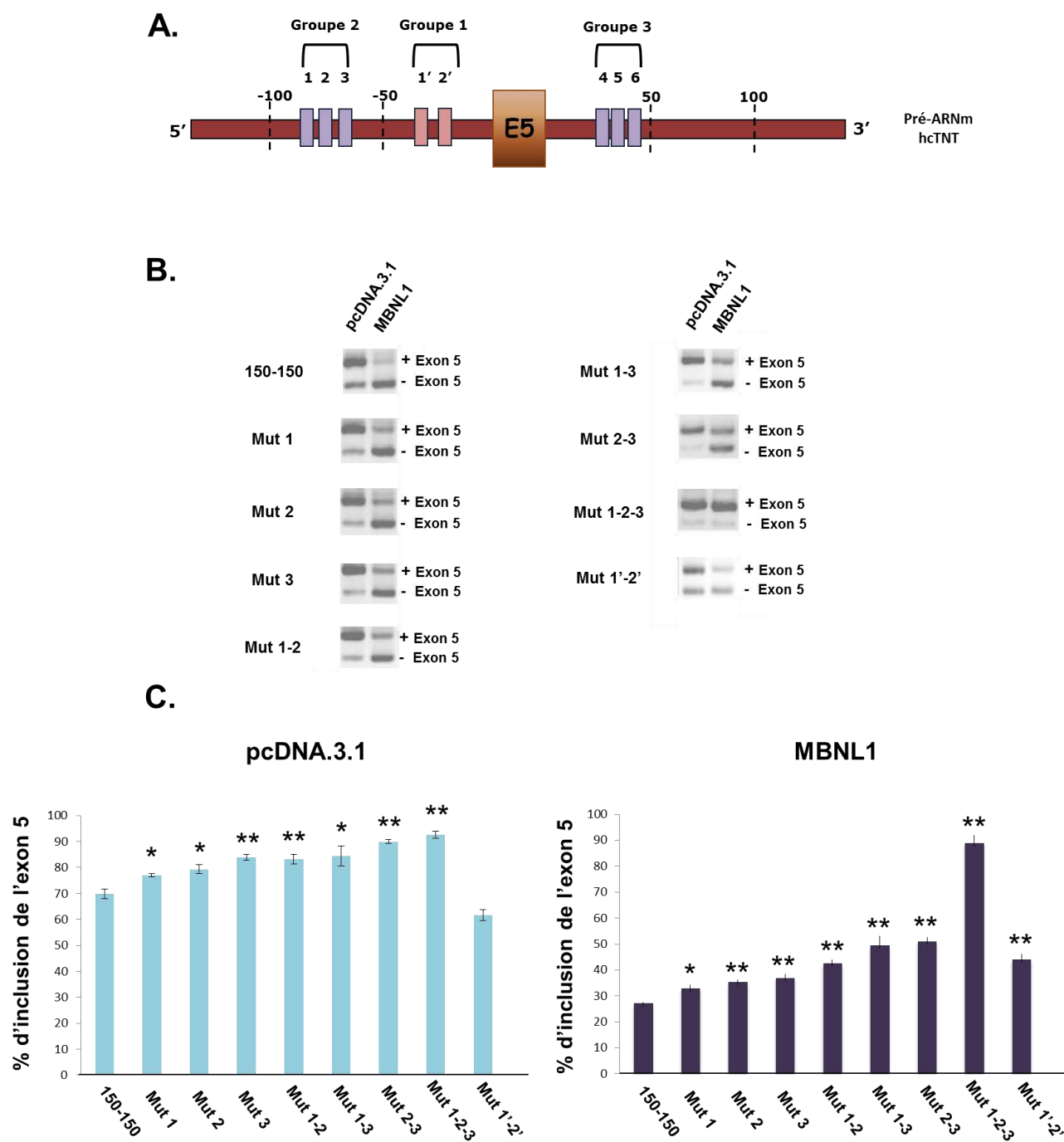


Figure III.2 : Analyse par RT-PCR de l'effet des mutations des sites 1, 2 et 3 du groupe 2 sur l'intégration de l'exon 5 dans l'ARNm produit à partir du minigène IR-hcTNT 150-E5-150. **A.** Représentation de l'organisation du pré-ARNm issu du minigène IR-hcTNT 150-E5-150. **B.** L'effet des mutations individuelles ou en combinaison des sites 1, 2 et 3 sur l'intégration de l'exon 5 dans les transcrits du minigène IR-hcTNT a été analysé en co-transfectant la construction 150-150 dans des cellules HEK293 en présence des plasmides pcDNA.3.1-Flag et pcDNA.3.1-MBNL1 (MBNL1). Les ARN totaux ont été extraits 48 h après transfection et traités à la DNase, puis rétro-transcrits en utilisant l'oligodésoxynucléotide dT15 complémentaire à la queue polyA des ARNm. Les ADNc ainsi générés ont servi de matrice pour une réaction de polymérisation en chaîne (PCR) réalisée à l'aide des oligonucléotides 5991 et 8078. Les amplicons ont été séparés sur gel d'agarose à bas point de fusion à 4%. **C.** Une quantification des produits d'amplification contenant ou non l'exon 5 a été réalisée. Les données sont exprimées en pourcentage des ARN contenant l'exon 5 par rapport au total des ARN épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01 par rapport au valeurs 150-150).

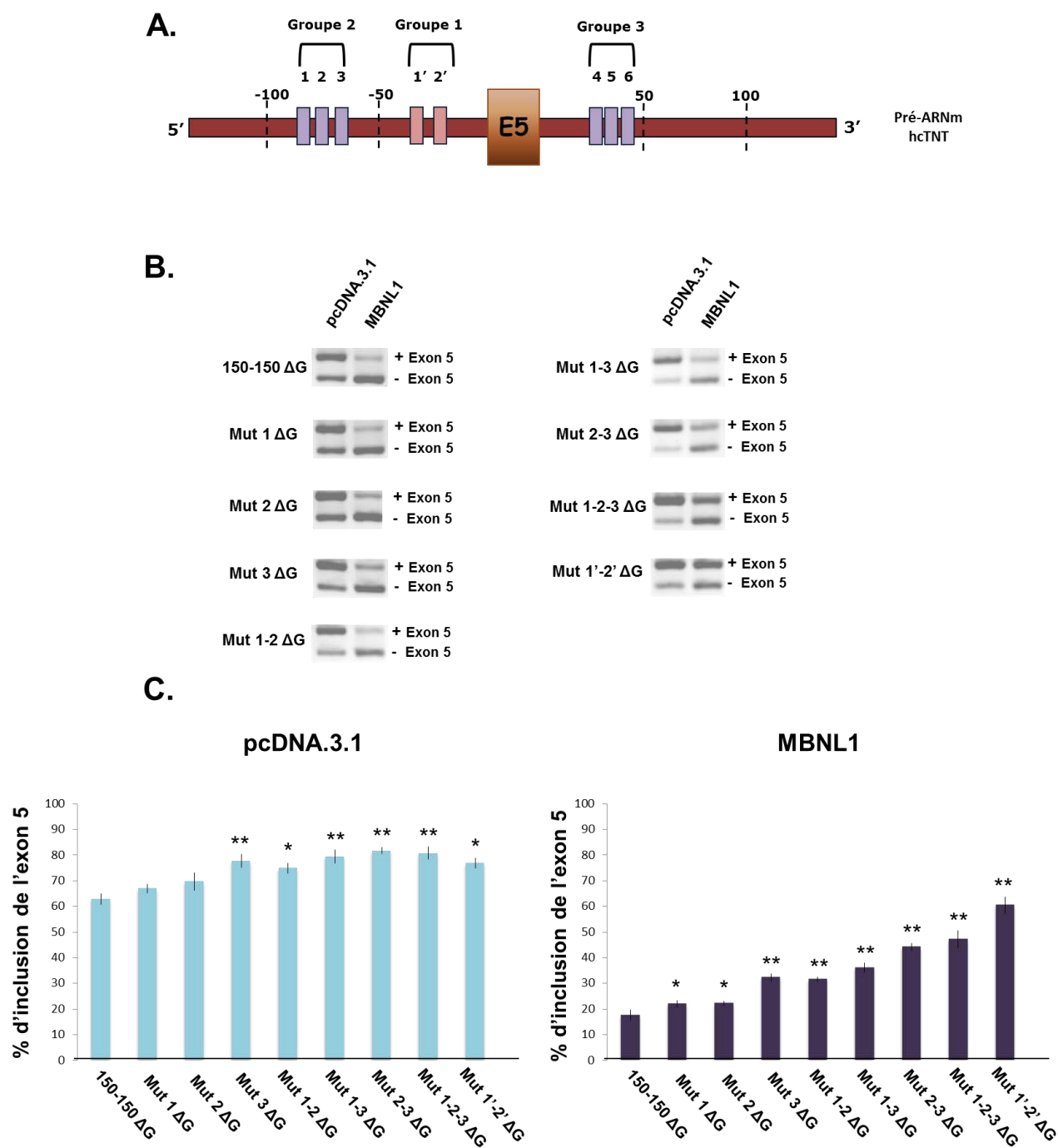


Figure III.3 : Analyse par RT-PCR de l'effet des mutations des sites 1, 2 et 3 du groupe 2 sur l'intégration de l'exon 5 dans l'ARNm produit à partir du minigène IR-hcTNT 150-E5-150 ΔG. **A.** Représentation de l'organisation du pré-ARNm issu du minigène IR-hcTNT 150-E5-150. **B.** L'effet des mutations individuelles ou en combinaison des sites 1, 2 et 3 sur l'intégration de l'exon 5 dans les transcrits du minigène IR-hcTNT a été analysé en co-transfectant la construction 150-150 ΔG dans des cellules HEK293 en présence des plasmides pcDNA.3.1-Flag et pcDNA.3.1-MBNL1 (MBNL1). Les ARN totaux ont été extraits 48 h après transfection et traités à la DNase, puis rétro-transcrits en utilisant l'oligodésoxynucléotide dT15 complémentaire à la queue polyA des ARNm. Les ADNc ainsi générés ont servi de matrice pour une réaction de polymérisation en chaîne (PCR) réalisée à l'aide des oligonucléotides 5991 et 8078. Les amplicons ont été séparés sur gel d'agarose à bas point de fusion à 4%. **C.** Une quantification des produits d'amplification contenant ou non l'exon 5 a été réalisée. Les données sont exprimées en pourcentage des ARN contenant l'exon 5 par rapport au total des ARN épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01 par rapport au valeurs 150-150 ΔG).

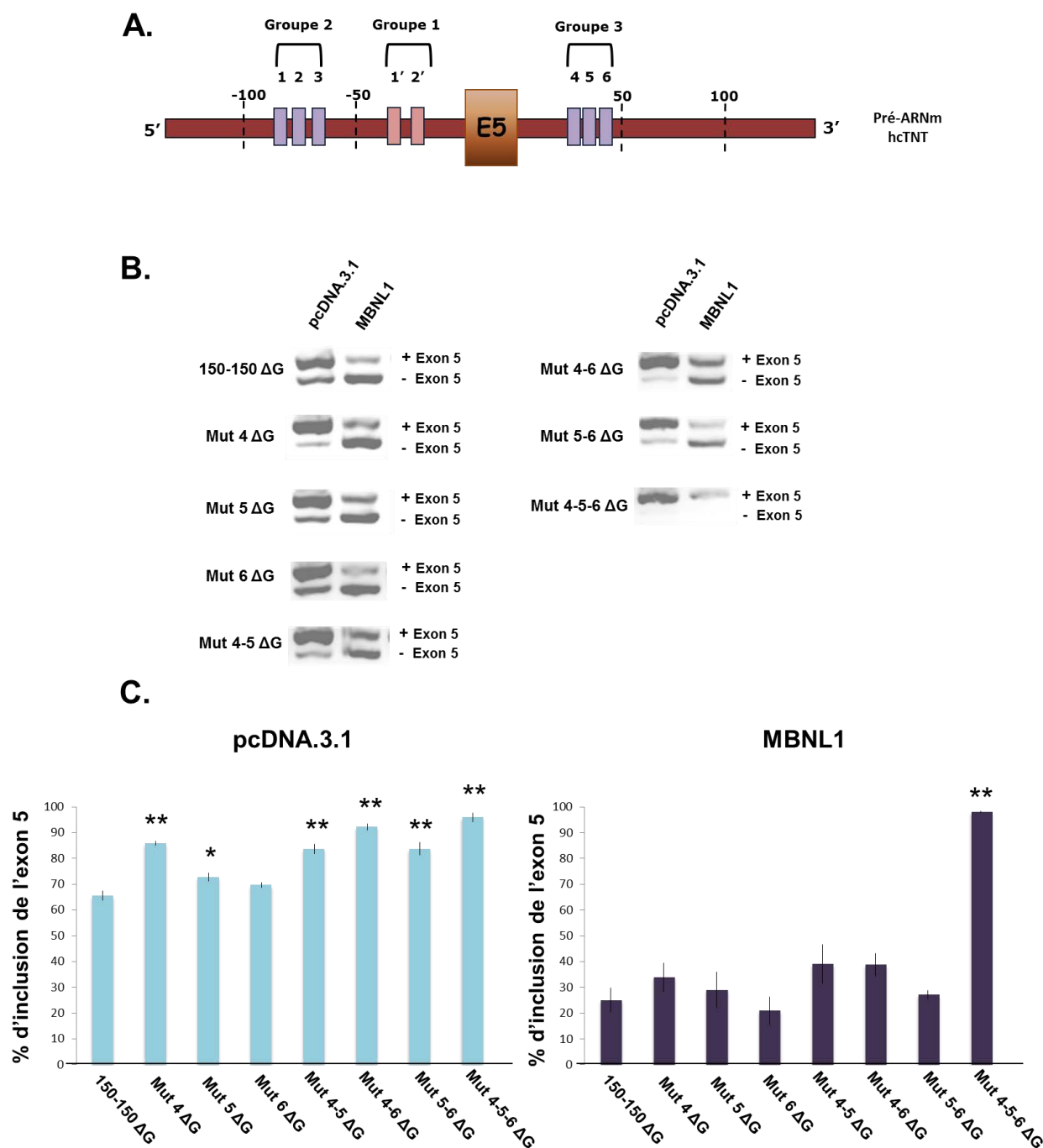


Figure III.4 : Analyse par RT-PCR de l'effet des mutations des sites 4, 5 et 6 du groupe 3 sur l'intégration de l'exon 5 dans l'ARNm produit à partir du minigène IR-hcTNT 150-E5-150 ΔG. **A.** Représentation de l'organisation du pré-ARNm issu du minigène IR-hcTNT 150-E5-150. **B.** L'effet des mutations individuelles ou en combinaison des sites 1, 2 et 3 sur l'intégration de l'exon 5 dans les transcrits du minigène IR-hcTNT a été analysé en co-transfectant la construction 150-150 ΔG dans des cellules HEK293 en présence des plasmides pcDNA.3.1-Flag et pcDNA.3.1-MBNL1 (MBNL1). Les ARN totaux ont été extraits 48 h après transfection et traités à la DNase, puis rétro-transcrits en utilisant l'oligodésoxynucléotide dT15 complémentaire à la queue polyA des ARNm. Les ADNc ainsi générés ont servi de matrice pour une réaction de polymérisation en chaîne (PCR) réalisée à l'aide des oligonucléotides 5991 et 8078. Les amplicons ont été séparés sur gel d'agarose à bas point de fusion à 4%. **C.** Une quantification des produits d'amplification contenant ou non l'exon 5 a été réalisée. Les données sont exprimées en pourcentage des ARN contenant l'exon 5 par rapport au total des ARN épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01 par rapport au valeurs 150-150 ΔG).

III.1.3. Les résultats obtenus *in vivo* par l'utilisation du minigène RTB300 différent de ceux obtenus avec le minigène IR-hcTNT.

Le minigène IR-hcTNT que nous avons utilisé lors des expériences précédemment décrites comporte les exons du récepteur de l'insuline de rat sous contrôle du promoteur SV40. Afin de vérifier que l'effet observé était indépendant du minigène utilisé, nous avons répété le même type d'expérience avec le minigène RTB300 (appelé aussi RTB300 300-E5-372) utilisé par les équipes de T. Cooper et A. Berglund. Ce minigène contient l'exon 5 hcTNT flanqué de séquences introniques plus longues (300 et 372 nts, respectivement, en amont et en aval de l'exon 5) qui sont bordées par des exons du gène *STNI* de poulet sous contrôle du promoteur RSV (Figure III.5A). Pour cette étude, nous avons demandé à T. Cooper de nous envoyer le minigène RTB300 et Christelle Aigueperse a introduit les triples mutations (YGCY→YACY) des 3 sites du groupe 2 (RTB300 mut 1-2-3) et des 3 sites du groupe 3 (RTB300 mut4-5-6 ΔG) dans ce minigène. En parallèle nous avons également testé les effets de la mutation (YGCY→YACY) des 2 sites du groupe 1 (RTB300 mut 1'2') au sein du minigène RTB300. Les différentes constructions utilisées pour cette étude sont représentées dans la figure III.5B.

L'expérience a été réalisée en utilisant le même protocole que pour l'étude réalisée avec le minigène IR-hcTNT, sauf que les oligonucléotides utilisés pour réaliser la PCR s'hybridaient dans les exons *STNI* de poulet et non dans les exons de l'insuline de rat. Comme attendu, les résultats de RT-PCR (Figure III.6) montraient que le minigène RTB300 300-E5-372 répondait de manière identique au minigène IR-hcTNT 150-E5-150 (IR 150-E5-150) dans les conditions de sur-expression de MBNL1. Par contre, nous observions que la mutation des sites des groupes 2 et 3 (RTB300 mut 1-2-3 et RTB300 mut 4-5-6 ΔG) dans le minigène RTB300 300-E5-372 n'affectait pas de façon significative le taux d'inclusion de l'exon 5 et n'inhibait pas la régulation de l'épissage par MBNL1. Il est à noter que l'effet de la mutation (YGCY→YACY) des sites 1' et 2' était partiel et similaire pour les deux minigènes, alors que la mutation (YGCY→YCCY) des sites 1' et 2' introduite par les équipes de T. Cooper et A. Berglund dans le minigène RTB300 300-E5-372 bloquait totalement la régulation. Ce qui renforçait l'idée de la contribution du renforcement de la séquence polypyrimidine de l'intron 4 hcTNT dans les effets observés par ces équipes.

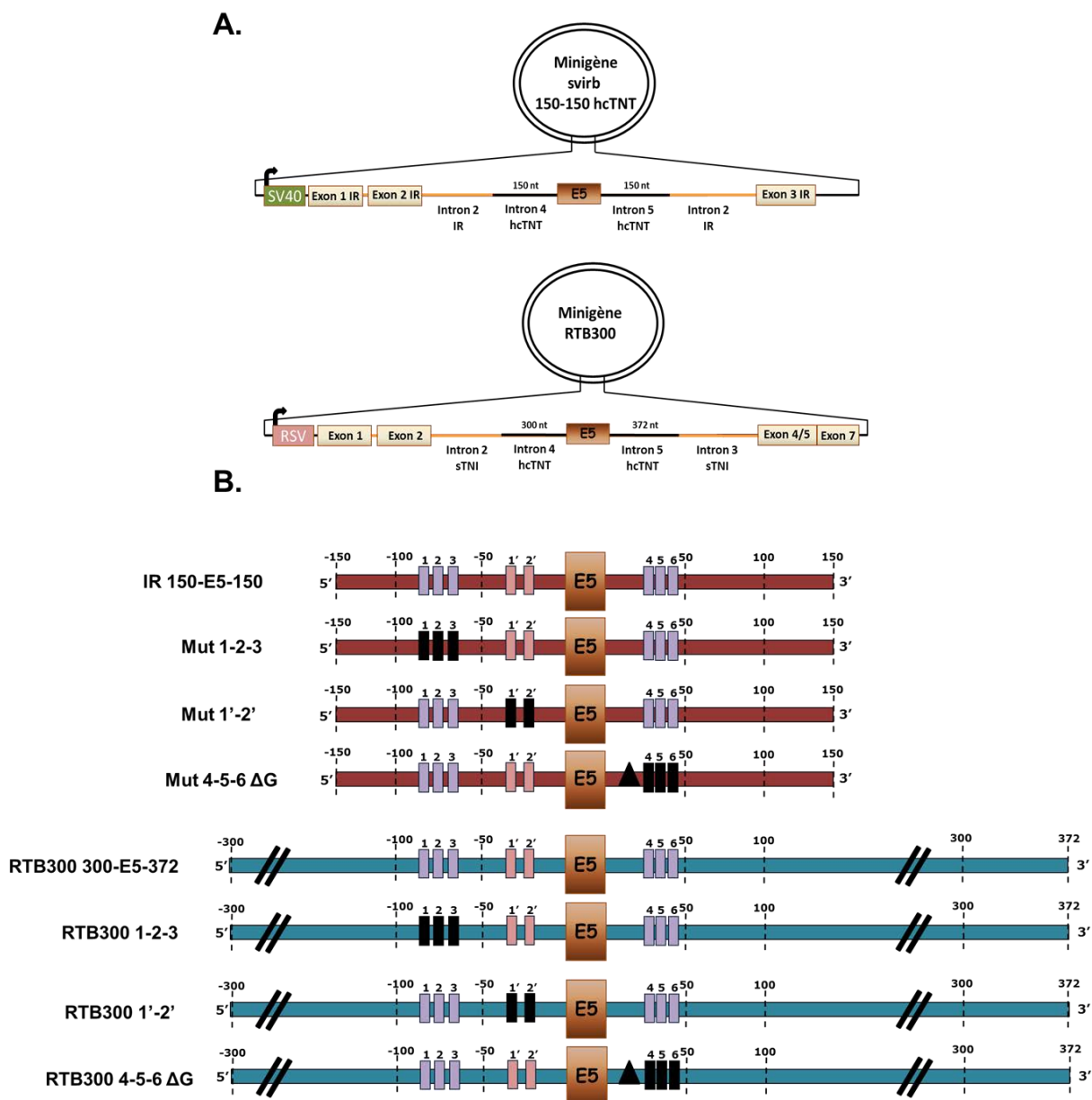
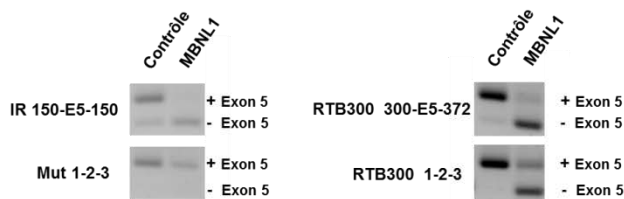


Figure III.5 : Description des différents mutants produits dans le minigène RTB300. A. Représentation des minigènes IR-hcTNT 150-E5-150 (svirb) et RTB300. Pour chaque construction les exons présents sont représentés en marron et la taille des séquences introniques bordant l'exon 5 du pré-ARNm hcTNT est indiquée. Le promoteur SV40 présent dans le minigène IR-hcTNT 150-E5-150 est indiqué en vert et le promoteur RSV présent dans le minigène RTB300 est indiqué en rose. **B.** Représentation des mutants ponctuels utilisés lors de cette étude. Les sites MBNL1 du groupe de sites 1 sont indiqués par des rectangles roses, et ceux des groupes 2 ou 3 par des rectangles violets. Les mutations générées au sein d'un site MBNL1 sont indiquées par un rectangle noir. La présence de la mutation ΔG est indiquée par un triangle noir. Le nom de la construction est indiqué à gauche.

A.



B.

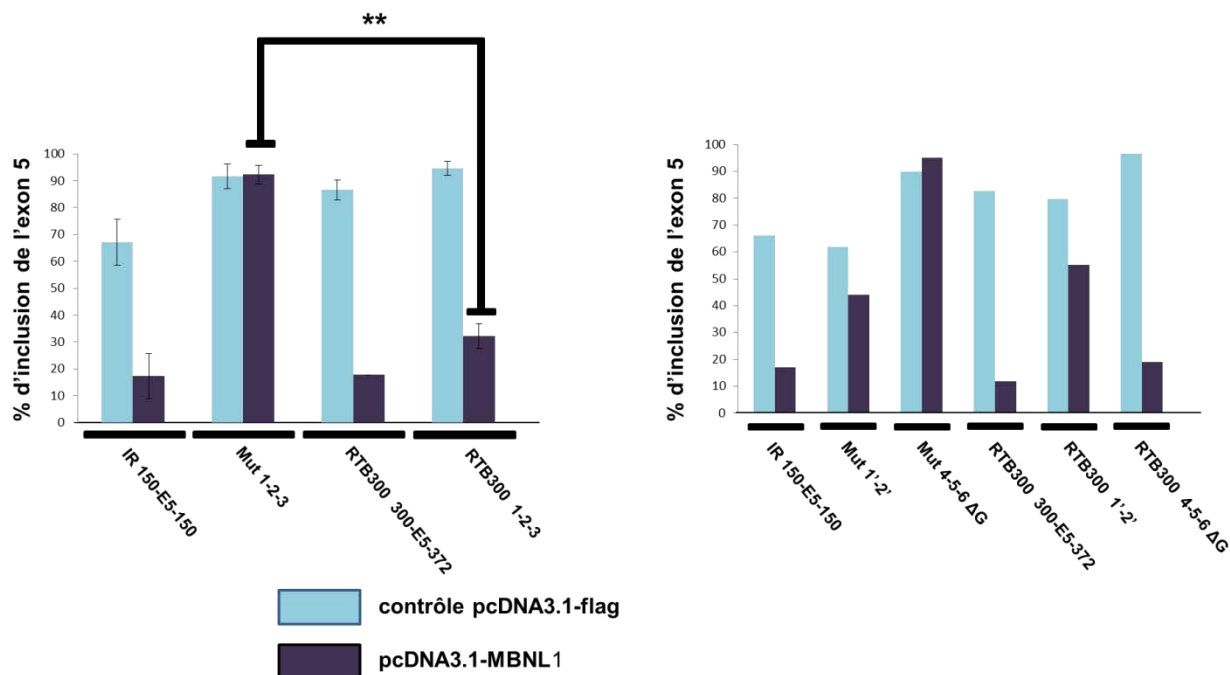


Figure III.6 : Analyse par RT-PCR de l'effet des mutations des sites MBNL1 des groupes 1, 2 et 3 sur l'intégration de l'exon 5 dans les ARNm produits à partir du minigène RTB300. A. L'effet des mutations de l'ensemble des sites MBNL1 de chaque groupe sur l'intégration de l'exon 5 dans les transcrits du minigène a été analysé en co-transfectant la construction RTB300 300-E5-372 ou RTB300 300-E5-372 ΔG dans des cellules HEK293 en présence des plasmides pcDNA.3.1-Flag (Contrôle en bleu) et pcDNA.3.1-MBNL1 (MBNL1 en violet). Les ARN totaux ont été extraits 48 h après transfection et traités à la DNase, puis rétro-transcrits en utilisant l'oligodésoxynucléotide dT15 complémentaire à la queue polyA des ARNm. Les ADNc ainsi générés ont servi de matrice pour une réaction de polymérisation en chaîne (PCR) réalisée à l'aide des oligonucléotides 8217 et 8218. Les amplicons ont été séparés sur gel d'agarose à bas point de fusion à 4%. Les résultats obtenus avec le minigène IR 150-E5-150 est aussi indiqué. **B.** Une quantification des produits d'amplification contenant ou non l'exon 5 a été exprimée en pourcentage du total des ARN épissés. Les résultats obtenus avec le minigène IR 150-E5-150 est aussi indiqué. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; ** :P<0.01 par rapport au valeurs de Mut 1-2-3).

La divergence de nos résultats suivant que nous utilisions notre construction ou celle de Cooper en ce qui concernait les sites des groupes 2 et 3 pouvait avoir plusieurs explications :

- le fait que les introns et exons bordants les séquences introniques et exoniques spécifiques de hcTNT au sein des 2 minigènes soient différentes et puissent présenter des sites de régulation supplémentaires qui altèrent la régulation de l'épissage de l'exon 5.
- le fait que le contexte génétique général (nature du promoteur transcriptionnel, de la région 3'-UTR, du site de polyadénylation..) des minigènes soit différent.
- le fait que la taille des séquences introniques hcTNT ne soit pas identique entre les deux constructions et qu'il soit possible que des séquences hcTNT supplémentaires présentes dans le minigène RTB300 interviennent dans la régulation de l'épissage de l'exon 5 dans des conditions de sur-expression de MBNL1.

Cette dernière hypothèse était d'autant plus plausible que nous avons observé que plusieurs sites de fixation potentiels de MBNL1 sont présents dans les séquences -300 à -150 du minigène RTB300 300-E5-372. Pour savoir si ces sites pouvaient être à l'origine des résultats obtenus avec le minigène RTB300, Christelle Aigueperse et Isabelle Behm-Ansmant ont préparé des constructions RTB300 placées dans le même contexte intronique que le minigène IR-hcTNT 150-E5-150, à savoir nous avons délété les séquences hcTNT introniques supplémentaires présentes dans le minigène RTB300 300-E5-372 afin d'obtenir une construction présentant 150 nucléotides hcTNT introniques en amont et en aval de l'exon 5 (RTB300 150-E5-150) (Figure III.7). En parallèle, une construction présentant 150 nucléotides en amont et 372 en aval a été réalisée (RTB300 150-E5-372). Pour chaque nouvelle construction, la triple mutation des sites de fixation de la protéine MBNL1 du groupe 2 (sites 1, 2 et 3) a été réalisée (RTB300 150-E5-150 1-2-3 et RTB300 150-E5-372 1-2-3). L'effet de cette triple mutation sur l'épissage des pré-ARNm issus des minigènes IR 150-E5-150, RTB300 150-E5-150 et RTB300 150-E5-372 a été comparé en absence ou en présence de sur-expression de la protéine MBNL1 (Figure III.8).

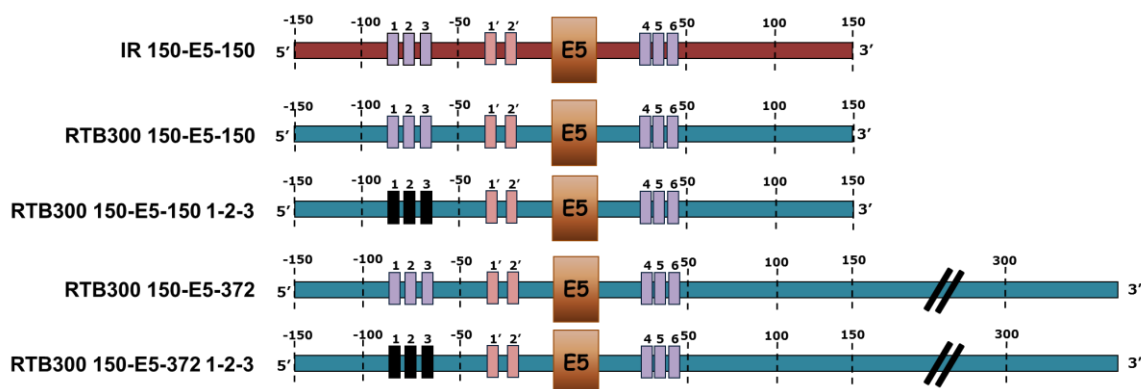


Figure III.7 : Description des différents mutants produits dans les minigènes RTB300 150-E5-150 et RTB 150-E5-372 utilisés lors de cette étude. Les sites MBNL1 du groupe de site 1 sont indiqués par des rectangles roses, et ceux des groupes 2 ou 3 par des rectangles violets. Les mutations générées au sein d'un site MBNL1 sont indiquées par un rectangle noir. Le nom de la construction est indiqué à gauche.

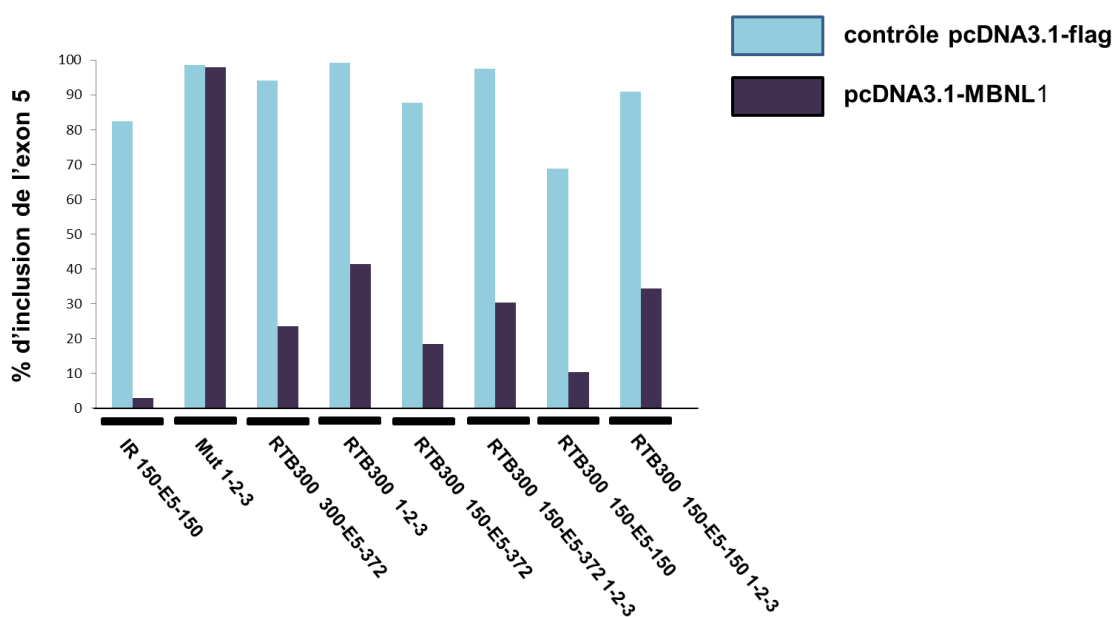


Figure III.8 : Analyse par RT-PCR de l'effet des mutations des sites MBNL1 des sites 1, 2 et 3 sur l'intégration de l'exon 5 dans les ARNm produits à partir des minigènes RTB300 150-150 et RTB300 150-372. L'effet des mutations des sites MBNL1 du groupe 2 sur l'intégration de l'exon 5 dans les transcrits du minigène a été analysé en co-transfectant la construction RTB300 150-150 ou RTB300 150-372 dans des cellules HEK293 avec des plasmides pcDNA.3.1-Flag (Contrôle en bleu) et pcDNA.3.1-MBNL1 (MBNL1 en violet). Les ARN totaux ont été extraits 48 h après transfection et traités à la DNase, puis rétro-transcrits en utilisant l'oligodésoxynucléotide dT15 complémentaire à la queue polyA des ARNm. Les ADNc ainsi générés ont servi de matrice pour une réaction de polymérisation en chaîne (PCR) réalisée à l'aide des oligonucléotides 8217 et 8218. Les amplicons ont été séparés sur gel d'agarose à bas point de fusion à 4%. Une quantification des produits d'amplification contenant ou non l'exon 5 a été exprimée en pourcentage du total des ARN épissés. Les résultats obtenus avec le minigène 150-150 sont aussi indiqués.

En absence de sur-expression de MBNL1 le taux d'inclusion de l'exon 5 était similaire pour les ARN produits à partir des constructions RTB300 300-E5-372 et RTB300 150-E5-372, par contre il était un peu plus faible pour la construction RTB300 150-E5-150. La mutation triple des 3 sites du groupe 2 avait peu d'effet sur l'inclusion de l'exon 5 dans le cas des ARN RTB300 300-E5-372 et RTB150-E5-372, tandis qu'elle diminuait légèrement l'inclusion dans la construction RTB300 150-E5-150 (15%). En présence d'une sur-expression de la protéine MBNL1, la triple mutation avait un effet légèrement positif sur l'inclusion de l'exon 5 pour les 3 constructions testées, en particulier pour l'ARN RTB300 150-E5-150. Néanmoins, nous n'observons pas un effet positif aussi marqué qu'avec l'ARN IR 150-E5-150 muté. Cette expérience n'a été répétée qu'une seule fois et devra être répétée deux fois pour confirmer les résultats. Elle montre d'ores et déjà un effet de la taille des séquences introniques dans l'ARN étudié, mais cet effet ne peut pas expliquer la forte différence de résultats que nous observons au niveau de l'effet de MBNL1 en présence des mutations des sites du groupe 2 entre les deux constructions. D'autres paramètres distinguant les deux constructions doivent intervenir (identité du promoteur transcriptionnel, exons hétérologues...) Nos observations soulignent à quel point des biais peuvent être générés par l'utilisation de minigènes. Afin de pouvoir conclure quant à l'importance des sites de fixation de MBNL1 que nous avons identifiés pour la régulation de l'inclusion de l'exon 5 dans les pré-ARNm hcTNT, nous avons décidé d'entreprendre une étude de l'inclusion de l'exon 5 au sein des ARNm hcTNT endogènes. Pour cela nous avons défini des oligos ARN antisens (ASO, pour AntiSense Oligonucleotides) chimiquement modifiés (2'O-méthylation des riboses et squelette phosphorothioate) permettant de cibler les sites MBNL1 des groupes 2 et 3. L'idée étant de bloquer l'utilisation de ces sites et de tester l'impact de ce blocage sur le taux d'inclusion de l'exon 5 des pré-ARNm hcTNT endogènes et ceci en présence ou non d'une sur-expression de la protéine MBNL1. Ces expériences sont actuellement en cours de réalisation et devraient permettre de conclure quant à l'impact réel des sites MBNL1 des groupes 2 et 3 dans la régulation de l'épissage de l'exon 5 du pré-ARNm hcTNT.

III.2. RECHERCHE DE NOUVEAUX REGULATEURS DE L'ÉPISSAGE ALTERNATIF DE L'EXON 5 DU pré-ARNm hcTNT.

Comme nous l'avons vu dans l'introduction de ce manuscrit en utilisant la construction IR-hcTNT 150-E5-150, notre équipe et celle de N. Sergeant avaient observé l'existence d'éléments activateurs et inhibiteurs d'épissage localisés dans les séquences introniques bordant l'exon 5 du pré-ARNm hcTNT et la nature des facteurs d'épissage se fixant sur ces éléments restait à déterminer. Afin d'essayer de les identifier, nous avons mis en œuvre une technique de chromatographie d'affinité des complexes ribonucléoprotéiques, qui avait été développée par l'équipe de R. Lurhmann du Max Planck Institut de Goettingen pour purifier des complexes pré-spliceosomiaux et largement utilisée au laboratoire pour l'étude de complexes ARN-protéines. Cette technique est basée sur l'utilisation d'une protéine de fusion MBP-MS2 et d'ARN hybrides contenant des motifs ARN MS2 ayant une haute affinité pour la protéine MS2.

III.2.1. Principe de la méthode de chromatographie d'affinité basé sur le système MS2/MBP.

Cette méthode repose sur l'insertion à l'extrémité 3' de l'ARN d'intérêt de trois motifs de l'ARN du bactériophage MS2 d'*E.coli* (motifs MS2). Ces motifs sont reconnus par le domaine MS2 de la protéine MBP-MS2 qui a une forte affinité pour ces motifs ARN (Kd de l'ordre du nM). Le domaine MBP a lui, une forte affinité pour le maltose et est de ce fait capable de se fixer de manière stable sur des billes d'amylose, ce qui permet ensuite d'immobiliser l'ARN hybride sur les billes auxquelles la protéine du fusion MBP-MS2 est fixée (Jurica et al., 2002). Nous avons donc produit un ARN hybride hcTNT-MS2 et ensuite formé des complexes entre cet ARN fixé sur les billes et ses partenaires présents dans un extrait nucléaire de cellules HeLa, ceci en se plaçant dans des conditions de tampon identiques à celles utilisées pour la réaction d'épissages *in vitro* (Figure III.9). Après plusieurs lavages destinés à éliminer les protéines liées de manière aspécifique, les protéines nucléaires liées à l'ARN ont été éluées des billes d'amylose par ajout de maltose. Elles ont ensuite été fractionnées par électrophorèse en gel SDS-PAGE. Après coloration au bleu de coomassie colloïdal, les bandes de gel ont été découpées et les protéines qu'elles contenaient analysées par spectrométrie de masse par l'équipe d'A. Van Dorsselaar (LSMBO, Starbourg), afin d'identifier les partenaires protéiques de l'ARN.

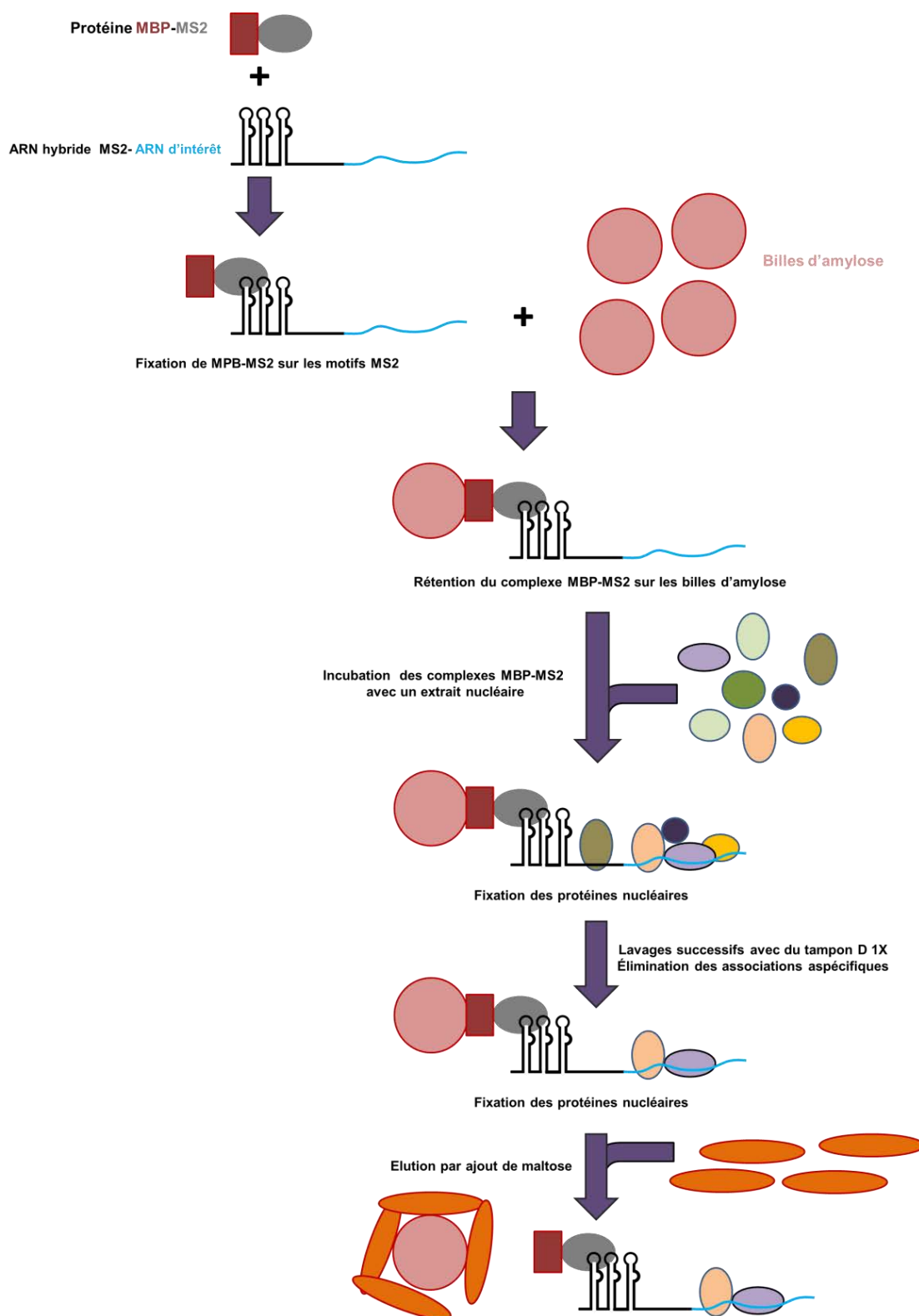


Figure III.9 : Schéma du principe de la purification des complexes ribonucléoprotéiques par chromatographie d'affinité basée sur l'emploi de la protéine de l'enveloppe du phage MS2 et de la protéine MBP. Un ARN hybride composé de 3 motifs MS2 et d'une séquence d'ARN d'intérêt est retenu sur des billes d'amylose par l'intermédiaire de la protéine MBP-MS2. L'ensemble est incubé en présence d'un extrait nucléaire de cellules HeLa. Les complexes générés entre l'ARN hybride et les protéines nucléaires sont ensuite lavés avec du tampon D 1X afin d'éliminer les interactions aspécifiques, puis élués par ajout de maltose.

III.2.2. Construction des plasmides puc18-hcTNT-MS2 permettant de produire par transcription *in vitro* des fragments de 50 nucléotides du pré-ARNm hcTNT fusionnés à 3 motifs MS2.

Afin de limiter les interactions aspécifiques avec les protéines nucléaires et identifier plus précisément la localisation des partenaires des séquences introniques hcTNT, nous avons généré des constructions produisant des ARN hybrides hcTNT-MS2 contenant en 5' des fragments de 50 nucléotides hcTNT et en 3' les motifs MS2. Les 6 fragments A, B, C, D, E et F couvraient donc l'ensemble des séquences introniques de l'ARN 150-E5-150 (Figure III.10). Comme explicité dans le chapitre Matériel et Méthodes, ces différents fragments ont été amplifiés par PCR en utilisant les couples d'oligodésoxynucléotides décrits dans le tableau 10 en annexe, puis clonés dans le plasmide pUC18-T7-MS2 immédiatement en amont de la séquence codant les 3 motifs MS2. La séquence des différents plasmides générés a été vérifiée par séquençage. La transcription *in vitro* des ARN hcTNT (A, B, C, D, E et F)-MS2 a été réalisée dans les conditions décrites dans la partie Matériel et Méthodes.

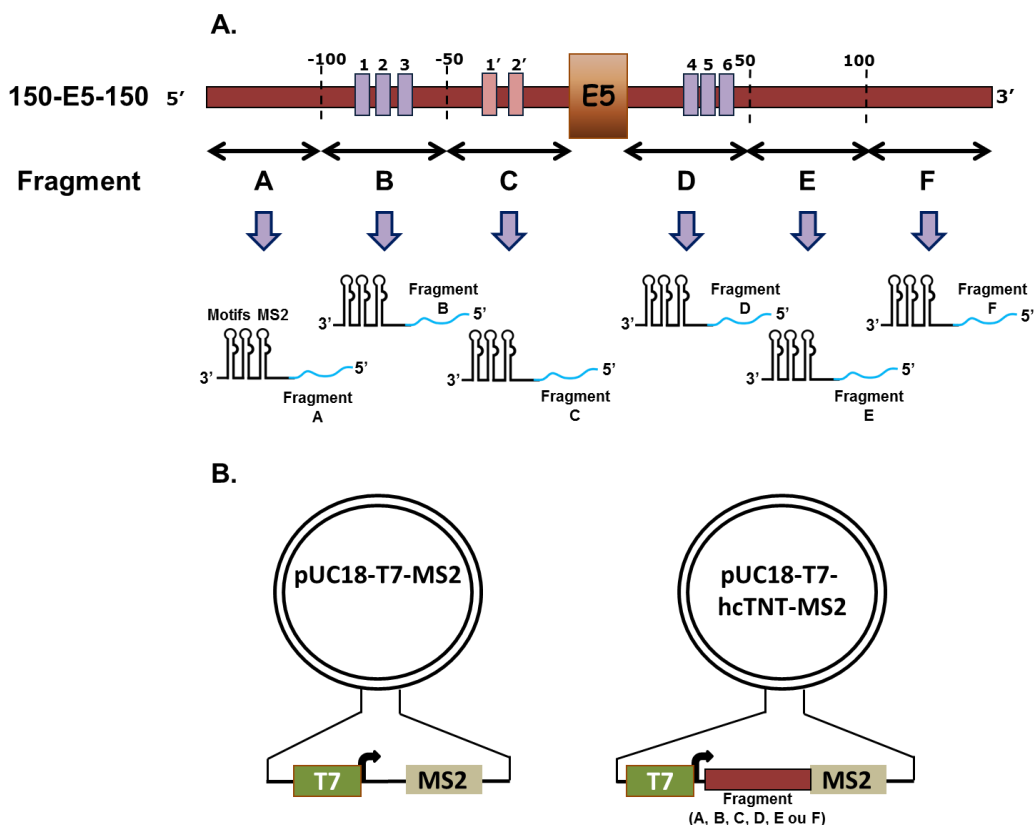


Figure III.10 : Schéma des constructions réalisées pour les expériences de purification par chromatographie d'affinité basée sur l'utilisation de la protéine MBP-MS2. A. Représentation schématique des fragments A, B, C, D, E et F, issus de l'ARN 150-E5-150, utilisés pour la purification. Chaque fragment a été inséré au sein du plasmide pUC18-T7-MS2 en amont d'une séquence codant les motifs MS2. Après transcription *in vitro* par l'ARN polymérase du phage T7, les différents ARN hybrides, composés des motifs MS2 associés aux différents fragments (A, B, C, D, E ou F), ont été purifiés sur colonne. **B.** Représentations schématiques des constructions pUC18-T7-MS2 et pUC-T7-hcTNT-MS2. Le promoteur T7 permettant la transcription des ARN d'intérêt est représenté par un rectangle vert et une flèche noire indique le début de la transcription. La séquence ADN permettant de générer les motifs MS2 est indiquée par un rectangle marron. Le plasmide pUC-T7-hcTNT-MS2 présente une séquence supplémentaire en amont de la séquence MS2 qui code l'un des fragments A, B, C, D, E ou F (rectangle rouge).

III.2.3. Purification des complexes ribonucléoprotéiques formés avec les ARN hcTNT (A, B, C, D, E et F)-MS2 et identification de la protéine hnRNP H.

Pour réaliser cette étude, 300 pmol de chaque ARN hcTNT-MS2 (A, B, C, D, E et F-MS2) ont été incubées avec 1250 pmol de la protéine MBP-MS2, puis fixées sur une résine d'amylose. Les complexes obtenus ont été ensuite incubés avec 250 μ L d'extrait nucléaire de cellule HeLa compétent pour l'épissage. Les complexes ribonucléoprotéiques (RNP) ainsi formés ont été purifiés par lavages successifs, puis élués en présence de maltose. Plusieurs contrôles ont été réalisés afin de s'assurer de la qualité de la fixation de la protéine MPB-MS2 sur les ARN hcTNT (A, B, C, D, E et F)-MS2, de l'efficacité de la rétention des ARN et protéines sur la résine amylose, de l'efficacité des lavages et de la qualité des éluions. Un exemple d'analyse est illustré dans la figure III.11A. En parallèle, une expérience contrôle a été réalisée avec l'ARN MS2 ne comportant que les trois motifs MS2, ceci afin d'identifier les protéines nucléaires capables de se lier à ces trois motifs et de définir lesquelles des protéines présentes dans les complexes se fixaient spécifiquement sur les fragments ARN d'intérêt. Le contenu en protéines des fractions d'éluion a été analysé par électrophorèse en conditions dénaturantes (SDS-PAGE), suivie d'une coloration au bleu de Coomassie colloidal. Les profils de fractionnement des protéines obtenus pour chaque ARN de fusion présentaient des différences nettes et suggéraient la présence de protéines spécifiquement associées à chacun des fragments d'ARN (Figure III.11B). Par conséquent, les bandes de gel ont été découpées, puis envoyées à Strasbourg pour être analysées par spectrophotométrie de masse par S. Sanglier et JM. Saliou de l'équipe d'A. Van Dorsselaar (LSMBO, Starbourg). Cependant, lors de l'analyse un problème technique sur le spectromètre de masse a entraîné une différence de traitement entre la condition témoin ARN MS2 et les conditions réalisées avec les ARN hydrides hcTNT (A, B, C, D, E et F)-MS2. De ce fait il n'a pas été possible d'interpréter les résultats. Etant donné que nous disposions au laboratoire d'une collection d'anticorps dirigés contre un grand nombre de protéines impliquées dans les régulations de l'épissage, nous avons décidé de les utiliser pour identifier les protéines présentes dans les fractions d'éluion de l'expérience de chromatographie d'affinité.

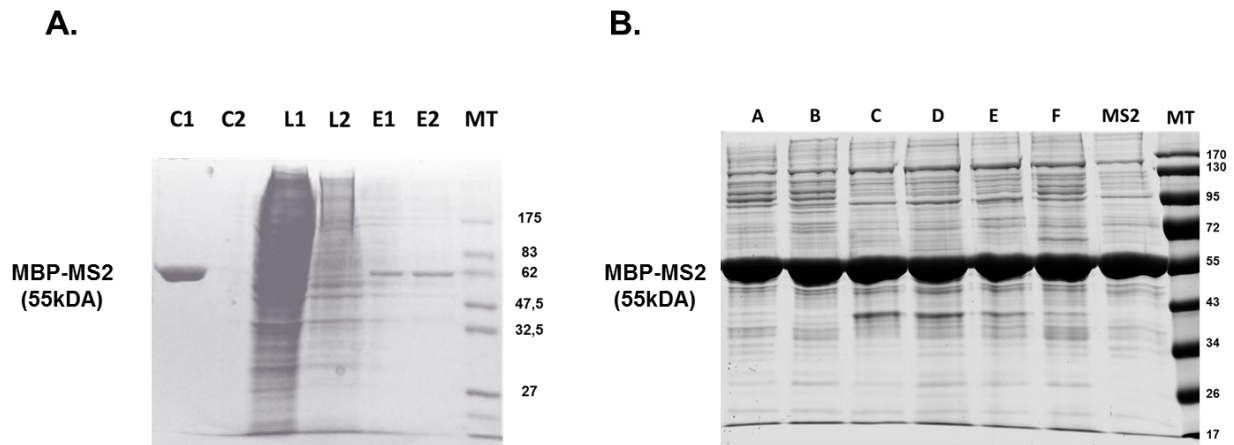


Figure III.11 : Purifications des RNP et des protéines par chromatographie d'affinité sur billes d'amylose. A. Vérification préalable réalisée avant la purification MBP-MS2. Les différentes fractions collectées durant l'expérience ont été fractionnées par électrophorèse en gel de polyacrylamide SDS-PAGE 10%. C1 représente la fraction obtenue après incubation de la protéine MBP-MS2 et de l'ARN hybride hcTNT-MS2. C2 représente la fraction non retenue après incubation avec les billes d'amylose. L1 et L2 sont les fractions issues des différents lavages des billes. E1 et E2, les fractions d'élution successives par une solution de maltose. **B.** Fractionnement par électrophorèse en gel SDS-PAGE 10% des protéines présentes dans les complexes ribonucléoprotéiques obtenus avec les ARN hybrides hcTNT (A, B, C, D, E et F)-MS2 et le contrôle correspondant à l'ARN MS2. Le gel a été coloré au bleu de Coomassie colloïdal. Le poids moléculaire des protéines du marqueur de taille (MT) est indiqué à droite en kDa. Le signal correspondant à la protéine MBP-MS2 est indiqué à gauche.

III.2.3.1. Identification par Western blot de la protéine hnRNP H en association avec des fragments hcTNT.

J'ai donc analysé les fractions d'élution obtenues avec chacun des ARN de fusion par expérience de Western blot en utilisant des anticorps dirigés contre des protéines SR et des protéines hnRNP. Pour cela, les complexes obtenus par élution au maltose après chromatographie par MBP-MS2 pour chacun des fragments d'ARN testés ont été fractionnés par électrophorèse en gel SDS-PAGE à 10%, puis transférés sur une membrane de nitrocellulose et traités selon le protocole décrit dans la partie Matériel et Méthodes. Les résultats de cette analyse sont présentés dans la Figure III.12. Ils montraient que les protéines hnRNP U et K interagissent avec chacun des ARN hcTNT-MS2 testés, mais également avec les trois motifs MS2 et ceci en absence de fusion avec des fragments hcTNT d'intérêt, montrant que les associations observées avec les fragments hcTNT sont probablement aspécifiques.

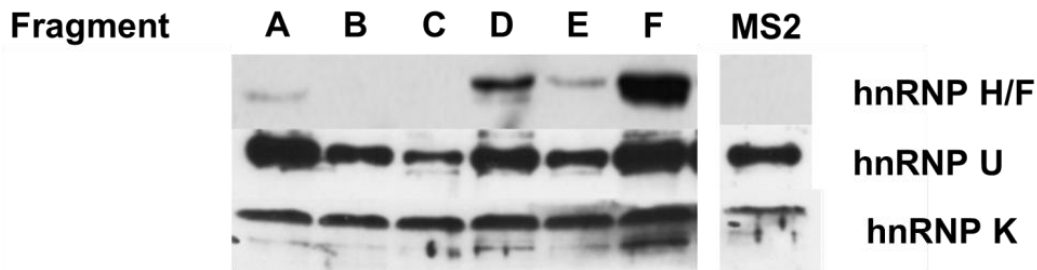


Figure III.12 : Analyse des fractions d'éluion obtenues avec chacun des ARN de fusion par Western blot. 10 μ L des éluats contenant les RNP formées sur les fragments A, B, C, D, E et F fusionnés à l'ARN MBP-MS2 et sur l'ARN MS2 seul ont été fractionnés sur un gel SDS-PAGE 10%. Les protéines ainsi fractionnées ont été transférées sur une membrane de nitrocellulose qui a été incubée avec des anticorps dirigés contre les facteurs d'épissage hnRNP H/F, U et K. MS2 représente le contrôle correspondant à l'ARN avec uniquement les trois motifs MS2.

Les résultats mettaient également en évidence une association du facteur d'épissage hnRNP H avec surtout les fragments D et F, le taux d'association étant très fort pour le fragment F. Dans la mesure où aucune association de hnRNP H n'était détectée avec l'ARN MS2, l'association de la protéine hnRNP H avec les fragments A, D, E et F semblait spécifique. De façon intéressante, la protéine hnRNP H avait été montrée être impliquée dans la rétention nucléaire des répétitions CUG présentes au sein de la région 3'UTR de l'ARNm DMPK, et a été trouvée surexprimée dans les myoblastes de patients atteints de DM1 (Kim et al., 2005; Paul et al., 2006). Par conséquent, dans la mesure où un lien semblait exister entre la protéine hnRNP H et la DM1 et que nous détectons une interaction de la protéine avec l'ARN hcTNT, nous avons décidé de focaliser notre étude cette protéine et entrepris de tester son implication possible dans la régulation de l'inclusion de l'exon 5 au des ARNm hcTNT.

III.2.3.2. Etude de l'interaction de la protéine hnRNP H avec l'ARN hcTNT 150-E5-150 et ses sous-fragments A, B, C, D, E et F.

Les expériences MBP-MS2 identifient des protéines associées à un fragment d'ARN, mais cela n'implique pas forcément une interaction directe entre l'ARN et les protéines, l'association détectée pouvait en effet résulter d'interactions protéine-protéine. Dans un premier temps, j'ai donc testé par des expériences de retard sur gel si la protéine hnRNP H était capable de se fixer directement sur les fragments A, D, E et F. Pour cela, j'ai transcrit *in vitro* les différents fragments A, B, C, D, E et F, ainsi qu'un fragment correspondant à l'exon 5

et un grand fragment correspondant à l'ensemble de l'insert hCNT 150-E5-150, ceci sans fusion avec l'ARN MS2. Les ARN ainsi transcrits ont été déphosphorylés, puis marqués radioactivement en 5' par emploi de [γ^{32} -P]ATP et de T4 polynucléotide kinase. L'affinité de la protéine hnRNP H recombinante a été testée en formant des complexes avec les ARN radiomarqués (20 cps) en présence de concentrations croissantes en protéine hnRNP H-Nter (concentrations allant de 0.21 μ M à 8.5 μ M). La protéine hnRNP H recombinante utilisée était disponible au laboratoire et ne comportait que la partie N-terminale de la protéine qui contient les 2 domaines qRRM nécessaires à la fixation sur les ARN. Les complexes formés entre les différents ARN étudiés et la protéine hnRNP H-Nter ont été fractionnés par électrophorèse en gel natif. Les résultats obtenus sont présentés dans la figure III.13.

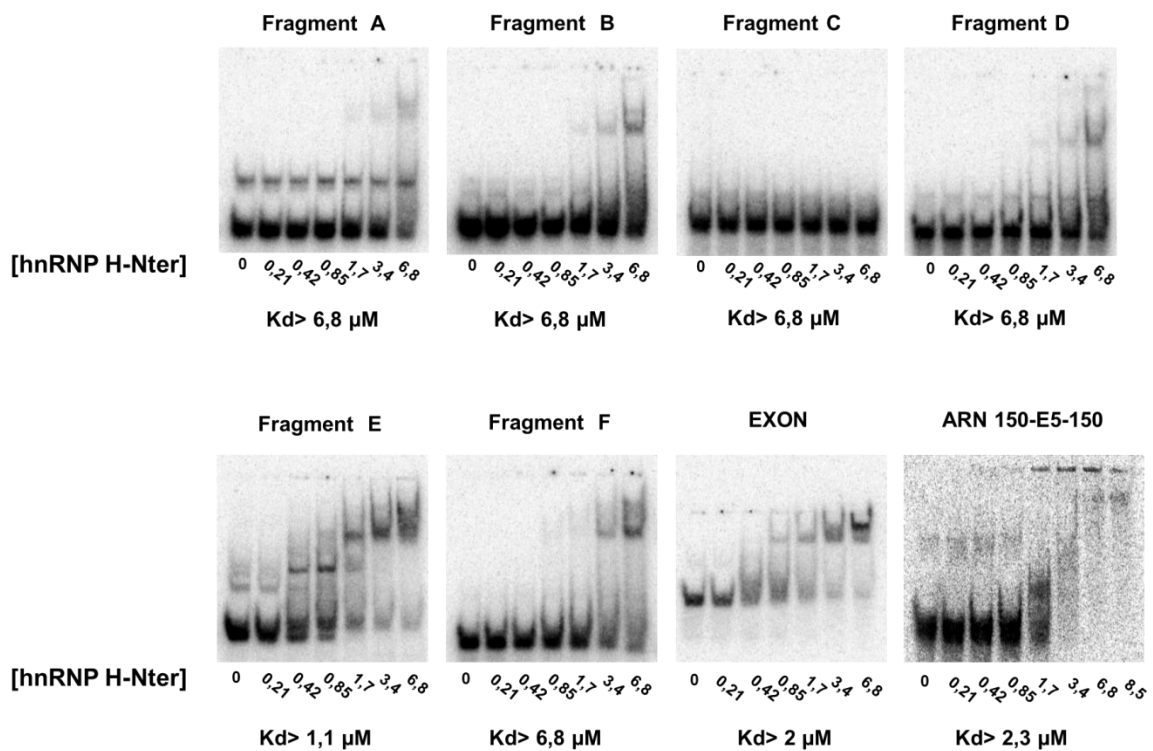


Figure III.13 : Etude de l'interaction de la protéine hnRNP H-Nter avec l'ARN 150-E5-150 et ses sous fragments. Test de fixation de la protéine hnRNP H-Nter par retard sur gel sur chacun des fragments A, B, C, D, E et F. La fixation de hnRNP H-Nter sur l'ARN 150-E5-150 et l'ARN correspondant à l'exon a également été testée. La protéine hnRNP H-Nter ne contenait que les domaines de fixation à l'ARN RRMH 1 et 2 de la protéine hnRNP H entière. Une gamme de concentration de protéine allant de 0 à 8.5 μ M a été testée avec 2 fmol d'ARN radiomarqué (20 cps).

L'analyse révélait qu'à forte concentration la protéine hnRNP H-Nter est capable de se lier à chacun des ARN A, B, D, E et F ainsi que sur l'exon. Seul le fragment d'ARN C n'interagissait pas du tout. Point intéressant, le fragment d'ARN E et l'exon 5 présentaient une affinité suggérant l'existence d'une interaction spécifique avec respectivement, un K_d de l'ordre de 1.5 et 2 μM . Dans le cas de l'ARN E plusieurs complexes ARN-protéines étaient clairement observés indiquant la fixation de plusieurs molécules de protéine hnRNP H-Nter sur cet ARN. La formation de ces différents complexes pouvait résulter soit de la présence de plusieurs sites de fixation au sein de l'ARN E, soit de l'oligomérisation de la protéine après reconnaissance d'un site de fixation unique au sein de l'ARN.

Comme mentionné précédemment, l'analyse de la structure 2D de l'ARN 150-E5-150 avait révélé un fort taux de structuration de l'ARN et l'existence d'appariements impliquant les séquences introniques situées en amont et en aval de l'exon 5. Il était par conséquent peu probable que la structure secondaire des fragments A, B, C, D, E et F soit la même que celle de ces mêmes segments au sein de l'ARN 150-E5-150. La protéine hnRNP H se fixe préférentiellement sur des séquences ARN simple-brin riches en résidus G pouvant se structurer en quadruplexes de G, par conséquent la fixation d'hnRNP H pouvait être affectée voire impossible au sein de l'ARN 150-E5-150. C'est pourquoi j'ai vérifié, toujours par des expériences de retard sur gel, que la protéine hnRNP H-Nter était capable de se fixer sur l'ARN 150-E5-150. Les résultats obtenus sont illustrés dans la figure III.13. Nous observons une fixation de la protéine hnRNP H-Nter sur l'ARN 150-E5-150 avec un K_d de 2.3 μM . Ce K_d est supérieur à celui observé pour le fragment E et l'exon 5, suggérant que la structure 2D du pré-ARNm hcTNT limite l'affinité de hnRNP H-Nter. Les sites de fixation de hnRNP H-Nter au sein de l'ARN 150-E5-150 ont été identifiés par des expériences d'empreintes en solution réalisées en présence de 2 ou 4 μM de protéine hnRNP H-Nter. Le protocole expérimental utilisé pour ces expériences est celui décrit dans le chapitre méthodologique que nous avons publié cette année dans *Method in Molecular Biology*, dont je suis premier auteur, et qui est inséré dans la partie Matériel et Méthodes de ce manuscrit (Philippe et al., 2015). Les résultats de cette analyse sont présentés dans la figure III.14. Une protection très forte du segment d'ARN allant des positions +116 à +125 (Fragment F) et un effet plus modéré sur l'accessibilité des nucléotides +95 à +97 (extrémité 3' du fragment E) étaient observés (Figure III.14A et B).

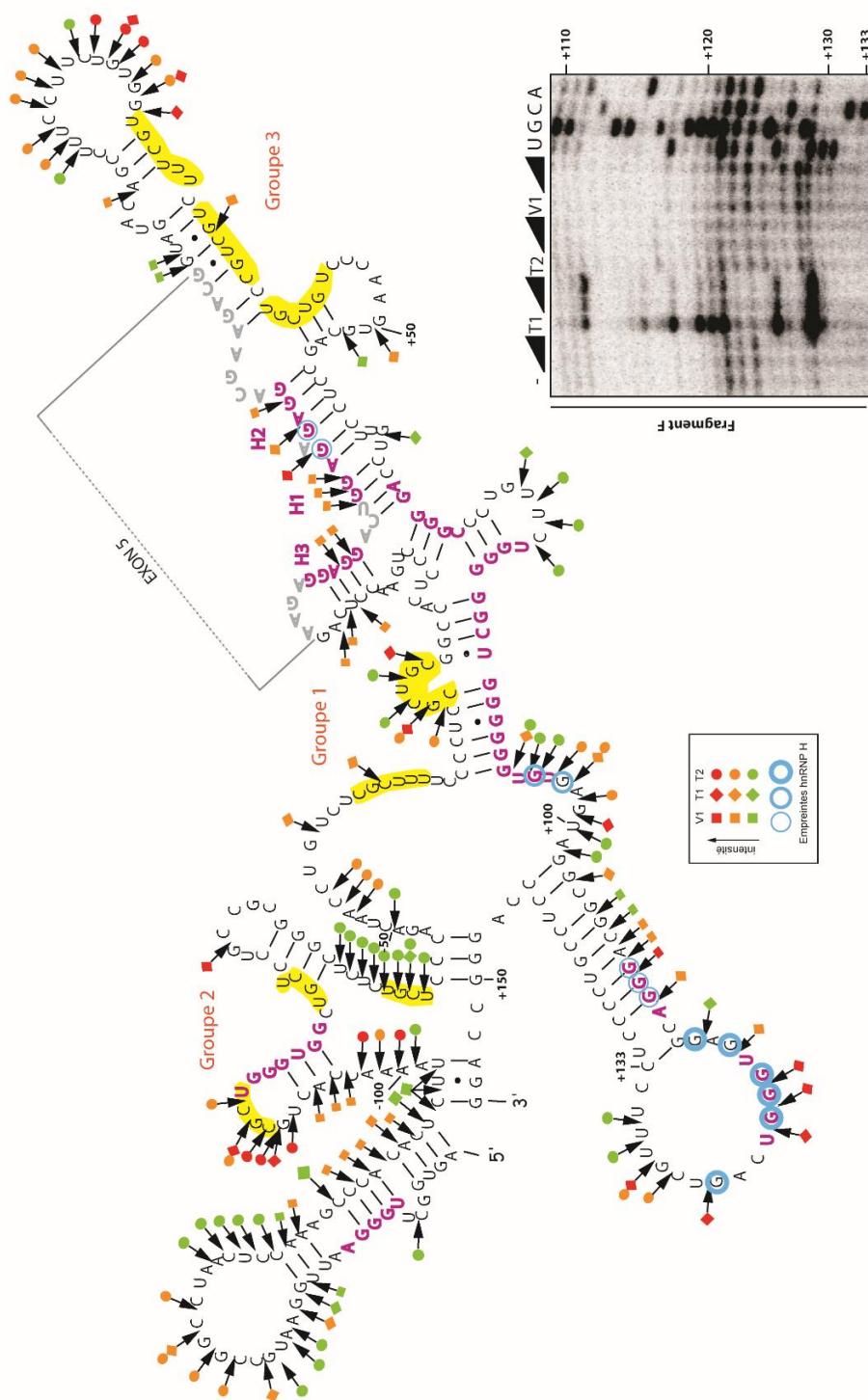


Figure III.14 : Analyse des protections générées par la fixation de la protéine hnRNP H-Nter sur l'ARN hCNT 150-E-150. Les clivages par les RNases V1, T1 et T2 sont représentés par des carrés, des losanges et des ronds, respectivement. La couleur des symboles représente l'intensité du clivage (faible en vert, moyenne en orange et forte en rouge). La séquence de l'exon 5 est indiquée en gris. Les huit sites MBNL1 sont surlignés en jaune et les sites potentiels de fixations de hnRNP H sont indiqués en violet. Les protections générées suite à l'incubation de l'ARN avec 2 ou 4 μM de protéine hnRNP H-Nter sont indiquées par des cercles dont l'épaisseur est proportionnelle à l'intensité de la protection. L'analyse d'un segment du fragment F est illustrée. Les positions des coupures ont été identifiées par extension d'amorce. Les arrêts d'extension ont été positionnés dans la séquence grâce à un séquençage du transcrit (pistes U, G, C et A). Les ADNc ont été fractionnés par électrophorèse en gel de polyacrylamide dénaturant à 7 %. Les positions des nucléotides sont indiquées à gauche de l'autoradiographie.

Aucune conclusion ne pouvait être tirée sur la fixation de hnRNP H-Nter au reste du fragment E car 90% de ce segment d'ARN présentait une très faible accessibilité aux RNases et il était par conséquent impossible de conclure quant à une éventuelle interférence de l'association de hnRNP H-Nter avec la digestion de ce segment d'ARN par les RNases T1, T2 et V1 (Figure III.14A et B). En accord avec les expériences de retard sur gel effectuées sur les fragments d'ARN hcTNT de 50 nucléotides de long, aucun effet de la fixation de hnRNP H-Nter sur le profil d'hydrolyse des segments d'ARN correspondant aux fragments A, B, C et D n'est observé, alors que l'accessibilité de quelques résidus appartenant à l'exon 5 était modérément affectée (Figure III.14). De façon intéressante et en accord avec la propriété de la protéine hnRNP H de se fixer spécifiquement sur des séquences contenant des motifs poly-G, les zones de protections détectées au sein de l'ARN 150-E5-150 correspondaient à des enchainements de résidus G. L'analyse complète de la séquence de l'ARN 150-E5-150 révélait la présence de motifs poly-G dans les fragments A, B, E et F (Figure III.14B). Aucun motif poly-G n'est par contre présent dans les fragments C et D et dans la séquence codant l'exon 5. Cependant, la présence d'enchaînements de type GGAGAGG au sein de la séquence codant l'exon 5 pourrait expliquer la fixation de hnRNP H, d'autant plus que de tels enchainements peuvent conduire à la formation de quadruplexes de G (Blice-Baum et Mihailescu, 2014). En conclusion, par ces expériences d'empreintes et de retard sur gel, nous confirmons la fixation directe de la protéine hnRNP H-Nter sur les segments d'ARN 150-E5-150 correspondants à l'exon 5 et au fragment F. La fixation sur le fragment E bien que très forte dans les expériences de retard sur gel n'a pu être confirmée par cette approche d'empreintes. La faible fixation de la protéine hnRNP H sur les fragments A et D, détectée lors de la chromatographie d'affinité ne semble donc pas directe mais médiée par d'autres protéines de l'extrait nucléaire.

III.3. ETUDE *IN CELLULO* DE L'EFFET DE LA VARIATION DE LA QUANTITE DE hnRNP H SUR L'INTEGRATION DE L'EXON 5 DU pré-ARNm hcTNT.

Une question importante était de savoir si la fixation de la protéine hnRNP H sur le pré-ARNm hcTNT pouvait intervenir dans la régulation de l'épissage de l'exon 5. Pour cela, nous avons réalisé des études *in cellulo* dans des cellules humaine HEK293, afin d'analyser l'effet d'une sur-expression ou d'une diminution de la quantité de protéine hnRNP H sur l'intégration de l'exon 5 dans le transcrit du minigène IR-hcTNT.

III.3.1. La sur-expression de la protéine hnRNP h dans des cellules HEK293T entraine une augmentation du taux d'inclusion de l'exon 5 du pré-ARNm hcTNT.

L'effet de la sur-expression de la protéine hnRNP H a été testé en co-transfectant des cellules HEK293 avec des quantités croissantes de plasmide pXJ41-hnRNP H (0, 0.05, 0.1, 0.2, 0.4 et 1 µg) avec 1,5 µg du minigène IR-hcTNT, ceci selon le protocole décrit dans la partie Matériel et Méthodes. Le plasmide pXJ41 vide qui ne permet pas la sur-expression de la protéine hnRNP H a été utilisé comme contrôle. L'intégration de l'exon 5 dans les produits d'épissage a été analysée par RT-PCR à partir des ARN totaux préparés 48 h après la transfection. Les résultats obtenus sont illustrés dans la figure III.15. La sur-expression de la protéine hnRNP H a été vérifiée par Western blot (Figure III.15A). Les résultats de RT-PCR montrent une augmentation de l'inclusion de l'exon 5 qui passe d'un taux d'inclusion de 49 % en absence de sur-expression de la protéine hnRNP H, à 62 % lors de sa sur-expression et ceci dès la co-transfection de la plus faible quantité de plasmide pXJ41-hnRNP H. L'augmentation observée lorsque des quantités plus importantes de plasmides pXJ41-hnRNP H étaient transfectées était faible, de l'ordre de quelques % (Figure III.15C). Ces résultats indiquaient que la quantité de protéine hnRNP H dans la cellule influence le taux d'intégration de l'exon 5 hcTNT dans le contexte du transcrit du minigène IR-hcTNT et que la protéine hnRNP H est un activateur de l'inclusion de cet exon.

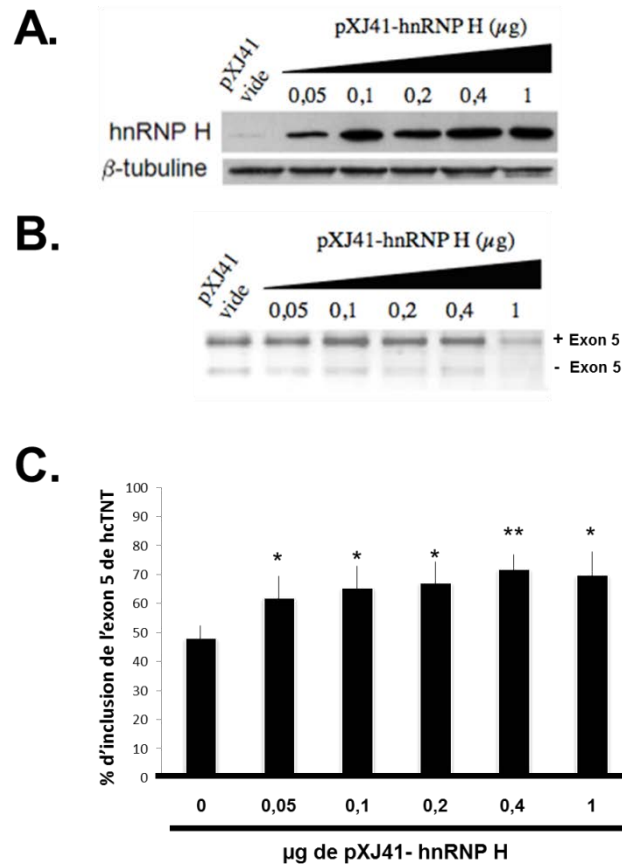


Figure III.15 : Etude de l'effet de la sur-expression de la protéine hnRNP H sur l'inclusion de l'exon 5 hcTNT dans le transcrit du minigène IR-hcTNT. 48 h après la co-transfection des constructions pXJ41-hnRNP H ou vide et du minigène IR-hcTNT 150-E5-150, les ARN totaux ont été extraits puis analysés par RT-PCR avec les oligodésoxynucléotides 5991 et 8078. Un contrôle a été réalisé en transfectant 1,5 μ g du plasmide pXJ41 vide (0). **A.** Vérification par Western blot de la sur-expression de la protéine hnRNP H. **B.** Fractionnement sur gel d'agarose à bas point de fusion 4% des produits de RT-PCR. L'identité des produits obtenus (+ ou – Exon 5) est indiquée à droite du gel. **C.** Quantification des produits d'amplification contenant ou non l'exon 5 exprimée en pourcentage du total des ARN épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01 par rapport au valeurs obtenues avec le contrôle pXJ41 vide).

Afin de vérifier que l'effet observé n'était pas lié au contexte du minigène utilisé, la même expérience a été répétée avec le minigène RTB300. Les résultats obtenus sont illustrés dans la figure III.16. On observe que la sur-expression de la protéine hnRNP H augmente également l'inclusion de l'exon 5 dans les produits d'épissage obtenus à partir du minigène RTB300. Dans ce cas, l'activation observée entraîne un taux d'inclusion de 90% de l'exon 5, contre 70% en absence de sur-expression de la protéine hnRNP H. Là encore, l'activation de l'intégration de l'exon 5 dans les produits d'épissage n'était pas proportionnelle à la quantité de plasmide pXJ41-hnRNP H transfectée. Ces résultats confirmaient donc l'effet activateur de la protéine hnRNP H sur l'inclusion de l'exon 5 hcTNT indépendamment du contexte du minigène utilisé.

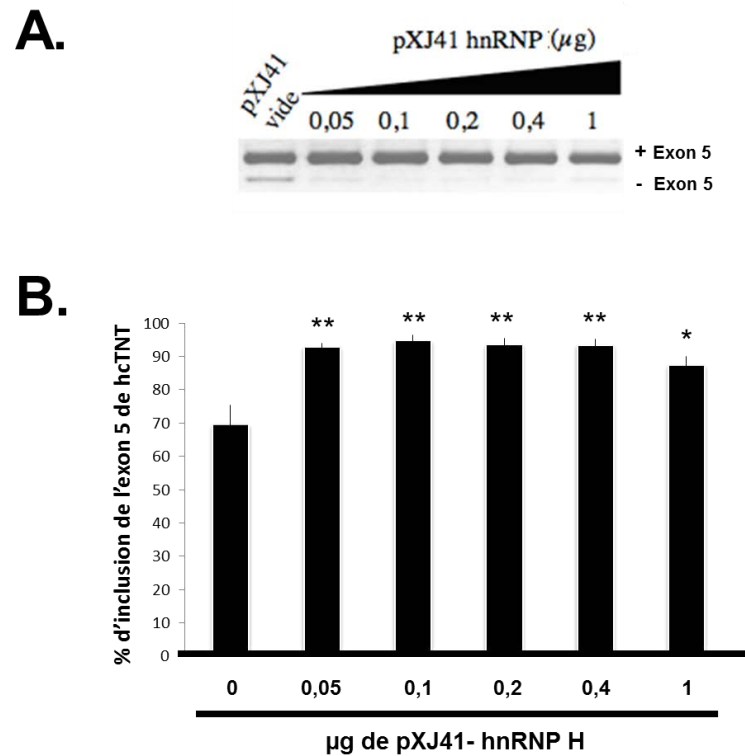


Figure III.16 : Etude de l'effet de la sur-expression de la protéine hnRNP H sur l'inclusion de l'exon 5 hcTNT dans le transcrit du minigène RTB300. 48 h après la transfection des constructions pXJ41-hnRNP H ou vide et du minigène RTB300, les ARN totaux ont été extraits puis analysés par RT-PCR. Un contrôle a été réalisé en transfectant 1,5 µg du plasmide pXJ41 vide (0). **A.** Fractionnement sur gel d'agarose à bas point de fusion 4% des produits de RT-PCR. L'identité des produits obtenus (+ ou - Exon 5) est indiquée à droite du gel. **B.** Quantification des produits d'amplification contenant ou non l'exon 5 exprimée en pourcentage du total des ARN épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01 par rapport au valeurs obtenues avec le contrôle pXJ41 vide).

III.3.2. L'absence de la protéine hnRNP H dans des cellules hek293 entraine une diminution du taux d'inclusion de l'exon 5 du pré-ARNm hcTNT.

Afin de confirmer le rôle de la protéine hnRNP H dans la régulation de l'inclusion de l'exon 5 hcTNT, nous avons diminué le niveau d'expression de la protéine hnRNP H dans des cellules HEK293T en utilisant des siRNA. Les déplétions ont été réalisées en utilisant un mélange de 4 siRNA H4-7 (Dharmacon), ceci selon le protocole décrit dans la partie Matériel et Méthodes. Le minigène IR-hcTNT a été co-transfecté avec les siRNA dans des cellules ayant déjà été traitées par les siRNA depuis 48 h. Un siRNA non target (NT) ciblant une séquence non exprimée dans les cellules HEK293 a été utilisé comme contrôle pour ces expériences. Les résultats obtenus sont présentés dans la figure III.17. L'efficacité du traitement par les siRNA H4-7 a été vérifiée par Western blot (Figure III.17A). Ce traitement conduit à une diminution

nette de la quantité de protéine hnRNP H. L'analyse par RT-PCR des produits d'épissage révèle une diminution de l'inclusion de l'exon 5 lors du traitement des cellules par les siRNA H4-7, avec un taux d'inclusion de l'exon 5 de 63% après traitement, contre 76% dans la condition contrôle (Figure III.17). Ces résultats confirment donc l'effet activateur de la protéine hnRNP H sur l'inclusion de l'exon 5 hcTNT.

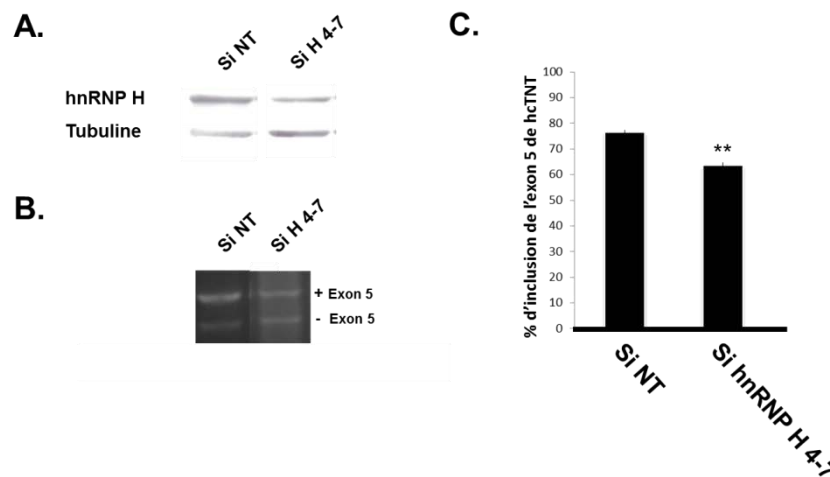


Figure III.17 : Etude de l'effet de la déplétion de la protéine hnRNP H sur l'inclusion de l'exon 5 hcTNT dans le transcrit du minigène IR-hcTNT. A. Vérification par Western blot de l'efficacité de la déplétion de la protéine hnRNP H suite au traitement avec les siRNA 4-7. B. Fractionnement sur gel d'agarose à bas point de fusion 4% des produits de RT-PCR obtenus sont fractionnés. L'identité des produits obtenus (+ ou - Exon 5) est indiquée à droite du gel C. Quantification des produits d'amplification contenant ou non l'exon 5 exprimée en pourcentage du total des ARN épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01 par rapport au valeurs obtenues NT).

III.3.3. Etude de l'importance des séquences introniques et exoniques pour l'activité de la protéine hnRNP H dans la régulation de l'épissage de l'exon 5 du pré-ARNm hcTNT.

Ayant démontré l'implication de la protéine hnRNP H dans la régulation de l'inclusion de l'exon 5 *in cellulo*, j'ai ensuite entrepris de démontrer que cet effet était direct et nécessitait la fixation de hnRNP H sur le pré-ARNm. Pour cela, j'ai testé l'effet de mutations ou de la délétion des segments du minigène IR-hcTNT 150-E5-150, capables de fixer hnRNP H *in vitro*, sur la régulation de l'inclusion de l'exon 5 par hnRNP H *in cellulo*. Afin de tester l'importance des séquences introniques correspondant aux fragments A (-150 à -100), E (+50 à +100) et F (+100 à +150), l'effet activateur de la protéine hnRNP H a été testé en utilisant des minigènes IR-hcTNT dans lesquels ces régions introniques étaient absentes (Figure III.18A).

Les variants 100-E5-150 et 150-E5-50 ont été co-transfectés dans des cellules HEK293 avec 1 µg de plasmide pXJ41-hnRNP H ou 1 µg de plasmide pXJ41 vide (Figure III.18A). Les résultats obtenus sont présentés dans la figure III.18B. Pour chacun des variants étudiés le taux d'inclusion de l'exon 5 en absence de sur-expression de la protéine hnRNP H était modifié du fait de la délétion de séquences régulatrices activatrices ou inhibitrices, même si dans l'expérience illustrée ici l'effet inhibiteur du segment -150 à -100 était très faible (Figure III.18B). Par ailleurs, pour chacun de ces variants, la sur-expression de la protéine hnRNP H entraînait toujours une augmentation de l'inclusion de l'exon 5 similaire à celle observée avec la construction 150-E5-150. Cela signifiait que les segments -150 à -100 et +50 à +150 ne sont pas individuellement nécessaires pour l'activation de l'inclusion de l'exon 5 par la protéine hnRNP H. Une explication possible à ces données était que l'activation par le facteur d'épissage hnRNP H soit dûe à la fixation de la protéine hnRNP H au niveau de l'exon 5. De ce fait, ne pouvant pas déléter l'exon 5, nous avons entrepris de réaliser des mutations des sites potentiels pour la fixation de la protéine hnRNP H au sein de l'exon 5.

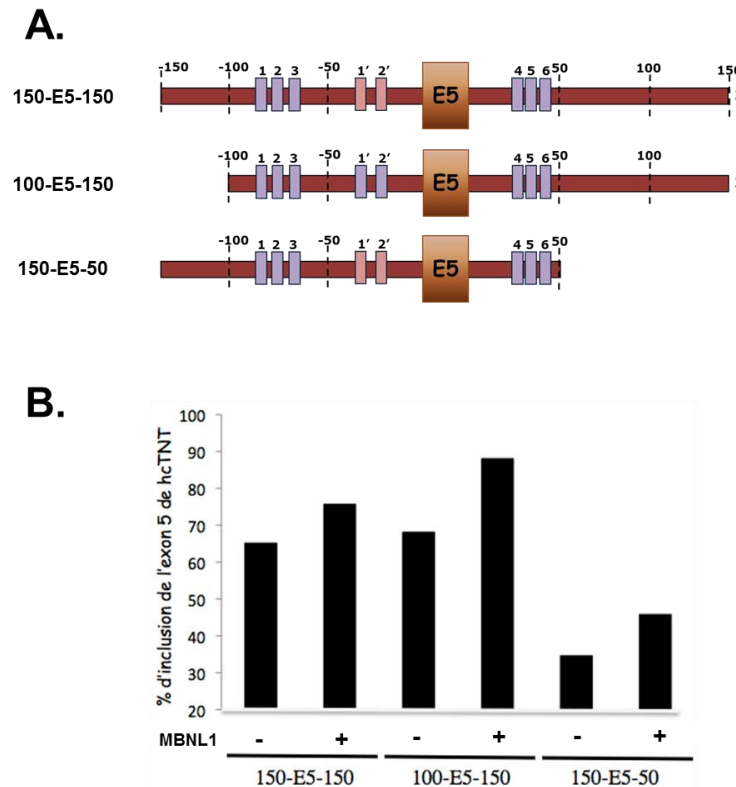


Figure III.18 : Etude de l'importance fonctionnelle des séquences introniques -150 à -100 et +50 à +150 dans la régulation de l'inclusion de l'exon 5 du pré-ARNm hcTNT par la protéine hnRNP H. **A.** Description des minigènes IR-hcTNT 150-E5-150, 100-E5-150 et 150-E5-50. Les sites MBNL1 du groupe de site 1 sont indiqués par des rectangles roses, et ceux des groupes 2 ou 3 par des rectangles violets. Le nom de la construction est indiqué à gauche. **B.** Effet de sur-expression de la protéine hnRNP H (+) ou du plasmide vide pXJ41-vide (-) sur l'épissage de l'exon 5 des minigènes 150-E5-150, 100-E5-150 et 150-E5-50. 48 h après la transfection des constructions pXJ41-hnRNP H ou vide et des différents minigènes, les ARN totaux ont été extraits puis analysés par RT-PCR avec les oligodésoxynucléotides 5991 et 8078. La quantification des produits d'amplification contenant ou non l'exon 5 a été exprimée en pourcentage du total des ARN épissés.

Trois sites, sites H1, H2 et H3, présentaient des séquences pouvant permettre une éventuelle interaction avec la protéine hnRNP H (Figure III.19A). Le site H1 correspond à la séquence TGGAG et le site H2 correspond à la séquence GAGGA. Ces sites ne sont espacés que par un seul nucléotide, un résidu A et nous observons lors des expériences d'empreintes de légères protections par la protéine hnRNP H-Nter au niveau du dernier G du site H1 et du 1^{er} G du site H2. Un autre site potentiel H3 correspondait à la séquence GGAGG. Afin de limiter la capacité de la protéine hnRNP H à se fixer sur la séquence exonique, j'ai réalisé des constructions à partir du minigène IR-150-E5-150 dans lesquelles j'ai muté simultanément ces 2 résidus G (Mut H1/2). Des mutations individuelles de chacun des G ont été également réalisées (Mut H1 et Mut H2), ainsi qu'une construction dans

laquelle j'ai muté 4 résidus G (Mut poly G), afin que la séquence exonique ne présente plus aucune répétition de résidus G et ne puisse plus former de quadruplexe de G. L'effet activateur de la protéine hnRNP H sur les ARN produits à partir de ces constructions a été testé dans des cellules HEK293. Les résultats obtenus sont présentés dans les figures III.19B et III.19C. En absence de sur-expression de la protéine hnRNP H, on remarque que chacune des mutations entraîne une augmentation de l'inclusion de l'exon 5 qui passe d'un taux d'inclusion de 60 % en absence de mutation à près de 82 % lorsque les différentes mutations sont introduites dans l'exon 5. Cela signifiait que les résidus G mutés participent à la réduction de l'inclusion de l'exon 5. Cependant, la sur-expression de la protéine hnRNP H entraînait toujours une activation de l'inclusion de l'exon 5 quelque soit la construction mutée considérée. Ces résultats montraient donc que les mutations H1, H2, H1/H2 et poly G n'inhibent pas l'effet activateur de la protéine hnRNP H sur l'inclusion de l'exon 5 du pré-ARNm hCNT. Pour mieux comprendre ces résultats nous avons entrepris de vérifier par des expériences de retard sur gel si ces mutations empêchaient l'interaction de la protéine hnRNP H avec la séquence exonique. Le résultat obtenu est présenté dans la figure III.19D. On constate que la présence de la mutation H1/H2 diminue certes l'affinité de la protéine hnRNP H-Nter pour l'exon 5, mais n'abolit pas l'interaction de hnRNP H-Nter avec l'exon 5. Des résultats similaires ont été obtenus avec le mutant poly G expliquant sans doute la persistance de l'effet activateur de hnRNP H en présence de ces mutations.

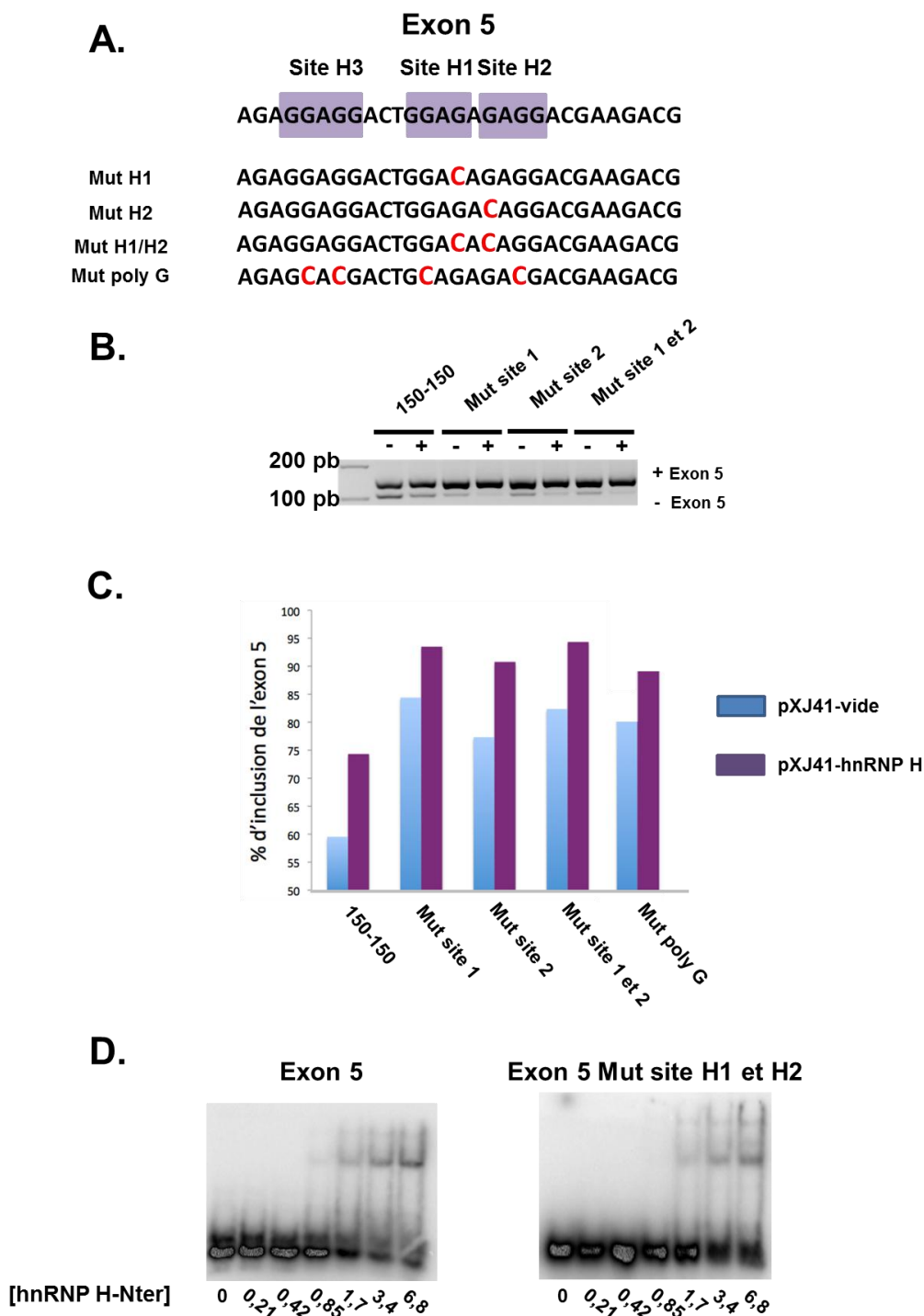


Figure III.19 : Etude de l'importance fonctionnelle de la séquence exonique dans la régulation de l'inclusion de l'exon 5 du pré-ARNm hCNT par la protéine hnRNP H. **A.** Description des différents sites potentiels d'interaction de la protéine hnRNP H au niveau de l'exon 5 du pré-ARNm hCNT et des mutations réalisées. Les sites potentiels de fixation de la protéine hnRNP H sont indiqués par un rectangle violet. **B.** Fractionnement sur gel d'agarose à bas point de fusion (4%) des produits de RT-PCR. L'identité des produits obtenus (+ ou - Exon 5) est indiquée à droite du gel. **C.** Quantification des produits d'amplification contenant ou non l'exon 5 exprimée en pourcentage d'inclusion de l'exon 5 par rapport au total des ARN épissés. **D.** Test de la fixation de la protéine hnRNP H-Nter sur l'exon 5 sauvage ou portant la mutation H1/2.

III.4. ETUDE IN VITRO DE L'EFFET DE LA VARIATION DE LA QUANTITE DE hnRNP H SUR L'INCLUSION DE L'EXON 5 DU pré-ARNm hcTNT.

Afin de démontrer que la régulation de l'inclusion de l'exon 5 par la protéine hnRNP H observée *in vivo* était bien directe et ne dépendait pas d'une cascade d'évènements cellulaires, il nous fallait démontrer la capacité de la protéine hnRNP H à réguler *in vitro* l'inclusion de l'exon 5. En effet, *in vitro* au sein d'un extrait cellulaire, la protéine hnRNP H garde sa capacité à interagir avec ses partenaires ARN ou protéines, mais aucune voie de signalisation ne peut être activée. Pour cela, nous avons transcrit *in vitro* un pré-ARNm du minigène IR-hcTNT portant une coiffe selon les conditions expérimentales décrites dans le Matériel et Méthodes. Ce pré-ARNm a été incubé dans un extrait nucléaire de cellules HeLa compétent pour l'épissage et le mélange a été complété par du tampon D ou par des quantités croissantes de protéine hnRNP H-Nter (0.5 µg et 1 µg). Après arrêt de la réaction d'épissage par ajout de protéinase K, les ARN épissés ont été extraits puis rétrotranscrits en ADNc et analysés par RT-PCR. Les résultats obtenus sont présentés dans la figure III.20. De manière inattendue, les résultats révélaient une diminution du taux d'inclusion de l'exon 5 qui passait de 49 % à 30 et 20 % suite à l'addition respective dans l'extrait nucléaire de 0.5 et 1 µg de protéine hnRNP H N-ter. Cette diminution semblait corrélée à la quantité de protéine hnRNP H-Nter ajoutée et indiquait un effet inhibiteur direct de la protéine hnRNP H N-ter sur la régulation de l'épissage alternatif de l'exon 5 du pré-ARNm hcTNT. L'observation de cet effet inhibiteur de la protéine hnRNP H-Nter était donc en contradiction avec les résultats obtenus *in cellulo* qui montraient un effet activateur de la protéine hnRNP H entière sur l'inclusion de l'exon 5 du pré-ARNm hcTNT.

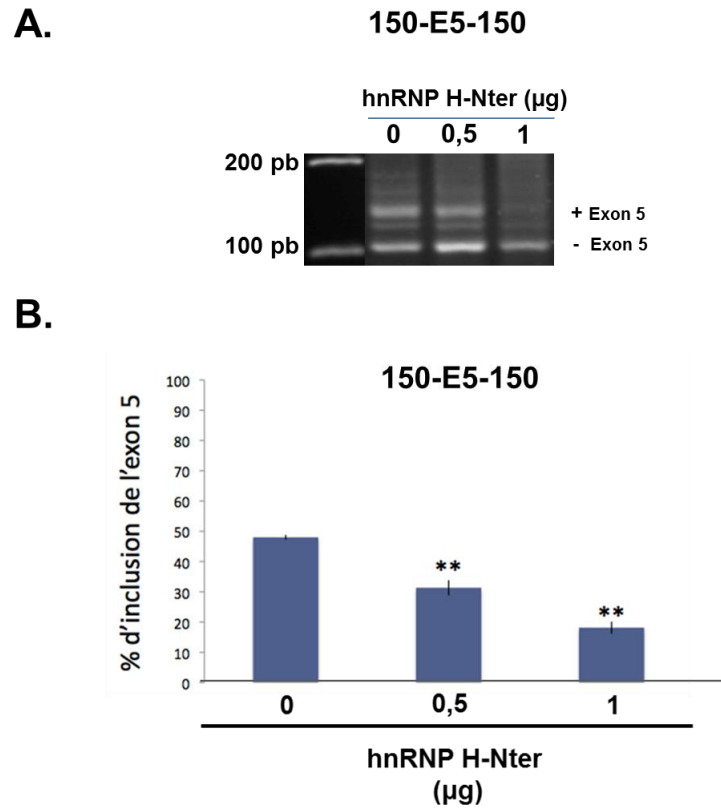


Figure III.20 : Effet de la variation de la quantité de hnRNP H-Nter sur le taux d'inclusion de l'exon 5 du pré-ARNm hCTNT *in vitro*. **A.** Fractionnement par électrophorèse en gel d'agarose à bas point de fusion 4 % des produits d'épissage détectés par RT-PCR en présence de quantités variables de protéine hnRNP H-Nter recombinante. **B.** Quantification des produits d'amplification contenant ou non l'exon 5 exprimée en pourcentage du total des ARN épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01 par rapport au valeurs obtenues sans addition de protéines hnRNP H-Nter).

Plusieurs hypothèses pouvaient être proposées pour expliquer ce résultat : i) le fait que la protéine hnRNP H recombinante utilisée pour réaliser l'étude *in vitro* soit tronquée pouvait affecter son activité. Il est en effet possible que le domaine RRMH 3 et le domaine riche en glycine manquant au sein de la protéine hnRNP H N-ter soient nécessaires pour l'activité activatrice de la protéine et notamment sa capacité d'interaction avec des partenaires protéiques ; ii) la perte d'un cofacteur nécessaire pour l'activité activatrice de la protéine hnRNP H lors de la préparation de l'extrait nucléaire ; iii) une inhibition aspécifique de l'épissage due à la quantité de protéine hnRNP H-Nter recombinante utilisée ; iv) la fixation de la protéine hnRNP H sur le pré-ARNm hCNT pourrait également inhiber l'inclusion de l'exon 5 *in cellulo*, mais être masquée par un effet indirect activateur de hnRNP H médié par une cascade d'événements cellulaires conduisant par exemple à une diminution de la quantité de protéine MBNL1 intracellulaire. Afin de tester ces différentes hypothèses, l'étude devra être recommencée en i) utilisant une protéine hnRNP H recombinante pleine longueur, ii) testant différentes préparations d'extraits nucléaires et une gamme plus large de quantité de protéine recombinante, et iii) incluant des pré-ARNm dont l'épissage alternatif est connu pour être ou ne pas être régulé par hnRNP H comme contrôles. Il nous paraissait également essentiel de tester si l'effet activateur de la sur-expression de hnRNP H *in cellulo* pouvait être lié à une diminution de la quantité intracellulaire de MBNL1.

III.5. ETUDE DE LA RELATION ENTRE LES PROTEINES MBNL1 ET hnRNP H.

Dans un premier temps, j'ai testé si l'inhibition de l'expression de hnRNP H par emploi de siRNA impactait le niveau d'expression de MBNL1. Le résultat est présenté dans la figure III.21. On constate que la déplétion de la protéine hnRNP H n'a aucun effet sur la quantité de MBNL1 intracellulaire. De même, la surexpression de hnRNP H n'affecte pas non plus la quantité de MBNL1. Les effets des variations de la quantité de hnRNP H cellulaire sur la régulation de l'inclusion de l'exon 5 hCNT ne sont donc pas dus à un effet indirect de hnRNP H sur le taux de MBNL1 cellulaire. Afin de mieux comprendre la relation existant entre les protéines MBNL1 et hnRNP H pour la régulation de l'inclusion de l'exon 5 hCNT, des expériences de compétition entre ces protéines ont alors été réalisées.

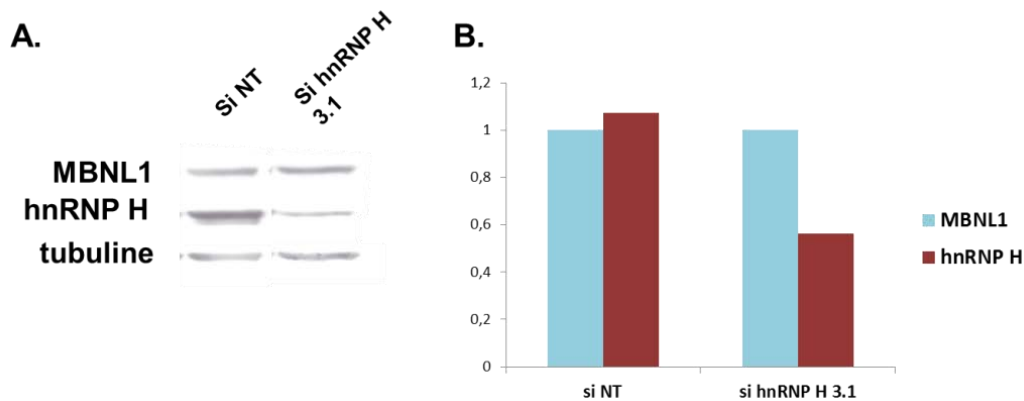


Figure III.21: Effet de la déplétion de la protéine hnRNP H sur la l'expression de la protéine MBNL1. A. Vérification par Western blot de l'efficacité de la déplétion de la protéine hnRNP H suite au traitement avec le siRNA 3.1 et de la quantité de protéine MBNL1. **B.** Quantification des niveaux d'expression des protéines hnRNP H (rouge) et MBNL1 (bleu) suite au traitement par siRNA NT et hnRNP H 1.3. Le signal obtenu par l'emploi d'un anticorps dirigé contre la tubuline a été utilisé comme rapporteur.

Pour cela, des cellules HEK293 ont été co-transfectées avec le minigène IR-hcTNT 150-E5-150 et différentes quantités de plasmides pcDNA3.1-MBNL1 et pXJ41-hnRNP H qui permettent respectivement la sur-expression des protéines MBNL1 et hnRNP H. Les résultats obtenus sont présentés dans la figure III.22. Comme attendu la transfection du plasmide pcDNA3.1-MBNL1 seul inhibait l'inclusion de l'exon 5, alors que la transfection du plasmide pXJ41-hnRNP H seul activait l'inclusion de l'exon 5. De manière intéressante, en présence des deux plasmides, on observait que plus le rapport en quantité, de chacun des plasmides transfectés (hnRNP H/MBNL1) augmentait, plus le pourcentage d'inclusion de l'exon 5 augmente et ceci de façon linéaire. Pour un rapport hnRNP H/MBNL1 égal à 3, on retrouve le pourcentage d'inclusion de l'exon 5 (65%) observé dans des cellules HEK293 en absence de sur-expression des protéines MBNL1 et hnRNP H. Ces résultats montraient que les protéines MBNL1 et hnRNP H entre en compétition pour la régulation de l'inclusion de l'exon 5 dans le transcrit produit à partir du minigène IR-hcTNT.

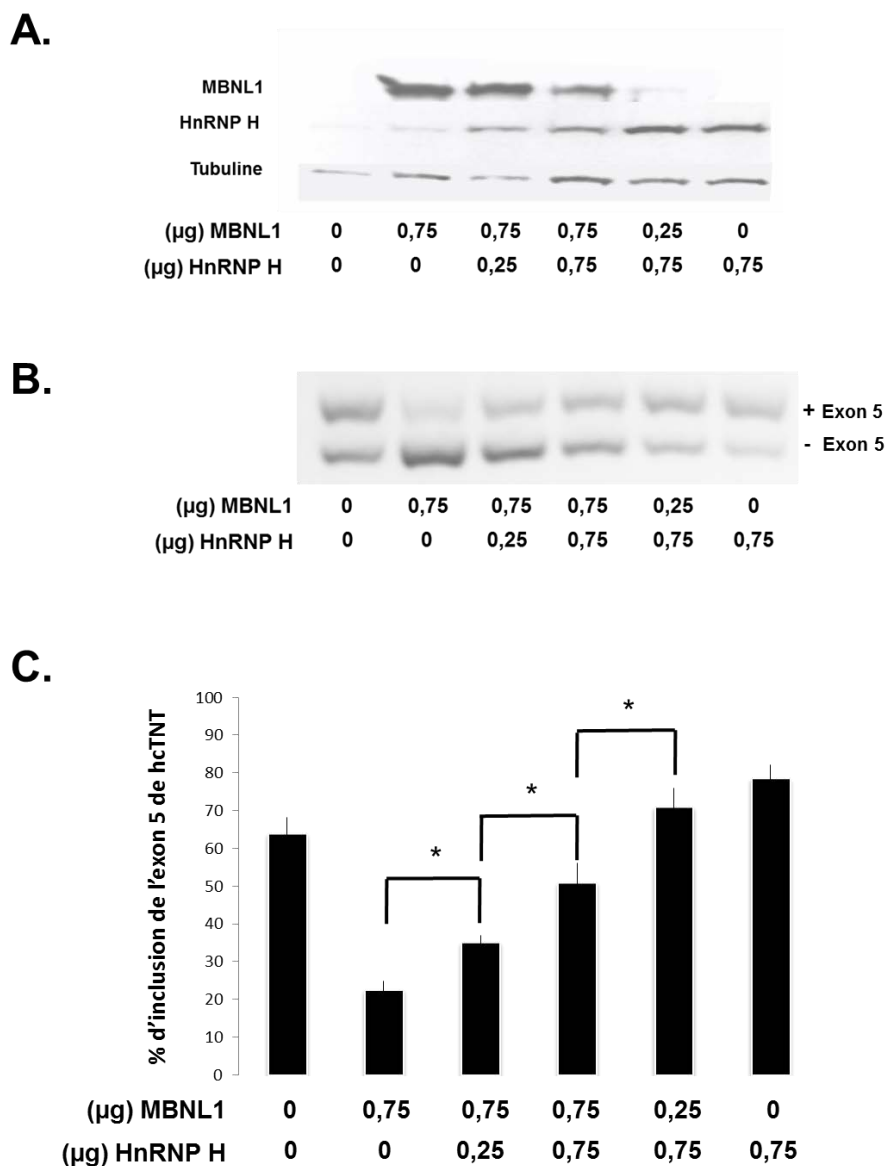
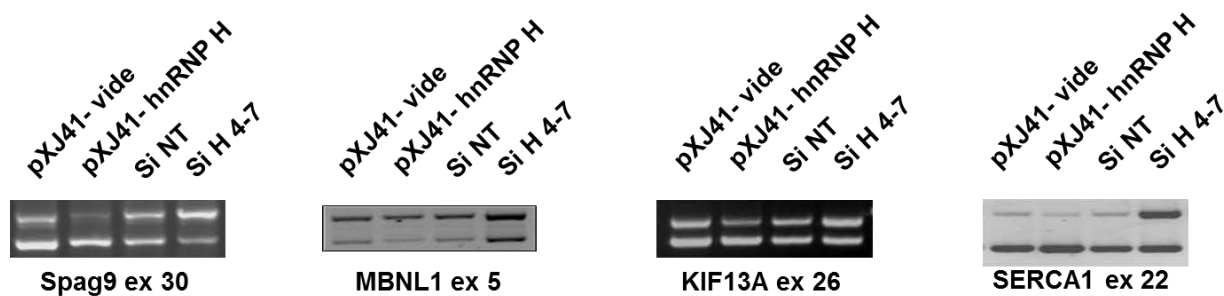


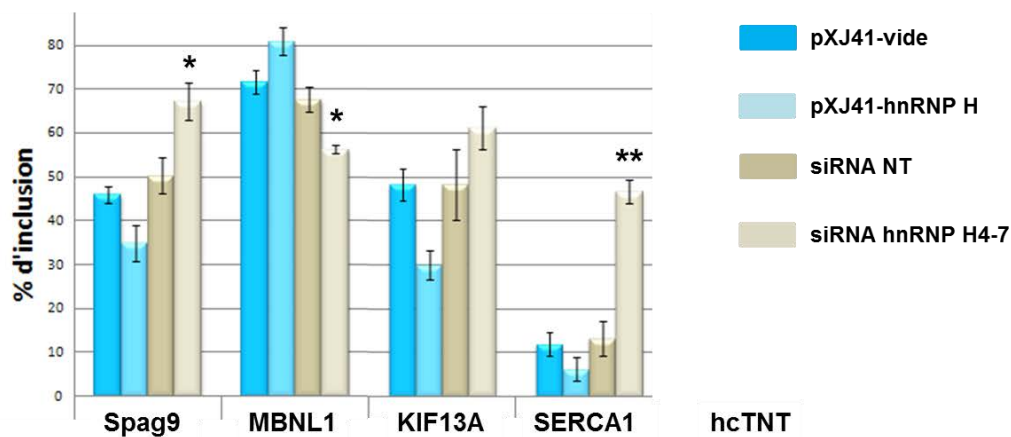
Figure III.22: Etude de l'effet de la compétition *in cellulo* entre les protéines hnRNP H et MBNL1 sur l'inclusion de l'exon 5 hcTNT dans le transcrit du minigène IR-hcTNT. 48 h après la co-transfection des constructions pXJ41-hnRNP H ou vide, pcDNA3.1-MBNL1 ou flag et du minigène 150-E5-150, les ARN totaux ont été extraits puis analysés par RT-PCR avec les oligodésoxynucléotides 5991 et 8078. Des contrôles ont été réalisés en transfectant 1,5 μ g du plasmide pcDNA3.1-Flag (quantité 0 pour MBNL1) ou pXJ41 vide (quantité 0 pour hnRNP H). **A.** Vérification de la sur-expression des protéines hnRNP H et MBNL1 par expérience de Western blot. Les protéines totales ont été extraites et analysées par Western blot en utilisant des anticorps dirigés contre les protéines MBNL1 et hnRNP H. La tubuline a été utilisée comme contrôle. **B.** Fractionnement sur gel d'agarose à bas point de fusion 4 % des produits d'amplification obtenus lors de l'analyse du taux d'inclusion de l'exon 5 au sein des ARNm produits à partir du minigène IR-hcTNT 150-E5-150 par RT-PCR. **C.** Quantification des produits d'amplification contenant ou non l'exon 5 exprimée en pourcentage des ARN totaux épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01).

Nous avons cherché ensuite à savoir si la compétition entre les protéines hnRNP H et MBNL1 était généralisable à d'autres ARNm dont l'épissage est altéré chez des patients atteints de DM1 ou était spécifique à la régulation de l'inclusion de l'exon 5 hCTNT. Pour cela, avec F.Oillo-Blanloeil (technicienne UL au sein de l'équipe), nous avons testé l'effet d'une sur-expression et d'une déplétion de la protéine hnRNP H sur l'épissage alternatif des exons 30, 5, 36 et 22 des pré-ARNm endogènes des gènes SPAG9, MBNL1, KIF13A et SERCA1, ceci dans des cellules HEK293. Les résultats obtenus sont illustrés dans la figure 46. L'analyse des résultats révèle que, tout comme dans le cas de l'exon 5 hCTNT, la quantité de protéine hnRNP H affectait la régulation de l'inclusion des différents exons analysés. En effet quelque soit l'ARNm analysé, l'effet sur l'inclusion de l'exon d'intérêt de la sur-expression de la protéine hnRNP H (pXJ41-hnRNP H) était opposé à celui de la déplétion de hnRNP H (si H 4-7). On notera également que hnRNP H peut soit activer l'inclusion d'exons alternatifs, comme c'est le cas pour l'exon 5 hCTNT et l'exon 5 de MBNL1, soit inhiber l'inclusion d'exons alternatifs, comme dans le cas des exons 30 SPAG9, 26 KIF13 et 22 SERCA1 (Figure III.23). De façon intéressante lorsque l'on compare l'effet que nous observons pour hnRNP H à celui déjà publié de MBNL1 pour la régulation de ces mêmes événements d'épissage, on trouve systématiquement un effet antagoniste des protéines MBNL1 et de hnRNP H (Figure III.23C), indiquant que l'antagonisme observé pour la régulation de l'inclusion de l'exon 5 hCTNT pourrait être généralisable à d'autres ARNm dont l'épissage est dérégulé chez les patients atteints de DM1.

A.



B.



C.



Figure III.23 : Effet de la modulation de la quantité de protéine hnRNP H intracellulaire sur la régulation de l'épissage de pré-ARNm dont l'épissage est régulé par MBNL1 et dérégulé dans le cadre de la DM1. **A.** Après transfection des plasmides pXJ41-hnRNP H ou pXJ41-vide et des siRNA hnRNP H4-7 ou NT, les ARN totaux ont été extraits puis analysés par RT-PCR. Les produits PCR obtenus ont été fractionnés par électrophorèse en gel agarose à bas point de fusion 4%. **B.** Quantification des produits d'amplification contenant ou non l'exon alternatif d'intérêt exprimée en pourcentage par rapport au total des ARN épissés. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01). **C.** L'effet, activateur (+) ou inhibiteur (-), des protéines MBNL1 et hnRNP H sur la régulation des différents événements d'épissage analysés est résumé.

III.6. CONCLUSION ET PERSPECTIVES.

III.6.1. Mise en évidence d'un grand nombre de sites de régulation de l'épissage de l'exon 5 de hcTNT par MBNL1 dans le contexte génétique du minigène IR-150-E5-150.

Le facteur d'épissage MBNL1 a été découvert récemment ceci à cause des défauts d'épissage alternatif induits lorsque des ARN contenant des répétitions CUG sont présents en quantités anormales dans les cellules humaines. Alors qu'il est clair que ces longues répétitions forment des structures tige-boucle très particulières qui reconnaissent spécifiquement MBNL1, le mode d'interaction de MBNL1 avec les ARN pré-messagers et de ce fait son mode de régulation de l'épissage alternatif est encore mal connu. Des séquences ayant une affinité pour MBNL1 ont été détectées dans quelques pré-ARNm dont l'épissage est régulé par MBNL1. En général au moins deux sites MBNL1 consécutifs sont présents, mais le mode d'organisation des molécules MBNL1 sur ces pré-ARNm n'a jamais été étudié expérimentalement de manière approfondie. Des hypothèses ont été formulées mais pas démontrées. Vu l'importance des dérégulations de l'épissage alternatif dans la pathologie DM1, il est impératif de comprendre quel est le mode d'action de MBNL1 afin de pouvoir proposer des stratégies thérapeutiques. Au début de ma thèse on savait que la région N-terminale de MBNL1 contenait deux couples de motifs à doigts de zinc fixant l'ARN et la structure 3D d'un de ces couples de doigts de zinc liés à un ARN avait été établie par cristallographie aux rayons X (Teplova et Patel, 2008). Par ailleurs, notre équipe, en collaboration avec l'équipe de N. Sergeant et M.L. Caillet-Boudin, avait montré l'importance de la séquence protéique codée par l'exon 3 qui sépare les deux couples de motifs de doigts de zinc dans l'affinité de MBNL1 pour les ARN et aussi que MBNL1 contient dans l'exon 7 alternatif un domaine permettant sa multimérisation (Tran et al., 2011). Nous avons également mis en évidence un grand nombre de sites potentiels de fixation de MBNL1 autour de l'exon 5 du pré-ARNm hcTNT. Les données obtenues au cours de ma thèse apportent des informations qui permettent de mieux comprendre comment MBNL1 peut limiter le taux d'inclusion de cet exon, limitation qui semble requise pour la fonction normale du muscle cardiaque adulte.

Nous montrons, dans le contexte des minigènes IR-150-E5-150 et RTB300, que la mutation des deux sites de fixation de MBNL1 1' et 2' identifiés par Ho *et al* n'abolit pas l'action négative de MBNL1 sur l'inclusion de l'exon 5. Par contre dans le contexte du minigène IR-150-E5-150, toute mutation d'un des trois sites sites du groupe 2 (sites 1, 2 et 3) ou du groupe 3 (sites 4, 5 et 6) limite l'effet négatif de MBNL1 aux concentrations normales de MBNL1.

III.6.2. Proposition d'un modèle de régulation de l'inclusion de l'exon 5 de hcTNT dans le contexte du minigène IR-150-E5-150.

De manière étonnante, nous avons vu des résultats très similaires en mutant les sites du groupe 2 ou les sites du groupe 3. Une explication possible de ces données est que la présence de trois sites de fixation de MBNL1 de part et d'autre de l'exon 5 est requise pour obtenir une limitation efficace de l'inclusion de l'exon 5 par MBNL1. Ceci pourrait s'expliquer par la nécessité de former une structure à longue distance au sein de l'ARN par interaction des groupes de sites auxquels plusieurs molécules de MBNL1 pourraient s'associer, ce qui isolerait l'exon 5 dans une boucle interne et limiterait la fixation des composants du spliceosome, comme cela a déjà été montré pour des facteurs régulant l'épissage tels que hnRNP A1 ou hnRNP H (Figure III.24). A plus forte concentration en MBNL1 (expérience de sur-expression), la présence d'un seul site dans un des groupes de site pourrait peut-être être suffisante pour permettre la régulation. D'où la nécessité dans ces conditions d'éliminer les trois sites 1, 2 et 3, soit les trois sites 4, 5 et 6 pour perdre complètement la régulation. Ce modèle de régulation est tout à fait nouveau et différent de toutes les hypothèses qui avaient été formulées jusque-là pour essayer d'expliquer l'action de MBNL1 sur l'exon 5 de hcTNT. L'hypothèse d'un rapprochement des groupes 2 et 3 dans l'espace pourrait expliquer les différences d'activité des sites 1' et 2', situés en amont de l'exon 5, lorsque la mutation ΔG est générée en aval de l'exon 5. En effet, en contexte sauvage, la mutation de ces sites n'avait aucun effet positif sur le taux d'inclusion 5. Par contre en présence de la mutation ΔG située en aval de l'exon 5, la mutation des sites 1' et 2', situés en amont de l'exon 5, augmentait fortement le taux d'inclusion de cet exon, mettant en évidence un lien possible entre séquences en amont et séquences en aval de l'exon 5. Un modèle 2D de la région 150-E5-150 du pré-ARNm hcTNT avait été établi au laboratoire sur la base de données d'empreintes chimiques et enzymatiques. Deux structures tige-boucle stables peuvent

d'après ce modèle être formées par les 50 nucléotides 5'-terminaux et 3'-terminaux. Le reste de la structure proposée est moins stable, en particulier la région contenant le groupe de sites 2 est peu structurée et la stabilité de la structure incluant les trois sites du groupe 3 est faible. Ceci pourrait permettre la formation d'une structure non canonique entre ces deux régions, structure qui serait stabilisée par la fixation de la protéine MBNL1. Les trois sites de fixations de MBNL1 du groupe 3 situés en aval de l'exon 5 sont localisés dans une région renfermant aussi des sites de fixation de CUG-BP1 et il avait été proposé que, chez le poulet, CUG-BP1 contrecarre l'effet négatif de MBNL1 sur l'inclusion de l'exon 54 (Ho et al., 2004). Si selon notre hypothèse, les trois sites MBNL1 des groupes 2 et 3 sont tous nécessaires à la formation d'une structure fixant MBNL1 et qui limite l'inclusion de l'exon 5, on peut imaginer que la fixation de CUG-BP1 en aval de l'exon 5 réduit de manière stérique l'accessibilité des sites du groupe 3 à MBNL1, et de ce fait empêche la formation de la structure à longue distance bloquant l'inclusion de l'exon 5. Cependant cette hypothèse repose sur les résultats obtenus dans le minigène IR-150-E5-150. De manière inattendue, l'étude réalisée sur l'impact des différents sites des groupes 2 et 3 identifiés par le laboratoire dans le contexte du minigène RTB300 a montré qu'aucune des mutations n'abolit l'effet de la protéine MBNL1. Nous avons pu montrer que ces résultats contraires à ceux obtenus dans le contexte du minigène IR-150-E5-150, n'étaient pas liés aux séquences introniques issues du pré-ARNm hCNT qui bordent l'exon 5. Les résultats obtenus lors de l'étude de l'effet de la sur-expression de la protéine hnRNP H montrent que l'effet activateur de la protéine hnRNP H est retrouvé dans les deux contextes de minigènes. Ainsi, ces données semblent suggérer que la différence de contexte génétique entre les deux minigènes n'influence que la régulation de l'épissage par MBNL1 et non l'activité de la protéine hnRNP H sur l'épissage de l'exon 5. Il sera donc nécessaire d'établir *in vivo* l'importance des sites des groupes 2 et 3 au sein du pré-ARNm hCNT issus de l'endogène pour déterminer le mode d'action de l'activité inhibitrice de la protéine MBNL1. Pour cela nous avons initié des expériences basées sur l'utilisation d'oligonucléotides antisens permettant de bloquer l'éventuelle reconnaissance des sites des groupes 2 et 3 par la protéine MBNL1 sur le pré-ARNm hCNT issus de l'endogène.

De manière intéressante, le modèle de régulation proposé peut permettre la formation d'ARN circulaire, dont la formation a déjà été montrée être favorisée en présence de MBNL1 et nous avons entrepris des mises aux points afin de pouvoir éventuellement détecter une circularisation de l'exon 5 dans des conditions de sur-expression de la protéine MBNL1.

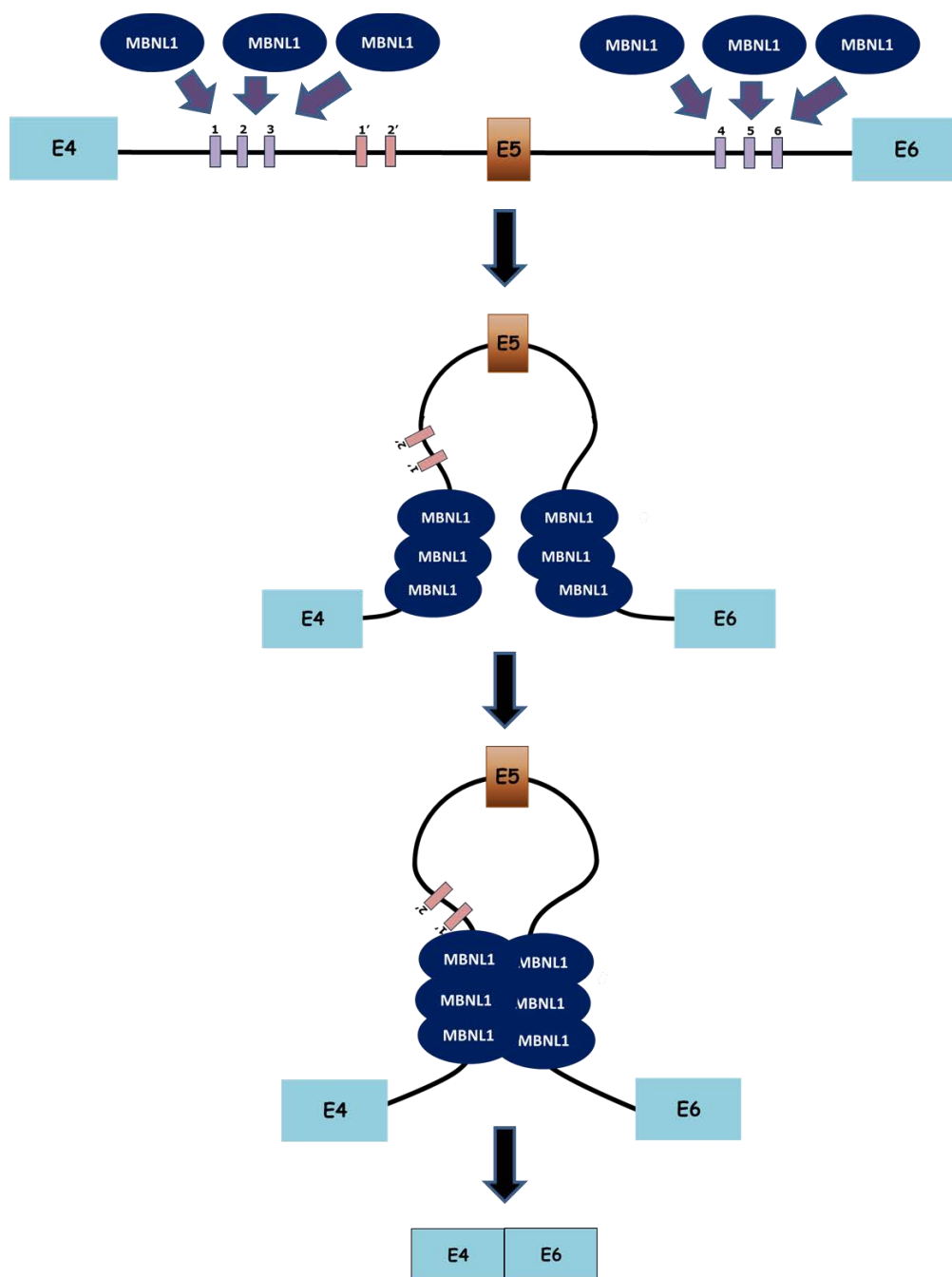


Figure III.24 : Nouveau modèle proposé pour la régulation de l'inclusion de l'exon 5 du pré-ARNm hCNT. La reconnaissance des sites 1, 2, 3, 4, 5 et 6 par la protéine MBNL1 entraîne la formation d'une structure à longue distance qui isole l'exon 5 dans une boucle, limitant ainsi la fixation des composant du spliceosome et l'inclusion de l'exon.

III.6.3. Identification de la protéine hnRNP H comme régulateur de l'inclusion de l'exon 5.

Notre équipe avait déjà montré que l'inclusion de l'exon 5 de hcTNT est soumise à des régulations supplémentaires par rapport à celles générées par MBNL1 et défini des segments du pré-ARNm hcTNT impliqués dans ces régulations. L'un de mes objectifs a été d'essayer d'identifier les facteurs nucléaires pouvant se lier sur les éléments de séquences en cis, ceci en appliquant à des fragments de 50 nucléotides de la région 150-E5-150 la technique de formation de complexes en extraits nucléaires, suivie de leur purification par l'intermédiaire d'une protéine de fusion MBP-MS2. J'ai ainsi pu montrer que la protéine hnRNP H peut s'associer en extrait nucléaire au segment D contenant les sites MBNL1 du groupe 3, au fragment F et E. Par ailleurs, en utilisant une protéine recombinante hnRNP H correspondant au domaine N-terminaux de cette protéine (hnRNP H-Nter), j'ai observé une capacité de fixation directe de cette protéine sur l'exon 5 et sur le fragment E et à un moindre degré sur les fragments A, B, D et F. Nous n'observons donc pas une corrélation directe entre la présence de hnRNP H dans les complexes formés en extrait nucléaire et la fixation de hnRNP H-Nter sur les segments d'ARN correspondants, ce qui peut être expliqué par l'existence d'associations médiées par d'autres protéines en extrait nucléaire. Un point important ressort néanmoins de l'étude, hnRNP H possède une affinité à la fois pour l'exon 5 et pour les séquences situées en aval de cet exon. Ceci peut s'expliquer par l'existence de nombreux sites potentiels de fixation de la protéine hnRNP H dans l'exon 5 et en aval de cet exon, tout particulièrement dans les fragments E et F. Autre résultat important de ces travaux, la protéine hnRNP H favorise, lorsqu'elle est sur-exprimée *in cellulo*, l'inclusion de l'exon 5 et ceci quelque soit le contexte hétérologue dans lequel cet exon est intégré. De manière opposée, la diminution du taux de hnRNP H *in cellulo* par action de siRNA diminue l'intégration de l'exon 5 en contexte sauvage et nous avons clairement mis en évidence une compétition entre les protéines hnRNP H et MBNL1 pour la régulation de l'inclusion de l'exon 5 par des expériences de co-transfection de quantités relatives variables des deux protéines. Ces données sont également compatibles avec le modèle proposé dans le contexte du minigène IR-150-E5-150 et la formation d'une structure à longue distance par fixation de la protéine MBNL1. En effet, il est possible que la fixation de la protéine hnRNP H sur l'exon 5 hcTNT empêche la formation d'une structure à longue distance et favorise la reconnaissance de l'exon par le spliceosome. En effet, nous avons mis en évidence que

l'absence des segments E et F au sein du pré-ARNm hcTNT n'abolissait pas l'effet activateur de la protéine hnRNP H sur l'inclusion de l'exon 5. Il semble donc évident que la séquence de l'exon 5 est nécessaire à la régulation de l'inclusion de l'exon 5 hcTNT par la protéine hnRNP H. Une étude récente a montré l'effet activateur de la protéine hnRNP H sur l'inclusion de l'exon hcTNT, ainsi qu'une interaction entre la protéine hnRNP H et MBNL1 (Paul et al., 2011). Cependant, cette étude propose un autre modèle de régulation de l'exon 5 hcTNT. En effet, les auteurs suggèrent un effet indirect de la protéine hnRNP H qui limiterait l'activité de la protéine MBNL1 en interagissant avec elle. Ce modèle ne rentre pas en contradiction avec le modèle que nous proposons et il est possible que la protéine hnRNP H agisse à la fois directement et indirectement sur l'épissage de l'exon 5 hcTNT. Cependant, bien que nous ayons montré la fixation directe de la protéine hnRNP H sur le pré-ARNm hcTNT, il sera nécessaire de comprendre le résultat obtenu *in vitro* qui montrait un effet direct, mais inhibiteur de la protéine hnRNP H-Nter sur l'inclusion de l'exon 5 hcTNT et de renouveler l'expérience avec l'utilisation d'une protéine hnRNP H de pleine taille.

PARTIE 2 :
ETUDE DU ROLE DE L'ÉPISSAGE
ALTERNATIF DES PRE-ARNm DANS LA
REPONSE DES CELLULES DE VERTEBRES
AU STRESS OXYDANT.

INTRODUCTION

CHAPITRE IV : ETUDE DU ROLE DE L'ÉPISSAGE ALTERNATIF DES pré-ARNm DANS LA REPONSE DES CELLULES DE VERTEBRES AU STRESS OXYDANT.

IV.1. DEFINITION ET ORIGINE D'UN STRESS OXYDANT.

IV.1.1. Définition d'un stress cellulaire.

Dans la nature, peu d'organismes vivent dans des conditions physico-chimiques qui leurs sont parfaitement adaptées. Les organismes sont soumis à des contraintes extérieures imposées par l'environnement dans lequel ils évoluent et auxquelles ils doivent s'adapter comme la température, le rayonnement UV et les espèces réactives de l'oxygène. De plus, ces contraintes peuvent très rapidement fluctuer dans le temps. On considère qu'une cellule est soumise à un stress cellulaire lorsqu'elle est soumise à une variation physico-chimique, imposée par son environnement, qui modifie son homéostasie en perturbant son métabolisme. Ces perturbations proviennent essentiellement de dommages ou altérations de différents constituants cellulaires comme les protéines, les lipides et les acides nucléiques. Au cours de l'évolution, les cellules et les organismes ont développé des mécanismes pour répondre à ces fluctuations. Ces mécanismes permettent à la cellule d'établir un équilibre physico-chimique afin de maintenir son intégrité. Les organismes possèdent plusieurs systèmes de réponse aux stress qui sont plus ou moins spécifiques. De manière générale, les cellules répondent à de faibles variations ou à des conditions sous optimales par des mécanismes liés à des activités spécifiques et dédiées. Tandis que la réponse à une variation importante et rapide de l'environnement met en jeu des systèmes de défense moins spécifiques et rapidement inductibles qui servent avant tout à limiter les dommages aux constituants cellulaires et les déséquilibres métaboliques. Cette diversité de réponses aux variations environnementales confère aux cellules une grande souplesse et efficacité dans la défense cellulaire. Néanmoins, la capacité d'une cellule à répondre à une modulation de l'environnement dépend de son état (densité de population, stade du cycle cellulaire...), de son type cellulaire, ainsi que de ces antécédents d'exposition à d'autres stress cellulaires. D'autre part, la réponse d'une cellule à un stress dépend aussi de la force et de la nature du stress. Dans le cas d'un stress trop important ou d'une variation physico-

chimique trop rapide, les mécanismes de réponses cellulaires peuvent être insuffisants pour rétablir l'homéostasie cellulaire, d'où une perturbation définitive du métabolisme et l'induction de l'apoptose ou de la nécrose cellulaire.

IV.1.2. Le stress oxydant et ses principaux acteurs.

Parmi les nombreux systèmes qui maintiennent l'homéostasie de la cellule, le système d'oxydo-réduction (système redox) est essentiel à la survie cellulaire. Toute vie sur terre, nécessite des réactions chimiques dans lesquelles des électrons sont transférés d'une molécule donneuse à une molécule acceptrice. Le système redox permet d'établir un équilibre entre les molécules donneuses et acceptrices d'électrons. Chez les organismes aérobies tout particulièrement, l'oxygène (O_2) est un des éléments clé de leur système redox dans la mesure où il est l'accepteur final d'électrons de la chaîne respiratoire mitochondriale, qui permet la synthèse d'énergie sous forme d'ATP. Néanmoins, l' O_2 peut également subir des réductions qui entraînent la formation d'espèces réactives de l'oxygène (ROS pour Reactive Oxygen Species) et de l'azote (RNS pour Reactive Nitrogen Species). Les ROS et les RNS sont formées naturellement au sein de la cellule et sont essentielles au fonctionnement de nombreux processus essentiels à la survie cellulaire. Cependant, par leur nature chimique, ces espèces sont très réactives et peuvent modifier la composition d'un grand nombre de biomolécules, altérant ainsi leurs fonctions au sein de la cellule. Dans ce contexte environnemental et dans des conditions physiologiques, les cellules doivent donc maintenir un équilibre redox à travers la génération et l'élimination des ROS et des RNS. Pour cela, il existe au sein de la cellule de nombreux systèmes antioxydants, afin de limiter la quantité intracellulaire de ROS et de RNS. Le stress oxydatif a été défini comme le déséquilibre entre l'apparition de ROS et de RNS, et la capacité de l'organisme à neutraliser leur action par les systèmes de protection antioxydants.

III.1.2.1. Les espèces réactives de l'oxygène et leur formation.

Les espèces réactives de l'oxygène et de l'azote (ROS et RNS) sont à la base du système redox des cellules aérobies. Les ROS incluent principalement les espèces radicalaires comme les dérivés des superoxydes (O_2^-), les radicaux hydroxyles (HO^\cdot), mais aussi des espèces non radicalaires comme le peroxyde d'hydrogène (H_2O_2). Les ROS sont naturellement produits dans les cellules au sein de plusieurs voies métaboliques dont principalement la chaîne respiratoire mitochondriale (Cadenas and Sies, 1998). Les mitochondries synthétisent l'ATP de façon oxygène-dépendante. Ce processus aboutit à la réduction de l'oxygène moléculaire en eau (H_2O) et met en jeu un flux d'électrons au travers de complexes protéiques, numérotés de I à V, de la membrane interne des mitochondries (Figure IV.1). Au cours de ce processus, il arrive que certains électrons s'échappent de la chaîne respiratoire au niveau des différents complexes (particulièrement au niveau des complexes I et III) et réduisent l'oxygène moléculaire en anion superoxyde ($O_2^{\cdot -}$), générant ainsi des ROS (Muller et al., 2004; Starkov, 2008).

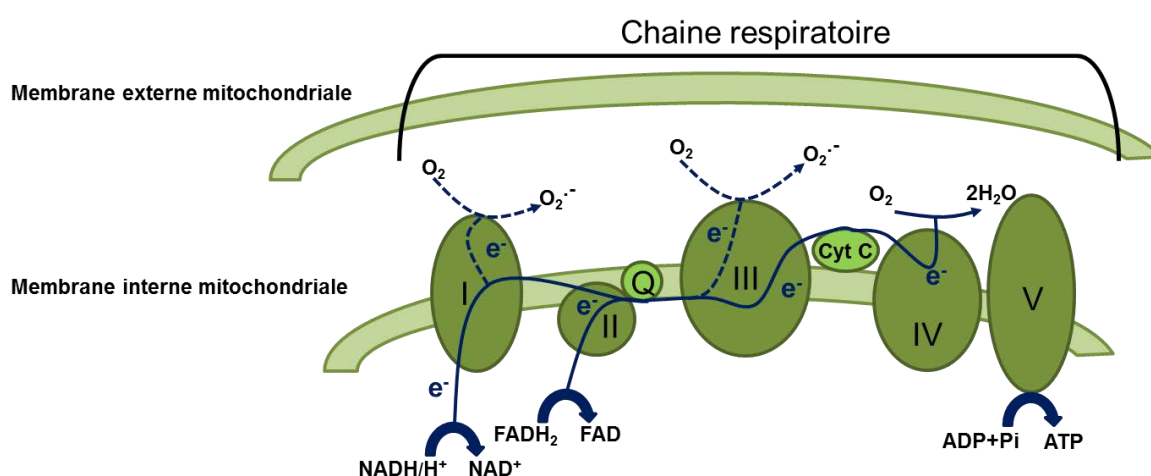


Figure IV.1 : Schéma de la chaîne respiratoire au sein de la membrane interne de la mitochondrie et de la production de ROS. La chaîne respiratoire permet de catalyser la phosphorylation de l'ADP en ATP par le complexe ATPase (complexe V) et implique un transfert d'électrons (e^-) au travers de 4 complexes (complexe I, II, III et IV), ainsi que le coenzyme Q (Q) et le cytochrome c (Cyt c). La plupart du temps, les électrons suivent la chaîne de transport entre les différents complexes (représentés par un trait bleu continu) jusqu'au complexe IV, mais il arrive que des électrons s'échappent de la chaîne, en particulier au niveau des complexes I et III (représentés par un trait bleu discontinu). Ces électrons réagissent rapidement avec du dioxygène (O_2) pour former un anion superoxyde ($O_2^{\cdot -}$).

Ce mécanisme de production de ROS semble être régulé par la cellule et il est proposé que le potentiel de la membrane mitochondriale et le niveau d'ATP ou de NADH/NAD⁺ soient des facteurs participants à cette régulation (Murphy, 2009; Mailloux and Harper, 2012). Une seconde source majeure de production de ROS est la NADPH oxydase qui catalyse la production de superoxydes à partir d'O₂ et de NADPH. Cependant, d'autres enzymes sont également impliquées dans la production de ROS, comme les enzymes du cytochrome P450, la xanthine oxydase ou des enzymes de la matrice mitochondriale telle que la pyruvate déshydrogénase (PDH) (pour revue Holmström and Finkel, 2014). Il existe également des sources exogènes de ROS (ou de RNS), tels que divers rayonnements (ultra-violets, rayons X) ou toxines (paraquat, rotenone...) (Sherer et al., 2002). Il est à noter le cas du peroxyde d'hydrogène (H₂O₂) qui est un ROS non radicalaire. Cette propriété confère à la molécule une plus faible réactivité mais une plus grande stabilité comparé à d'autres ROS comme O₂⁻. De plus, H₂O₂ diffuse facilement à travers les membranes et il est proposé que ce processus pourrait être facilité par les aquaporines (Bienert et al., 2014). De ce fait, l'H₂O₂ a la particularité de pouvoir affecter des biomolécules distantes de son origine de production. La toxicité de la molécule H₂O₂ passe principalement par sa réaction avec le fer ferreux (Fe²⁺) selon la réaction de fenton et de Haber-Weiss, qui entraîne la formation de radicaux hydroxyle (HO[·]), qui sont les ROS les plus réactifs causant des dommages oxydatifs irréversibles à pratiquement toutes les macromolécules biologiques (Figure IV.2) (Leonard et al., 2004).

Réaction de Fenton



Réaction de Haber-Weiss



Figure IV.2 : Réactions de Fenton et de Haber-Weiss qui permettent à partir du peroxyde d'hydrogène (H₂O₂) de générer des radicaux hydroxyles (HO[·]).

La plus grande stabilité de l' H_2O_2 par rapport à d'autres ROS confère également à la molécule un rôle central dans l'équilibre redox des cellules. En effet, l'un des principaux systèmes antioxydant présents dans les cellules a pour effet de favoriser la conversion de l' $\text{O}_2^{\bullet-}$ mitochondriale en H_2O_2 qui est le substrat d'un ensemble d'enzymes antioxydantes. Cependant H_2O_2 est produit également par les peroxysomes et le réticulum endoplasmique qui sont également des sites majeurs de la consommation d' O_2 , grâce à des enzymes tissu spécifique comme ERO1 (endoplasmic reticulum oxydase 1) dans le cas du réticulum endoplasmique. Cette production de H_2O_2 par ces organites est nécessaire afin de réaliser certaines de leurs fonctions. Néanmoins, le peroxysome et le réticulum endoplasmique maintiennent un équilibre dans la production de ROS par l'intermédiaire de molécules antioxydantes qui empêchent l'accumulation des ROS dans l'organite (Fransen et al., 2012).

IV.1.2.2. Les espèces réactives de l'azote et leur formation.

Les RNS incluent l'oxyde nitrique ou monoxyde d'azote (NO^\bullet), le dioxyde de nitrate NO_2^\bullet et le peroxyde nitrique (ONOO^-), qui sont aussi produits naturellement dans la cellule comme second messenger ou neurotransmetteur et interviennent également dans plusieurs processus cellulaires dont l'inflammation (Moncada and Bolaños, 2006; Valko et al., 2007a). L'oxyde nitrique est produit majoritairement par une famille de protéine, les NOS (pour nitric oxide synthase). Cette famille de protéines est composée de trois membres dont l'expression est tissu-spécifique et comprend nNOS/NOS1 (neuronal NOS) et eNOS/NOS3 (endothelial NOS) qui sont exprimées constitutivement, mais aussi iNOS/NOS2 (inducible NOS) dont l'expression nécessite un stimulus. Ces enzymes nécessitent pour leur activité de la L-Arginine et de l' O_2 comme substrat afin de produire lors de la réaction du NO^\bullet . De plus ces enzymes nécessitent des cofacteurs qui peuvent être le NADPH (Smith and Lassmann, 2002). Il est à noter que NO^\bullet peut être produit également dans la cellule à partir de NO_2^- . Ainsi, une partie de l'oxyde nitrique produit peut ensuite réagir avec $\text{O}_2^{\bullet-}$ pour générer ONOO^- , qui est un puissant oxydant présentant une chimie complexe avec les molécules biologiques (Beckman et al., 1990). Le peroxyde nitrique réagit avec les liposomes au niveau des acides gras insaturés et initie la peroxydation lipidique (Darley-Usmar et al., 1992; Radi et al., 1991). ONOO^- une fois protoné, peut subir un clivage homolytique et générer les radicaux OH^\bullet et NO_2^\bullet (pour revue Hogg and Kalyanaraman, 1999).

IV.1.2.3. Les systèmes antioxydants.

Bien que les ROS soient indispensables pour le fonctionnement cellulaire, une surproduction de ROS ou de RNS a des effets néfastes pour la cellule. Il est alors essentiel pour la cellule de maintenir un équilibre entre la production de ROS/RNS et leur élimination par des systèmes antioxydants. Il existe deux grands systèmes antioxydants dans la cellule, le système enzymatique et le système non enzymatique. Le système enzymatique fait intervenir de nombreuses enzymes qui agissent à différents niveaux dans la production des ROS et dans différents compartiments cellulaires avec des fonctions distinctes. Le processus enzymatique antioxydant le plus étudié met en jeu la superoxyde dismutase 2 (SOD2) qui intervient au sein de la mitochondrie et convertit l' $O_2^{\cdot-}$ généré par l'activité de la chaîne respiratoire en H_2O_2 . Il est à noter qu'il existe la SOD1 et 3, qui réalisent la même réaction enzymatique, mais qui sont respectivement localisées dans le cytoplasme et la matrice extracellulaire. L' H_2O_2 ainsi formé est en partie dégradé par différentes enzymes comme la catalase qui catalyse la dismutation de la molécule en H_2O et O_2 . D'autres enzymes sont essentielles dans ce système antioxydant enzymatique dont notamment les peroxyrédoxines (Prx) et les glutathions peroxydases (GPx) qui sont retrouvées dans différents compartiments cellulaires (Espinosa-Diez et al., 2015). Les GPx sont des enzymes majeures dans la réponse antioxydante de la cellule dans la mesure où ces protéines catalysent la réduction de H_2O_2 en H_2O par l'oxydation du glutathion (GSH) en disulfure de glutathion (GSSG) (Brigelius-Flohé, 2006). Le GSH est un tripeptide constitué d'un glutamate, d'une cystéine ainsi que d'une glycine et représente le principal composé antioxydant non enzymatique de la cellule. Il est à noter que la réduction du GSSG par la GSH réductase, afin de régénérer le GSH, nécessite l'utilisation de NADPH comme donneur d'électrons (Schafer and Buettner, 2001; Pastore et al., 2003). Ainsi, le NADPH joue également un rôle critique dans le système redox où il participe à la fois à la formation et à la dégradation des ROS. Cependant, il existe d'autres molécules essentielles qui appartiennent au système antioxydant non enzymatique de la cellule. Ces molécules, qui sont représentées essentiellement par des vitamines et des cofacteurs, peuvent agir comme pièges à espèces radicalaires. Ainsi, il est proposé que certaines vitamines (vitamine E et C par exemple) aient une action directe sur les ROS et neutralisent la capacité de ces espèces réactives à oxyder des molécules (Mangge et al., 2014). D'autre part, certains antioxydants non enzymatiques favorisent l'activation de voies

métaboliques impliquées dans la réponse antioxydante. C'est notamment le cas de la vitamine D et du cofacteur Q10 (Indo et al., 2015). En effet, une carence en vitamine D entraîne une dérégulation de l'équilibre redox au sein de la cellule (Indo et al., 2015). Ainsi, les systèmes antioxydants enzymatique et non enzymatique agissent en synergie afin de maintenir l'équilibre redox dans la cellule. Cependant, lorsque le système antioxydant ne peut plus maintenir la quantité de ROS à un niveau physiologique du fait soit d'une diminution de l'activité des systèmes antioxydants, soit d'une augmentation trop importante de la quantité de ROS, il s'établit alors un stress oxydant.

IV.2. DOMMAGES DES ROS ET DES RNS DANS LA CELLULE.

L'établissement d'un stress oxydant dans une cellule peut impacter de nombreux processus cellulaires et avoir des conséquences catastrophiques voire létales pour la cellule, selon l'importance du déséquilibre et la force du stress. En effet, dans des conditions physiologiques, les ROS et les RNS participent au bon fonctionnement du métabolisme cellulaire et sont intégrés dans la structure de macromolécules de manière contrôlée. Cependant, lors d'un stress oxydant, les ROS et les RNS, de par leur nature réactive, attaquent de manière incontrôlée les constituants cellulaires, en particulier les lipides, les protéines et l'ADN entraînant des dommages importants qui peuvent être irréversibles. De plus, l'augmentation du taux de ROS et de RNS entraîne également un déséquilibre de la balance redox qui intervient dans la régulation de nombreux processus biologiques qui sont essentiels au métabolisme cellulaire (Fulda et al., 2010; Genestra, 2007; Trachootham et al., 2008). Ainsi, le stress oxydant entraîne non seulement des dommages directs sur les constituants cellulaires, mais aussi une dérégulation globale du métabolisme cellulaire.

IV.2.1. La peroxydation des lipides.

IV.2.1.1. Les réactions de la peroxydation des lipides.

Les lipides et les phospholipides jouent de nombreux rôles dans la cellule, en particulier dans la constitution des membranes biologiques dont ils sont les principaux composants. L'une des conséquences du stress oxydant est l'oxydation des lipides et plus particulièrement la peroxydation des lipides insaturés et polyinsaturés de la cellule au niveau des membranes. Les ROS et le RNS peuvent directement réagir au niveau des carbones insaturés des lipides pour former une grande variété de produits d'oxydation selon un processus en trois étapes : l'initiation, la propagation et la terminaison (Figure IV.3). Lors de la première étape une molécule pro-oxydante comme $O_2^{\cdot-}$ réagit avec l'hydrogène d'un lipide (LH) pour former un radical lipidique (L^{\cdot}) qui a son tour réagit rapidement avec de l' O_2 pour former un lipide peroxydé (LOO^{\cdot}). Ce lipide peroxydé va ensuite réagir avec l'hydrogène d'un autre LH, générant ainsi un lipide hydroperoxyde (LOOH) et un nouveau L^{\cdot} qui va poursuivre la réaction en chaîne, c'est la phase de propagation. La phase de terminaison correspond à la phase de production de produit non radicalaire où de nombreux produits secondaires vont être formés à partir du produit primaire LOOH issus de la phase d'élongation, notamment le 4-hydroxynonanal (4-HNE) et le malondialdéhyde (MDA) (pour revue Valko et al., 2007b; Yin et al., 2011). Le 4-HNE et le MDA sont les produits de la peroxydation des lipides dont les effets sur la cellule sont les plus étudiés, dans la mesure où ces molécules sont extrêmement toxiques et mutagènes (pour revue Ayala et al., 2014; Zhong and Yin, 2015).

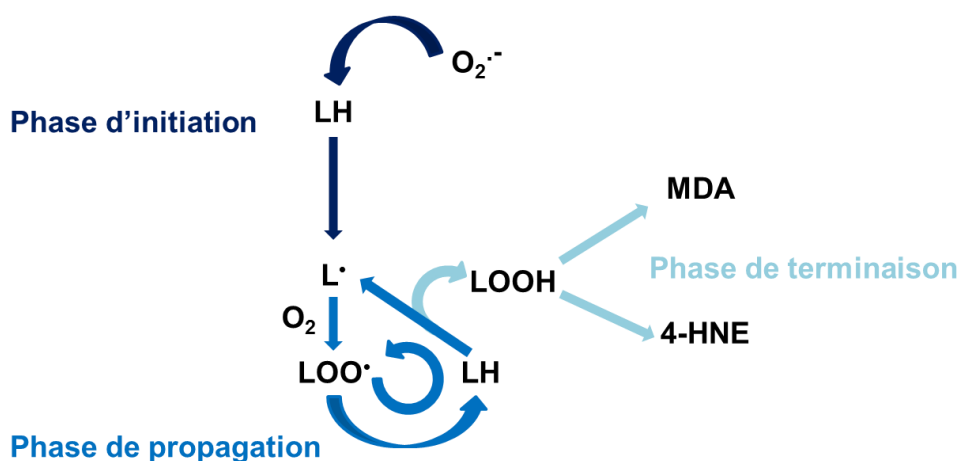


Figure IV.3 : Schéma du processus de peroxydation des lipides. Lors de la phase d'initiation (bleu foncé), une espèce radicalaire de l'oxygène réagit avec un lipide (LH) pour former un radical lipidique (L^{\bullet}). Dans la phase de propagation (bleu), le L^{\bullet} réagit rapidement avec de l' O_2 pour former un lipide peroxydé (LOO^{\bullet}) qui peut à son tour interagir avec un autre lipide pour générer un nouveau L^{\bullet} et un hydroperoxyde (LOOH). Le nouveau L^{\bullet} va à son tour réagir avec de l' O_2 et interagir avec un autre LH. Lors de la phase de terminaison les LOOH formés vont permettre la production de produits non radicalaires dont le hydroxynonanal (4-HNE) et le malondialdéhyde (MDA).

Il est à noter que deux LOO^{\bullet} peuvent former LOOL avec de l' O_2 et que deux L^{\bullet} peuvent interagir ensemble pour former deux L lors de la phase de terminaison. Par conséquent, la peroxydation lipidique induit non seulement une altération de la composition en lipides des membranes et une modification des propriétés membranaires, mais aussi la formation de molécules secondaires qui vont perturber le métabolisme cellulaire, ainsi que les voies de transduction du signal et générer ultimement des mutations génétiques.

IV.2.1.2. Conséquences de la peroxydation lipidique sur les membranes.

Les membranes permettent de délimiter les cellules du milieu extérieur en agissant comme une barrière, mais permettent aussi de délimiter les différents organites cellulaires. Chaque type d'organite a une fonction précise qui nécessite un environnement adapté pour le bon fonctionnement de l'activité métabolique associée au compartiment cellulaire. Ainsi, les membranes jouent un rôle essentiel dans l'établissement et le maintien d'un équilibre dans la composition interne d'un compartiment cellulaire. De plus, les membranes abritent de nombreuses protéines qui participent à l'activité et au bon fonctionnement du compartiment et, par conséquent, chaque compartiment est délimité par une membrane qui lui est spécifique avec une composition différente en protéines, en sucres, mais également

en lipides. La modification de la composition lipidique des membranes entraîne donc des changements dans la fluidité et le dynamisme des membranes cellulaires, qui peuvent perturber la perméabilité membranaire ainsi que de la composition des membranes en protéines et en sucres. D'où des modifications des équilibres ioniques normalement maintenus dans chaque compartiment cellulaire et des altérations des fonctions de chaque organite (Catalá, 2014; Janshoff and Steinem, 2015; Wong-ekkabut et al., 2007). C'est notamment le cas au niveau de la membrane interne mitochondriale avec la peroxydation de la cardiolipine qui est un phospholipide riche en acides gras polyinsaturés (Paradies et al., 2009). La cardiolipine est retrouvée essentiellement dans la membrane interne mitochondriale et participe à l'organisation structurale de nombreux complexes protéiques dont le cytochrome c et les complexes III/IV de la chaîne respiratoire (Paradies et al., 2011; Zhang et al., 2005). La peroxydation de ce phospholipide déstabilise la structuration de ces complexes, mais aussi la perméabilité de la membrane mitochondriale, notamment au Ca^{2+} qui est impliqué dans la régulation de nombreux processus biologiques au sein de la mitochondrie (Brookes et al., 2004). L'afflux de Ca^{2+} dans la mitochondrie, associé à l'altération membranaire, favorise le relâchement du cytochrome c dans le cytoplasme qui peut induire la voie de l'apoptose et la mort cellulaire (Ott et al., 2007; Paradies et al., 2009; Petrosillo et al., 2006). De plus, la déstabilisation des complexes de la chaîne respiratoire favorise la fuite des électrons et la production de ROS, ce qui peut amplifier l'intensité du stress (Paradies et al., 2001, 2002, 2014). Cependant l'impact des ROS et des RNS sur les membranes dépend de la force du stress, une faible augmentation des ROS/RNS entraîne une réponse de la cellule qui compense l'effet de la peroxydation. L'établissement d'un stress plus important mais modéré entraîne des altérations du métabolisme et favorise l'activation de l'apoptose. Par contre, un stress oxydant trop important associé à un niveau de peroxydation des lipides élevé entraîne la lyse des membranes et la nécrose cellulaire (Figure IV.4) (Girotti, 1998).

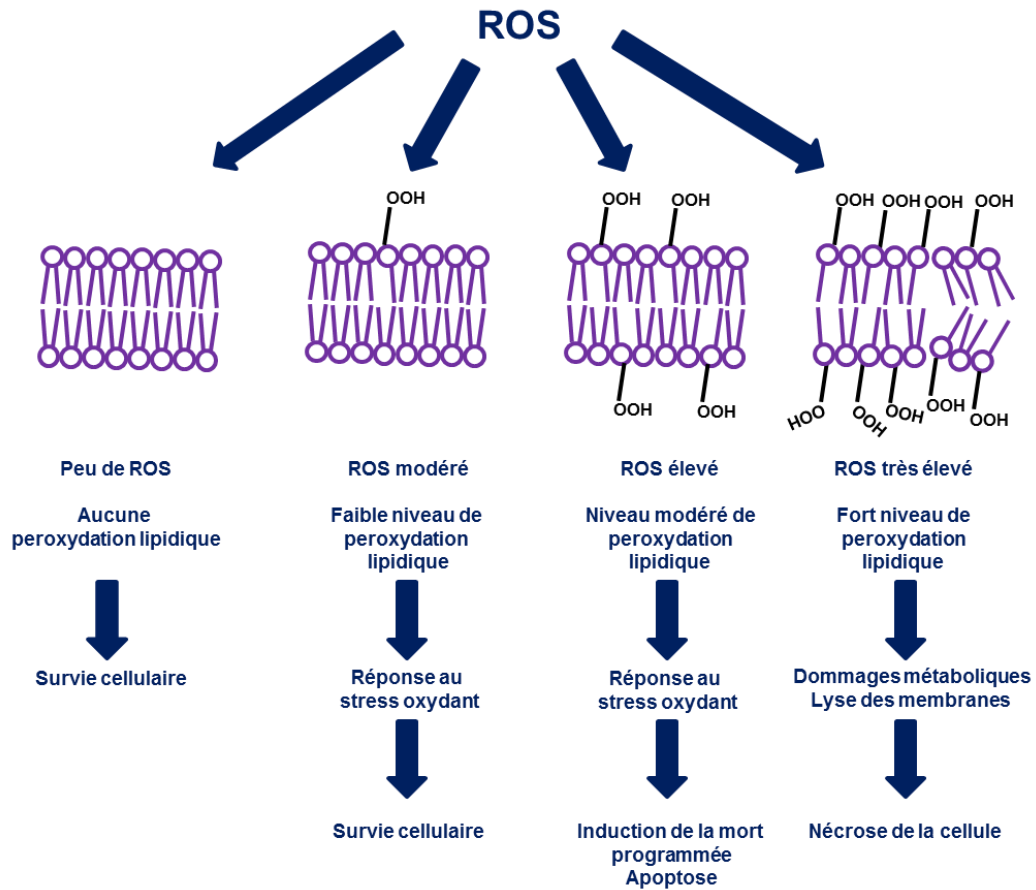


Figure IV.4 : Représentation de l'impact d'un stress oxydant selon le niveau de ROS sur la peroxydation lipidique et ses conséquences cellulaires. (d'après Girotti, 1998)

IV.2.2. Impact du stress oxydant sur l'ADN.

Les ROS sont naturellement capables de traverser l'enveloppe nucléaire et d'atteindre l'ADN qui peut être soumis à des oxydations au niveau de la plupart de ses composants. En effet, les bases purines et pyrimidine ainsi que les sucres qui composent l'ADN présentent de nombreux atomes d'azote et d'oxygène qui sont sensibles aux électrophiles comme HO[•]. Par conséquent, les ROS peuvent conduire à différents types de lésions à l'ADN en provoquant des cassures de l'hélice d'ADN (simple ou double-brin) et des modifications de bases dont la plus fréquente est la 8-hydroxyguanine (8-OH-G) (figure IV.5). Bien qu'il existe un système de réparation de l'ADN, certaines modifications de bases peuvent perdurer et conduire à la formation de mutations au sein de l'ADN (Dizdaroglu et al., 2002; Shigenaga et al., 1989; pour revue Cooke et al., 2003a).

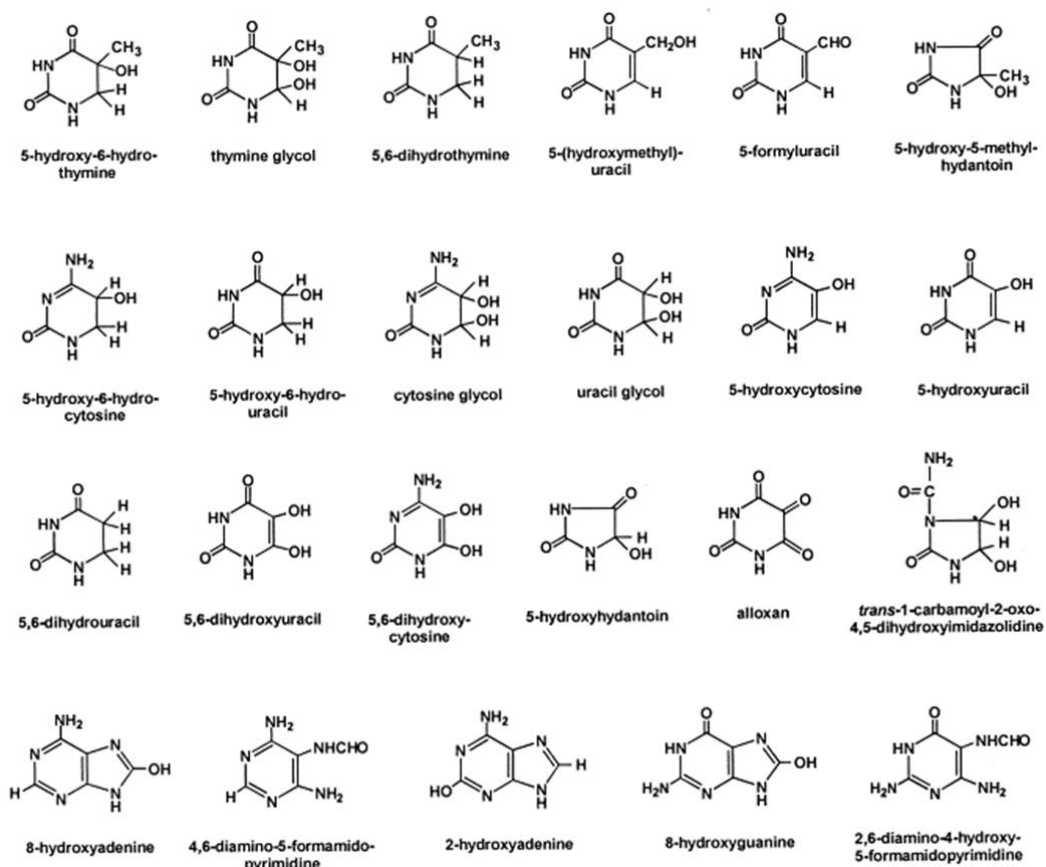


Figure IV.5 : Représentation des différentes bases de l'ADN pouvant être produites lors d'un stress oxydant (tirée de Cooke et al., 2003a).

En parallèle, la production de 4-HNE et de MDA affecte également l'ADN, suite à la peroxydation lipidique. En effet, ces composés peuvent interagir directement avec l'ADN au niveau des bases azotées. De plus, il a été montré que le 4-HNE peut affecter la télomérase, une enzyme impliquée dans la réplication des extrémités chromosomique (Pizzimenti et al., 2006; Winczura et al., 2012). Selon leur localisation dans le génome, ces mutations peuvent inhiber ou favoriser la transcription d'un gène, mais aussi l'inclusion d'un exon. Ces mutations peuvent également induire des erreurs de réplication et une instabilité génomique qui peuvent conduire à l'établissement de cancer (Cooke et al., 2003b; Valko et al., 2006).

IV.2.3. Effet des ROS et des RNS sur les protéines.

Au même titre que les lipides, les protéines peuvent subir également des oxydations qui peuvent être réversibles ou irréversibles selon la réaction d'oxydation réalisée et le niveau du stress oxydant. Bien que tous les acides aminés qui composent une protéine puissent être oxydés, certains résidus sont plus sensibles à l'oxydation, particulièrement les résidus contenant au sein de leur chaîne latérale un groupement amine ($-NH_2$), hydroxyle ($-OH$) ou un soufre (Levine et al., 1996). Il est à noter que ces résidus sont également des cibles préférentielles pour les modifications post-traductionnelles des protéines, telles que la phosphorylation, et sont généralement impliqués dans les sites actifs des enzymes (Spickett et al., 2006). Les principales modifications d'acides aminés générées lors d'un stress oxydant sont récapitulées dans la Figure IV.6. Les RNS peuvent également oxyder les chaînes latérales des acides aminés. Cependant, la nitrosylation par $ONOO^-$ entraîne une oxydation irréversible des résidus cystéines, tyrosines, méthionines et tryptophanes qui ne peuvent plus être modifiés post-traductionnellement, altérant ainsi la localisation et la fonction de certaines protéines (Radi et al., 2001). De plus, les dérivés carbonyles issus de la peroxydation lipidique comme le 4-HNE et le MDA peuvent également réagir avec les protéines en réalisant un pontage avec les chaînes latérales de certains acides aminés (Friguet and Szweda, 1997; Szweda et al., 1993). L'établissement de ces pontages conduit à des produits protéiques inactifs qui peuvent s'agréger en complexes qui inhibent l'activité du protéasome (Zhang et al., 2008b).

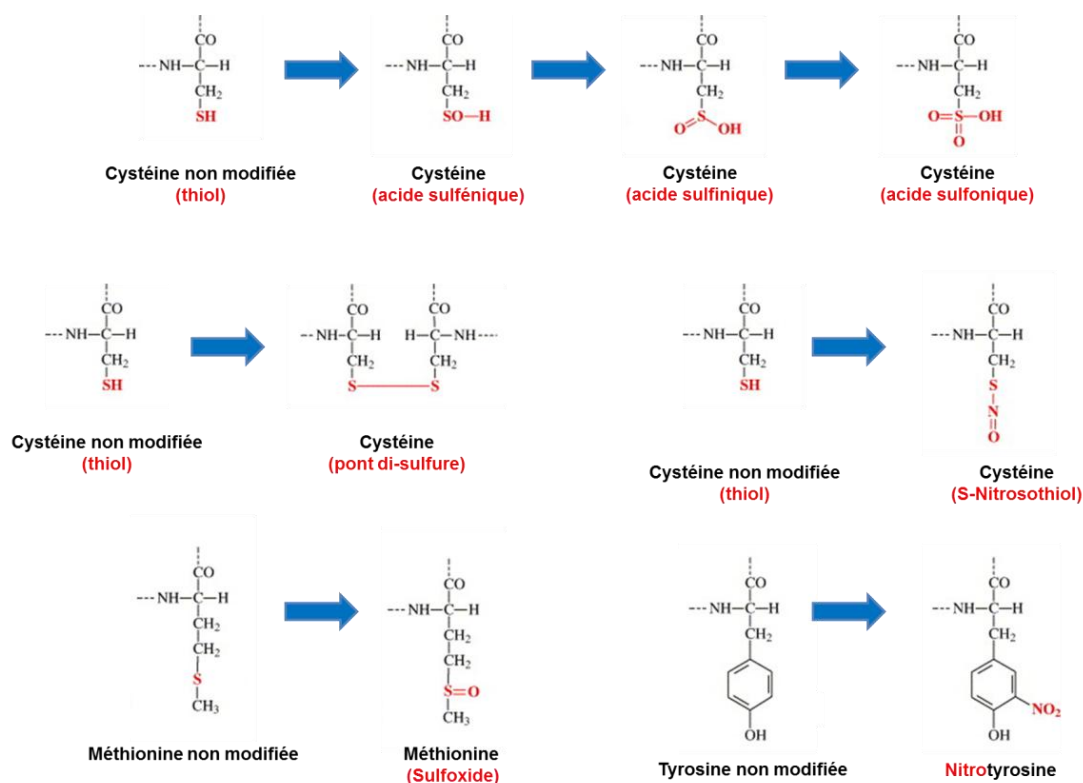


Figure IV.6 : Principales modifications des acides aminés cystéines, méthionine et tyrosine par des espèces réactives de l'oxygène. (tirée de Spickett et al., 2006)

L'impact du stress oxydant sur les fonctions thiols (-SH) des résidus cystéines (Cys-SH) a été particulièrement étudié. Ces résidus sont souvent impliqués dans les domaines fonctionnels ou régulateurs de nombreuses protéines dont l'activité dépend de l'état redox de la cellule. Il est par ailleurs proposé que l'état réversible d'oxydation de ces résidus joue un rôle important dans les voies de signalisation cellulaire. C'est notamment le cas de certaines tyrosines phosphatases qui sont impliquées dans la régulation des modifications post-traductionnelles des protéines. Ces enzymes possèdent deux Cys-SH catalytiques qui sont oxydées par les ROS en acide sulfénique (Cys-SOH). Les Cys-SOH sont particulièrement réactives et l'oxydation d'une des cystéines peut entraîner une réaction avec la fonction -SH de la seconde, ce qui entraîne la formation d'un pont disulfure (S-S). L'oxydation des cystéines en acides sulféniques ou la formation d'un pont disulfure entre les deux résidus cystéines est réversible mais entraîne une inhibition de l'activité enzymatique de la protéine (pour revue Meng et al., 2002; Reczek and Chandel, 2015). Il est cependant possible que les résidus cystéines soient suroxydés de manière irréversible en Cys-SO₂H puis en Cys-SO₃H (Marnett et al., 2003). De plus, les résidus cystéines peuvent être impliqués dans la

structuration des domaines en doigt de zinc et l'oxydation de ces résidus déstabilisent la conformation de ces domaines. Ainsi, lors d'un fort stress oxydant, les ROS favorisent l'activation des protéines kinases C (PKC) par la perte de la conformation des domaines en doigts de zinc de ces protéines (Giorgi et al., 2010). De manière intéressante, l'oxydation des résidus cystéines peut également entraîner la formation de ponts disulfures entre des complexes protéiques et entraîner leur inactivation. C'est le cas du complexe SMN qui participe à la maturation des snRNP. En effet, il a été mis en évidence la formation d'un pontage entre deux protéines SMN par l'intermédiaire d'un pont disulfure lors d'un stress oxydant qui entraîne l'inactivité du complexe SMN (Wan et al., 2008). Par conséquent, au-delà des altérations oxydatives des protéines qui altèrent leur activité localement, l'oxydation des protéines peut engendrer des modifications de mécanismes plus généraux comme la maturation des snRNA, mais aussi les mécanismes de régulation de la cellule. En effet, les tyrosines phosphatases et les PKC sont des protéines impliquées dans de nombreuses voies métaboliques et dans la transduction du signal. Il est cependant à noter que ces modulations dans la transduction du signal sont également nécessaires pour activer la réponse au stress oxydant et permettre la survie cellulaire (pour revue Cosentino-Gomes et al., 2012).

IV.3. LES PRINCIPAUX MECANISMES DE REPONSE AU STRESS OXYDANT.

Les modifications réversibles comme irréversibles des composants cellulaires induites par l'augmentation des ROS participent à la transduction du signal. En effet, dans la signalisation redox, l'oxydation de molécules cibles par les ROS et les RNS agit comme un signal interrupteur pour l'activation ou l'inhibition de certaines voies de régulation. Dans la mesure où les bio-particules n'ont pas la même sensibilité au stress oxydant, il est suggéré que le niveau d'activation de certaines voies de régulation dépende de l'intensité du stress et de ses effets sur les composants cellulaires. Ainsi, le niveau de modifications générées au sein des protéines de la transduction du signal entraîne une réponse cellulaire adaptée aux différents stress afin de permettre la survie cellulaire. Le principal mécanisme de réponse au stress oxydant est l'augmentation des capacités antioxydantes de la cellule et nécessite l'activation de la protéine Nuclear factor-E2-related factor 2 (Nrf2) (Chen et al., 2015a).

IV.3.1. La voie Nrf2 : la voie antioxydante.

IV.3.1.1. La protéine Nrf2 et son activation.

La protéine Nrf2 est un facteur de régulation transcriptionnel exprimé constitutivement dans la cellule. Cette protéine est composée de 6 domaines Neh (Neh1 à 6) qui sont hautement conservés dans de nombreuses espèces (Figure IV.7). Les domaines Neh1 et 2 contiennent différents motifs qui permettent à la protéine Nrf2 d'interagir avec plusieurs protéines dont notamment Keap1 (pour Kelch-like ECH-associated protein 1) au niveau des motifs ETGE and DLG du domaine Neh2 (McMahon et al., 2006; Tong et al., 2006). Le domaine Neh3 correspond au domaine C-terminal de la protéine et sa fonction n'a pas encore été clairement élucidée. Les domaines Neh4 et 5 sont essentiels à la régulation de la transcription par la protéine Nrf2. Enfin le domaine Neh6 est impliqué dans la régulation de la stabilité de la protéine (pour revue Chen et al., 2015)(Figure IV.7).

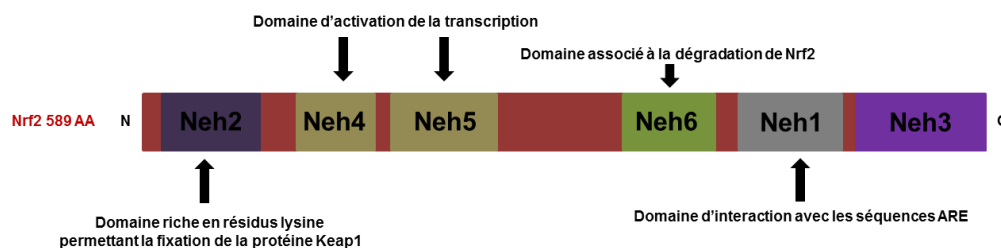


Figure IV.7 : Schéma de l'organisation des domaines fonctionnels de la protéine Nrf2

Cette protéine est capable d'interagir avec l'ADN au niveau de promoteurs transcriptionnels appelés antioxidant-response elements (AREs) qui sont associés aux gènes codant la plupart des protéines des systèmes antioxydants, mais aussi d'autres systèmes impliqués dans la survie de la cellule. Cependant, dans des conditions physiologiques, cette protéine est localisée dans le cytoplasme de la cellule en interaction avec la protéine Keap1. Keap1 agit comme un adaptateur pour la protéine cullin 3 (CUL3) qui va favoriser l'ubiquitination et la dégradation de Nrf2 par ubiquitination (Kobayashi and Yamamoto, 2006). Par conséquent, la protéine Nrf2 est continuellement dégradée dans des conditions physiologiques et ne possède qu'une faible activité. Lors d'un stress oxydant, les résidus Cys de la protéine Keap1

sont oxydés, ce qui entraîne la dissociation de la protéine avec Nrf2 qui est alors activée et se relocalise dans le noyau (Kobayashi and Yamamoto, 2006) (Figure IV.8). En addition à l'oxydation de Keap1, Nrf2 peut également être dissociée de Keap1 au travers de sa phosphorylation par différentes kinase dont les PKC qui sont activées lors d'un stress oxydant (Huang et al., 2002). Dans le noyau, Nrf2 va former un hétérodimère avec différentes protéines qui vont moduler son activité, comme les protéines c-Jun, Maf et ATF-4 au niveau des AREs (pour revue Hayes and McMahon, 2001). La fixation de Nrf2 sur les AREs va permettre de réguler la transcription de gènes impliqués dans de nombreuses fonctions cellulaires dont en particulier l'expression de protéines à activités antioxydantes et/ou anti-apoptotiques (pour revue Huang et al., 2015b).

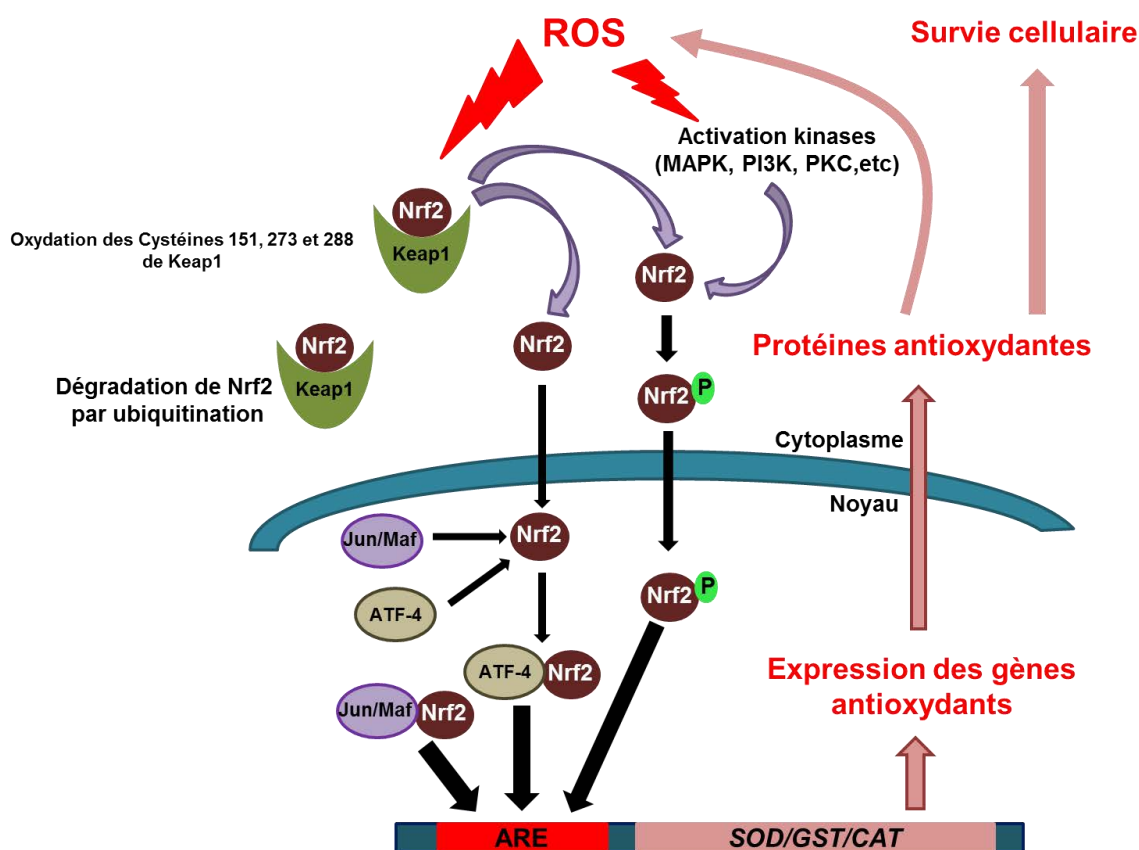


Figure IV.8: Schéma du mécanisme d'activation de la voie Nrf2 suite à un stress oxydant et de l'activation de l'expression des gènes ARE-dépendant. L'augmentation du taux de ROS dans le cytoplasme favorise l'oxydation de la protéine Keap1 qui ne peut plus interagir avec la protéine Nrf2 et entraîne sa dégradation par ubiquitination. Parallèlement Nrf2 peut également être phosphorylée (P) par l'activation de kinases suite au stress oxydant. Nrf2 se relocalise dans le noyau où il peut former un hétérodimère avec les protéines Jun/Maf où ATF-4. La reconnaissance des séquences ARE par Nrf2 active la transcription des gènes de réponse au stress oxydant comme notamment les superoxydes dismutases (SOD), la glutathione S-transferase (GST) et la catalase (CAT).

IV.3.1.2. Protéines dont l'expression est modulée par l'activation de la voie Nrf2.

Des études chez la souris ont montré que Nrf2 est impliquée dans la régulation de plusieurs centaines de gènes codant des protéines dont les fonctions suggèrent que l'activation de Nrf2 a un impact majeur dans la réponse au stress oxydant et dans la survie cellulaire (Kwak et al., 2003; Malhotra et al., 2010). Ces études ont également montré que l'activation de la protéine Nrf2 conduit à la surexpression de certaines catégories de protéines, telles que les protéines de détoxification (la catalase, les SODs, la glutathione-S-transférase et les glutathions peroxydases) ou de la réparation de l'ADN comme PARP1, et à une diminution de l'expression des protéines impliquées dans la synthèse des ROS et RNS (NOX et iNOS) (pour revue Chen et al., 2015). Ainsi, l'activation de la protéine Nrf2 conduit à limiter et à réparer les dégâts oxydatifs lors d'un stress oxydant. En parallèle, de nombreuses protéines associées aux voies de transduction du signal sont également modulées afin d'inhiber la voie apoptotique et favoriser la survie cellulaire (Kumar et al., 2013b). De plus, certaines heat shock proteins (HSP), qui sont des protéines chaperonnes impliquées dans la réponse au stress thermique, sont également surexprimées, ce qui suggère des voies communes dans la réponse à différents stress cellulaires (Alani et al., 2015; Dayalan Naidu et al., 2015; Hensen et al., 2013). La réponse au stress oxydant au travers de l'activation de la protéine Nrf2 nécessite donc un remodelage du profil d'expression de l'ensemble du génome.

IV.3.2. Adaptation métabolique au stress oxydant : les granules de stress cytoplasmiques.

Le stress oxydant induit de nombreux dommages cellulaires et affecte la plupart des processus cellulaires. Par conséquent, la cellule doit mobiliser de nombreuses ressources afin de produire en masse les protéines nécessaires à sa survie comme les enzymes de détoxification et de réparation, mais aussi permettre leur activité afin de rétablir son homéostasie. Dans ce contexte, les cellules de mammifères ont maintenu au cours de l'évolution des mécanismes qui permettent l'arrêt des processus dispensables, comme la traduction d'ARNm non essentiels, et une redistribution des ressources vers les processus indispensables à la survie cellulaire.

IV.3.2.1. Formation des granules de stress cytoplasmique et blocage transitoire de la traduction de certains ARNm.

Suite à divers stress dont le stress oxydant, des granules de stress (SG) se forment au sein du cytosol des cellules. La formation de ces granules est liée à un évènement clé qui est la phosphorylation induite du facteur d'initiation de la traduction eIF2 α . Cet évènement inhibe la traduction coiffe-dépendante des ARNm en limitant l'action du complexe eIF2-GTP-tRNAiMet sur les ARNm, qui ne peut plus charger l'ARNt méthionine initiateur. Il en résulte un blocage dans l'initiation de la traduction et une dissociation des polysomes. Les ARNm ainsi non traduits, présents sous forme de ribonucléoprotéines messagers (RNPm), se chargent alors avec plusieurs partenaires protéiques spécifiques notamment TIA1 et G3BP, ce qui modifie la nature des RNPm. Les RNPm vont ensuite s'agrèger par oligomérisation pour former les granules de stress cytoplasmiques. La formation des SG favorise la relocalisation transitoire d'un grand nombre de molécules, en particulier des protéines impliquées dans la stabilité et la traduction des ARNm tel que PABP, eIF3, eIF4G et hnRNP A1 mais aussi de l'ARNr 40S. Ainsi, la formation des granules de stress cytoplasmiques permet de limiter la traduction coiffe-dépendante des ARNm qui ne sont pas essentiels à la survie cellulaire (Bley et al., 2015). De plus, la formation des SG permet également de relocaliser transitoirement les protéines associées à la prolifération cellulaire et les protéines pro-apoptotiques ou impliquées dans les voies de signalisation de l'apoptose contribuant ainsi à la survie cellulaire (Arimoto et al., 2008). Fait intéressant, il est proposé que les SG participent également à l'inhibition de l'apoptose en réduisant la production de ROS au cours d'un stress oxydant (Takahashi et al., 2013).

La plupart des SG sont retrouvés à proximité ou en interaction avec les P-bodies (PB) qui sont des granules cytoplasmiques distincts des SG et qui ont pour principale fonction la dégradation des ARNm et/ou la répression des ARNm. Les PB peuvent être naturellement présents dans le cytoplasme des cellules mais leur nombre augmente suite à un stress oxydant. Les PB sont constitués de protéines spécifiques, mais il existe entre les PB et les SG des liens tant au niveau de leur composition que de leurs fonctions qui suggèrent la mise en place d'un cycle des ARNm depuis l'inhibition de leur traduction jusqu'à leur dégradation (Kedersha et al., 2005; pour revue Buchan and Parker, 2009). De plus, il a été mis en évidence que les SG se forment au niveau des PB lors d'un stress oxydant. Cependant, des

SG peuvent également se former de manière PB-indépendante (Mollet et al., 2008). La cinétique de formation de ces granules n'est pas encore complètement élucidée et des études suggèrent qu'elle serait dépendante de la nature du stress et du type cellulaire (Buchan et al., 2008; Kedersha et al., 2005). Néanmoins, une étude récente, basée sur des prédictions informatiques et des résultats expérimentaux, a proposé un modèle d'assemblage des SG suite à un stress oxydant montrant une phase de latence après l'initiation du stress, puis une forte augmentation du nombre de SG dont la taille augmente au cours du temps. De plus, ce modèle met en évidence un dynamisme dans l'espace des SG qui se déplacent dans le cytoplasme par le biais des microtubules et d'un transport antérograde vers la région périnucléaire (Loschi et al., 2009; Ohshima et al., 2015). Il existe également un dynamisme important dans la composition des SG et des PB qui est essentiellement dû à des échanges de RNPM entre les SG et les PB, mais aussi entre ces granules et le cytoplasme (Parker and Sheth, 2007). Ainsi, la séquestration de certains ARNm n'est qu'un état transitoire et la traduction des ARNm peut être réactivée.

IV.3.2.2. Traduction indépendante de la coiffe.

La formation des granules de stress cytoplasmiques (SG) régule l'expression des ARNm en inhibant la traduction coiffe-dépendante. Cependant, les SG peuvent également favoriser la traduction de certains ARNm par des processus non-canoniques tels que l'utilisation de sites d'initiation internes d'entrée du ribosome (IRES pour internal ribosome entry site). Les ARNm dont la traduction est IRES-dépendante recrutent les facteurs de traduction et le ribosome de manière coiffe-indépendante et sont préférentiellement traduits lors d'un stress (Spriggs et al., 2008). La protéine hnRNP A1 qui est retrouvée dans les SG interagit avec les IRES de certains ARNm et favorise ainsi leur traduction (Damiano et al., 2013). Il est estimé que 10 à 15% des ARNm possèdent un IRES au sein de leur région 5'-UTR. Ces ARNm codent généralement des protéines impliquées dans la réponse aux stress, ainsi que dans la survie cellulaire (Spriggs et al., 2008). C'est notamment le cas de l'ARNm codant la protéine Nrf2 (Li et al., 2010).

C'est la présence de cet IRES qui permet à la cellule de maintenir la réponse antioxydante par l'expression constante de la protéine Nrf2 tout au long d'un stress oxydant. Ainsi, la formation des SG et la traduction IRES-dépendante des protéines essentielles à la survie cellulaire participent à la reprogrammation de l'expression des protéines. Néanmoins, des études suggèrent que la reprogrammation du profil d'expression des protéines se réalise également en amont lors de l'épissage alternatif et du transport des ARNm.

IV.3.3. Adaptation métabolique au stress oxydant : les granules nucléaires.

La littérature reporte un effet du stress oxydant sur différentes structures nucléaires, cependant beaucoup de ces structures, bien que portant des noms différents, réfèrent en fait à un même type de granules. Par souci de clarté, je distinguerai ici deux types de structures nucléaires : les corps nucléaires PML (PML NBs), appelés ainsi du fait de la présence au sein de ces granules de la protéine promyelocytic leukemia (PML) et les granules de stress nucléaires ou nSBs (pour nuclear stress bodies).

III.3.3.1. Les corps nucléaires PML (PML NBs).

Les PML NBs sont des structures de petite taille (0,1-1 μ m), au nombre de 5 à 15 par cellule (Lallemand-Breitenbach and de Thé, 2010). Ils ne contiennent ni ARN, ni ADN et sont principalement constitués de protéines, une centaine répertoriées à ce jour, dont la seule propriété commune est leur capacité à être sumoylées (Bernardi and Pandolfi, 2007). La formation des PML NBs est un processus en 2 étapes : une étape initiale, redox-dépendante, de multimérisation de la protéine PML, suivie du recrutement SIM-SUMO-dépendant des autres composants protéiques des PML NBs (Sahin et al., 2014) (Figure IV.9). Ces granules, à la différence des nSB, sont déjà présents dans les cellules en absence de stress, cependant leur morphologie et leur nombre sont affectés par les différents stress cellulaires ou infections virales subis par la cellule. Ainsi, un stress oxydant à l'arsenite induit une diminution du nombre de PML NBs et une augmentation de leur taille (Lallemand-Breitenbach et al., 2001; Zhu et al., 1997), alors qu'un choc thermique va conduire à la dispersion des PML NBs (Eskiw et al., 2003). Les processus cellulaires pouvant être impactés par les PML NBs sont assez diverses et comprennent aussi bien la régulation de la stabilité du génome, que la réparation de l'ADN, le contrôle de la transcription ou la défense virale

(pour revue Dundr and Misteli, 2010). Dans le cadre d'un stress oxydant à l'arsenite il est proposé que l'augmentation du degré d'oxydation de PML favorise la formation de PML NBs plus grands et augmente la sumoylation, la poly-ubiquitination et la dégradation subséquente de certaines protéines présentes dans les PML NBs via le protéasome dont les sous-unités sont relocalisées dans les PML NBs suite à un stress oxydant (Lallemand-Breitenbach et al., 2008; Sahin et al., 2014). Beaucoup des protéines présentes au sein des PML NBs étant impliquées dans la modification post-traductionnelle des protéines, il est proposé que ce mécanisme permette de moduler le degré global de modifications post-traductionnelles des protéines en fonction de l'intensité du stress oxydant subi par la cellule et constituerait une étape clé dans la réponse des cellules à un stress oxydant, notamment au travers de la régulation de la phosphorylation de p53 (Sahin et al., 2014).

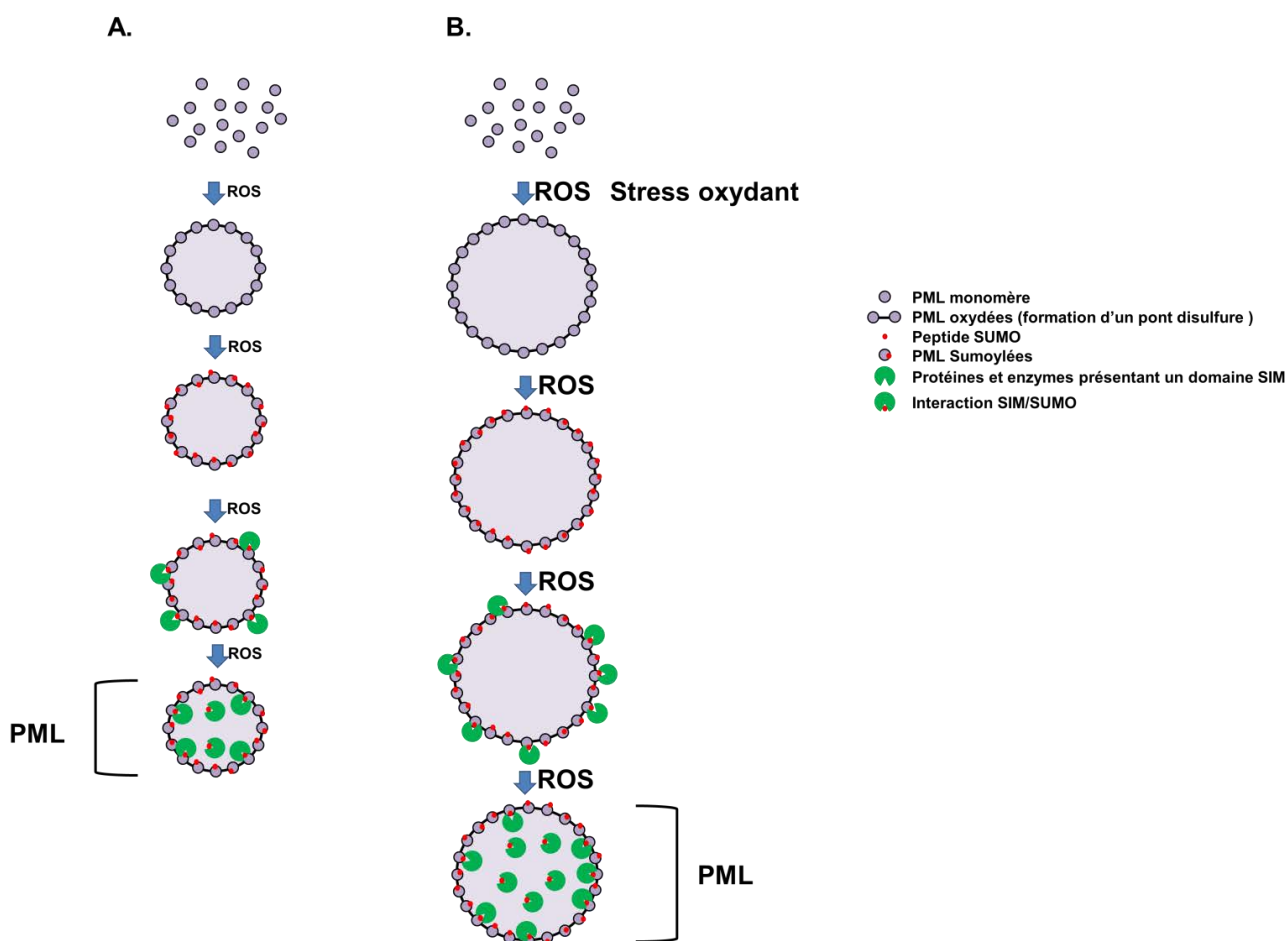


Figure IV.9 : Schématisation de la formation des PML (Adaptée de Sahin et al., 2014). A. Formation redox-dépendante des PML en absence de stress. B. Formation redox-dépendante des PML en condition de stress oxydant. Les PML générés en conditions de stress ont une taille plus importante mais sont moins nombreux.

III.3.3.2. Les granules de stress nucléaires (nSBs) et les ARN satellites III.

Les nSBs sont des structures plus grosses que les PML NBs (1-2 μ m) et ils sont généralement moins nombreux dans les cellules (2 à 6 nSBs par cellule), cependant leur taille et nombre peuvent varier selon la nature du stress subi par la cellule (Dundr and Misteli, 2010). Les nSBs sont également connus sous le nom de granules HSF1, granules HAP ou corps nucléaires SAM68 induits par le stress (stress-induced Sam68 nuclear bodies ou stress-induced SNB) (Biamonti, 2004). Contrairement aux PML NBs, leur formation requiert l'expression d'ARN spécifiques, les ARN satellites III. Les ARN Sat III sont de long ARN non codants de taille variable, mais présentant une organisation similaire (Figure IV.10). Ils sont tous constitués de deux types de séquences répétées. D'une part, des séquences pentanucléotidiques répétées en tandem caractérisées par une séquence consensus GGAAU, qui peut présenter des variations au niveau d'une des positions par rapport à la séquence consensus. D'autre part, des décanucléotides, de consensus CAACCCGAGU, pouvant présenter une plus grande variabilité dans les ARN Sat III. Au sein des ARN Sat III ces deux types de séquences sont organisés sous forme d'unités structurales composées de 1 à 39 pentanucléotides, suivis d'un seul décanucléotide. Cette unité structurale est répétée un certain nombre de fois au sein des ARN Sat III et le nombre de répétitions détermine la taille de l'ARN (Valgardsdottir et al., 2005) (Figure IV.10). Les ARN Sat III sont transcrits par l'ARN polymérase II à partir de régions centromériques et péricentromériques des chromosomes 9, 12 et 15 (Denegri et al., 2002; Jolly et al., 2002). Initialement, ces régions ont été décrites comme transcriptionnellement inertes dans la mesure où elles sont structurées sous forme d'hétérochromatine. Cependant, différents stress cellulaires peuvent induire leur transcription et leur expression serait essentielle non seulement pour l'adaptation des cellules à différents stress, mais également pour la formation des kinétochores lors de la division cellulaire chez la drosophile (Valgardsdottir et al., 2008; Rošić et al., 2014).

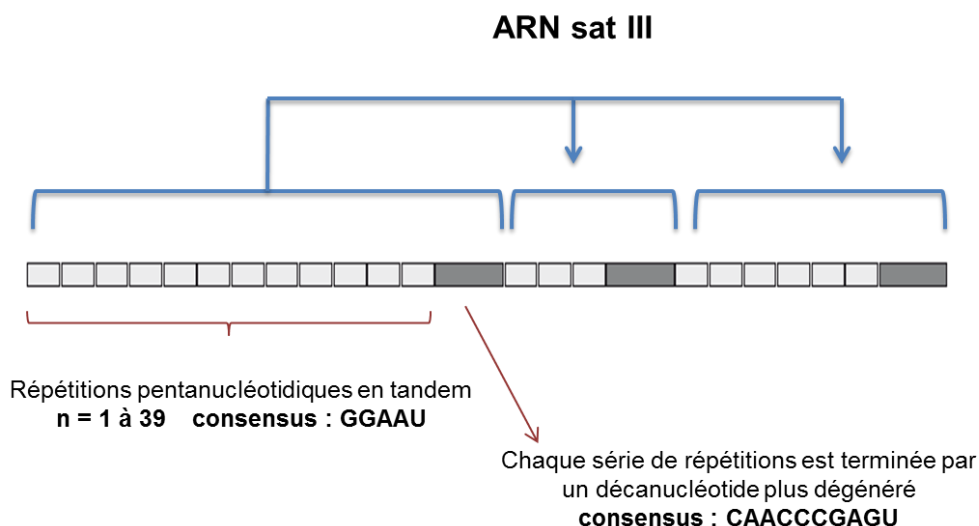


Figure IV.10 : Organisation des séquences d'ARN Sat III (d'après Valgardsdottir et al., 2005). Chaque unité structurelle est constituée de répétitions de pentanucléotides de consensus GGAAU (rectangles claires) suivies d'un décanucléotide (rectangles foncés). Chaque ARN Sat III est constitué d'un nombre variable d'unités structurelles d'où la taille variable des ARN Sat III.

III.3.3.3. Transcription des ARN Sat III et formation des nSBs lors d'un stress oxydant.

Les ARN Sat III sont transcrits suite à différents stress cellulaires, cependant leur niveau d'expression et la cinétique de leur transcription au cours de la réponse au stress diffèrent selon les stress (Figure IV.11). Ainsi, lors d'un choc thermique modéré (1 à 2h à 42°C, puis 1h de récupération à 37°C, cellules HeLa) l'induction de l'expression des ARN Sat III est très rapide et très forte (induction >1000 fois dès 30 minutes après le début du choc thermique) et conduit à la formation de nombreux granules dans près de 90 % des cellules, alors qu'un stress hyperosmotique (8 h en présence de 0.8M sorbitol, cellules HeLa) induit une expression plus tardive et plus modérée des ARN Sat III : induction de 10 fois 2h après la fin du stress et de 100 à 1000 fois 16h après la fin du stress associée à la formation de granules moins nombreux et plus petits dans 30 % des cellules (Valgardsdottir et al., 2008). Deux facteurs de transcription ont été décrits pouvoir activer la transcription des ARN Sat III : le facteur de transcription spécifique du choc thermique HSF1 (pour Heat Shock Factor 1) et le facteur de transcription TonEBP (pour Tonicity Element Binding protein), également appelé OREBP (pour Osmotic Response Element Binding Protein) ou NFAT5 (Ferraris et al., 2002; Valgardsdottir et al., 2008; Sengupta et al., 2009). Des sites de fixation pour chacun de ces

facteurs de transcription, HSE (pour Heat Shock Element), de consensus nGAAn, pour HSF1 et TonE (pour Tonicity Element), de consensus TGGAAANN(C/T)N(C/T) pour TonEBP, ont été trouvés dans plusieurs clones d'ADN satellite III. Bien qu'aucune étude exhaustive de l'identité du facteur de transcription selon le stress subi n'ait été réalisée, on notera que HSF1 est utilisée dans le cadre de la réponse au choc thermique, alors que TonEBP est spécifique du stress osmotique, d'où la question d'une corrélation possible entre la cinétique et le niveau d'expression des ARN Sat III et la nature du facteur de transcription responsable de leur expression. Lors d'un stress oxydant (20 min d'exposition à 200 μM H_2O_2 , cellules HeLa) l'induction de l'expression des ARN Sat III est faible et tardive (<100 fois, 16 heures après la fin du stress) et conduit à la formation de nSBs de très petite taille et en faible nombre (1 à 2 micro-nSBs par cellule) dans seulement 5 % des cellules. Le facteur de transcription responsable de la transcription des ARN Sat III dans le cadre d'un stress oxydant n'a pas été recherché.

Niveau d'induction Temps d'induction	Pas d'induction	Induction faible	Induction moyenne	Induction forte
Précoce	Hypoxie Cobalt Thymidine Hypo-osmotique	MMS	Cadmium (1h+régénération)	Choc thermique
Moyenne		Etoposide Aphidicoline		Cadmium (6h constant)
Tardif		Oxydant (20 min + 16h de Régénération)	UV-C Hyper-osmotique	

Figure IV.11 : Effets de différents stress sur la production d'ARN sat III et la cinétique des nSB (adapté de Valgardsdottir et al., 2008). Trois niveaux d'induction des ARN Sat III sont distingués, faible (<100x), moyen (entre 100 et 1000x) et fort (>1000x). Les inductions aux temps précoces correspondent à l'apparition des nSB entre 2h et 4h après la fin du stress, entre 6h et 8h pour les temps moyens et plus de 10h pour les temps d'inductions tardifs. Les différents stress ont été réalisés pendant 1h sauf indication dans les conditions suivantes : cadmium : 50 μM de cadmium ; hyper-osmotique : 0.8M sorbitol ou urée ou 0.1 NaCl ; UV-C : 80 J/m^2 ; oxydant 200 μM H_2O_2 ; méthyl méthanosulfate (MMS) : 3.54 mM ; aphidicoline : 2 $\mu\text{g}/\text{mL}$; étoposide : 100 μM ; thymidine : 2mM ; Cobalt : 21% d' O_2 + 100 μM COCl_2 ; hypoxie : 3% O_2 ; hypo-osmotique : 30% DMEM+ 70% eau et choc thermique : 42°C.

De façon intéressante, une augmentation du taux de ROS intracellulaires peut contribuer à l'activation de HSF1 et TonEBP. En effet, un stress oxydant peut induire la formation d'un pont disulfure impliquant les résidus Cys C35 et C105 de HSF1, restructurant ainsi HSF1 et favorisant son activation et sa relocalisation dans le noyau par multimérisation (Ahn and Thiele, 2003). Par ailleurs, une augmentation du taux de ROS intracellulaires est nécessaire pour l'activation de TonEBP lors d'un stress hyperosmotique, mais ne suffit pas, à elle seule, à activer TonEBP (Zhou et al., 2006a). Il est par conséquent tout à fait envisageable que ces deux facteurs de transcription puissent être impliqués dans l'expression des ARN Sat III lors d'un stress oxydant. La biogenèse des nSBs a été plus particulièrement étudiée dans la cadre d'un stress thermique. Un choc thermique modéré entraîne la relocalisation de la protéine HSF1 et la formation de granule HSF1 au niveau des régions péri-centromérique des chromosomes contenant les séquences Sat III qui sont alors transcrites par l'ARN polymérase II (Eymery et al., 2010; Jolly et al., 2002, 2004). Les ARN Sat III restent associés à leurs sites de transcription au sein des granules HSF1 par un mécanisme encore non élucidé et provoque la formation de nSBs par la relocalisation de plusieurs facteurs protéiques. De nombreuses protéines associées au nSBs ont été identifiées suite à un choc thermique, dont notamment des protéines de régulation de l'épissage alternatif comme SRSF1, SRSF7 et SRSF9 suggérant un rôle des nSBs dans la régulation de l'épissage alternatif (Metz et al., 2004 ; Biamonti and Vourc'h, 2010). Dans le cas de SRSF1 cette relocalisation au sein des nSBs est quasi-totale alors que pour certaines protéines, telles que SRSF7, cette relocalisation est partielle ou transitoire dans la mesure où SRSF7 est également présente dans les speckles. Il a été établi que certaines protéines comme hnRNP HAP sont présentes dans les nSBs suites à différents stress, suggérant que la composition des nSBs pourrait être identique indépendamment de la nature du stress subi par la cellule (Valgardsdottir et al., 2008). Néanmoins, aucune démonstration formelle de ce point n'a été publiée. Concernant le rôle des nSBs, la présence de nombreux facteurs de régulation de l'épissage au sein de ces granules, a conduit à proposer un rôle majeur de ces granules et des ARN Sat III dans la réponse et l'adaptation cellulaires au stress via la régulation de l'épissage alternatif. Ainsi, la séquestration de plusieurs facteurs de régulation de l'épissage dans les nSB diminuerait leur concentration effective dans le nucléoplasme et affecterait le profil d'épissage alternatif global dans la cellule après le stress (Figure IV.12).

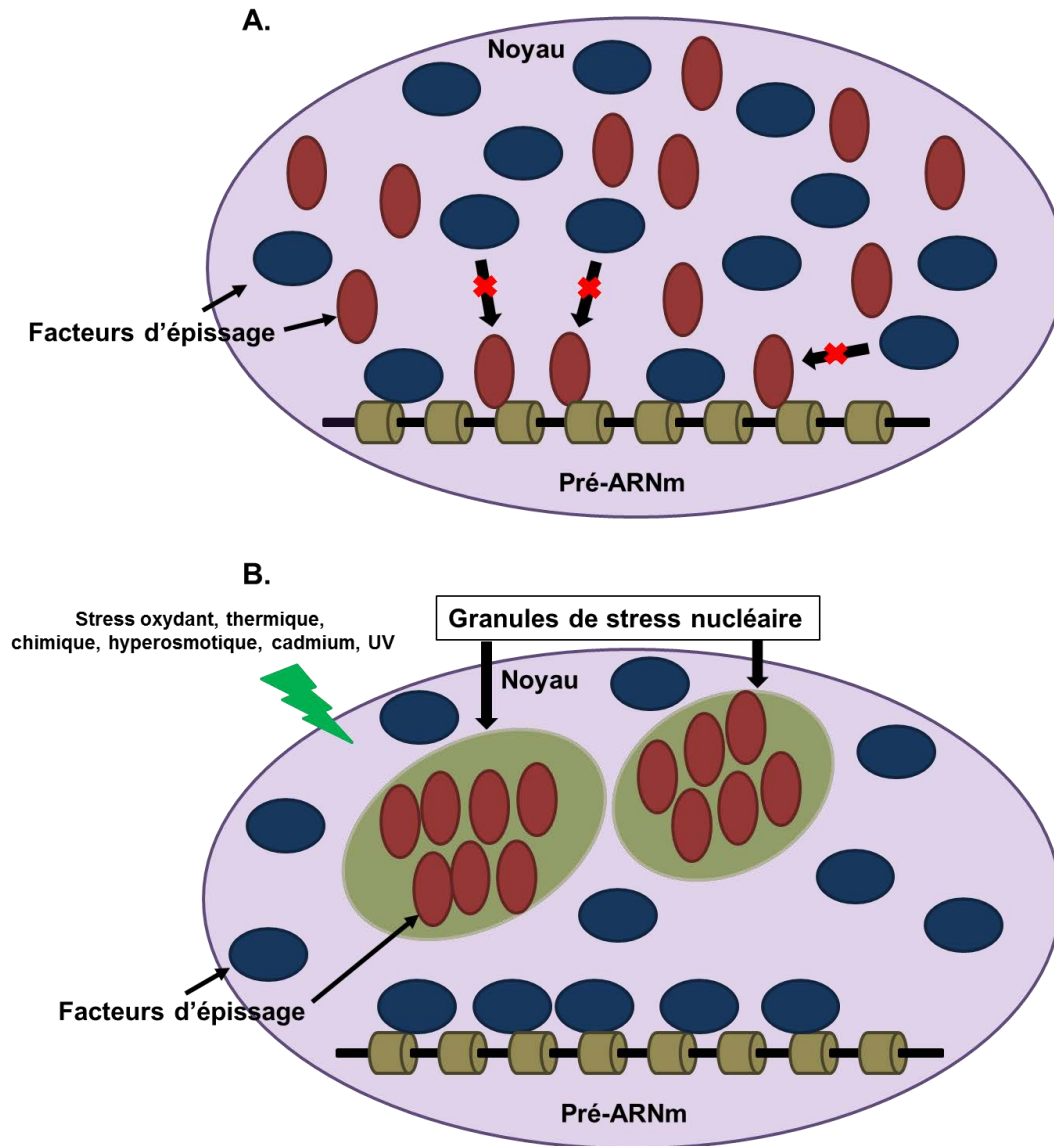


Figure IV.12 : Modèle proposé actuellement pour la régulation globale de l'épissage alternatif dans le noyau par formation des granules de stress nucléaires. A. Cellules non stressées où l'ensemble des facteurs de régulation de l'épissage intervient dans la maturation des pré-ARNm. **B.** Cellules stressées avec formation de granules de stress nucléaires qui entraînent la relocalisation de certaines protéines de régulation de l'épissage et altère de ce fait la régulation de l'épissage alternatif.

IV.3.4. Régulation de l'épissage et de l'épissage alternatif suite à différents stress.

Comme nous venons de le voir, les stress induisent de nombreuses perturbations du métabolisme cellulaire, ainsi qu'une reprogrammation de l'expression génique au travers de la régulation de divers processus cellulaires, dont l'épissage alternatif. Ainsi, un stress génotoxique entraîne des modifications de l'épissage alternatif de nombreux pré-ARNm et notamment de ceux codant différentes protéines de la famille des caspases. Ces modifications conduisent à la production d'isoformes protéiques pro-apoptotiques favorisant la mort cellulaire (Dutertre et al., 2010 et 2011; Muñoz et al., 2009). De même, une reprogrammation globale de l'épissage alternatif suite à un stress thermique a été mise en évidence chez la souris avec plus de 1000 modulations d'épissage alternatif détectées en plus d'une inhibition massive de l'épissage post-transcriptionnel (Shalgi et al., 2014). L'épissage alternatif permet donc de moduler la quantité (couplage avec le NMD), la localisation et l'activité des protéines et de produire de nouvelles isoformes avec une activité différente ou plus adaptée au contexte cellulaire (Biamonti and Caceres, 2009). Les pré-ARNm des protéines régulatrices de l'épissage étant généralement eux-mêmes soumis à épissage alternatif, il est possible en réponse à différents stress de modifier très rapidement le profil global d'expression et d'épissage des ARNm. Par ailleurs, l'observation d'une relocalisation de plusieurs protéines impliquées dans la régulation de l'épissage alternatif au sein de structures nucléaires ou cytoplasmiques spécifiquement formées en réponse aux stress pourrait participer à la modification des profils globaux d'épissage alternatif (Biamonti and Caceres, 2009; Busà et al., 2010).

A l'heure actuelle aucune étude globale de l'effet d'un stress oxydant sur le profil global d'épissage des ARNm n'a été publiée, cependant des variations du profil d'épissage alternatif de certains ARNm lors d'un stress oxydant ont été reportées et ceci dans différentes lignées cellulaires humaines (Kano et al., 2014; Cote et al., 2012; Maracchioni et al., 2007). Parmi ces variations, la modulation de l'épissage alternatif du pré-ARNm codant la protéine TRA2 β est l'une des plus étudiées. Le gène TRA2 β humain se compose de 10 exons et génère 5 isoformes d'ARNm (TRA2 β 1 à 5) par épissage alternatif. L'isoforme TRA2 β 1 est l'isoforme majoritairement exprimée. Elle contient 9 exons et permet la production d'une

protéine pleine longueur. L'isoforme TRA2 β 4 inclue un exon supplémentaire, l'exon 2, dont l'inclusion introduit des codons stop prématurés. Par conséquent, l'isoforme TRA2 β 4 n'est pas traduite et est dégradée par le système NMD. L'exposition de cellules AGS (lignée gastrique cancéreuse humaine) à un stress oxydant généré par 100 μ M d'arsenite, conduit, dans les premières heures suivant l'initiation du stress à une augmentation de l'inclusion de l'exon 2 et de l'isoforme TRA2 β 4 au dépend de la formation de l'isoforme TRA2 β 1 (Takeo et al., 2009). Les bases moléculaires de cette augmentation de l'inclusion de l'exon 2 ont été étudiées. Il s'avère qu'elle est liée à une augmentation de la phosphorylation des résidus S88 et T118 de la protéine HuR (phosphorylation Chk2- and p38^{MAPK}-dépendante), qui favorise l'interaction directe de HuR avec une séquence de 39 nucléotides en 5' de l'exon 2 stimulant ainsi l'inclusion de l'exon 2 (Akaike et al., 2014). La conséquence première de cette modulation est une diminution d'expression et d'activité de la protéine TRA2 β . Cette protéine étant un facteur de régulation de l'épissage alternatif, cette diminution de son activité a des répercussions sur la régulation de l'épissage alternatif d'autres pré-ARNm, notamment les pré-ARNm codant les protéines CD44, SMN1, SMN2, RIPK2 et Tau (Takeo et al., 2009; Akaike et al., 2014). Cet exemple illustre bien le rôle clé de l'épissage alternatif dans la réponse cellulaire au stress oxydant.

De manière intéressante, des études suggèrent l'établissement d'un épissage alternatif dit "aberrant" suite à un stress oxydant. Ainsi, une équipe a mis en évidence des modulations de l'épissage du pré-ARNm codant l'ADN polymérase β (POLB) conduisant à la formation de nombreuses isoformes qui ne sont pas connues pour être générés dans des conditions physiologique ou en réponse à d'autres événements. Ce type de modulation de l'épissage alternatif peut être également retrouvé dans la réponse au stress génotoxique où une modulation similaire de l'épissage du pré-ARNm codant la protéine MDM2, une ubiquitine ligase impliquée dans la régulation de la protéine p53, a été observée (Dutertre et al., 2010). Dans ce dernier cas, la diminution de l'inclusion de 8 des 12 exons de ce transcrit conduit à la formation de plusieurs isoformes d'ARNm qui ne permettent pas d'aboutir à une forme active de la protéine MDM2 et lève la répression de p53 par MDM2 (Dutertre et al., 2010). Ces résultats posent la question de la nature des variants d'épissages alternatifs, à savoir s'ils participent à la réponse cellulaire ou s'ils sont issus d'un épissage aberrant qui doit être corrigé (Disher and Skandalis, 2007; Skandalis and Uribe, 2004).

IV.4. Objectifs.

Comme il a été mentionné dans l'introduction de cette partie, le stress oxydant représente l'une des plus importantes conditions de stress auxquelles les cellules de vertèbres doivent s'adapter. De nombreuses études ont mis en avant l'impact du stress oxydant sur le métabolisme cellulaire et les mécanismes de réponse au stress. L'épissage alternatif est un processus essentiel dans la réponse à différents stress et de plus en plus de travaux mettent en évidence l'intérêt de déterminer les profils d'épissage des cellules, notamment cancéreuses à des fins thérapeutiques. Cependant, seules quelques altérations de l'épissage alternatif ont été observées et étudiées suite à un stress oxydant. C'est pourquoi nous avons entrepris une étude globale à l'échelle du génome humain de l'effet d'un stress oxydant sur l'épissage alternatif des pré-ARNm.

Cette étude avait plusieurs objectifs :

- i) évaluer le degré de modulation de l'épissage alternatif lors de la réponse des cellules de vertèbres au stress oxydant,
- ii) rechercher si les ARNm dont l'épissage est altéré correspondent à des fonctions cellulaires bien définies,
- iii) étudier les mécanismes cellulaires à l'origine des variations d'épissage observées après un stress oxydant au travers, en particulier, de l'étude de l'impact d'un stress oxydant sur l'expression et la localisation des composants du spliceosome et des facteurs de régulation de l'épissage.

RESULTATS DE LA PARTIE 2

CHAPITRE V : RESULTATS DE L'ETUDE DES CONSEQUENCES D'UN STRESS OXYDANT SUR L'EPISSAGE ALTERNATIF.

V.1. MISE AU POINT DES DIFFERENTES CONDITIONS EXPERIMENTALES POUR LA GENERATION D'UN STRESS OXYDANT.

Afin de réaliser les différents objectifs de la thèse, il était nécessaire de mettre au point des conditions expérimentales pour la génération d'un stress oxydant dans des cellules HeLa. Sur la base de données de la littérature, nous avons décidé dans un premier temps de tester différents agents inducteurs d'un stress oxydant qui sont classiquement utilisés et que nous disposons au laboratoire. Ainsi nous avons testé l'effet de l' H_2O_2 , du paraquat et du DNP (dinitrophénol) à différentes doses sur des cellules HeLa. Comme il a été dit dans l'introduction l' H_2O_2 est un ROS non radicalaire naturellement présent dans les cellules humaine et dont la toxicité passe principalement par sa réaction avec le fer ferreux (Fe_2^+) selon la réaction de fenton et de Haber-Weiss, qui entraîne la formation de radicaux hydroxyle (HO^{\cdot}) qui sont les ROS les plus réactifs. Le paraquat est connu pour être un puissant poison pour l'homme qui a la propriété d'inhiber la chaîne respiratoire et favoriser une faible production de ROS. Enfin le DNP découple la chaîne respiratoire et entraîne une augmentation de ROS intracellulaire avec une diminution de l'activité mitochondriale. Afin de déterminer les meilleures conditions expérimentales pour la réalisation de notre étude, nous avons analysé l'impact de ces différents agents sur le taux de ROS intracellulaire et la survie cellulaire en utilisant respectivement les kits OxySelect ROS Assay kit (Euromedex) basé sur l'utilisation de la molécule H₂DCFDA (dichlorodihydrofluorescein diacetate) et le kit Vybrant MTT cell proliferation Assay Kit (Invitrogen). L'H₂DCFDA est une molécule non fluorescente qui peut traverser la membrane des cellules, qui une fois oxydée par des ROS devient fluorescente. Il est alors possible de mesurer le niveau de ROS intracellulaire par fluorométrie. Le dosage de la viabilité cellulaire repose sur un test colorimétrique par ajout d'un sel de tétrazolium (MTT) (jaune) qui est réduit par la succinate déshydrogénase mitochondriale qui est active seulement dans les cellules viables en formazan (violet). L'intensité de la coloration violette est proportionnelle à l'activité métabolique des cellules et à leur nombre.

V.1.1. Importance du lot de cellules HeLa utilisé et du nombre de passage des cellules pour la réalisation de l'étude.

Afin de trouver les conditions expérimentales les plus adéquates pour notre étude et permettant d'obtenir des résultats reproductibles j'ai testé l'effet du peroxyde d'hydrogène (100 μ M, 200 μ M et 500 μ M), du paraquat (0,4 mM, 0,8 mM et 1mM) et du DNP (50 μ M, 100 μ M et 200 μ M) à des doses différentes sur le taux de ROS intracellulaires et sur la viabilité des cellules HeLa traitées à 90 % de confluence. Les doses utilisées ont été choisies au regard de la littérature et différents paramètres ont été pris en compte pour s'assurer de la reproductibilité future des expériences tels que le lot de cellules HeLa utilisé ou le nombre de passages des cellules après décongélation. Ainsi j'ai induit les différents stress oxydants sur 2 lots de cellules HeLa différents (lot 1 et 2) ayant subi 3 ou 5 passages avant le début des stress (P+3 et P+5). Chaque lot de cellules est issu de la même lignée de cellules HeLa mais se distingue par leur historique d'utilisation et de protocoles de conditions de culture par différents expérimentateurs avant le début de notre étude. Pour chacune des conditions nous avons quantifié les ROS et la viabilité cellulaire 24h après le début des stress. En parallèle nous avons réalisé une condition sans ajout d'agent oxydant (T) comme contrôle négatif de l'effet du peroxyde d'hydrogène et du paraquat et une condition en ajoutant de méthanol comme contrôle négatif du DNP (M) puisque ce dernier, pour des raisons de solubilité, est dilué dans du méthanol. Les résultats obtenus sont illustrés dans la figure V.1. De manière générale, on constate que quelque soit le traitement réalisé, chaque agent oxydant entraîne une augmentation du taux de ROS intracellulaires et une diminution de la viabilité cellulaire et cela indépendamment du lot de cellules ou du nombre de passages réalisés (Figure V.1 A et B). Néanmoins on constate une différence de sensibilité aux agents oxydants dans les cellules HeLa du lot 1 en fonction du nombre de passages, où pour chaque traitement testé on observe pour les doses les plus faibles une augmentation moins importante du taux de ROS intracellulaire dans les cellules ayant subi plus de passages. A l'inverse, ces cellules montrent un plus grand % de mortalité aux doses les plus importantes.

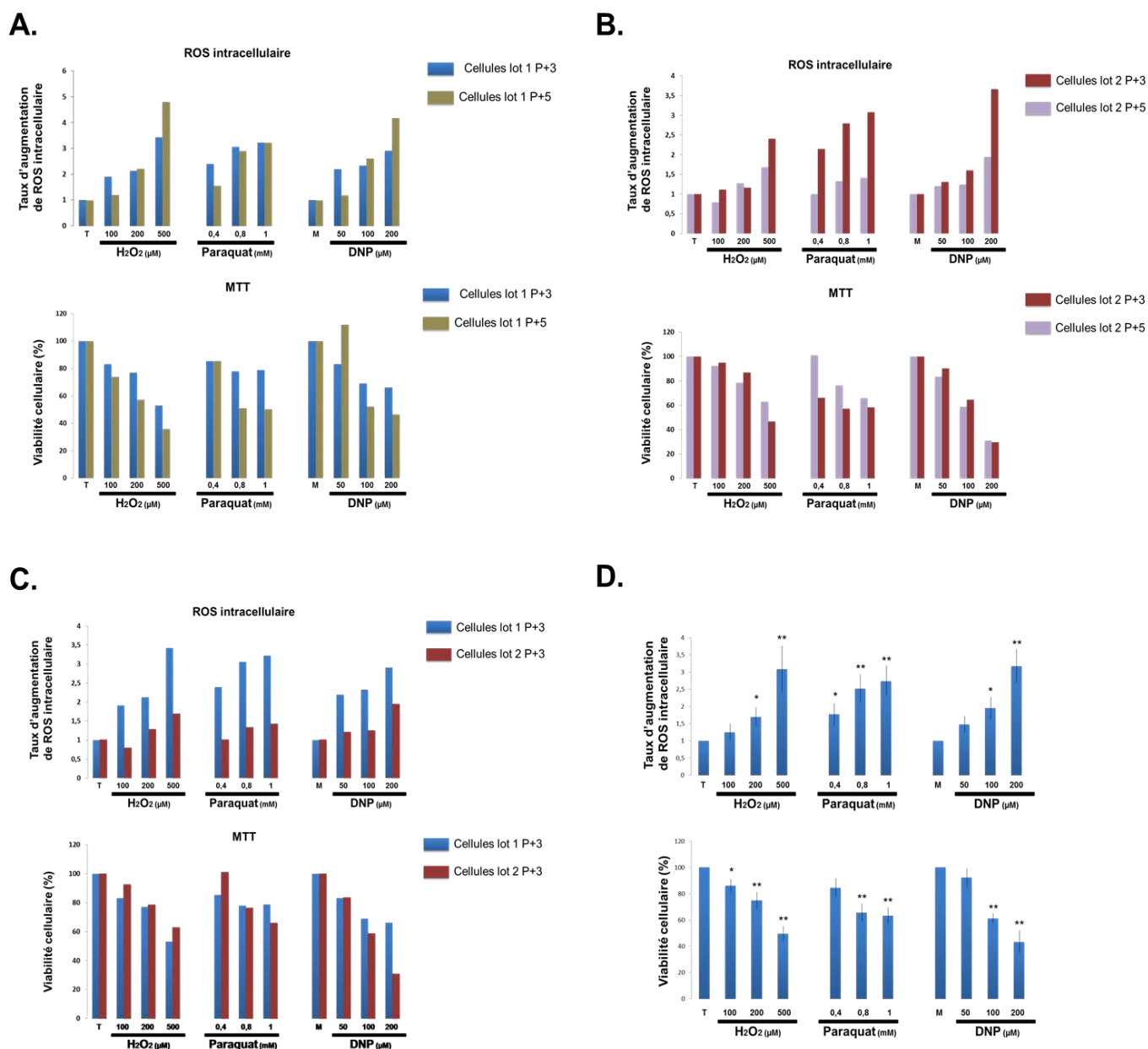


Figure V.1 : Analyse de la génération de ROS intracellulaires suite à un traitement de 24h par le peroxyde d'hydrogène (H₂O₂, 100 μM, 200 μM et 500 μM), le paraquat (0,4 mM, 0,8 mM et 1mM) et le DNP (50 μM, 100 μM et 200 μM) en fonction du nombre de passages et de l'origine des cellules HeLa. Une condition témoin eau (T) a été réalisée ainsi qu'un témoin méthanol (M) qui est utilisé pour dissoudre le DNP. La quantité de ROS intracellulaires générée a été rapportée au nombre de cellules vivantes. **A. Résultats obtenus pour les cellules du lot 1 traitées par les différents agents après 3 ou 5 passages (respectivement P+3 et P+5). **B.** Résultats obtenus pour les cellules du lot 2 traitées par les différents agents après 3 ou 5 passages (respectivement P+3 et P+5). **C.** Résultats obtenus pour les cellules du lot 1 et 2 et traités par les différents agents après 3 passages (P+3). **D.** Résultats obtenus pour les cellules du lot 1 traitées par les différents agents après 3 passages. Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé (n=3 ; ** : P<0.01 ; * : P<0.05).**

Cette constatation se retrouve également dans les résultats obtenus avec le lot 2 de cellules HeLa, en particulier suite au traitement au paraquat. En effet, on observe un pourcentage de viabilité équivalent entre les cellules ayant subi 3 ou 5 passages, alors que le paraquat a un effet plus de 2x plus important sur la production de ROS intracellulaires sur les cellules avec moins de passage par rapport aux cellules avec le plus de passage. Ces résultats mettent en évidence une sensibilité différente aux différents agents oxydants en fonction du nombre de passages réalisés avant traitement. Lorsqu'on compare les résultats obtenus entre les cellules des lots 1 et 2 avec le même nombre de passage avant traitement, on constate également une forte variabilité entre les cellules (Figure V.1 C). Bien qu'on observe un même pourcentage de viabilité cellulaire, on remarque que les différents agents oxydants entraînent une plus forte augmentation du taux de ROS intracellulaires dans les cellules du lot 1 par rapport à celles du lot 2. A partir de ces résultats, des conditions optimisées de génération d'un stress oxydant sur des cellules HeLa ont été définies en utilisant le lot 1 de cellules HeLa. L'analyse des ROS et de la survie cellulaire dans ces nouvelles conditions montrait un effet statistiquement reproductible des différents agents oxydants (Figure V.1D).

V.1.2. Vérification des conditions expérimentales sur la modulation de l'épissage alternatif liée au stress oxydant.

Afin de déterminer si le stress oxydant généré sur les cellules HeLa était efficace en termes de modulation de l'épissage alternatif, j'ai testé par RT-PCR si je pouvais détecter des variants d'épissage d'ARNm dont l'épissage était connu être modulé par le stress oxydant dans des cellules humaines. Un phénomène marqué d'épissage illégitime du pré-ARNm POLB suite à un stress oxydant, généré par un traitement avec 500 μM d' H_2O_2 , avait été décrit (Skandalis et al., 2004). J'ai donc entrepris d'analyser l'épissage du pré-ARNm POLB 24 h après le début du traitement des cellules HeLa par différentes doses de nos trois agents oxydants. Les oligonucléotides utilisés pour l'amplification et les conditions de PCR étaient celles publiées (Disher et Skandalis, 2007). Les données obtenues montrent que le stress

oxydatif généré par les différents agents oxydants entraîne une variation de l'épissage alternatif du pré-ARNm POLB (Figure V.2). En effet, pour chacune des conditions de stress oxydant, l'apparition de variants d'épissage aberrants est observée aux plus fortes concentrations des agents oxydants confirmant un phénomène d'épissage illégitime du pré-ARNm POL B suite au stress oxydant. Cependant, on remarque que le profil d'épissage observé pour chacun des stress est différent, indiquant une modulation de l'épissage du pré-ARNm POLB spécifique à chacun des stress, voire même à chacune des concentrations d'agent oxydant puisque pour un même agent oxydant on peut observer des profils différents selon la concentration d'agent utilisée (Figure V.2) On note que c'est l'ajout de 500 μM d' H_2O_2 qui entraîne l'effet le plus marqué, mais également que le seul ajout de méthanol module l'épissage alternatif du pré-ARNm POLB, rendant l'étude d'un stress oxydant généré par le DNP délicate et que le paraquat, pesticide hautement toxique pour l'homme, n'affecte que modérément cet événement d'épissage. C'est pourquoi nous avons décidé de ne garder que la condition d'un stress oxydant par ajout de 500 μM d' H_2O_2 sur des cellules HeLa du lot 1 après trois passages pour réaliser notre analyse globale des profils d'épissage suite à un stress oxydant.

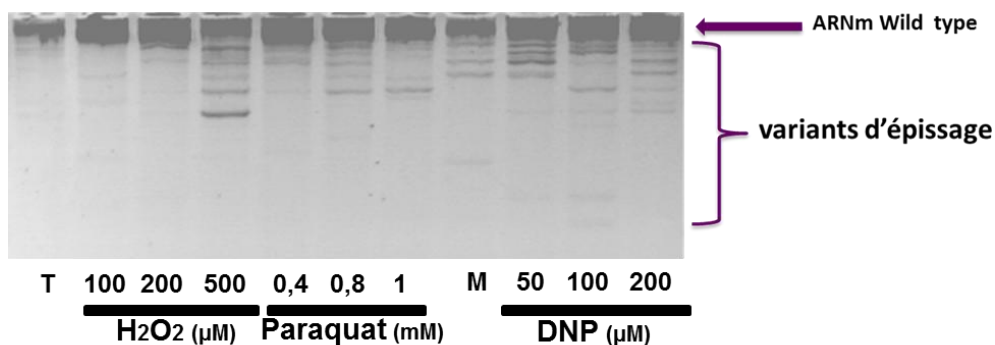


Figure V.2: Analyse par RT-PCR de l'effet sur l'épissage alternatif du pré-ARNm POLB du peroxyde d'hydrogène (H_2O_2 100 μM , 200 μM et 500 μM), du paraquat (0,4 mM, 0,8 mM et 1mM) et du DNP (50 μM , 100 μM et 200 μM) dans des cellules Hela. Une condition témoin eau (T) a été réalisée ainsi qu'un témoin méthanol (M) qui est utilisé pour dissoudre le DNP.

V.2. CARACTERISATION DE LA REPONSE DES CELLULES HELA A UN STRESS OXYDANT GENERE PAR H₂O₂.

V.2.1. Impact du stress oxydant sur la viabilité cellulaire et le taux de ROS intracellulaires.

Ayant fait le choix d'utiliser H₂O₂ comme agent oxydant pour réaliser notre étude. Il nous fallait caractériser la réponse de nos cellules HeLa à cet agent afin de déterminer les conditions expérimentales les plus pertinentes pour notre analyse globale de l'impact d'un stress oxydant sur l'épissage alternatif. Les analyses préliminaires que je viens de vous présenter avaient toutes été réalisées 24 heures après l'induction du stress oxydant, cependant, comme nous l'avons vu dans l'introduction, le facteur temps est un paramètre important dans la réponse des cellules à un stress. C'est pourquoi j'ai entrepris d'analyser l'impact d'un stress oxydant généré par l'ajout de 500 µM d'H₂O₂ sur le taux de ROS intracellulaires et la viabilité cellulaire à différents temps après l'induction du stress et non plus uniquement 24 h après le début du stress. Pour cette expérience, nous avons choisi de réaliser nos analyses 20 minutes, 1, 2, 4, 8, 16 ou 24 heures après l'induction du stress. Les résultats obtenus sont illustrés dans la figure V.3. Le suivi de la viabilité cellulaire au cours de la réponse au stress oxydant montre une très forte diminution de la viabilité au cours des 20 premières minutes après l'ajout de 500 µM d'H₂O₂, avec une perte de près de 40 % des cellules viables au cours de ces 20 minutes, suivie d'une diminution plus modérée de la viabilité jusque 16 heures après le début du stress, avec à 16 h seulement 18 % de cellules viables (Figure V.3 A). Un doublement de la viabilité cellulaire est observé entre les temps 16 et 24 h, avec à 24 h 40 % de cellules viables (Figure V.3 A). On notera que le taux de cellules viables 24 h après le début du traitement est tout à fait dans la norme de ce qui est généralement observé après un traitement similaire d'autres lignées cellulaires humaines (Disher et Skandalis, 2007 ; Cote et al., 2012). Ces données révèlent donc un effet marqué du stress oxydant généré par 500 µM d'H₂O₂ sur la viabilité des cellules lors des 16 premières heures suivant l'ajout de l'agent oxydant dans le milieu de culture, suivi d'une phase de récupération des cellules à partir de 16 h. De façon surprenante nous notons un maintien de la viabilité cellulaire entre les temps 1 h et 2 h (Figure V.3 A).

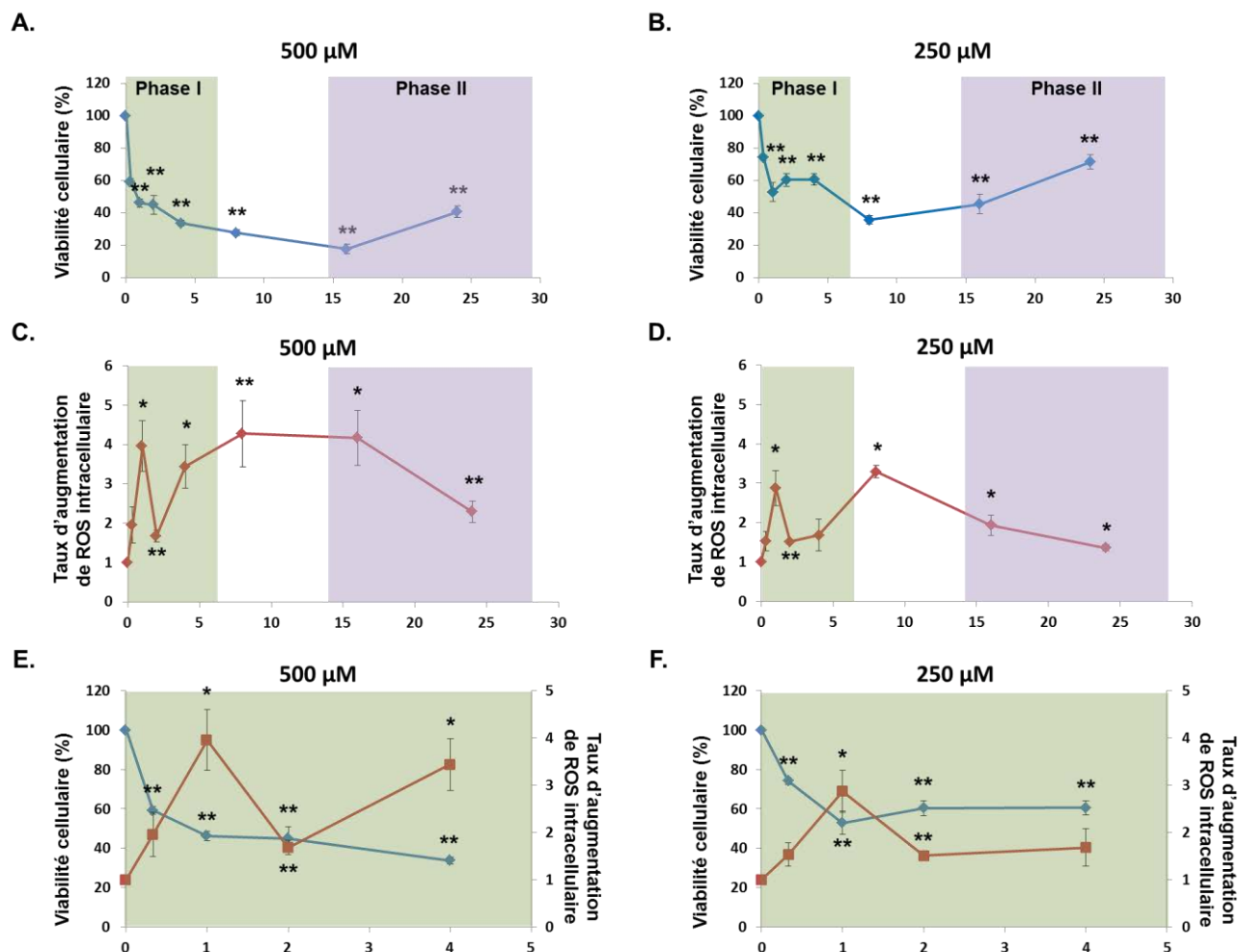


Figure V.3 : Analyse cinétique de la cytotoxicité et de l'induction de ROS par l'H₂O₂ dans des cellules HeLa. Le dosage de la viabilité cellulaire et des ROS intracellulaire a été réalisé par l'utilisation respective du MTT et du H₂DCFDA. La viabilité cellulaire est exprimée en % de cellules vivantes par rapport aux cellules contrôles non traitées. Le niveau de ROS dans les cellules traitées est utilisé comme référence pour déterminer le taux d'induction des ROS dans les cellules exposées à l'H₂O₂. **A.** Cinétique de la viabilité cellulaire suite à un stress oxydant réalisé avec une dose de 500 μM H₂O₂. **B.** Cinétique de la viabilité cellulaire suite à un stress oxydant réalisé avec une dose de 250 μM H₂O₂. **C.** Cinétique d'induction de ROS intracellulaire suite à un stress oxydant réalisé avec une dose de 500 μM H₂O₂. **D.** Cinétique d'induction de ROS intracellulaire suite à un stress oxydant réalisé avec une dose de 250 μM H₂O₂. **E.** Représentation à une échelle différente de la cinétique d'induction de ROS intracellulaire et de la viabilité cellulaire suite à un stress oxydant réalisé avec une dose de 500 μM H₂O₂ aux temps 20 min, 1 h, 2 h et 4 h. **F.** Représentation à une échelle différente cinétique d'induction de ROS intracellulaire et de la viabilité cellulaire suite à un stress oxydant réalisé avec une dose de 250 μM H₂O₂ aux temps 20 min, 1 h, 2 h et 4 h. Chaque barre d'erreur représente la valeur +/- de l'erreur standard (n=3 ; * :P<0.05 ; ** :P<0.01 par rapport aux cellules non stressées).

Ce résultat a été observé dans chacun des trois réplicats biologiques réalisés et suggérait une récupération partielle, mais non suffisante des cellules, puisque suivie ensuite à nouveau d'une diminution de la viabilité cellulaire.

Afin de confirmer cette hypothèse, nous avons réalisé la même cinétique sur des cellules HeLa traitées avec seulement 250 μM d' H_2O_2 . L'idée étant que si cette phase de stagnation représentait effectivement une phase de récupération partielle des cellules, elle serait probablement accentuée lors d'un traitement moins agressif des cellules. Et c'est effectivement ce que nous observons, puisque lors d'un traitement des cellules HeLa avec 250 μM d' H_2O_2 la viabilité cellulaire diminue au cours de la 1^{ère} heure après ajout de l'agent oxydant, puis se maintient, voire augmente légèrement entre 1 et 4h, pour ensuite chuter de nouveau entre 4 et 8 h, puis reprendre définitivement jusqu'à 24 h où l'on retrouve une viabilité cellulaire d'environ 70 % (Figure V.3 B). Ces données nous ont conduits à proposer que la réponse des cellules HeLa à un stress oxydant généré par H_2O_2 se ferait en 2 phases : une phase de réponse précoce se produisant entre 1 et 4 h après l'induction du stress (durée de la phase variable selon l'intensité du stress) et une phase de réponse tardive se produisant entre 16 et 24 h après l'induction du stress.

En parallèle de l'effet sur la viabilité cellulaire, j'ai analysé l'effet du stress oxydant généré par l'addition de 500 ou 250 μM d' H_2O_2 sur le taux de ROS intracellulaires. Les résultats obtenus sont présentés dans les panneaux C et D de la figure V.3. On constate que la diminution de la viabilité cellulaire est généralement corrélée à une augmentation du taux de ROS intracellulaires et que les phases de récupération des cellules sont associées à une diminution du taux de ROS. Ces données sont donc en parfait accord avec notre hypothèse d'une réponse en 2 phases et l'observation d'une diminution marquée du taux de ROS 2 h après l'induction d'un stress oxydant avec 500 μM d' H_2O_2 renforce cette hypothèse.

V.2.2. Impact du stress oxydant sur l'expression des ARN sat III.

Comme mentionné dans l'introduction, l'équipe de G. Biamonti avait observé une surexpression faible (induction < 100 fois) des ARN Sat III 16 h après la fin d'un stress oxydant généré par l'ajout de 200 μM d' H_2O_2 durant 20 minutes dans le milieu de culture de cellules HeLa (Valgadorsttir et al., 2008). L'expression des ARN Sat III et la formation des granules de stress nucléaires étant proposées contribuer aux modulations d'épissage alternatif au cours de la réponse à différents stress cellulaires, j'ai donc analysé par RT-qPCR l'expression des ARN Sat III au cours de la réponse au stress oxydant. Les oligonucléotides et le protocole utilisés sont ceux initialement publiés par l'équipe de G. Biamonti et implantés au laboratoire par V. Vautrot lors de sa thèse. Les résultats obtenus sont illustrés dans la figure V.4.

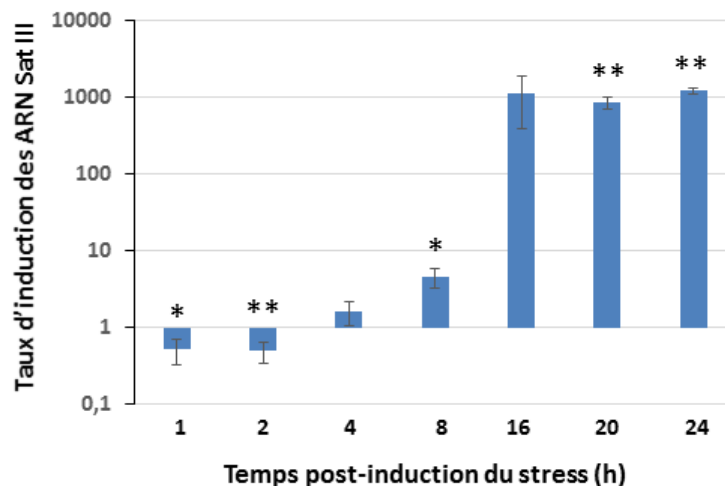


Figure V.4 : Détermination par RT-qPCR du taux d'induction des ARN Sat III au cours de la réponse à un stress oxydant généré par l'ajout de 500 μM d' H_2O_2 sur des cellules HeLa. Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé (n \geq 3 ; ** : P<0.01 ; * : P<0.05).

De façon intéressante, alors qu'aucune induction de l'expression des ARN Sat III n'est détectée dans les premières heures suivant l'induction du stress, une forte induction (facteur \geq 800) est observée à partir de 16 h après l'induction du stress et est maintenue jusqu'à au moins 24 h post-induction du stress. En accord avec les données de l'équipe de G. Biamonti, nous observons donc une induction tardive de l'expression des ARN Sat III, mais celle-ci est beaucoup plus forte que celle décrite par l'équipe de G. Biamonti. Cette différence est probablement liée au fait que nous utilisons une concentration d' H_2O_2 2,5 fois

supérieure à celle utilisée par G. Biamonti et que nous ne changeons pas le milieu de culture 20 minutes après le début du stress. Nous observons donc dans nos conditions expérimentales une induction de l'expression des ARN Sat III quasiment équivalente à celle observée suite à un choc thermique et concomitante avec la phase tardive d'augmentation de la viabilité cellulaire.

Nous avons ensuite testé si cette induction de l'expression des ARN Sat III était associée à la formation de foyers nucléaires enrichis en ARN Sat III. Pour cela I. Behm-Ansmant et G. Heckler (stagiaire Master 1) ont réalisé des expériences de RNA-FISH, permettant de visualiser les ARN Sat III, à différents temps après l'induction d'un stress oxydant. Le protocole utilisé a été mis au point au sein du laboratoire par V. Vautrot et C. Aigueperse et est décrit dans le Matériel et Méthodes (Vautrot et al., 2015). Les résultats obtenus sont présentés dans la figure V.5.

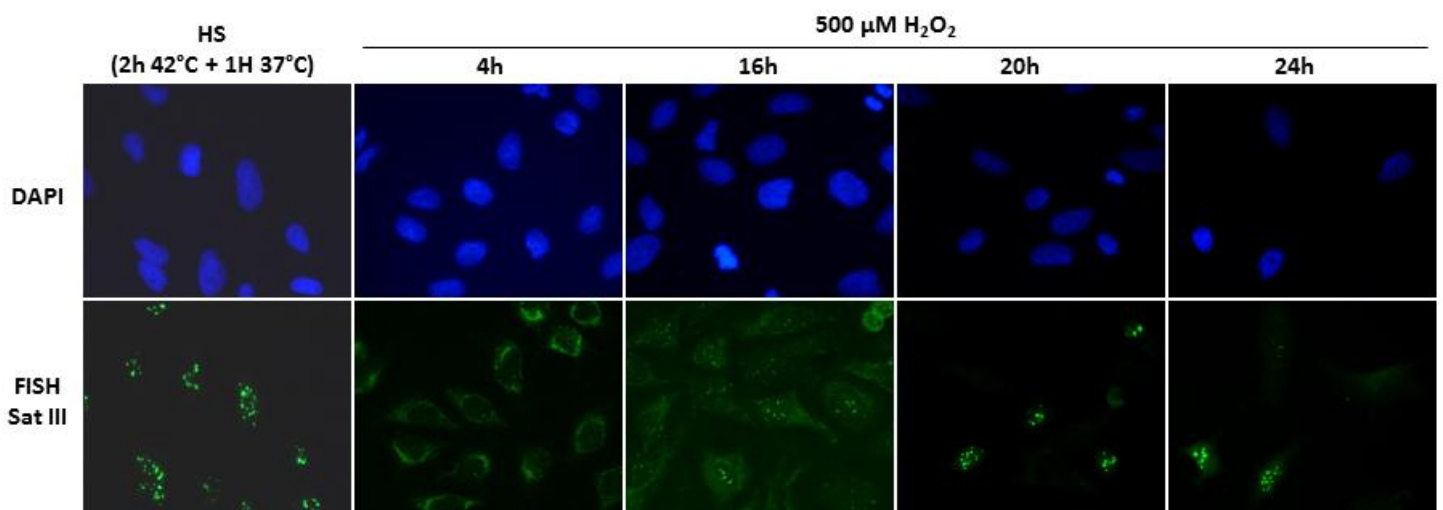


Figure V.5 : Suivi par RNA-FISH de la localisation des ARN Sat III au cours de la réponse de cellules HeLa S3 à un stress oxydant généré par l'ajout de 500 μM d' H_2O_2 . La position des noyaux est indiquée par DAPI. La localisation des ARN Sat III lors d'un choc thermique généré par 2 h d'incubation à 42°C suivies d'1 h de récupération à 37°C est illustrée pour comparaison.

On observe que 4 h après l'induction du stress oxydant, le signal de FISH des ARN Sat III est faible et présente une localisation à prédominance cytoplasmique (Fig V.5, 4 h). Ce profil de localisation est équivalent à celui obtenu avec une sonde contrôle ne reconnaissant aucune séquence humaine et utilisée comme contrôle négatif (scramble), laissant suggérer que le signal détecté serait aspécifique. Seize heures après l'induction du stress oxydant, un voile cytoplasmique est toujours observé, cependant des foyers nucléaires commencent à être détectés dans près de 75 % des cellules. Ces foyers sont en moyenne au nombre de 5 par cellule (Fig V.5). A 20 et 24 h, les ARN Sat III sont uniquement détectés au sein de foci nucléaires très similaires, en termes d'aspect, aux granules de stress nucléaires formés suite à la réalisation d'un choc thermique. Cependant, alors que ces granules sont observés dans plus de 90 % des cellules lors d'un choc thermique, seules 50 à 60 % des cellules soumises à un stress oxydant présentent ces foyers nucléaires 20 à 24 h après l'induction du stress.

Afin de vérifier que ces foyers nucléaires correspondaient bien à des granules de stress nucléaires, des expériences de RNA FISH couplées à des expériences d'immunofluorescence ont été réalisées pour analyser simultanément la localisation des ARN Sat III et de la protéine hnRNP HAP (SafB), un composant cœur de ces granules et ceci 24 h après l'induction du stress oxydant (Figure V.6). On observe clairement une colocalisation des ARN Sat III et de la protéine hnRNP HAP dans les cellules soumises au stress oxydant, indiquant que les foyers d'ARN Sat III, observés en phase tardive de réponse au stress oxydant, sont très probablement des granules de stress nucléaires.

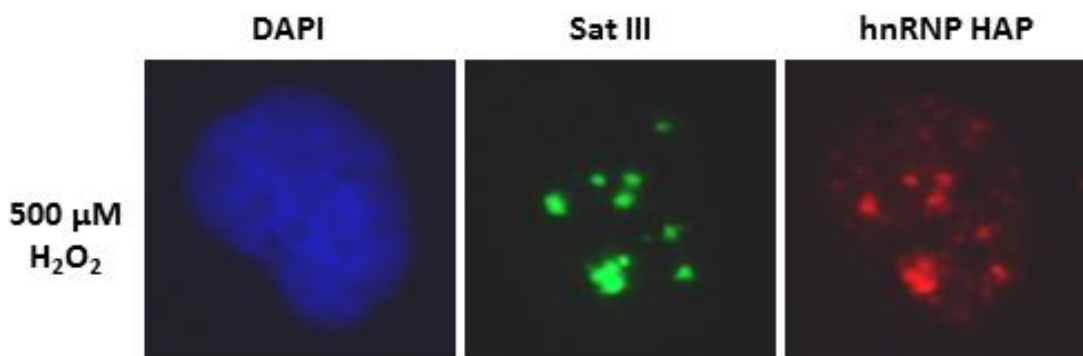


Figure V.6 : Visualisation de la co-localisation entre les ARN sat III et la protéine hnRNP HAP dans des cellules HeLa S3 soumises depuis 24 h à un stress oxydant par ajout de 500 µM d'H₂O₂. Le noyau est indiqué par DAPI.

Nos données mettaient donc en évidence qu'un stress oxydant pouvait induire, tout comme le choc thermique, une surexpression forte des ARN Sat III et la formation de foyers nucléaires contenant à la fois les ARN Sat III et la protéine hnRNP HAP. Leur taille, aspect, nombre et la présence de la protéine hnRNP HAP au sein de ces foyers nucléaires suggéraient qu'ils correspondaient à des granules de stress nucléaires. Se posait alors la question de la nature du facteur de transcription à l'origine de l'induction de la surexpression des ARN Sat III.

V.2.3. Identification du facteur de transcription responsable de l'expression des ARN sat III.

Comme mentionné dans l'introduction, deux facteurs de transcription avaient été décrits pouvoir activer la transcription des ARN Sat III : le facteur de transcription spécifique du choc thermique HSF1 et le facteur de transcription spécifique du stress osmotique TonEBP (Ferraris et al., 2002; Valgardsdottir et al., 2008; Sengupta et al., 2009). Pour rappel, lors d'un choc thermique modéré l'induction de l'expression des ARN Sat III par HSF1 est très rapide et très forte (induction >1000 fois dès 30 minutes après le début du choc thermique, cellules HeLa) et conduit à la formation de nombreux granules dans près de 90 % des cellules. Au contraire, un stress hyperosmotique (8 h en présence de 0.8M sorbitol, cellules HeLa) induit une expression plus tardive et plus modérée des ARN Sat III, induction de 10 fois 2 h après la fin du stress et de 100 à 1000 fois 16 h après la fin du stress, associée à la formation de granules moins nombreux et plus petits dans seulement 30 % des cellules. Les observations de l'équipe de G. Biamonti reportaient une induction faible et tardive (<100 fois, 16 heures après la fin du stress) de l'expression des ARN Sat III et la formation de nSBs de très petite taille et en faible nombre (1 à 2 micro-nSBs par cellule) dans seulement 5 % des cellules suite à un stress oxydant (20 min d'exposition à 200 μ M H₂O₂). Si l'on fait l'hypothèse d'une corrélation directe entre la nature du facteur de transcription activant l'expression des ARN Sat III et la cinétique et le niveau d'expression des ARN Sat III, ainsi que le nombre et l'aspect des granules de stress nucléaires formés, ces données initiales de l'équipe de G. Biamonti étaient plutôt en faveur d'un rôle de TonEBP dans l'activation de l'expression des ARN Sat III lors d'un stress oxydant. A l'opposé, nos données, en termes de nombre et taille de granules, sont plus en faveur de l'implication de HSF1 dans ce processus,

même si l'induction de l'expression des ARN Sat III demeure tardive. Le facteur de transcription responsable de l'expression des ARN Sat III étant généralement présent dans les granules de stress nucléaires, nous avons entrepris d'étudier la localisation de HSF1 et TonEBP lors de la réponse à un stress oxydant généré par l'ajout de 500 μM d' H_2O_2 .

Pour réaliser cette étude, nous avons entrepris de tester par immunofluorescence (IF) une colocalisation éventuelle de HSF1 et/ou TonEBP avec la protéine hnRNP HAP, utilisée comme marqueur protéique des granules de stress nucléaires. Nous disposions déjà au laboratoire d'anticorps, fonctionnant en IF, dirigés contre HSF1, cependant nous n'en avons pas pour TonEBP. Nous avons commandé deux anticorps différents dont un était censé fonctionner en IF, mais nous n'avons jamais réussi à obtenir de signal d'IF satisfaisant, ni même réussi à détecter TonEBP par Western blot dans des cellules HeLa soumises à un stress oxydant. En l'absence d'un contrôle « stress osmotique », que nous mettons actuellement au point, nous ne sommes donc pas en mesure de conclure quant à l'origine de ce problème : niveau d'expression trop faible de TonEBP et absence d'effet du stress oxydant sur l'expression et la localisation de TonEBP ou anticorps non fonctionnels. Concernant HSF1, les données de localisation sont illustrées dans la figure V.7. Nous observons, 1 h après l'induction du stress oxydant par 500 μM d' H_2O_2 , un signal HSF1 faible et diffus dans le noyau, alors que 2 h post-induction le signal HSF1 augmente dans environ 15 % des cellules et 2 à 5 foci nucléaires de petite taille apparaissent. Quatre heures post-induction HSF1 est présent sous formes de foci nucléaires (3 à 6 par cellule) dans environ 60 % des cellules, mais la protéine hnRNP HAP n'est pas enrichie au niveau de ces foci, ni d'ailleurs au niveau des foci observés 2 h post-induction. Les ARN Sat III n'étant pas non plus surexprimés 2 h et 4 h après le début du stress oxydant (Figure V.4). Ces foyers HSF1 tendent à disparaître à 8 h post-induction et d'autres foci nucléaires, contenant cette fois-ci la protéine hnRNP HAP, sont détectés à partir de 16 h post-induction dans 40 %. Ces foci sont maintenus jusque 24 h post-induction dans environ 60 % des cellules, même si à 24 h on observe de plus en plus fréquemment des foci de protéine hnRNP HAP ne contenant plus HSF1. Il apparaît donc que le facteur de transcription HSF1 est présent au sein des granules de stress nucléaires observés entre 16 h et 24 h après l'induction du stress oxydant, suggérant un rôle d'HSF1 dans l'induction de l'expression des ARN Sat III en phase tardive de réponse au stress oxydant.

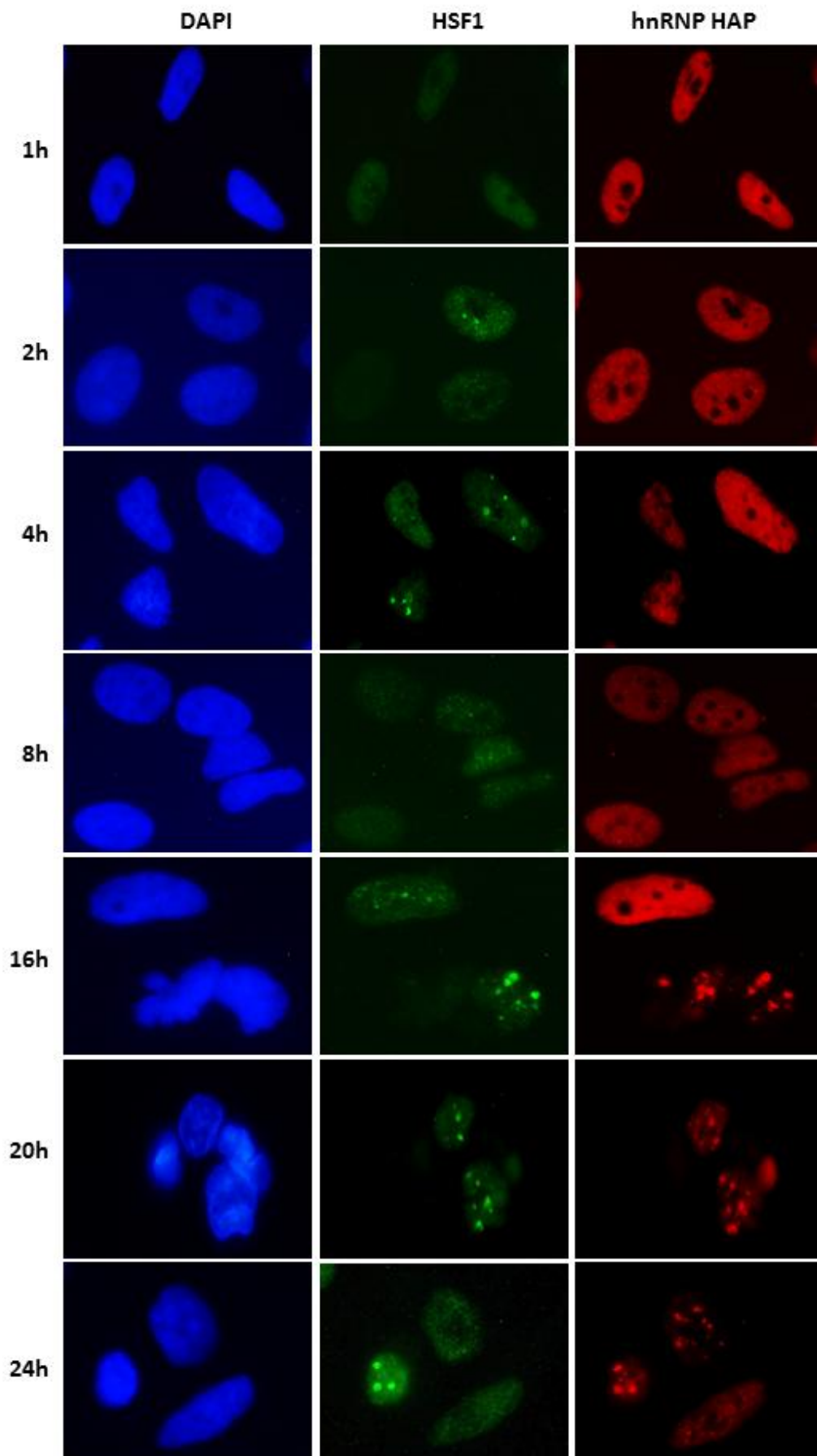


Figure V.7 : Suivi par immunofluorescence de la localisation des protéines HSF1 et hnRNP HAP au cours de la réponse de cellules HeLa S3 à un stress oxydant généré par l'ajout de 500 μM d' H_2O_2 .

Afin de confirmer le rôle d'HSF1 dans cette induction, j'ai déplété HSF1 dans des cellules HeLa S3 par emploi d'un cocktail de siRNA (GE Healthcare, Dharmacon) et testé l'induction de l'expression des ARN Sat III dans ces cellules 24 h après l'induction d'un stress oxydant (500 μ M d' H_2O_2). Les conditions de déplétion de HSF1 avaient été mises au point au laboratoire par C. Aigueperse dans le cadre d'un autre projet et sont décrites dans le Matériel et Méthodes. L'efficacité de la déplétion d'HSF1 a été vérifiée par Western blot et les ARN Sat III ont été dosés par RT-qPCR. Comme illustré dans la figure V.8, nos conditions de déplétion d'HSF1 sont efficaces et permettent de réduire le niveau d'expression de HSF1 à moins de 5 % de son niveau initial. Dans ces conditions, l'induction de l'expression des ARN Sat III 24 h après le début du stress oxydant passe de 756 fois (siNT) à 33 fois (siHSF1), soit une diminution de près de 95 % de l'induction des ARN Sat III (Figure V.8). Ces données établissent donc clairement le rôle central joué par le facteur de transcription HSF1 dans l'expression des ARN Sat III en réponse à un stress oxydant.

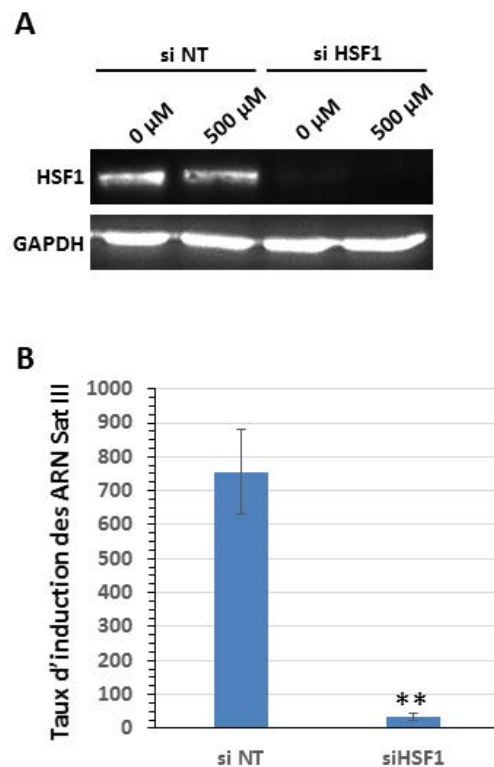


Figure V.8 : Effet de la déplétion du facteur de transcription HSF1 sur l'induction de l'expression des ARN Sat III en réponse à un stress oxydant généré par l'ajout de 500 μ M d' H_2O_2 . **A.** Vérification de l'efficacité de la déplétion de la protéine HSF1 par Western blot. La protéine GAPDH est utilisée comme contrôle de charge. **B.** Impact de la déplétion de la protéine HSF1 sur l'expression des ARN Sat III 24 h après l'induction d'un stress oxydant par ajout de 500 μ M d' H_2O_2 dans le milieu de culture. La quantité d'ARN Sat III au sein des cellules a été estimée par RT-qPCR. Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé ($n \geq 3$; ** : $P < 0.01$).

Plusieurs études avaient déjà établi un lien entre activation du facteur de transcription HSF1 et réponse au stress oxydant. Ainsi, comme mentionné dans l'introduction, il avait été montré qu'un stress oxydant pouvait favoriser l'activation d'HSF1 et sa relocalisation dans le noyau par multimérisation du fait d'une restructuration d'HSF1 consécutive à la formation d'un pont disulfure impliquant les résidus Cys C35 et C105 (Ahn and Thiele, 2003). Par ailleurs, plus récemment, il a été montré que la déplétion d'HSF1 affecte la réponse des cellules humaines au stress oxydant et conduit à une augmentation du taux de ROS intracellulaires lors de ce stress, indiquant un rôle crucial d'HSF1 dans la réponse des cellules humaines au stress oxydant (Yoo et al., 2014). Notre observation de l'induction de l'expression des ARN Sat III de façon HSF1-dépendante en phase tardive de réponse au stress oxydant et de façon concomitante à la reprise de la viabilité cellulaire et à une diminution du taux de ROS intracellulaires pourrait constituer une des explications quant à l'importance de HSF1 dans la réponse au stress oxydant.

Les données relatives à la caractérisation de la réponse des cellules HeLa au stress oxydant généré par H₂O₂ présentées dans ce chapitre, font l'objet de la préparation d'un article, dont je suis 1^{er} auteur et qui sera soumis d'ici ma soutenance de thèse.

V.3. ANALYSE GLOBALE DES EFFETS D'UN STRESS OXYDANT SUR L'EXPRESSION ET L'ÉPISSAGE ALTERNATIF DES PRE-ARNM DE CELLULES HELA PAR UTILISATION DE PUCES A ADN DETECTANT LES EXONS.

Ayant mis en évidence que la réponse des cellules HeLa S3 à un stress oxydant, généré par l'ajout de 500 μM d' H_2O_2 dans le milieu de culture, impliquait une phase précoce (1 à 4 h) et une phase tardive (24 h) de réponse, nous avons alors entrepris une analyse globale des effets de ce stress sur l'expression des gènes et l'épissage alternatif des pré-ARNm à 4 temps différents : 1 h, 2 h, 4 h et 24 h. Cette analyse à l'échelle du génome a été effectuée par emploi de puces à ADN spécifiques des exons (Affymetrix Human Exon ST 1.0 Arrays) dont la description suit.

V.3.1. Description des puces à ADN utilisées lors de cette étude.

Les puces Affymetrix Human Exon ST 1.0 Arrays sont des puces à ADN qui permettent de détecter les exons des ARNm humains. Ces puces contiennent des sondes spécifiques de tous les exons identifiées dans les ARNm humains dans la banque de données Ensembl Transcripts, soit 306 583 exons. Le principe général des puces à exons est schématisé dans la figure V.9. Chaque exon, dont la taille est supérieure à 100 pb, est ciblé au minimum par un ensemble (« probeset ») de 4 sondes. Les exons de taille inférieure sont également détectés, mais avec un nombre de sondes inférieur qui dépend à la fois de la longueur et de la séquence de l'exon. Selon le procédé Affymetrix, plusieurs sondes identiques sont directement synthétisées sur un point donné de la puce. Les ADNc d'un échantillon marqués par un fluorophore sont ensuite hybridés sur la puce et chaque point renvoie un signal fluorescent proportionnel à la quantité d'ADNc cibles hybridés. Par ce procédé, il est ainsi possible de déterminer la fréquence d'utilisation de chaque exon, ainsi que le niveau d'expression de chaque ARNm. La manipulation des puces, ainsi que l'analyse des données ont été réalisées en collaboration avec l'équipe de L.Vallar (CRP Luxembourg).

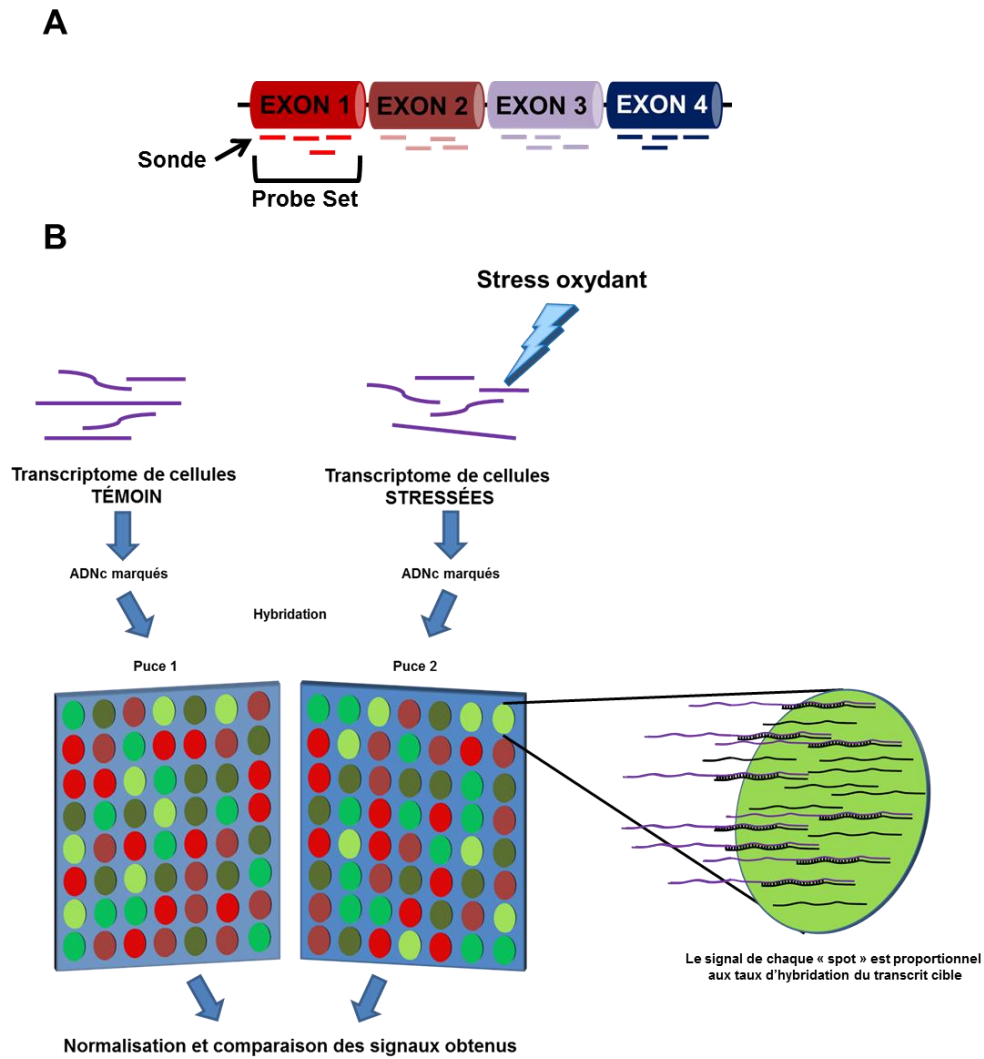


Figure V.9 : Principe des puces à exons utilisées. **A** : Représentation des sondes, spécifiques à chaque exon, utilisées dans les puces exons. **B** : Stratégie générale pour la comparaison des transcriptomes de cellules non stressées et stressées par emploi de puces à ADN. Les ARN totaux à comparer sont rétrotranscrits puis marqués grâce à un fluorophore, puis hybridés sur les micropuces. Chaque « 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 du taux d'expression des différents segments pour un transcrit donné.

V.3.2. Caractéristiques des échantillons analysés sur les puces à exons.

Les ARN utilisés pour l'étude provenaient de 3 expériences distinctes incluant chacune l'isolement d'ARN à partir de cellules HeLa S3 contrôle et de cellules HeLa S3 soumises à un stress oxydant. Le stress oxydant a été réalisé par addition de 500 μM d' H_2O_2 dans les conditions expérimentales établies dans le chapitre précédent avec différents temps de traitement : 1 h, 2 h, 4 h et 24 h. Pour chaque temps de traitement testé, un contrôle négatif sans ajout d' H_2O_2 a été réalisé et l'intégrité des ARN totaux extraits à partir des cellules a été vérifiée par électrophorèse capillaire (voir Matériel et Méthodes). N. Nicot, de l'équipe de L. Vallar au CRP Luxembourg, a ensuite soumis ces ARN totaux à différents traitements afin de les convertir en ADNc marqués. Les ADNc ainsi obtenus ont été hybridés aux puces. Une fois l'hybridation réalisée et les lavages effectués, le signal lumineux émis par chacun des spots a été scanné, puis quantifié. Pour chaque expérience, les données obtenues dans les cellules non stressées ont été comparées à celles obtenues dans les cellules stressées. Une différence de signal indiquait une différence du taux d'expression du segment d'ARNm correspondant à la sonde contenue dans le spot. Les différences obtenues au sein de chacune des 3 expériences et ceci pour chacun des temps considérés ont été moyennées, puis une analyse statistique a été réalisée par T. Kaoma de l'équipe de L. Vallar. À l'issue de ces analyses, un tableau de synthèse des résultats nous a été fourni.

V.3.3. Données concernant les variations d'expression des gènes.

V.3.3.1. Gènes dont l'expression est induite ou réprimée suite à un stress oxydant.

L'analyse globale de l'expression des gènes montre une expression différentielle significative (False Discovery Rate ou $\text{FDR} \leq 10^{-3}$) de 51, 947, 1870 et 1376 gènes, 1 h, 2 h, 4 h et 24 h respectivement après l'induction du stress oxydant (voir Tableau V.1). De manière intéressante on constate une augmentation croissante du nombre de gènes dont l'expression est modulée suite au stress entre les temps 1h et 4h et ensuite une diminution à 24 h. De plus, on remarque qu'aux temps 1 h et 2 h près de 80 % des gènes modulés le sont positivement, tandis qu'au temps 4 h on retrouve autant de gènes positivement que négativement régulés et que 24 h après le début du stress 61 % des gènes dont l'expression

est modulée le sont négativement. L'expression de nombreux gènes est donc préférentiellement activée au cours de la phase précoce de réponse au stress oxydant et réprimée lors de phase de réponse tardive.

	1 H	2 H	4 H	24 H
Gènes totaux dont l'expression varie	51	947	1870	1376
Gènes négativement régulés	8	203	938	834
Gènes positivement régulés	43	744	932	542

Tableau V.1 : Nombre de gènes dont l'expression varie de manière significative ($p\text{-value} \leq 10^{-3}$) 1 h, 2 h, 4 h et 24 h après le début du stress oxydant. Pour chacun des temps, le nombre de gènes dont l'expression est négativement ou positivement régulée est indiqué.

En ce qui concerne les variations d'expression de gènes observées au cours de la réponse précoce, on constate que tous les gènes dont l'expression est modulée au temps 1 h voient également leur expression régulée aux temps 2 h et 4 h, ce qui est en accord avec une implication des produits de ces gènes dans la réponse précoce au stress oxydant. Parmi, ces gènes, celui présentant la plus forte induction au temps 1 h, avec un taux d'induction de 3,4x, est le gène NGLY1 qui code la protéine N-glycanase 1, une protéine associée au protéasome et qui participe à la dégradation des glycoprotéines mal repliées. L'expression de ce gène est induite 5,5x et 2,6x, respectivement, aux temps 2 h et 4 h. De même, on retrouve des modulations positives de l'expression de gènes codant des protéines de réponse au stress oxydant tel que les gènes POR (P450 oxydoréductase) et DDIT3 (DNA-dommage_inductible transcript 3) qui interviennent, respectivement, dans les systèmes de détoxification de la cellule et de réparation de l'ADN. De manière attendue, l'expression de nombreux gènes impliqués dans l'apoptose et la mort cellulaire est positivement régulée, en particulier les gènes FOS et FOSB (FBJ murine osteosarcoma viral oncogene homolog B) dont l'expression est augmentée respectivement de 3.2x et 4.4x aux temps 4 h, ainsi que p53 (TP53) de 1,6x. Les gènes codant des protéines, dont l'expression est connue être induite lors d'un stress oxydant, sont également présents parmi la liste de gènes dont l'expression est modulée au temps 4 h comme GADD45 α et β (Growth Arrest And DNA-Damage-Inducible), qui régulent la transition de » la phase G1 vers la phase S du cycle cellulaire. Il est

à noter que le gène le plus surexprimé au temps 4 h code la protéine FBXW10, qui participe au phénomène d'ubiquitination impliquant la protéine E3 ubiquitine ligase en association avec le protéasome. Cependant, l'expression de ce gène ne semble pas être modulée aux temps 1 h et 2 h.

Le gène le plus réprimé au temps 1 h est le gène NEXN qui code la protéine nexilin qui est proposée intervenir dans la régulation de l'adhésion et la migration cellulaire au travers de son association avec le cytosquelette d'actine. Comme nous le verrons par la suite, les gènes dont l'expression est régulée au cours de la réponse cellulaire peuvent être regroupés selon leur fonction cellulaire et l'adhésion des cellules constituent l'un de ces groupes. Parmi les gènes les plus négativement régulés, on retrouve de nombreux gènes d'histones dont HIST1H3F (histone cluster 1, H3f) qui est négativement régulé de 2.5x au temps 2 h et 3.6x au temps 4 h, mais aussi HIST1H4D (histone cluster 1, H4d) qui est négativement régulé aux temps 1 h (-1,7x), 2 h (-1,9x) et 4 h (-2,9x). Cette régulation négative de l'expression de certains gènes d'histones a déjà été observée dans d'autres stress, notamment suite à un stress thermique (Kato et al., 2009; Takahashi et al., 2008).

En ce qui concerne les variations de l'expression des gènes lors de la phase tardive, près de 34 % des gènes positivement régulés sont communs aux différents temps précoces testés et 16 % en ce qui concerne les gènes négativement régulés. Ce résultat suggère une continuité dans la réponse cellulaire au stress entre les deux phases. Cependant, on retrouve également des gènes dont l'expression est positivement régulée dans les temps précoces de réponse et négativement régulée en phase tardive. Ainsi, l'expression du gène codant la protéine TNFSF10 (pour Tumor Necrosis Factor Superfamily 10, appelé également Trail), qui participe à l'induction de l'apoptose, est positivement régulée au temps 4 h (1,6x), puis négativement régulée au temps 24 h (-1,6x). Parmi, les gènes les plus positivement régulés spécifiques au temps 24 h, on trouve le gène codant la protéine MMP1 (matrix metalloproteinase 1), qui joue un rôle essentiel dans le développement des tumeurs métastatiques, avec un taux d'induction de 4x.

V.3.3.2. Analyse bio-informatique des voies de régulation et des groupes fonctionnels auxquels appartiennent les gènes dont l'expression est modifiée lors d'un stress oxydant.

Afin d'avoir une vision plus intégrée de l'effet du stress oxydant sur l'expression des gènes de cellules HeLa, nous avons entrepris de rechercher si certaines voies de régulation ou certaines fonctions étaient plus particulièrement affectées au cours de la réponse au stress oxydant. Plusieurs outils informatiques permettent de réaliser ce type d'analyse à partir de données d'expériences à haut débit. Dans notre cas, en collaboration avec A. Muller de l'équipe de L. Vallar au CRP de Luxembourg, nous avons réalisé cette analyse en utilisant les logiciels DAVID et Ingenuity.

V.3.3.2.1. Analyse des données par le logiciel DAVID.

Le logiciel David (Database Annotation, Visualization and integrated Discory) utilise simultanément plusieurs bases de données de termes ontologiques en particulier les bases de données Interpro, Pir superfamily, ainsi que les termes GO (Gene Ontology) issus de la base de données <http://geneontology.org> (Ashburner et al., 2000). Ces termes GO correspondent à des dénominations standardisées attribuées à chaque gène suivant les caractéristiques des produits qu'il exprime. Ainsi, chaque gène entré dans le logiciel David est associé à un ou plusieurs GO caractérisant le processus biologique dans lequel intervient la protéine associée à ce gène (prolifération cellulaire, cycle cellulaire, métabolisme des protéines, apoptose...), sa fonction moléculaire (interaction avec l'ADN, activité de transduction du signal, interaction avec les protéines..) et sa localisation cellulaire. Le logiciel David regroupe ensuite les différents gènes en fonction des termes ontologiques qui leurs sont associés et identifie d'éventuel enrichissement de termes biologiques. Pour chaque GO identifié, un taux d'enrichissement est calculé et des tests statistiques sont effectués pour évaluer le niveau de significativité de l'enrichissement du terme d'intérêt. A partir de ces résultats, le logiciel regroupe les GO biologiquement reliées entre eux en cluster et associe un score d'enrichissement. Ce score d'enrichissement est directement lié à la p-value calculée pour le groupe fonctionnel selon la formule suivante : $\text{score} = -\log(p\text{-value})$, par conséquent un score d'enrichissement ≥ 2 correspond à une p-value ≤ 0.01 et plus le score d'enrichissement est élevé, plus l'enrichissement de ces termes est significatif. Les listes de gènes dont l'expression variait 1 h, 2 h, 4 h et 24 h après l'induction du stress oxydant ont été soumises au logiciel DAVID.

Du fait du faible nombre de gènes dont l'expression variait 1 h post-induction, l'analyse par le logiciel DAVID n'a pas permis de mettre en évidence un enrichissement significatif d'un quelconque processus cellulaire à ce temps. Cependant, les données obtenues aux temps 2 h et 4 h ont permis de faire ressortir un enrichissement de plusieurs groupes fonctionnels qui sont identiques à ces deux temps. Les trois groupes fonctionnels, possédant les plus hauts scores d'enrichissement DAVID aux temps 4 h sont présentés dans le tableau V.2.

On constate que 4 h après le début du stress oxydant le groupe fonctionnel présentant le plus haut score d'enrichissement est associé à la régulation de l'apoptose et à la survie cellulaire. Ceci est tout à fait en accord avec la mortalité cellulaire observée au cours de la réponse au stress oxydant. L'analyse par DAVID met également en évidence un effet marqué du stress oxydant sur l'expression de gènes dont la fonction est liée à l'actine et à l'organisation du cytosquelette. Il est à noter que l'analyse par DAVID fait ressortir également d'autres groupes fonctionnels associés au cytosquelette, mais avec un score d'enrichissement moins important, mais néanmoins significatif, ce qui renforce le résultat (données présentées en Annexe 1). L'analyse met également en évidence un enrichissement de termes GO associés à la régulation de la transcription, indiquant une forte modulation de la transcription lors de la réponse au stress oxydant.

L'analyse par DAVID des données obtenues au temps 24 h a été également réalisée et les principaux résultats sont présentés dans le tableau V.3. La liste complète des groupes fonctionnels identifiés et des scores d'enrichissement associés est présentée en Annexe 2. Comme pour les temps précoces, l'analyse montre un enrichissement de GO associés au cytosquelette, à la régulation de l'apoptose, mais aussi à la migration des cellules qui représente le plus haut score d'enrichissement DAVID. De manière attendue, les groupes fonctionnels présentant les 2^{ème} et 3^{ème} plus hauts scores d'enrichissement correspondent à la réponse aux vitamines, avec notamment la réponse à la vitamine A qui appartient aux systèmes antioxydants non enzymatiques, et à la réponse au stress oxydant, avec des gènes associés à la réponse au peroxyde d'hydrogène. Enfin, l'analyse met également en évidence un groupe fonctionnel associé à l'angiogenèse parmi les plus hauts scores d'enrichissement DAVID, ce qui, en association avec le groupe fonctionnel associé à la migration des cellules, est en adéquation avec l'implication du stress oxydant dans le développement tumoral.

Groupe fonctionnel : Apoptose		Score d'enrichissement : 5,53				
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0043065~positive regulation of apoptosis	71	4.6	2,0E+00	5,5E-08	2,0E-04
terme GO BP	GO:0043068~positive regulation of programmed cell death	71	4.6	1,9E+00	7,3E-08	1,3E-04
terme GO BP	GO:0010942~positive regulation of cell death	71	4.6	1,9E+00	8,7E-08	1,1E-04
terme GO BP	GO:0042981~regulation of apoptosis	110	7.1	1,6E+00	3,1E-07	1,2E-04
terme GO BP	GO:0043067~regulation of programmed cell death	110	7.1	1,6E+00	5,1E-07	1,6E-04
terme GO BP	GO:0010941~regulation of cell death	110	7.1	1,6E+00	6,1E-07	1,5E-04
terme GO BP	GO:0006917~induction of apoptosis	53	3.4	2,0E+00	3,0E-06	5,9E-04
terme GO BP	GO:0012502~induction of programmed cell death	53	3.4	2,0E+00	3,3E-06	6,1E-04
terme GO BP	GO:0043066~negative regulation of apoptosis	52	3.4	1,7E+00	1,0E-04	1,2E-02
terme GO BP	GO:0043069~negative regulation of programmed cell death	52	3.4	1,7E+00	1,5E-04	1,6E-02
terme GO BP	GO:0060548~negative regulation of cell death	52	3.4	1,7E+00	1,6E-04	1,6E-02
terme GO BP	GO:0006916~anti-apoptosis	33	2.1	1,9E+00	5,3E-04	3,2E-02

Groupe fonctionnel : Cytosquelette		Score d'enrichissement : 5,47				
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mot-clé Pir Superfamily	actin-binding	45	2.9	2,4E+00	1,0E-07	7,3E-06
terme GO MF	GO:0008092~cytoskeletal protein binding	75	4.9	1,8E+00	1,6E-06	1,7E-03
terme GO MF	GO:0003779~actin binding	54	3.5	2,0E+00	2,5E-06	1,4E-03
terme GO CC	GO:0015629~actin cytoskeleton	38	2.4	1,9E+00	3,0E-04	1,3E-02

Groupe fonctionnel : Régulation de la transcription				Score d'enrichissement : 5,18		
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mot-clé Pir Superfamily	nucleus	435	28.4	1,3E+00	7,4E-12	2,3E-09
terme GO BP	GO:0031328~positive regulation of cellular biosynthetic process	99	6.4	1,7E+00	1,0E-07	9,5E-05
terme GO BP	GO:0010557~positive regulation of macromolecule biosynthetic process	95	6.2	1,7E+00	1,5E-07	1,1E-04
terme GO BP	GO:0045893~positive regulation of transcription, DNA-dependent	75	4.9	1,9E+00	1,7E-07	1,1E-04
terme GO BP	GO:0009891~positive regulation of biosynthetic process	99	6.4	1,7E+00	2,1E-07	1,1E-04
terme GO BP	GO:0051254~positive regulation of RNA metabolic process	75	4.9	1,8E+00	2,4E-07	1,1E-04
terme GO BP	GO:0045941~positive regulation of transcription	84	5.4	1,8E+00	3,0E-07	1,3E-04
terme GO BP	GO:0006357~regulation of transcription from RNA polymerase II promoter	101	6.6	1,6E+00	5,1E-07	1,7E-04
terme GO BP	GO:0051173~positive regulation of nitrogen compound metabolic process	92	6.0	1,7E+00	5,2E-07	1,5E-04
terme GO BP	GO:0010628~positive regulation of gene expression	85	5.5	1,7E+00	5,5E-07	1,5E-04
terme GO BP	GO:0045935~positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	89	5.8	1,7E+00	8,7E-07	2,0E-04
terme GO BP	GO:0010604~positive regulation of macromolecule metabolic process	113	7.3	1,6E+00	1,4E-06	3,0E-04
Mot-clé Pir Superfamily	Transcription	215	14.0	1,4E+00	1,5E-06	7,8E-05
Mot-clé Pir Superfamily	transcription regulation	211	13.7	1,4E+00	1,6E-06	7,6E-05
terme GO BP	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	59	3.8	1,9E+00	2,9E-06	6,0E-04

Tableau V.2 : Groupes fonctionnels de gènes pour lesquels des variations significatives d'expression sont détectées avec les meilleurs scores statistiques par le logiciel DAVID au temps 4 h. Pour chaque groupe fonctionnel sont présentés les différents termes associés et leurs origines (catégorie), le nombre de gènes concernés dans nos données (nombre) et le pourcentage de gènes que cela représente (%) sur la totalité des données analysées. L'enrichissement par rapport à la proportion que ces termes représentent dans la totalité du génome humain est également indiqué. Le niveau de significativité des différents enrichissements est déterminé par deux tests statistiques (P-value, Benjamini). A partir des p-values obtenues pour chacun des termes, DAVID calcule un score d'enrichissement pour le groupe fonctionnel selon la formule " score = $-\log(p\text{-value})$.

Groupe fonctionnel : Migration cellulaire		Score d'enrichissement : 6,7				
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0030334~regulation of cell migration	35	2,8	3,0E+00	1,6E-08	1,1E-05
terme GO BP	GO:0040012~regulation of locomotion	37	3,0	2,8E+00	4,1E-08	1,3E-05
terme GO BP	GO:0051270~regulation of cell motion	37	3,0	2,7E+00	4,7E-08	1,2E-05
terme GO BP	GO:0030335~positive regulation of cell migration	22	1,8	3,5E+00	5,7E-07	9,6E-05
terme GO BP	GO:0040017~positive regulation of locomotion	23	1,9	3,4E+00	7,7E-07	1,2E-04
terme GO BP	GO:0051272~positive regulation of cell motion	22	1,8	3,2E+00	3,1E-06	3,7E-04

Groupe fonctionnel : Réponse aux vitamines		Score d'enrichissement : 6,2				
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0033273~response to vitamin	35	1,7	3,0E+00	1,6E-08	1,1E-05
terme GO BP	GO:0009991~response to extracellular stimulus	37	3,2	2,8E+00	4,1E-08	1,3E-05
terme GO BP	GO:0007584~response to nutrient	37	2,3	2,7E+00	4,7E-08	1,2E-05
terme GO BP	GO:0031667~response to nutrient levels	22	2,8	3,5E+00	5,7E-07	9,6E-05
terme GO BP	GO:0033189~response to vitamin A	23	0,9	3,4E+00	7,7E-07	1,2E-04

Groupe fonctionnel : Réponse au stress oxydant		Score d'enrichissement : 6,13				
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0010035~response to inorganic substance	44	3,6	3,1E+00	4,9E-11	8,6E-08
terme GO BP	GO:0006979~response to oxidative stress	33	2,7	2,9E+00	8,9E-08	1,7E-05
terme GO BP	GO:0000302~response to reactive oxygen species	17	1,4	3,2E+00	4,8E-05	3,2E-03
terme GO BP	GO:0042542~response to hydrogen peroxide	12	1,0	3,1E+00	1,5E-03	4,4E-02

Groupe fonctionnel : Angiogenèse		Score d'enrichissement : 5,76				
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0001568~blood vessel development	43	3,5	2,5E+00	4,7E-08	1,3E-05
terme GO BP	GO:0001944~vasculature development	43	3,5	2,5E+00	9,4E-08	1,7E-05
terme GO BP	GO:0001525~angiogenesis	26	2,1	2,5E+00	3,1E-05	2,2E-03
terme GO BP	GO:0048514~blood vessel morphogenesis	32	2,6	2,2E+00	6,3E-05	3,9E-03

Groupe fonctionnel : Cytosquelette		Score d'enrichissement : 5,58				
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mot-clé Pir Superfamily	actin-binding	38	3,1	2,4E+00	7,8E-07	9,6E-05
terme GO MF	GO:0008092~cytoskeletal protein binding	65	5,3	1,8E+00	2,5E-06	2,7E-03
terme GO MF	GO:0003779~actin binding	46	3,7	2,0E+00	9,5E-06	5,2E-03

Groupe fonctionnel : Apoptose		Score d'enrichissement : 5,0				
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0042981~regulation of apoptosis	99	8,0	1,8E+00	2,6E-08	1,3E-05
terme GO BP	GO:0043067~regulation of programmed cell death	99	8,0	1,7E+00	4,3E-08	1,3E-05
terme GO BP	GO:0010941~regulation of cell death	99	8,0	1,7E+00	5,1E-08	1,2E-05
terme GO BP	GO:0043068~positive regulation of programmed cell death	57	4,6	1,9E+00	5,0E-06	5,7E-04
terme GO BP	GO:0010942~positive regulation of cell death	57	4,6	1,9E+00	5,7E-06	6,3E-04
terme GO BP	GO:0043065~positive regulation of apoptosis	56	4,5	1,9E+00	8,4E-06	7,8E-04
terme GO BP	GO:0043066~negative regulation of apoptosis	46	3,7	1,9E+00	6,4E-05	3,9E-03
terme GO BP	GO:0043069~negative regulation of programmed cell death	46	3,7	1,8E+00	8,9E-05	4,7E-03
terme GO BP	GO:0060548~negative regulation of cell death	46	3,7	1,8E+00	9,6E-05	5,0E-03
terme GO BP	GO:0012502~induction of programmed cell death	41	3,3	1,8E+00	2,4E-04	1,0E-02
terme GO BP	GO:0006917~induction of apoptosis	40	3,2	1,8E+00	4,6E-04	1,8E-02
terme GO BP	GO:0006916~anti-apoptosis	28	2,3	1,9E+00	1,1E-03	3,7E-02

Tableau V.3 : Groupes fonctionnels de gènes pour lesquels des variations significatives d'expression sont détectées avec les meilleurs scores statistiques par le logiciel DAVID au temps 24 h. Pour chaque groupe fonctionnel sont présentés les différents termes associés et leurs origines (catégorie), le nombre de gènes concernés dans nos données (nombre) et le pourcentage de gènes que cela représente (%) sur la totalité des données analysées. L'enrichissement par rapport à la proportion que ces termes représentent dans la totalité du génome humain est également indiqué. Le niveau de significativité des différents enrichissements est déterminé par deux tests statistiques (P-value, Benjamini). A partir des p-values obtenues pour chacun des termes, DAVID calcule un score d'enrichissement pour le groupe fonctionnel selon la formule " score = $-\log(p\text{-value})$.

V.3.3.2.2. Analyse des données par le logiciel Ingenuity.

Afin de compléter les analyses effectuées avec le logiciel DAVID, nous avons, toujours en collaboration avec l'équipe de L.Vallar au CRP Luxembourg, analysé nos résultats en utilisant le logiciel Ingenuity. Ce logiciel, tout comme DAVID, permet d'identifier des processus cellulaires et des voies métaboliques associés à un nombre significatif de gènes dont l'expression varie suite à un traitement. Cependant, la particularité du logiciel Ingenuity est qu'il permet également de réaliser une analyse dite en amont ("upstream analysis"). En effet, Ingenuity prend en considération le niveau de variation des différents gènes modulés, c'est-à-dire s'ils sont négativement ou positivement régulés, et à partir d'une base de données incluant des résultats expérimentaux publiés, le logiciel permet de mettre en évidence et d'identifier l'activation ou l'inhibition de voies métaboliques, mais aussi de potentiels facteurs de régulations impliqués dans les variations d'expression observées dans les données. Les éléments régulateurs considérés peuvent être des facteurs de transcription, mais également des microARN, des kinases ou des enzymes. Pour chacun des régulateurs potentiels identifiés, des tests statistiques sont réalisés et une p-value, reflétant 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, est calculée.

V.3.3.2.2.a. Analyse des voies métaboliques par le logiciel Ingenuity

Les listes de gènes dont l'expression variait aux temps 1 h, 2 h, 4 h et 24 h ont été soumises au logiciel Ingenuity afin de mettre en évidence des voies métabolique impliquées dans la réponse au stress oxydant. Tout comme avec le logiciel DAVID aucun résultat significatif n'a pu être obtenu pour les données du temps 1 h du fait du faible nombre de gènes dont l'expression variait. Les principaux résultats obtenus pour chacun des autres temps sont présentés dans le tableau V.4. L'analyse des résultats des temps 2 h et 4 h met en évidence un enrichissement de gènes impliqués dans divers voies métaboliques associées au système immunitaire, mais aussi, comme attendu, dans la réponse au stress oxydant via le facteur de transcription Nrf2 et l'action antioxydante de la vitamine C. Concernant les résultats au temps 24 h, on observe un très fort enrichissement de gènes impliqués dans la biosynthèse

des cholestérols, mais aussi de gènes impliqués dans la différenciation cellulaire. Ces résultats, en accord avec ceux obtenus par le logiciel DAVID, mettent également en évidence des voies de régulation associées à la survie cellulaire (p53 Signaling) et la transduction du signal (Protein Kinase A Signaling). On notera également que, en accord avec nos données concernant l'expression des ARN Sat III de façon tardive au cours de la réponse au stress oxydant, la réponse aux protéines dénaturées, qui repose en grande partie sur l'activité du facteur de transcription HSF1, fait partie des voies métaboliques spécifiquement enrichies à 24 h.

Les résultats des analyses par les logiciels DAVID et Ingenuity, bien que nécessitant d'être confirmés expérimentalement, nous permettaient donc d'avoir un aperçu global de l'impact d'un stress oxydant généré par l'ajout de 500 μM d' H_2O_2 sur les cellules HeLa.

2h		4h	
Voie métabolique	p-value	Voie métabolique	p-value
CD27 Signaling in Lymphocytes	1,5E-05	IL-17A Signaling in Gastric Cells	2,8E-05
TNFR1 Signaling	2,9E-04	Activation of IRF by Cytosolic Pattern Recognition Receptors	1,6E-04
GNRH Signaling	4,1E-04	Protein Kinase A Signaling	1,7E-04
NRF2-mediated Oxidative Stress Response	5,6E-04	Antioxidant Action of Vitamin C	4,8E-04
LPS-stimulated MAPK Signaling	6,6E-04	EGF Signaling	7,6E-04
CD40 Signaling	1,5E-03	Cholecystokinin/Gastrin-mediated Signaling	9,8E-04
Trans, trans-farnesyl Diphosphate Biosynthesis	1,9E-03	Role of RIG1-like Receptors in Antiviral Innate Immunity	1,1E-03
PPAR Signaling	2,8E-03	NRF2-mediated Oxidative Stress Response	1,5E-03
IL-10 Signaling	2,8E-03	RAN Signaling	1,9E-03
RANK Signaling in Osteoclasts	4,1E-03	Role of JAK1, JAK2 and TYK2 in Interferon Signaling	2,0E-03
TNFR2 Signaling	4,1E-03	April Mediated Signaling	2,1E-03
ILK Signaling	4,6E-03	Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes	2,5E-03
HGF Signaling	4,6E-03	Role of PKR in Interferon Induction and Antiviral Response	2,7E-03
p53 Signaling	5,2E-03	B Cell Activating Factor Signaling	2,7E-03
4-1BB Signaling in T Lymphocytes	5,8E-03	iNOS Signaling	2,7E-03

24h	
Voie métabolique	p-value
Hepatic Fibrosis / Hepatic Stellate Cell Activation	1,5E-07
Cholesterol Biosynthesis I	5,6E-07
Cholesterol Biosynthesis II (via 24,25-dihydrostanosterol)	5,6E-07
Cholesterol Biosynthesis III (via Desmosterol)	5,6E-07
Superpathway of Cholesterol Biosynthesis	3,0E-05
Superpathway of Serine and Glycine Biosynthesis I	1,8E-04
p53 Signaling	6,2E-04
Unfolded protein response	8,3E-04
Folate Transformations I	1,1E-03
Regulation of the Epithelial-Mesenchymal Transition Pathway	1,4E-03
TR/RXR Activation	2,1E-03
IL-8 Signaling	2,2E-03
ILK Signaling	2,7E-03
Protein Kinase A Signaling	3,0E-03
Ethanol Degradation IV	4,6E-03

Tableau V.4 : Voies métaboliques identifiées par le logiciel Ingenuity pour lesquelles l'expression d'un nombre significatif de gènes est modulée aux temps 2 h, 4 h et 24 h. Pour chaque temps sont listées les différentes voies métaboliques et leur p-value associée.

V.3.3.2.2.b. Identification de potentiel facteur de régulations impliquées dans le stress oxydant par le logiciel Ingenuity.

Comme mentionné précédemment, le logiciel Ingenuity permet aussi d'identifier de potentiels facteurs régulateurs susceptibles d'avoir provoqué l'induction ou l'inhibition d'un ensemble de gènes dont l'expression est modulée suite à un traitement, ceci en se basant sur des réseaux de régulation établis à partir de données expérimentales. Pour chaque régulateur potentiel identifié comment pouvant contribuer à la modulation d'expression d'un ensemble de gènes, une p-value est déterminée afin de vérifier statistiquement le niveau de significativité du résultat. De plus, en se basant sur le niveau d'activation ou d'inhibition des gènes cibles du régulateur potentiel identifié, Ingenuity prédit également l'impact du traitement effectué sur les cellules sur le niveau d'activation de la protéine régulatrice, à savoir si celle-ci est activée ou inhibée suite au traitement. Le niveau de significativité de cette prédiction est estimé par le calcul d'un Z-score, une valeur de Z-score élevée étant en faveur d'une activation du régulateur et une valeur de Z-score faible en faveur de son inhibition. Ainsi une valeur de Z-score supérieure à 2, indique que le régulateur potentiel identifié est probablement activé et plus la valeur de Z_score est importante, plus la probabilité que l'activation soit réelle est importante. A l'inverse, une valeur de Z-score inférieure à -2 indique une inhibition de l'activité du régulateur identifié et plus la valeur de Z_score est faible, plus la probabilité que l'inhibition soit réelle est importante.

Ainsi, l'utilisation d'Ingenuity a permis l'identification aux temps 2 h, 4 h et 24 h de respectivement 190, 179 et 298 régulateurs en amont potentiels dont l'expression des gènes cibles variait significativement dans nos données expérimentales ($p\text{-value} \leq 0,01$; listes disponibles en Annexe 3). Les résultats obtenus au temps 2 h et 4 h sont, une fois encore, assez semblables avec 63 régulateurs potentiels aux temps 4 h qui présentent un Z-score significatif, c'est-à-dire ≥ 2 ou ≤ -2 . Les régulateurs potentiels ont été classés par Ingenuity en différentes catégories selon leur fonction : régulateur de la transcription, cytokine, récepteur membranaire, kinase, miRNA, enzyme, facteur de croissance... Les principaux résultats obtenus au temps 4 h sont résumés dans le tableau V.5

Le régulateur potentiel qui présente le Z-score le plus élevé est le facteur de transcription NUPR1 (=p8 ou COM1, p-value = $1,45 \cdot 10^{-14}$, Z-score = 5,69) dont Ingenuity prédit une activation suite au stress oxydant. Le fait que seule une augmentation faible, mais néanmoins significative, du taux d'ARNm NUPR1 ait été détectée sur nos puces (augmentation de 1,26 x, p-value = $6,5 \cdot 10^{-3}$), n'est pas en contradiction avec cette prédiction, puisque l'activité de NUPR1, tout comme celle de nombreux facteurs de transcription, peut être régulée au niveau traductionnel ou post-traductionnel. Par ailleurs, NUPR1 a récemment été montré être impliqué dans le maintien de l'homéostasie redox et sa déplétion entraîne une augmentation de la quantité de ROS intracellulaires (Weis et al., 2015). NUPR1 est également connu pour être associé à la réponse à de nombreux stress, notamment thermique et génotoxique, de part sa capacité à réguler l'expression de gènes de réponse aux stress, mais aussi de part son interaction avec la protéine p53 (TP53) qui est également présente dans la liste des facteurs de régulation proposée par Ingenuity, mais avec un Z-score non significatif (p-value = $1,9 \cdot 10^{-3}$, Z-score = -0,37). Il est à noter que NUPR1 peut être activé par ATF4 une autre protéine proposée par le logiciel comme étant activée (Clark et al., 2008; Jin et al., 2009). D'autres régulateurs de la transcription comme FOS, JUN, RELA et les complexe NFkB et Ap1 qui sont impliqués dans l'apoptose et la survie cellulaire sont proposés être activés lors de la réponse précoce au stress oxydant, ce qui est cohérent avec les résultats obtenus lors de l'analyse avec le logiciel DAVID. Les résultats suggèrent également l'activation d'ERK1/2 qui est impliqué dans la stabilisation de la protéine Nrf2 et son activation dans la réponse au stress oxydant ainsi que dans la survie cellulaire. De façon surprenante cependant, la protéine Nrf2 ne ressort pas au niveau de cette analyse, alors que l'expression de certains de ses gènes cibles avérés, tels que HMOX1 (heme oxygenase (decycling) 1) et TXNDR1 (thioredoxin reductase 1), est positivement et significativement régulée 4 h après l'induction du stress (1, 5x avec une p-value de $1,03 \cdot 10^{-4}$ pour HMOX1 et 2,28x avec une p-value de $3,3 \cdot 10^{-3}$ pour TXNDR1). De nombreuses cytokines semblent également être activées et sont à mettre en relation avec la protéine TREM1 qui stimule la réponse inflammatoire et qui est également prédite pour être activée. Il est à noter que l'activité de TREM1 peut également être régulée par la protéine Nrf2 qui régule la transcription du gène TREM1 lors de processus inflammatoires non contrôlés (Thimmulappa et al., 2006). Le logiciel Ingenuity permet également de mettre en évidence l'implication potentielle de certains miRNA et notamment, ici à 4 h, du miR-155-5p dont l'activité est

prédite inhibée (p -value = $1,34 \cdot 10^{-5}$, Z-core = -3,123). Ce miRNA est proposé être impliqué dans certaines pathologies cardiaques et certains cancers et intervient dans plusieurs processus cellulaires comme la réponse inflammatoire et l'apoptose en régulant, par exemple, l'expression de FOS, FOXO3 et STAT3 (Huang et al., 2015a; Lin et al., 2015; Zhang et al., 2015). A 4 h, le Z-score le plus bas (-3,276) correspond à l'inhibition des enzymes Hdac (Histone déacétylase). Une diminution de l'expression de ces enzymes est généralement associée au passage de l'hétérochromatine péricentromérique à un état d'euchromatine et pourrait contribuer à l'expression des ARN Sat III cependant à ce temps aucune induction de l'expression des ARN Sat III n'est détectée. On notera par ailleurs, que le facteur HSF1 n'est pas retrouvé au sein de la liste des régulateurs potentiels à 4 h et que parmi ces cibles connues pour être exprimées lors d'un stress oxydant seule l'expression du gène BAG3 (pour BCL2-Associated Athanogene 3 ou BIS ou CAIR-1 ou MCM6) est significativement activée (2,5x avec une p -value de $2,9 \cdot 10^{-5}$). Il est par conséquent possible que l'absence d'expression des ARN Sat III 4 h après le début du stress soit liée à une faible activité du facteur de transcription HSF1 à ce temps. Les données obtenues au temps 24 h sont assez similaires à celles du temps 4 h avec notamment l'activation prédite de NUPR1, mais aussi des protéines impliquées dans la survie cellulaire comme FOS, RELA et le complexe NFKB (voir Tableau V.6). De même, on observe également l'activation de nombreuses cytokines et de TREM1. L'expression de miR-155-5p est également prédite être inhibée au temps 24 h (p -value de $8,9 \cdot 10^{-4}$ et Z-score de -2,621), ainsi que celle de miR-34 (p -value de $2,7 \cdot 10^{-2}$ et Z-score de -2,219) qui est proposé favoriser la transcription du gène codant la protéine p53. On constate cependant, que certaines des prédictions faites au temps 4 h ne le sont plus au temps 24 h, suggérant des régulations spécifiques aux temps précoces et à la phase tardive. C'est le cas d'ERK1/2 qui n'est plus prédit comme activé au temps 24 h. A l'inverse, d'autres protéines prédites comme activées sont spécifiques aux temps 24 h comme la protéine KDM5B (lysine (K)-specific demethylase 5B) qui régule la déméthylation des histones H3K4. Il est proposé que la protéine KDM5B contribue à l'inhibition de l'expression de certains gènes en favorisant le compactage de la chromatine, en particulier au niveau de gènes impliqués dans la sénescence cellulaire (Dai et al., 2014; Li et al., 2013). C'est également le cas du facteur de transcription HSF1 qui, bien que non présenté dans le tableau V.6 où seuls les facteurs de régulation avec des Z-score ≥ 2 ou ≤ -2 sont listés, est présent dans la liste des régulateurs potentiels à 24 h avec une p -value de $7,6 \cdot 10^{-3}$, alors qu'il ne l'était pas à 4h

(Annexe 3). Cette dernière observation est en accord avec l'observation de la surexpression des ARN Sat III en phase tardive de réponse au stress oxydant. Il est nécessaire de rappeler que les prédictions faites par Ingenuity et décrites dans ces derniers paragraphes demandent à être confirmées expérimentalement. Néanmoins, ces analyses nous ont permis d'avoir un aperçu des processus et des mécanismes qui peuvent être mis en jeu dans les différents de temps de réponse au stress oxydant.

4h				
Régulateur de la transcription				
Régulateur	taux d'induction	prédiction	z-score	p-value
NUPR1	1,259	Activé	5,696	1,5E-14
FOXL2	-1,013	Activé	3,441	5,7E-05
RELA	1,015	Activé	3,293	6,7E-07
EZH2	1,172	Activé	3,03	6,7E-03
PPRC1	-1,053	Activé	2,722	1,5E-04
CEBPA	1,101	Activé	2,514	9,1E-03
CREB1	-1,096	Activé	2,416	1,4E-05
HIF1A	-1,1	Activé	2,25	2,4E-05
FOS	3,242	Activé	2,236	5,6E-03
ATF4	1,21	Activé	2,2	1,2E-05
FOXO3	1,229	Activé	2,146	8,0E-08
JUN	1,842	Activé	2,131	3,8E-05
NFKB1	2,293	Activé	2,129	3,6E-05
ERG	1,05	inhibé	-2	6,9E-04
MNT	1,09	inhibé	-2	2,3E-03
MAX	1,21	inhibé	-2	2,9E-03
SIN3A	-1,135	inhibé	-2	4,4E-02

4h				
Complexe				
Régulateur	taux d'induction	prédiction	z-score	p-value
PDGF BB		Activé	4,566	2,8E-20
NFKB (complex)		Activé	3,742	8,2E-04
Lh		Activé	2,915	5,3E-05
Cg		Activé	2,749	5,6E-03
Ap1		Activé	2,219	1,2E-02
Fcer1		Activé	2,202	7,6E-03
Immunoglobulin		Activé	2,138	1,7E-03
FSH		Activé	2,049	9,6E-07
IgG		inhibé	-2,399	4,6E-04
CD3		inhibé	-2,529	1,2E-03

4h				
Enzyme				
Régulateur	taux d'induction	prédiction	z-score	p-value
TGM2	-1,273	Activé	3,311	1,2E-05
ERK1/2		Activé	2,25	1,4E-06
MAP2K1/2		Activé	2,066	3,9E-05
PTGS2	2,833	Activé	2,061	9,8E-04
TAB1	-2,027	inhibé	-2	2,9E-02
SCD	-1,187	inhibé	-2,378	4,8E-06
Hdac		inhibé	-3,276	1,1E-06

4h				
Kinase				
Régulateur	taux d'induction	prédiction	z-score	p-value
PRKCA	1,002	Activé	2,382	1,9E-03
MAP2K1	1,095	Activé	2,213	1,8E-02
TRIB3	1,082	inhibé	-2,186	3,5E-03

4h				
miRNA				
Régulateur	taux d'induction	prédiction	z-score	p-value
miR-155-5p		inhibé	-3,123	1,3E-05

4h				
Cytokine				
Régulateur	taux d'induction	prédiction	z-score	p-value
TNF	1,582	Activé	4,05	1,7E-09
IL1B	1,014	Activé	3,749	1,5E-03
IL17A		Activé	3,341	2,0E-05
IFNG	-1,161	Activé	2,946	6,3E-03
IL1A	1,596	Activé	2,71	8,9E-03
CD40LG	1,039	Activé	2,665	3,2E-04
IL12B	1,096	Activé	2,219	3,5E-03
IL1RN	1,111	inhibé	-2,714	9,8E-04

4h				
Recepteur membranaire				
Régulateur	taux d'induction	prédiction	z-score	p-value
TREM1	1,077	Activé	4,105	7,1E-12
TLR7	-1,1	Activé	3,263	1,0E-05
CD28	1,113	inhibé	-2,159	1,8E-02

Tableau V.5 : Principaux résultats des analyses et prédictions en amont des données d'expression des puces par le logiciel Ingenuity au temps 4 h. Les protéines régulatrices prédites par le logiciel Ingenuity sont regroupées selon leur fonction. Pour chaque protéine sont indiqués le taux d'induction de son expression d'après les données de puces, la prédiction de son activation par Ingenuity (activé ou inhibé), le Z-score correspondant à cette activation et le niveau de significativité de la protéine en tant que possible régulateur (p-value).

24h				
Régulateur de la transcription				
Régulateur	taux d'induction	prédiction	z-score	p-value
NUPR1	1,459	Activated	7,988	2,1E-16
KDM5B	-1,083	Activated	3,626	4,4E-05
PPRC1	1,184	Activated	3,347	4,2E-09
FOXL2	1,021	Activated	2,506	2,0E-04
TP53	-1,282	Activated	2,364	2,4E-12
EZH2	-1,164	Activated	2,224	2,2E-02
RELA	1,219	Activated	2,07	9,5E-05
EGR1	1,59	Activated	2,067	4,5E-06
YBX1	1,034	Activated	2	1,9E-03
CDKN2A	-1,159	Activated	2	3,6E-02
EHF	1,154	Inhibited	-2,333	1,4E-02

24h				
Complexe				
Régulateur	taux d'induction	prédiction	z-score	p-value
PDGF BB		Activated	2,987	9,3E-15
Cg		Activated	2,965	3,2E-11
NFkB (complex)		Activated	2,441	6,9E-07
IL12 (complex)		Activated	2,402	6,3E-03

24h				
Enzyme				
Régulateur	taux d'induction	prédiction	z-score	p-value
DNMT3B	1,042	Activated	2,538	3,9E-03
SCD	-2,037	Inhibited	-2,183	4,7E-05

24h				
Cytokine				
Régulateur	taux d'induction	prédiction	z-score	p-value
TNF	1,092	Activated	3,22	9,0E-14
IFNG	-1,048	Activated	2,882	1,7E-03
IL1B	1,107	Activated	2,777	3,8E-04
IL17A		Activated	2,265	4,7E-04
CD40LG	-1,015	Activated	2,05	4,4E-03
IL12B	1,071	Activated	2	1,0E-02

24h				
Kinase				
Régulateur	taux d'induction	prédiction	z-score	p-value
EGFR	2,473	Activated	2,458	7,6E-05
IKBKE	-1	Activated	2,429	1,2E-04
CHUK	1,051	Activated	2,2	1,3E-03
TRIB3	2,172	Inhibited	-2,339	1,5E-08

24h				
miRNA				
Régulateur	taux d'induction	prédiction	z-score	p-value
mir-34	1,06	Inhibited	-2,219	2,7E-02
miR-155-5p		Inhibited	-2,621	8,9E-04

24h				
Recepteur membranaire				
Régulateur	taux d'induction	prédiction	z-score	p-value
TREM1	1,093	Activated	3,767	6,8E-13
TLR7	-1,178	Activated	2,46	2,6E-03

Tableau V.6 : Principaux résultats des analyses et prédictions en amont des données d'expression des puces par le logiciel Ingenuity au temps 24 h. Les protéines régulatrices prédites par le logiciel Ingenuity sont regroupées selon leur fonction. Pour chaque protéine sont indiqués le taux d'induction de son expression d'après les données de puces, la prédiction de son activation par Ingenuity (activé ou inhibé), le Z-score correspondant à cette activation et le niveau de significativité de la protéine en tant que possible régulateur (p-value).

V.3.4. Données concernant les variations d'épissage alternatif.

V.3.4.1. Analyse globale des transcrits qui présentent une modulation de l'épissage alternatif.

L'analyse des données des puces pour la détection d'évènements d'épissage modulés en réponse au stress oxydant a nécessité de nombreux traitements statistiques différents, réalisés par T. Kaoma de l'équipe de L. Vallar, afin d'obtenir les résultats les plus significatifs. Seuls les résultats ayant répondu de manière satisfaisante aux différents tests statistiques ($p < 0.05$) ont été retenus pour l'analyse. Il est à noter qu'une analyse de données entre les conditions témoins aux temps 24 h et 1 h a été réalisée afin d'identifier des modulations d'épissage associées à la croissance cellulaire ou à l'appauvrissement du milieu de culture dans le temps pour le point 24 h. Cette analyse, nous a permis de conserver seulement les évènements qui de manière significative sont associés à la réponse au stress oxydant au temps 24 h.

Le tableau V.7 récapitule le nombre d'évènements de modulation d'épissage significativement détectés pour chacun des temps d'analyse, ainsi que le nombre de transcrits associés à ces modulations. De manière générale, on observe que c'est aux temps 2 h et 4 h que l'on détecte le plus de modulations d'épissage avec respectivement 399 et 388 évènements détectés contre 151 évènements au temps 24 h et 43 évènements au temps 1 h. Les évènements d'épissage modifiés au cours de la réponse au stress oxydant et repérés grâce aux puces ont été classés en 3 catégories : les exons cassette, les exons 5' alternatifs et les exons 3' alternatifs. L'analyse de la distribution des évènements régulés à 24 h au sein de ces 3 catégories révèle une proportion de 47.6 % d'exons cassette, 36.2 % d'exons 3' alternatifs et 16,2 % d'exons 5' alternatifs.

	1 H	2 H	4 H	24 H
Evènements d'épissage détecté	43	399	388	151
Gènes dont l'épissage est modulé	43	347	291	101

Tableau V.7 : Nombre d'évènements d'épissage modulé de manière significative ($p\text{-value} \leq 10^{-3}$) 1 h, 2 h, 4 h et 24 h après le début du stress oxydant. Pour chacun des temps, le nombre d'évènements d'épissage modulé au cours du stress ainsi que le nombre de gènes concernés par ces modulations est indiqué.

De façon intéressante, plusieurs exons d'un même ARNm peuvent être différentiellement régulés au cours de la réponse au stress oxydant. Dans certains cas, cela est dû à la régulation simultanée d'un ensemble d'exons consécutifs, c'est notamment le cas lors d'un choix alternatif de site de polyadénylation, par contre dans d'autres cas aucune corrélation évidente ne semble exister entre les différents événements touchant un même ARNm. Il faut cependant garder à l'esprit que nous parlons ici de populations hétérogènes d'ARNm et que chacun des évènements détectés peut en fait être spécifique à une sous-population d'ARNm et que le fait que 2 modulations différentielles d'épissage soient détectées pour un même ARNm ne signifie pas que ces 2 modulations se produisent simultanément au sein d'une même molécule d'ARNm.

Dans un premier temps, nous avons cherché à savoir si des processus biologiques étaient spécifiquement associés aux transcrits dont l'épissage était modulé en utilisant à nouveau le logiciel DAVID (Annexe 4). Une fois de plus, l'analyse n'a pas pu être effectuée sur les données du temps 1 h du fait du trop faible nombre de modulations d'épissage détectées à ce temps. Les principaux groupes fonctionnels présentant un score DAVID supérieur à 1,3 (ce qui correspond à une p-value $\leq 0,05$) aux temps 2 h, 4 h et 24 h sont présentés dans le tableau V.8. Chaque groupe fonctionnel est nommé en fonction du terme GO le plus représentatif au sein du groupe. De manière générale, les groupes fonctionnels identifiés présentent des scores DAVID plus faibles que ceux obtenus lors de l'analyse de l'expression des gènes. Cette différence s'explique en partie par le fait que moins de gènes ont été entrés dans le logiciel d'analyse. Le meilleur score obtenu au temps 2h concerne la régulation de la transcription avec un score de 3.63. Ce groupe fonctionnel est également détecté au temps 4 h, mais avec un Z-score d'enrichissement inférieur à 1,3 et est donc considéré comme non significatif. Les groupes fonctionnels présentant les meilleurs scores d'enrichissement au temps 4 h sont des processus associés à l'activité des protéines kinases et aux ATPases avec des scores respectifs de 2.07 et 1.63. Fait intéressant, l'analyse semble mettre en évidence un lien entre les modulations d'épissage détectées et le métabolisme des lipides ainsi que la structure du cytosquelette. De manière générale, les résultats des analyses aux temps 2 h et 4 h sont assez proches et en accord avec les résultats obtenus lors de l'analyse de l'expression des gènes.

Cette correspondance entre les deux analyses peut s'expliquer par le fait que 30 % des transcrits présentant une modulation d'épissage alternatif au temps 2 h sont exprimés de manière différentielle au temps 4 h. Les résultats obtenus au temps 24 h sont différents de ceux obtenus aux temps précoces et les meilleurs scores d'enrichissement DAVID concernent des processus de différenciation cellulaire. De manière intéressante, parmi les meilleurs scores, on observe de nombreux processus associés à la migration et à l'adhésion cellulaires. Par conséquent, il est raisonnable de penser que suite à un stress oxydant et une première réponse anti-oxydante dans les temps précoces afin de limiter les dégâts cellulaires, la cellule par le biais de l'épissage alternatif favorise sa migration vers un nouvel environnement.

2 h		24 h	
Processus Biologique	Score David	Processus Biologique	Score David
Régulation de la transcription	3,63	Différenciation neuronale	3,53
Régulation des protéines d'oxydo/réduction	2,39	Contraction musculaire	2,18
Cycle cellulaire	2,18	Migration cellulaire	2,02
Réponse aux stress	2,03	Développement cardiaque	1,9
Processus Catabolique des protéines	1,99	Jonction cellulaire	1,83
Régulation de l'apoptose	1,99	Adhésion des cellules	1,68
Métabolismes des lipides	1,84	Cytosquelette/microtubules	1,63
Adhérences des cellules	1,81	Collagène	1,61
Cytosquelette/microtubules	1,59	Activité des cytokines	1,59
Cytosquelette/actine	1,44	Matrice extracellulaire	1,58

4 h	
Processus Biologique	Score David
Activité des protéines kinases	2,07
Activité des protéines ATPase	1,63
Organisation de l'actine/cytosquelette	1,54
Métabolisme des sucres	1,43
Fixation du Ca ²⁺	1,43
Métabolisme des phospholipases	1,4

Tableau V.8: Groupes fonctionnels établis par le logiciel DAVID à partir des listes de gènes pour lesquels des variations d'épissage significatives ont été détectées 2 h, 4 h ou 24 h après induction du stress oxydant. Chacun des groupes est nommé en fonction du terme GO le plus représentatif du groupe. Le score d'enrichissement DAVID est calculé pour l'ensemble du groupe fonctionnel selon la formule " score = $-\log(p\text{-value})$

V.4. CONFIRMATION PAR RT-PCR OU RT-QPCR DES EVENEMENTS D'ÉPISSAGE ALTERNATIF DETECTES APRES UN STRESS OXYDANT.

Comme il a été dit précédemment, les données obtenues par les puces sont soumises à différents traitements statistiques qui peuvent entraîner dans certains cas l'apparition de faux positifs. Bien que l'utilisation de plusieurs outils statistiques tende à réduire la possibilité d'un tel évènement, il demeure nécessaire de confirmer expérimentalement les variations d'épissage alternatif détectées par les puces.

Au travers de cette étude, notre but ultime n'était pas de valider de façon exhaustive toutes les variations d'épissage observées au cours de la réponse au stress oxydant, mais plutôt de constituer une liste d'évènements d'épissage différentiels communs ou spécifiques à chacun des temps en vue de caractériser par la suite les effets du stress oxydant sur les composants du spliceosome et la régulation de l'épissage. C'est pourquoi nous avons choisi, dans un premier temps, de ne valider qu'un nombre restreint d'évènements, en basant nos choix sur l'intensité de l'effet observé, sa spécificité à un temps donné ou au contraire sa généralisation à tous les temps et enfin sur l'intérêt biologique potentiel de la variation observée. Il est à noter que cette étape de validation est actuellement toujours en cours.

Pour réaliser ces validations, nous avons réalisé, selon la localisation de l'exon d'intérêt au sein de son transcrit (exon cassette, exon 5'- ou 3'-alternatif) soit des expériences de RT-PCR semi-quantitatives, soit des expériences de RT-qPCR (RT-PCR quantitatives). Dans les deux cas, nous avons utilisé les ARN totaux utilisés pour réaliser l'hybridation des puces, mais également des ARN totaux extraits à partir de séries supplémentaires de cellules HeLa soumises ou non à un stress oxydant réalisé par ajout de 250 μM ou 500 μM d' H_2O_2 dans le milieu de culture. Les ARN totaux ont été ensuite rétrotranscrits en utilisant un oligonucléotide dT₁₅ s'hybridant au niveau de la queue polyA des ARNm ou un oligonucléotide hexamérique dégénéré (N₆) s'hybridant aléatoirement tout au long de la séquence des ARNm. Dans le cas de la validation de l'inclusion différentielle d'un exon cassette, j'ai réalisé une PCR semi-quantitative en utilisant des oligonucléotides s'hybridant au sein des exons flanquants l'exon d'intérêt. La PCR réalisée permettait de détecter simultanément les isoformes présentant ou non l'inclusion de l'exon d'intérêt au sein de la même amplification. Les différents produits PCR obtenus ont été ensuite séparés dans un gel

d'agarose, puis quantifiés grâce au logiciel SynGene et les résultats sont exprimés en proportion relative de chacune des isoformes. Dans le cas de la validation de l'inclusion différentielle d'un exon 5'- ou 3-terminal, il n'était pas possible d'utiliser la même stratégie puisque les exons terminaux ne sont pas encadrés d'exons flanquants et j'ai donc du réaliser une PCR quantitative en utilisant des oligonucléotides spécifiques de l'exon d'intérêt et des oligonucléotides rapporteurs qui ciblent un autre exon du transcrit auquel appartient l'exon d'intérêt et qui est prédit pour être inclus de façon constitutive et donc être invariant. Pour chaque validation, les expériences ont été réalisées trois fois minimum et les résultats obtenus ont été moyennés et soumis à un test statistique (t-student) afin de déterminer si les variations observées étaient significatives.

V.4.1. Validation de la modulation de l'épissage de l'exon 3 du pré-ARNm codant la protéine TRA2B, un facteur de régulation de l'épissage alternatif.

Comme présenté dans le chapitre de l'introduction dédié à l'épissage, Tra2 β est une protéine dite SR-like qui participe à de nombreux processus cellulaires dont la régulation de l'épissage alternatif de plusieurs pré-ARNm et qui est connue pour être impliquée dans certaines pathologies comme l'artériosclérose et certains cancers. Pour rappel, le gène tra2 β est composé de 10 exons qui permettent la production de 5 isoformes d'ARNm dont tra2 β 1 qui est l'isoforme majoritaire. Tra2 β 1 comporte l'ensemble des exons à l'exception de l'exon 2 et permet la synthèse d'une protéine active composée de deux domaines SR et d'un domaine RRM (Figure V.10).

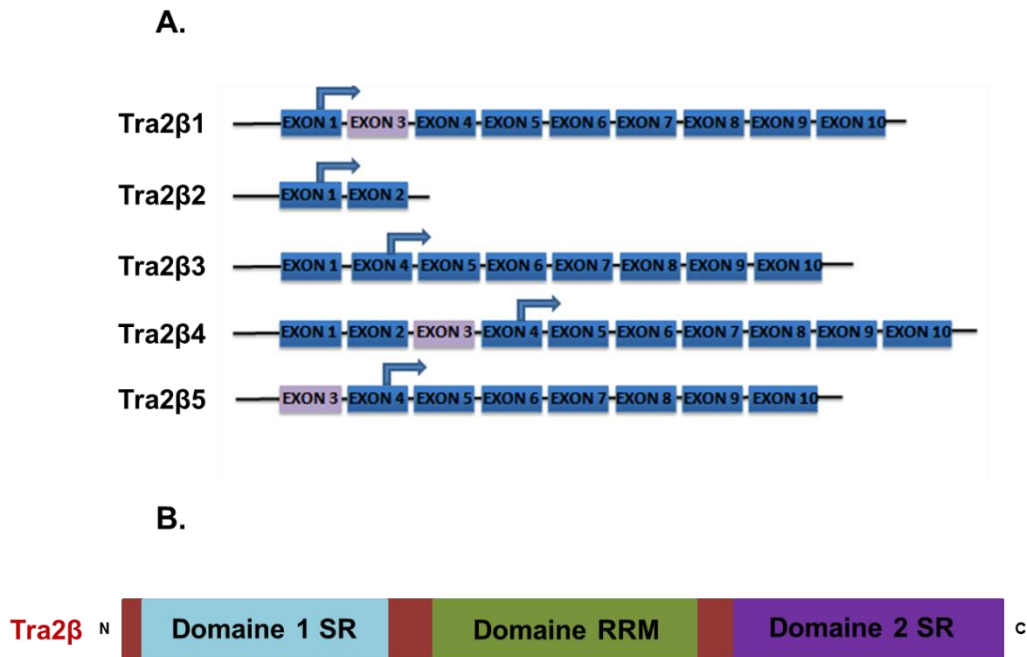


Figure V.10 : Représentation schématique des différentes isoformes d'ARNm du gène *Tra2β* et de la protéine *tra2β* de taille entière (adapté de Takeo et al., 2009). A. Schéma des isoformes d'ARNm *tra2β*1 à 5. Les exons sont représentés par un rectangle bleu hormis l'exon 3 pour lequel une inclusion différentielle a été détectée sur puces et qui est représenté par un rectangle violet. Pour chaque isoforme, la localisation du codon AUG initiateur de la traduction est représentée par une flèche bleue. **B.** Schéma des domaines structuraux principaux de la protéine *tra2β*.

Des études, présentées dans le chapitre d'introduction dédié au stress oxydant, avaient montré qu'un stress oxydant à l'arsenite induisait une augmentation de l'inclusion de l'exon 2 au sein des ARNm *tra2β* favorisant ainsi la formation de l'isoforme *tra2β*4 (Takeo et al., 2009). Dans notre cas, c'est une exclusion préférentielle de l'exon 3 que nous détectons sur nos puces aux temps précoces de réponse à un stress oxydant. L'exon 2 pouvant être aussi alternatif, la validation de cet évènement d'épissage a été réalisée par RT-PCR semi-quantitative en utilisant des oligonucléotides s'hybridant au niveau des exons 1 et 4 afin d'observer l'impact du stress oxydant sur l'inclusion des exons 2 et 3. Les résultats obtenus sont illustrés dans la figure V.11. Dans les cellules non stressées nous détectons une production majoritaire de l'isoforme *tra2β*1 (+ Exon 3) et minoritaire (environ 10%) de l'isoforme *tra2β*3 (- Exon 3), alors que dans les cellules stressées avec 500 μ M d' H_2O_2 des quantités équivalentes des isoformes *tra2β*1 et *tra2β*3 sont détectées. Cette augmentation de l'isoforme *tra2β*3 est détectée durant les 16 premières heures après le début du stress oxydant, mais n'est plus visible au temps 24 h. Il est à noter que nous n'avons pas réussi à détecter l'isoforme *tra2β*4 qui présente l'inclusion de l'exon 2. L'inclusion de l'exon 2 au sein

de l'ARNm conduit à la dégradation de l'ARN par le système NMD. Par conséquent, cette isoforme n'est pas stable dans le temps et la méthode de détection par RT-PCR ne permet pas de détecter son augmentation. Dans les études qui avaient permis de détecter son augmentation au cours d'un stress oxydant les auteurs avaient utilisé la technique de Northern blot et à partir de 20 µg d'ARN total, ils avaient comparé les signaux obtenus en utilisant une sonde correspondant à l'ORF entière ou une sonde spécifique de l'exon 2. Cette approche, bien que nécessitant une quantité d'ARN 20 fois supérieure à celle que nous utilisons pour les expériences de RT-PCR semi-quantitative, semble plus adaptée pour la détection des ARNm contenant l'exon 2 (Takeo et al., 2009). La perte de l'exon 3 a été décrite dans la littérature et entraîne l'introduction d'un codon stop prématuré par décalage du cadre de lecture. Cependant, il existe un codon d'initiation de la traduction au sein de l'exon 4, dont l'utilisation est dans ce cas favorisée et entraîne la production d'une protéine qui ne possède pas le premier domaine RS de la protéine Tra2β. Bien que cette isoforme soit peu exprimée dans les cellules HeLa, il a été observé qu'elle peut, dans certains types cellulaires tels que les cellules nerveuses et le foie, être autant exprimée que l'isoforme tra2β1 (Nayler et al., 1998). A l'heure actuelle, la fonction de la protéine traduite à partir de l'isoforme tra2β3 n'est pas connue, mais il est raisonnable de penser que l'absence du 1er domaine RS modifie l'affinité et la reconnaissance de la protéine vis-à-vis de ses cibles. Les conséquences fonctionnelles de cette modulation d'épissage alternatif, lors d'un stress oxydant reste à être déterminées expérimentalement. Néanmoins, une expérience de siRNA dirigée contre l'exon 3 de l'ARNm codant tra2β a été réalisée et une modulation subséquente de l'épissage alternatif d'un ensemble d'exon du pré-ARNm codant la protéine CD44 a pu être mise en évidence, modulation qui est également détectée suite à un stress oxydant a été observée (Takeo et al., 2009).

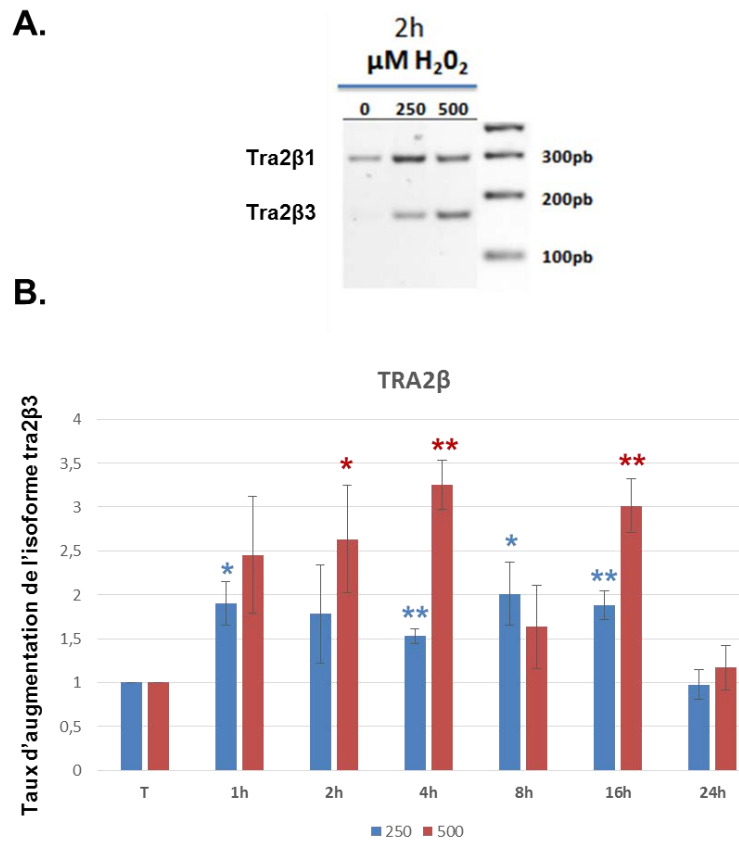


Figure V.11 : Analyse du taux d'augmentation de l'exclusion de l'exon 3 au sein des ARNm tra2β au cours de la réponse à un stress oxydant. A. Analyse par RT-PCR des quantités relatives des isoformes tra2β1 et 3 2 h après l'induction d'un stress oxydant par ajout de 250 ou 500 μM d'H₂O₂. Les amplifications ont été réalisées en utilisant des oligonucléotides spécifiques s'hybridant au sein des exons 1 et 4. Les amplicons ont été séparés sur un gel d'agarose à bas point de fusion à 2 % et identifiés à partir de leur taille. **B.** Quantification par RT-PCR du taux d'augmentation de l'isoforme tra2β3 différents temps (1 h, 2 h, 4 h, 8 h, 16 h, et 24 h) suite à l'induction d'un stress oxydant par ajout de 250 μM d'H₂O₂ (bleu) ou 500 μM d'H₂O₂ (rouge). Le contrôle (T) a été réalisé pour chaque temps sans ajout d'H₂O₂. Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé (n=3 ; ** : P<0.01 ; * : P<0.05)

V.4.2. Validation de la modulation de l'épissage de l'exon 22 du pré-ARNm codant la protéine SERCA2, un transporteur d'ions Ca²⁺

De nombreux événements différentiels d'épissage détectés par les puces sont associés à des ARNm codant des transporteurs d'ions. Parmi ces transporteurs, la protéine SERCA2 est une pompe calcique du réticulum endoplasmique chargée du maintien du taux de calcium intracytoplasmique. Le Ca²⁺ joue un rôle essentiel pour la contraction cardiaque, mais aussi comme second messenger dans la transduction du signal à travers la protéine kinase C et permet l'activation du facteur NF-KB (Giorgi et al., 2010; Zhou et al., 2006b). Des études récentes sur les cellules dendritiques suggèrent même qu'une diminution de l'expression de

SERCA2 pourrait favoriser les capacités migratoires de ces cellules en raison de l'augmentation du niveau de calcium dans le cytoplasme (Choi et al., 2015). A l'inverse, une augmentation de l'expression de SERCA2 a pu être corrélée au développement du cancer colorectal et à un stress du réticulum endoplasmique (Caspersen et al., 2000; Fan et al., 2014). Par conséquent des modulations de l'expression et de l'épissage des ARNm codant cette protéine, pourraient jouer un rôle important dans l'adaptation des cellules à un stress oxydant. Le gène SERCA2 (ATP2A2) est composé de 22 exons, dont plusieurs étaient décrits comme alternatifs, et 4 isoformes différentes, nommées Serca2a à Serca2d, produites par épissage alternatif et pouvant présenter une localisation tissu-spécifique, avaient été décrites (Figure V.12). L'épissage alternatif du pré-ARNm SERCA2 se distingue par des modulations d'épissage d'exons non traduits dont l'inclusion modifie la séquence de la région 3'-UTR de l'ARNm. C'est le cas, en particulier de l'exon 22, dont une augmentation de l'exclusion était détectée sur nos puces principalement 24 h après l'induction du stress oxydant. Cet exon étant terminal, les validations ont été effectuées par RT-qPCR en utilisant un couple d'oligonucléotides spécifique de l'exon 22 et un second couple spécifique de l'exon 19. Ce dernier exon, présent au sein de toutes les isoformes d'ARNm SERCA2, a été utilisé comme référence interne. Les analyses par RT-qPCR confirment une diminution significative de l'inclusion de l'exon 22 d'environ 65 %, 24 h après le début du stress oxydant (Figure V.13). Cette diminution, bien que retrouvée également aux temps précoces 1 h, 2 h et 4 h suite à un stress oxydant généré par 500 μ M d' H_2O_2 est cependant plus marquée à 24 h (Figure V.13).

Cette variation du taux d'inclusion de l'exon 22 modifie la région 3'-UTR de l'ARNm SERCA2 et notamment conduit à la perte de 3 des 5 motifs AU-rich présents au sein de cette région. Ceci pourrait affecter la fixation de la protéine AUF1 connue pour se fixer sur cette région et altérer la régulation de la stabilité de l'ARNm SERCA2 par AUF1 (Misquitta et al., 2005; Blum et al., 2005). Etant donné qu'une augmentation de l'expression de SERCA2 a pu être corrélée au développement du cancer colorectal et que l'expression de SERCA2 a été proposée comme marqueur pronostique potentiel de ce cancer, cette variation d'épissage pourrait impacter la progression tumorale et il sera important de tester si la variation observée dans les cellules Hela est observée avec des lignées cellulaires de cancer colorectal (Chung et al., 2006 ; Huang et al., 2012).

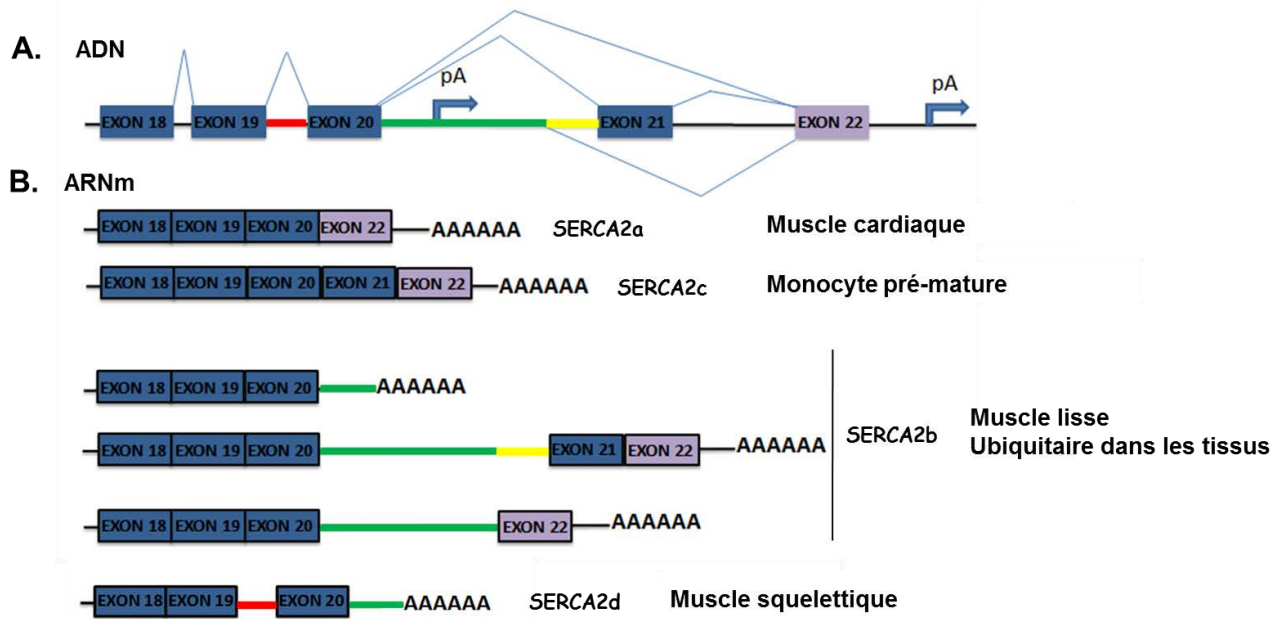


Figure V.12: Représentation schématique des séquences exoniques et introniques alternatives de la séquence 3'-UTR du gène SERCA2 et des différentes isoformes connues de l'ARNm SERCA2 (adaptée de Kimura et al., 2005). A. Représentation des séquences exoniques et introniques alternatives de la séquence 3'-UTR du gène SERCA2. Les exons sont représentés par des rectangles bleus et l'exon 22 est représenté par un rectangle violet. Les traits rouge, vert et jaune représentent les séquences introniques qui peuvent être retenues au sein de l'ARNm. Les sites de polyadénylation sont indiqués par une flèche bleue. B. Schéma des différentes isoformes connues de l'ARNm SERCA2. Pour chaque isoforme est indiqué le type cellulaire où cette isoforme est préférentiellement exprimée. On notera que l'isoforme SERCA2b regroupe en réalité 3 isoformes différentes.

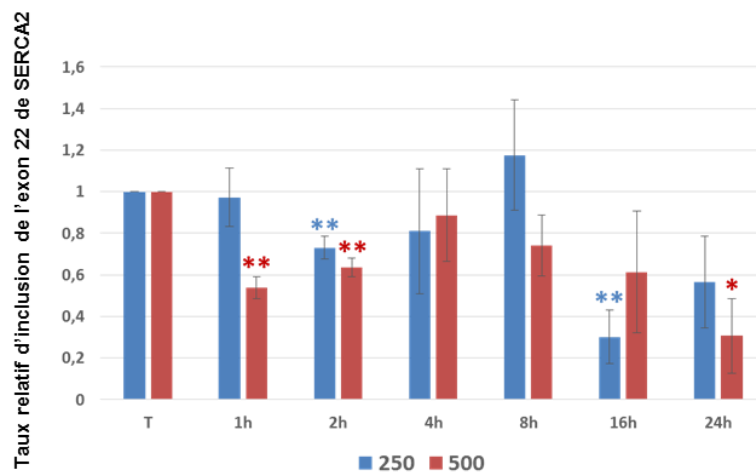


Figure V.13 : Analyse du taux d'inclusion de l'exon 22 au sein de l'ARNm SERCA2 au cours de la réponse à un stress oxydant. L'analyse par RT-qPCR a été réalisée en utilisant des oligonucléotides spécifiques de l'exon 22 et le signal obtenu a été rapporté au signal obtenu en utilisant des oligonucléotides spécifiques de l'exon 19 constitutif et prédit invariant. La quantification par RT-PCR du taux d'inclusion de l'exon 22 a été réalisée aux temps 1 h, 2 h, 4 h, 8 h, 16 h, 24 h suite à l'induction d'un stress oxydant par ajout de 250 µM H₂O₂ (bleu) ou 500 µM H₂O₂ (rouge). Le contrôle (T) a été réalisé pour chaque temps sans ajout H₂O₂. Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé (n=3 ; ** : P<0.01 ; * : P<0.05).

V.4.3. Validation de la modulation de l'épissage des exons 7 et 10 du pré-ARNm codant la protéine ECM2.

Les analyses bio-informatiques des données des puces ont montré un impact du stress oxydant sur l'épissage alternatif d'un nombre significatif de pré-ARNm codant des protéines impliquées dans l'adhésion et la migration des cellules. Parmi ces variations, les variations des taux d'inclusion des exons 7 et 10 du pré-ARNm codant la protéine ECM2 sont les plus importantes détectées au temps 24 h. La protéine ECM2 est une protéine de la matrice extracellulaire qui promeut l'adhésion des cellules et qui régule négativement la croissance et la prolifération cellulaires (Claeskens et al., 2000). Son expression est significativement diminuée dans de nombreux cancers et l'absence d'expression intratumorale de ce gène est associée à la récurrence du cancer de la prostate (Hurley et al., 2012, Hurley et al., 2015; Xiang et al., 2013). C'est pourquoi nous avons cherché à confirmer l'augmentation de l'inclusion des exons 7 et 10 par RT-qPCR. L'exon 5, constitutif et prédit invariant, a été utilisé comme référence interne. Les résultats nous ont permis de confirmer l'augmentation de l'inclusion de ces deux exons, avec une augmentation de 1,5x de l'inclusion de l'exon 7 et de 6x de celle de l'exon 10, 24 h après le début du stress oxydant (Figure V.14).

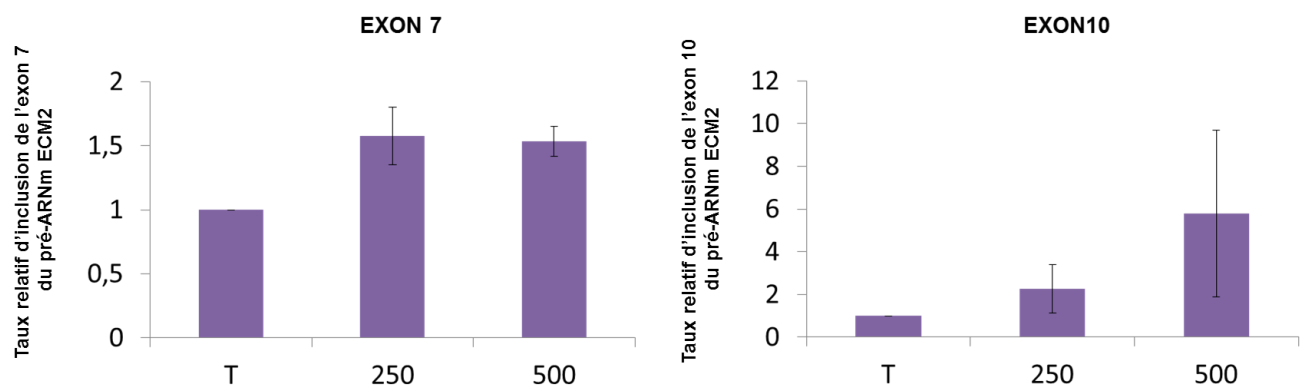


Figure V.14: Analyse du taux d'inclusion des exons 7 et 10 au sein de l'ARNm ECM2 24 h après l'induction d'un stress oxydant. L'analyse par RT-qPCR a été réalisée en utilisant des oligonucléotides spécifiques des exons 7 et 10 et le signal obtenu a été rapporté au signal obtenu en utilisant des oligonucléotides spécifiques de l'exon 5 constitutif et prédit invariant. La quantification par RT-PCR du taux d'inclusion de l'exon 22 a été réalisée au temps 24 h suite à l'induction d'un stress oxydant par ajout de 250 ou 500 μM d' H_2O_2 . Le contrôle (T) a été réalisé sans ajout d' H_2O_2 . Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé (n=3 ; ** : $P < 0.01$; * : $P < 0.05$).

D'après les banques de données transcriptomiques, il existe une isoforme d'ARNm d'ECM2 produite par modulation de l'épissage de l'exon 10, mais dont la fonction n'a pas été étudiée. L'exon 7, quant à lui, n'est pas décrit pour être un exon alternatif et notre observation d'une inclusion différentielle de cet exon au cours de la réponse à un stress oxydant pourrait constituer la 1ère description d'un épissage alternatif de cet exon. Par conséquent, il pourrait être intéressant d'étudier les fonctions de ces différents exons et comprendre leur implication dans la mobilité cellulaire.

V.4.4. Validation de la modulation de l'épissage de l'exon 1 du pré-ARNm codant la protéine FBXW10.

FBXW10 (F-box and WD Repeat Containing 10) est une protéine de 120 kDa, contenant un motif F-box de 40 aa, impliquée dans l'activité de l'E3 ubiquitine ligase. En association avec les protéines SKP1 et cullin, elle forme le complexe SCF qui joue un rôle essentiel de la spécificité de reconnaissance des substrats de l'E3 ubiquitine ligase lors du processus de dégradation des protéines. Il existe peu d'études concernant la fonction de la protéine FBXW10. Cependant, il a été montré que son activation peut entraîner une altération des protéines lamines et de l'enveloppe nucléaire et favoriserait la dégradation de protéines spécifiques de l'hétérochromatine (Chaturvedi and Parnaik, 2010). En lien avec notre étude, les données de puces ont révélées que le gène le plus exprimé au temps 4 h était FBXW10. De plus, l'une des plus fortes modulations d'épissage détectées par les puces au temps 24 h correspond à la diminution de l'inclusion de l'exon 1 du pré-ARNm codant FBXW10. Une diminution de près de 60 % du taux d'inclusion de l'exon 1 de FBXW10 a pu être expérimentalement confirmée par RT-qPCR au temps 24 h et ceci que le stress soit induit avec 250 μ M ou 500 μ M d' H_2O_2 (Figure V.15) Il est à noter qu'il existe une isoforme alternative de FBXW10 comportant un 1^{er} exon alternatif mais la fonction de cette isoforme n'a pas été étudiée.

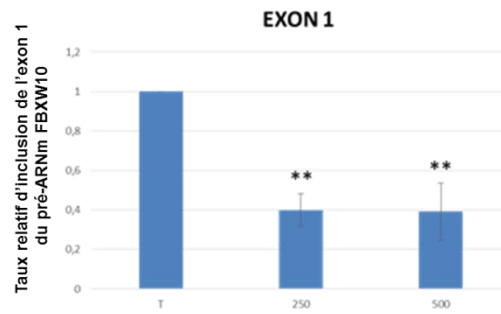


Figure V.15: Analyse du taux d'inclusion de l'exon 1 au sein de l'ARNm FBXW10 24 h après l'induction d'un stress oxydant. L'analyse par RT-qPCR a été réalisée en utilisant des oligonucléotides spécifiques de l'exon 1 et le signal obtenu a été rapporté au signal obtenu en utilisant des oligonucléotides spécifiques de l'exon 14 constitutif et prédit invariant. La quantification par RT-PCR du taux d'inclusion de l'exon 1 a été réalisée au temps 24 h suite à l'induction d'un stress oxydant par ajout de 250 ou 500 μM d' H_2O_2 . Le contrôle (T) a été réalisé sans ajout d' H_2O_2 . Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé (n=3 ; ** : $P < 0.01$; * : $P < 0.05$).

V.4.5. Validation de la modulation de l'épissage de l'exon 3 du pré-ARNm codant la protéine NR4A3.

La famille de protéine NR4A (pour nuclear receptor subfamily 4) est composée de trois protéines, NR4A1, NR4A2 et NR4A3 (appelée aussi Nor-1). Ces protéines sont répertoriées comme récepteurs orphelins dans la mesure où leurs ligands ne sont pas encore connus. Elles sont impliquées dans de nombreux processus biologiques en réponse à différents stimuli, en particulier lors de la réponse précoce à différents stress (Safe et al., 2015). Plusieurs études ont mis en évidence un rôle de ces protéines dans les cancers notamment en tant que suppresseurs de tumeurs, mais aussi dans l'angiogenèse, la prolifération cellulaire et dans la régulation de l'homéostasie cellulaire. Bien que la plupart des fonctions associées aux protéines NR4A soient redondantes, ces protéines ont également des fonctions et des partenaires spécifiques, comme NRA43 qui est impliqué plus spécifiquement dans le métabolisme oxydatif des muscles squelettiques et qui permet l'induction d'AMPC dans les hépatocytes et dans le foie (Beard et al., 2015; Safe et al., 2015). Les résultats des puces mettent en évidence la surexpression de ces 3 protéines, mais aussi la modulation de l'épissage alternatif de l'exon 3 du pré-ARNm codant NR4A3 dans les temps précoces de réponse à un stress oxydant, en particulier au temps 4 h. Le gène codant la protéine NR4A3 est composé de 9 exons qui permettent de générer 3 isoformes connues (Figure V.16), NR4A3 α et NR4A3 β qui sont les isoformes majoritaires, et NR4A3 α +exon3 qui est la seule isoforme incluant l'exon 3.

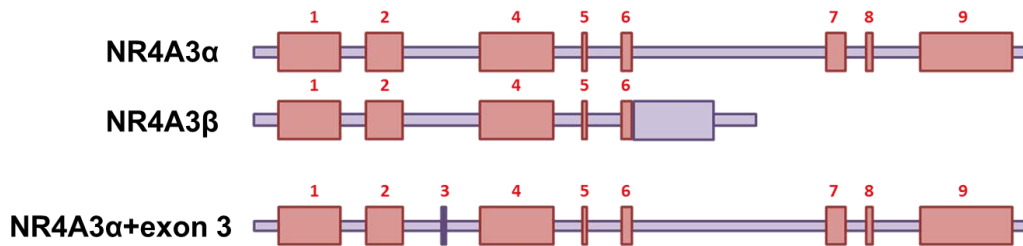


Figure V.16 : Représentation schématique des isoformes NR4A3α, NR4A3β et NR4A3α+exon3 produites à partir du gène NR4A3. Les rectangles roses représentent les exons et l'exon 3 alternatif est représenté par un rectangle gris, le rectangle violet représente la séquence intronique retenue au sein de l'ARNm codant NR4A3β.

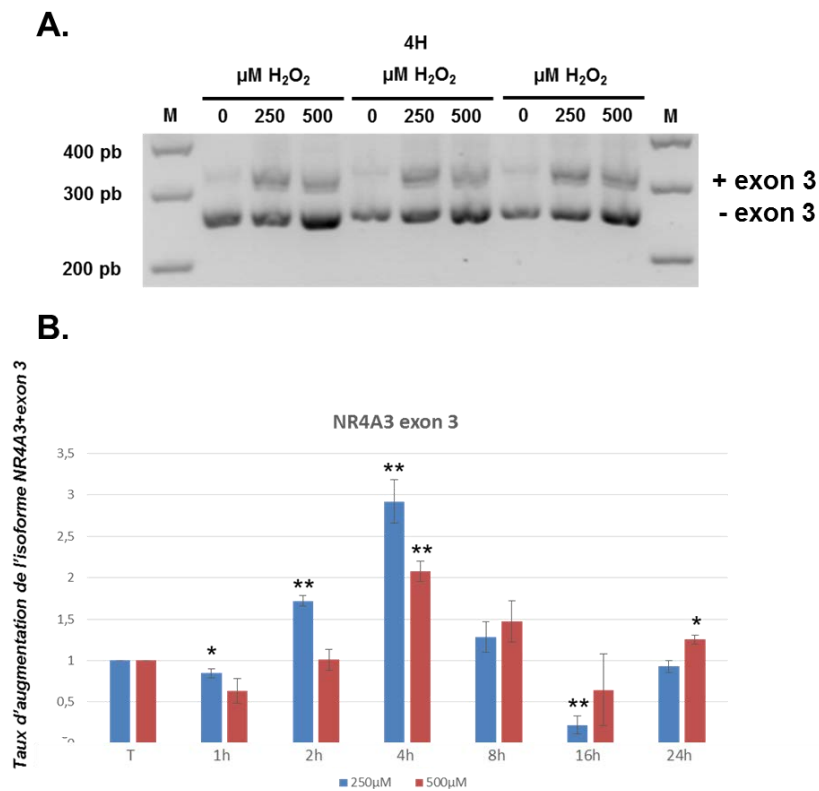


Figure V.17: Analyse du taux d'augmentation de l'exclusion de l'exon 3 au sein de l'ARNm NR4A3 au cours de la réponse à un stress oxydant. **A.** Analyse par RT-PCR des quantités relatives des différentes isoformes NR4A3 4 h après l'induction d'un stress oxydant par ajout de 250 ou 500 μM d'H₂O₂. Les amplifications ont été réalisées à partir de 3 séries biologiques indépendantes en utilisant des oligonucléotides spécifiques s'hybridant au sein des exons 1 et 4. Les amplicons ont été séparés sur un gel d'agarose à bas point de fusion à 2 % et identifiés à partir de leur taille. **B.** Quantification du taux d'augmentation de l'isoforme NR4A3α+exon 3 à différents temps (1 h, 2 h, 4 h, 8 h, 16 h, et 24 h) suite à l'induction d'un stress oxydant par ajout de 250 μM d'H₂O₂ (bleu) ou 500 μM d'H₂O₂ (rouge). Le contrôle (T) a été réalisé pour chaque temps sans ajout d'H₂O₂. Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé (n=3 ; ** : P<0.01 ; * : P<0.05).

Nous avons cherché à valider l'augmentation de l'inclusion de l'exon 3 au cours de la réponse précoce au stress oxydant par RT-PCR semi-quantitative en utilisant des oligonucléotides s'hybridant au sein des exons 1 et 4. Les résultats obtenus confirment une augmentation significative de l'isoforme NR4A3 α +exon3 d'un facteur 2 en moyenne aux temps 2 h et 4 h, ainsi qu'une plus faible augmentation au temps 24 h (Figure V.17). De façon intéressante le codon d'initiation de la traduction est normalement présent au tout début de l'exon 4 et l'inclusion de l'exon 3 entraîne l'utilisation d'un site d'initiation de la traduction alternatif, en phase avec le site d'initiation normalement utilisé, et génère une protéine NR4A3 possédant 11 acides aminés supplémentaires en N-terminal. Il est à noter qu'une étude récente a mis en évidence qu'une inhibition des HDAC entraînait une activation rapide de la transcription de NR4A3 en favorisant le recrutement d'un régulateur transcriptionnel, la protéine CREB phosphorylée et que cette activation s'accompagne d'une augmentation de la stabilité de la protéine NR4A3 (Zhao et al., 2014). En lien avec cette observation, nos données montrent que le stress oxydant impacte significativement l'expression de nombreuses histones, mais aussi des HDAC notamment aux temps précoces de réponse et la protéine CREB est également prédite par Ingenuity être activée lors de la phase précoce de réponse. On notera que parmi les acides aminés supplémentaires introduits du fait de l'inclusion de l'exon 3, on retrouve un résidu serine qui peut être phosphorylé par des kinases. Il est probable qu'une telle modification post-traductionnelle puisse influencer la stabilité de la protéine. En effet, il a été montré que NR4A3 peut être phosphorylée par la voie des MAPKs et de PI3K-AKT (Masuyama et al., 2001) pour le maintien de l'homéostasie cellulaire, mais l'impact de ces phosphorylations n'a été étudié et il serait intéressant de vérifier expérimentalement si l'inclusion de l'exon 3 favorise la stabilité de la protéine en augmentant son degré de phosphorylation.

V.5. ETUDE DES MECANISMES CELLULAIRE A L'ORIGINE DES VARIATIONS D'ÉPISSAGE OBSERVEES APRES UN STRESS OXYDANT.

Le 3^{ème} objectif de mon projet était d'étudier les mécanismes cellulaires à l'origine des variations d'épissage observées après un stress oxydant au travers, en particulier, de l'étude de l'impact du stress oxydant sur l'expression et la localisation des composants du spliceosome et des facteurs de régulation de l'épissage. Comme mentionné dans l'introduction, le stress oxydant peut affecter l'activité du complexe SMN, complexe qui joue un rôle essentiel dans la biogenèse des UsnRNA spliceosomaux transcrits par l'ARN polymérase II. Ainsi, l'inactivation du complexe SMN, bien que n'inhibant ni la transcription, ni l'épissage, affecte de façon tissu dépendante la quantité de certains UsnRNA et peut conduire à la modulation de l'épissage alternatif de nombreux pré-ARNm. C'est pourquoi nous avons entrepris d'analyser les effets du stress oxydant sur la quantité des UsnRNA spliceosomaux et sur l'expression et la localisation de leurs partenaires protéiques spécifiques. Pour cela j'ai dosé par RT-qPCR la quantité des UsnRNA spliceosomaux au cours de la réponse au stress et j'estimé par Western blot la quantité de certaines protéines cœur des spliceosomes majeurs et mineurs, ainsi que des protéines auxiliaires qui modulent l'activité de ces spliceosomes.

V.5.1. Etudes de la variation de la quantité de snRNA spliceosomaux au cours de la réponse à un stress oxydant.

Pour réaliser cette étude nous avons extrait les ARN totaux de cellules non stressées et de cellules stressées par ajout de 250 ou 500 μM d' H_2O_2 à différentes temps après l'induction du stress oxydant. Les ARN totaux ont été ensuite rétrotranscrits à l'aide d'un oligonucléotide hexamérique dégénéré. Les ADNc obtenus ont été ensuite amplifiés par qPCR en présence de couple d'oligonucléotides spécifiques de chacun des UsnRNA des spliceosomes majeur et mineur. Pour chaque couple d'oligonucléotides l'efficacité et la spécificité des couples ont été déterminées et vérifiées. Seul le couple d'oligonucléotide permettant la détection spécifique du snRNA U5 ne satisfaisait pas les critères de spécificité et l'étude concernant ce snRNA n'a pas été réalisée. Afin de normaliser les données de qPCR, nous avons utilisé deux couples d'oligonucléotides permettant la détection spécifique de l'ARNr 5S et de l'ARNm GAPDH. La quantité des différents UsnRNA a été déterminée au

cours de la réponse au stress oxydant en utilisant systématiquement au minimum trois essais biologiques aux différents temps testés.

Les résultats obtenus ont été rapportés sur chacun des deux rapporteurs, puis moyennés. Les valeurs déterminées à partir de chacun des rapporteurs ne présentaient pas de différences significatives. Les résultats sont présentés dans la figure V.18 et représentent les valeurs obtenues en utilisant la GAPDH comme rapporteur. Concernant les UsnRNA associés au spliceosome majeur, on observe une augmentation du taux du snRNA U1 de 1,4x au temps 2 h de manière significative, ainsi que du snRNA U4 suite à un stress oxydant à une dose de 500 μM H_2O_2 . Les autres UsnRNA associé au spliceosome majeur semblent être invariant dans les temps précoces de réponse au stress oxydant. Durant la phase tardive de réponse au stress, on observe que les UsnRNA du spliceosome majeur sont globalement surexprimés aux temps 16 h et qu'au temps 24 h, seul le snRNA U6 diminue de manière significative. Bien que les résultats obtenus au temps 16 h ne répondent pas aux critères de significativité, il est à noter que pour chaque UsnRNA testé, une augmentation a été détectée dans toutes les séries biologiques testées. Les résultats obtenus pour les UsnRNA associés au spliceosome mineur montrent une plus forte augmentation des snRNA U4atac et U11 aux temps 2 h et 4 h que leurs homologues U1 et U4, ainsi qu'une très faible augmentation du snRNA U12, ce qui suggère un impact important du stress oxydant sur le spliceosome mineur. Tout comme pour les UsnRNA du spliceosome majeur, on observe une augmentation de l'ensemble des UsnRNA du spliceosome mineur au temps 16 h, mais avec une diminution significative du snRNA U12 au temps 24 h. Ces résultats mettent donc en évidence une augmentation de la quantité des snRNA U1, U4, U11, U4atac et U12 lors de la phase précoce de réponse au stress oxydant, ainsi qu'une augmentation générale des UsnRNA au début de la phase tardive (16 h), mais qui n'est pas retrouvée au temps 24 h.

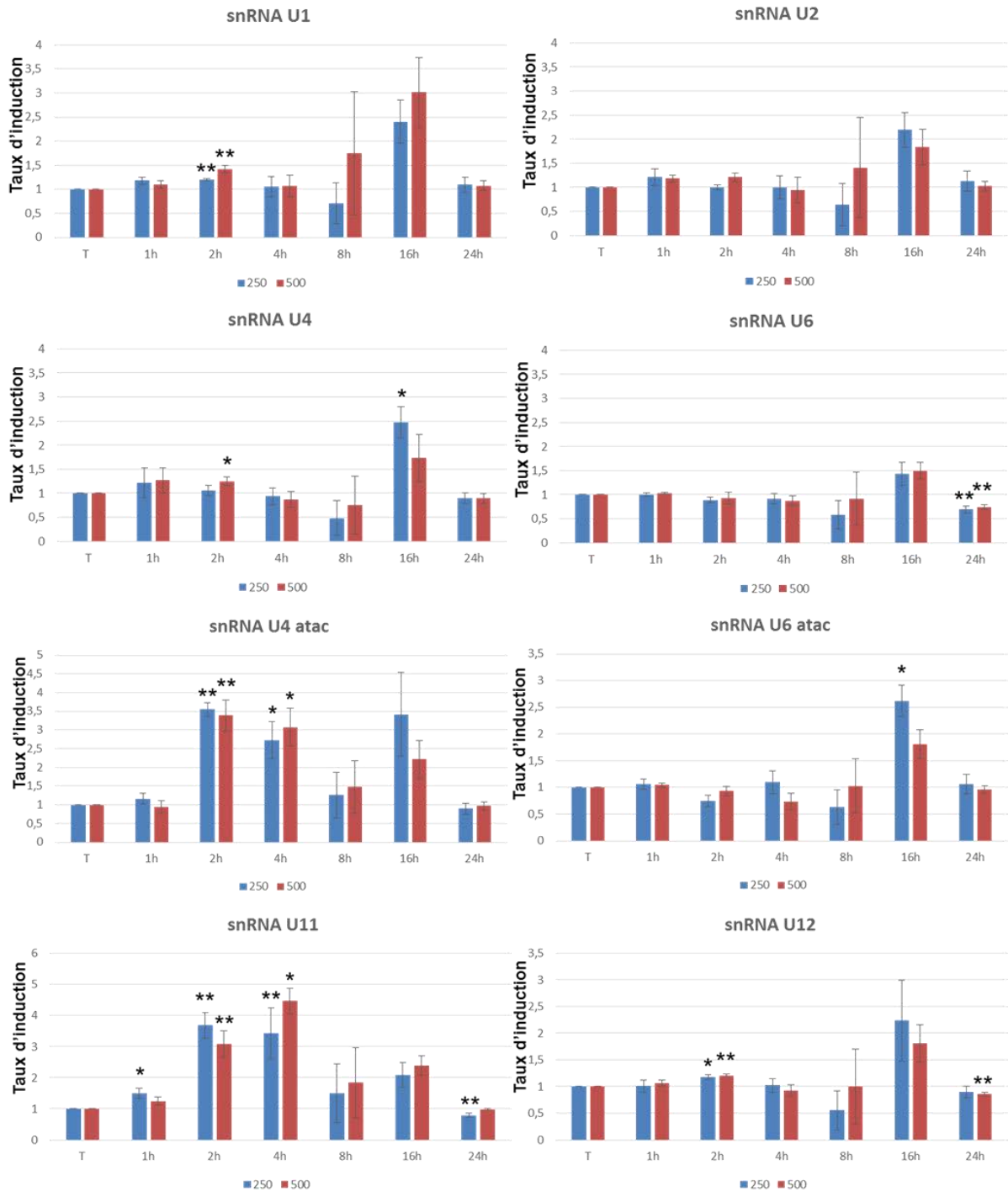


Figure V.18 : Analyse de la quantité des UsnRNA des spliceosomes majeur et mineur au cours de la réponse au stress oxydant. Les expériences ont été réalisées sur 3 séries biologiques pour chacun des stress oxydant. Les cellules ont été prélevées aux temps 1 h, 2 h, 4 h, 8 h, 16 h et 24 h suite à l'induction d'un stress oxydant par ajout de 250 μM H₂O₂ (bleu) ou 500 μM H₂O₂ (rouge). Pour chaque UsnRNA testé, le taux d'induction de l'ARN a été déterminé et un contrôle (T) a été réalisé pour chaque temps sans ajout H₂O₂. Les barres d'erreur représentent la valeur +/- de l'erreur standard et un test de student a été réalisé (n=3 ; ** : P<0.01 ; * : P<0.05).

Des études avaient déjà mis en évidence une augmentation de près de 2x du taux de snRNA U4atac chez des patients atteints d'un déficit en hormone de croissance lié à une mutation du gène *RNPC3* qui code une protéine essentielle pour la formation de la di-snRNP U11/U12 (Argente et al., 2014). Dans le cas de cette pathologie l'augmentation du taux de snRNA U4atac était associée à une réduction de la stabilité de la di-snRNP U11/U12 et, contrairement à ce que nous observons lors du stress oxydant, la quantité de snRNA U11 n'était pas affectée et une faible diminution de la quantité de snRNA U12 était observée. Ces modulations de la quantité des UsnRNA mineurs et de la stabilité de la di-snRNP U11/U12 avaient pour conséquences des défauts d'épissage d'environ 4 % des introns U12-dépendants se traduisant par des phénomènes de rétention d'introns (Argente et al., 2014). A notre connaissance aucune étude à ce jour n'a reporté une augmentation de la quantité de snRNA U11 dans le cadre d'une pathologie ou d'un stress cellulaire. C'est pourquoi, afin de mieux caractériser l'effet de l'augmentation de la quantité des snRNA U11 et U4atac, j'ai recherché si des modulations d'évènement d'épissage U12-dépendant pouvaient être observées suite à un stress oxydant.

V.5.2. Analyse d'évènement d'épissage U12-dépendant par RT-PCR.

Pour savoir si les variations de la quantité des snRNA U4atac et U11 détectées aux temps 2 h et 4 h pouvait entraîner un défaut d'épissage des introns U12 dépendants, nous avons recherché dans un premier temps l'ensemble des introns de type U12 dans la U12 intron DataBase (<http://genome.crg.es/cgi-bin/u12db/u12db.cgi>). Une des questions qui se posait était de savoir si l'augmentation de la quantité des snRNA U11 et U4atac était plutôt espérée activer ou inhiber l'épissage des introns U12-dépendant. En effet plusieurs cas de figures étaient envisageables : i) ces augmentations pouvaient induire une augmentation de la quantité de snRNP U11 et U4atac et augmenter l'efficacité de l'épissage, ou ii) ces augmentations pouvaient n'avoir aucun effet sur l'épissage des intron U12-dépendant du fait que les UsnRNA fonctionnent souvent sous forme de di- ou tri-snRNP composées de quantités stoechiométriques de chacune des UsnRNP et qu'une augmentation du taux de snRNA U4atac non associée à une augmentation du taux de snRNA U6atac n'allait pas forcément donner lieu à une augmentation de la quantité de di-snRNP U4atac/U6atac, de

même pour l'augmentation du taux de snRNA U11 et la formation de la disnRNP U11/U12, ou enfin iii) l'augmentation du taux des snRNA U4atac et U11 pouvaient conduire à l'agrégation dans la cellule de snRNP U12 et U6atac non fonctionnelles et au final inhiber l'épissage des introns U12-dépendant. Dans la mesure où les puces que nous avons utilisées ne permettaient pas d'analyser directement les phénomènes de rétention d'intron, nous avons fait l'hypothèse qu'un phénomène de rétention d'intron avait toutes les chances d'introduire un codon STOP prématuré et de conduire soit à la dégradation nucléaire de l'ARNm par l'exosome, soit à sa dégradation cytoplasmique par la machinerie du NMD. C'est pourquoi nous avons recherché si, parmi les 726 gènes contenant un ou plusieurs introns de type U12, certains voyaient leur expression diminuer aux temps 2 h et 4 h, temps auxquels la surexpression des snRNA U11 et U4atac est la plus forte. Nous avons ainsi listé 59 gènes négativement régulés aux temps 2 h et 4 h et présentant au sein de leurs pré-ARNm un ou plusieurs introns de type U12. Pour chacun des 59 gènes, nous avons analysés les séquences et la taille des introns afin de déterminer des cibles potentielles dont la rétention d'intron pouvait être observable par RT-PCR semi-quantitative. Les critères utilisés pour réaliser cette sélection étaient les suivants :1) une expression avérée dans des cellules HeLa, 2) la taille de l'intron d'intérêt devait être inférieure à 1kB afin de pouvoir détecter simultanément les produits d'amplifications contenant ou non l'intron U12 et 3) les exons bordant l'intron d'intérêt devait permettre l'élaboration d'un couple d'oligonucléotides efficace dont un spécifique d'une jonction exon-exon afin de détecter uniquement les ARN épissés. A l'issue de cette analyse, nous avons retenus 6 introns U12-dépendants qui respectaient nos critères (Tableau V.9).

Symbole du gène	Intron U12-dépendant
GBP1	Intron 4
MORC3	Intron 4
MORC3	Intron 7
MARCH6	Intron 13
SBNO1	Intron 10
RNF38	Intron 13

Tableau V.9 : Gènes dont l'expression diminue sur les puces, qui contiennent un ou plusieurs introns de type U12 et qui respectent les critères minima pour une analyse par RT-PCR.

L'analyse par RT-PCR des ARNm produits à partir de ces gènes aux temps de la phases précoce de réponse au stress oxydant n'a pas permis d'observer de phénomènes de rétentions d'introns à l'exception de l'intron 4 du pré-ARNm GBP1 au temps 2 h où l'on observe une augmentation de 2x de l'isoforme présentant l'intron au sein de l'ARNm (Figure V.19). Dans la mesure où la rétention d'un intron entraîne une dégradation rapide des ARN par l'exosome ou par la machinerie de NMD, il est difficile de pouvoir observer par RT-PCR la présence d'ARNm présentant une rétention d'intron. Pour cette raison, nous avons entrepris d'essayer de stabiliser ces espèces en bloquant leur dégradation en déplétant par emploi de siRNA soit la protéine Upf1, un composant essentiel de la machinerie de NMD, soit la protéine Rrp41, une sous-unité de l'exosome, dont la déplétion a été montrée ralentir la cinétique de dégradation des ARNm ayant retenus des introns de type U12 (Niemela et al., 2014). Les conditions de déplétion de ces protéines sont actuellement en cours de mises au point et par conséquent les résultats de ces expériences ne peuvent pas être présentés dans ce manuscrit.

On notera que nos données obtenues au temps 4 h permettent néanmoins de confirmer la diminution de l'expression de l'ensemble des gènes testés, avec des diminutions allant de 20 à 70 %. Il est proposé que le spliceosome mineur participe à la régulation de l'expression des gènes présentant des introns de type U12 en favorisant l'inclusion d'intron et la dégradation d'ARNm par l'exosome afin de limiter l'expression de certaines gènes en réponse à des stimuli extracellulaires. Des analyses complémentaires de la fonction de ces gènes devront être réalisées afin de voir si une diminution programmée de l'expression de ces gènes par le spliceosome mineur aurait un sens physiologique.

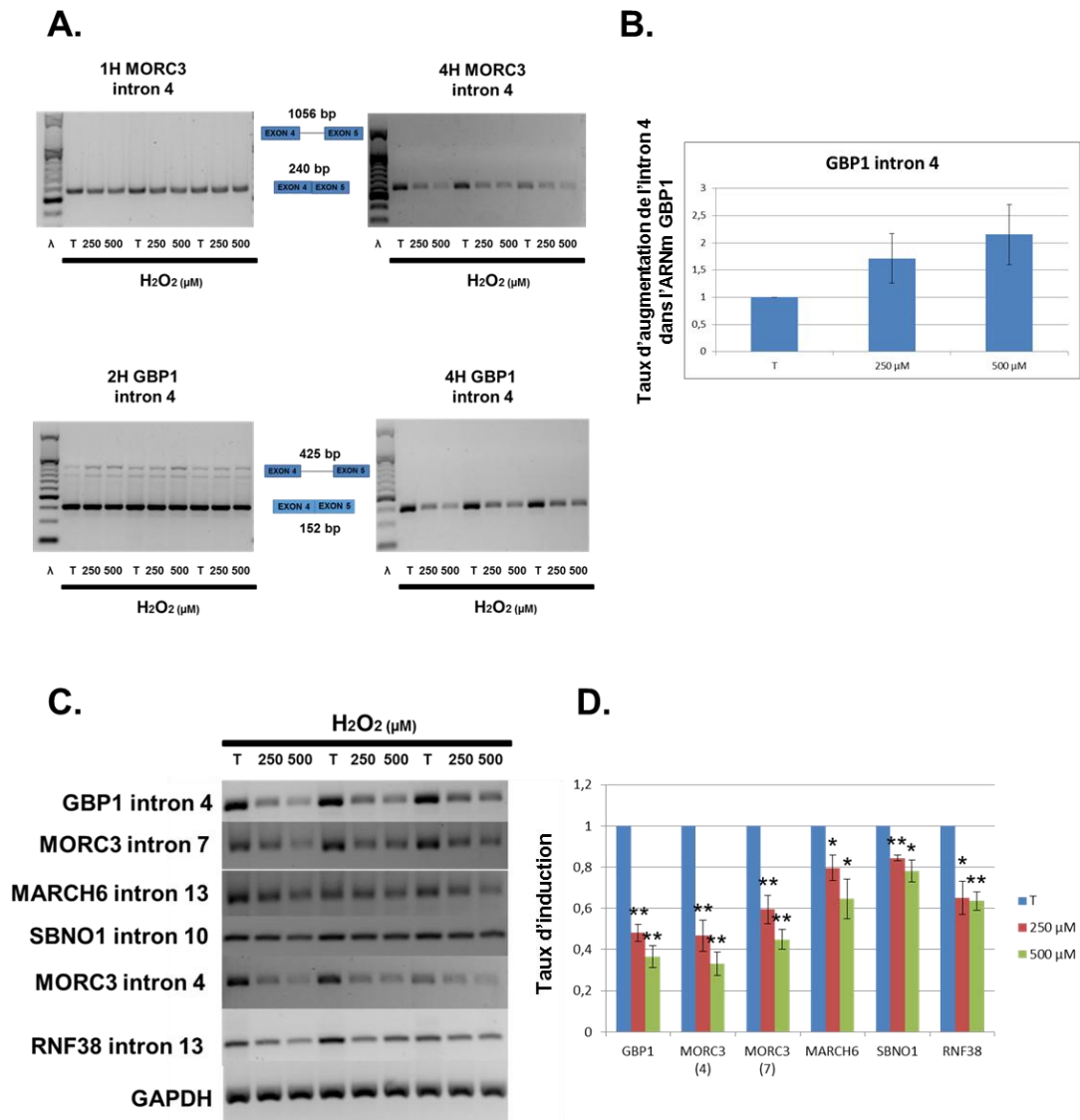


Figure V.19 : Analyse de la rétention d'introns U12 –dépendants au sein d'ARNm dont l'expression diminue lors de la phase précoce de la réponse au stress oxydant. **A.** Analyse par RT-PCR des différentes isoformes MORC3 aux temps 1h et 4h et de GBP1 aux temps 2h et 4h. Après traitement à différentes doses de peroxyde d'hydrogène, les ARN totaux ont été extraits, traités à la DNase, puis rétro-transcrits. Les ADNc ainsi générés ont servi de matrice pour une réaction de polymérisation en chaîne (PCR). Les amplicons ont été séparés sur un gel agarose à bas point de fusion à 2%. Trois séries biologiques distinctes sont présentées aux doses de 0, 250 et 500 μM et un marqueur de taille a été déposé (λ). **B.** Une quantification des produits d'amplification contenant ou non l'intron 4 GBP1 a été exprimée en taux d'augmentation de l'inclusion de l'intron. Les barres d'erreur représentent la valeur +/- de l'erreur standard. **C.** Analyse RT-PCR de la quantité des différents ARNm (GBP1, MORC3, MARCH6, SBNO1, RNF38) au temps 4 h. La quantité d'ARNm GAPDH a été analysée comme contrôle. **D.** Une quantification des produits d'amplification des différents ARNm (GBP1, MORC3, MARCH6, SBNO1, RNF38) a été réalisée en utilisant la quantité de GAPDH comme valeur rapportrice. Les barres d'erreur représentent la valeur +/- de l'erreur standard).

V.5.3. Etude de la relocalisation des UsnRNP et UsnRNA spliceosomaux dans le cytoplasme.

A l'heure actuelle, bien que la localisation nucléaire des UsnRNA du spliceosome majeur soit établie, il a été suggéré, bien que ces résultats sont contestés, que certains UsnRNA, associés au spliceosome mineur, puissent avoir également une localisation cytoplasmique (Konig et al., 2008). Par ailleurs, il a été mis en évidence chez la drosophile l'existence de foci cytoplasmiques, enrichis en UsnRNP et en facteur d'assemblage de ces UsnRNP dont le complexe SMN, qui fonctionneraient comme lieu de stockage des UsnRNP avant leur réimport dans le noyau. Ces foci ont été nommés U-bodies et semblent liés au P-bodies (Liu and Gall, 2007). Par conséquent, nous avons cherché à savoir s'il était possible d'observer une relocalisation des snRNA vers le cytoplasme suite à un stress oxydant, mais aussi si les protéines associées aux différentes UsnRNP étaient également soumises à une relocalisation cytoplasmique lors d'un stress oxydant, pouvant ainsi expliquer éventuellement les variations d'épissage observées au cours du stress oxydant. Afin de réaliser cette étude, nous avons entrepris de réaliser des expériences de fractionnement cytoplasme/noyau sur des cellules HeLa traitées ou non par de l' H_2O_2 aux doses de 250 μM et 500 μM pendant 1 h, 2 h et 4 h, temps auxquels nous retrouvons les plus fortes variations des taux de snRNA U4atac et U11. A l'issue des différents stress nous avons réalisé des fractionnements des compartiments cytoplasmique et nucléaire selon le protocole décrit dans la partie Matériel et Méthodes. Pour chaque compartiment nous avons réalisé une extraction des protéines afin de vérifier la qualité des fractionnements. Les protéines ont été fractionnées sur un gel SDS-PAGE 10% et transférées sur une membrane de nitrocellulose qui a été coloré au rouge ponceau. Le signal obtenu est présenté dans la figure V.20 et met en évidence une diminution significative, dans la fraction cytoplasmique, des protéines de bas poids moléculaire qui à l'inverse sont enrichies dans la fraction nucléaire. Afin de confirmer la qualité du fractionnement et l'absence de contamination de la fraction nucléaire par des protéines cytoplasmiques, nous avons également réalisé un western blot en utilisant un anticorps dirigé contre la protéine GAPDH qui est majoritairement cytoplasmique. Le résultat nous permet de valider que la fraction nucléaire n'est pas contaminée par des protéines du cytoplasme.

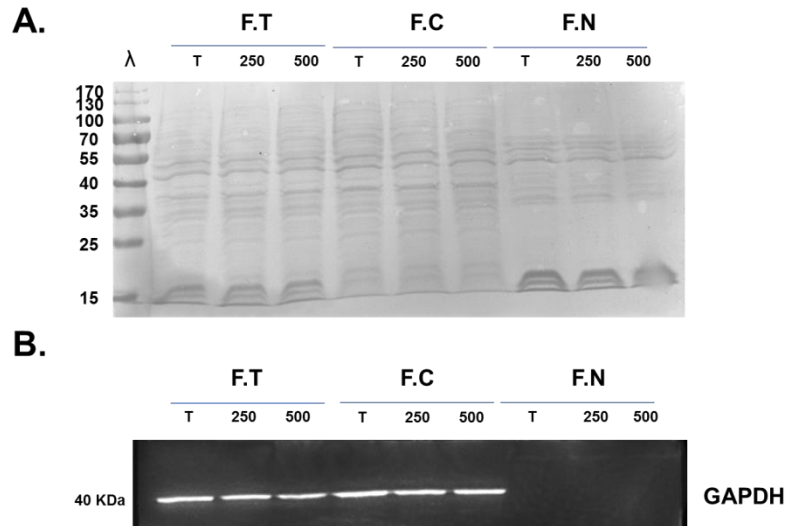


Figure V.20 : Vérification de la qualité du fractionnement des compartiments cytoplasmiques et nucléaires. Pour chaque fraction obtenue, les protéines ont été fractionnées dans un gel SDS-PAGE 10% puis transférées sur une membrane de nitrocellulose. **A.** Une coloration au rouge ponceau a été réalisée pour observer le profil général des protéines pour chaque échantillon déposé. Un marqueur de taille a été déposé et le poids moléculaire est indiqué à gauche en kDa. **B.** Western blot réalisé avec l'utilisation d'anticorps dirigés contre la protéine GAPDH.

En parallèle, nous avons également réalisé une extraction des ARN totaux afin d'analyser par RT-qPCR la quantité relative des U snRNA au sein de chaque compartiment. Cependant, nous n'avons pas été en mesure de pouvoir réaliser une quantification des snRNA à l'heure actuelle, dans la mesure où les oligodésoxynucléotides utilisés lors de l'analyse précédente sur des ARN totaux ne permettent pas d'obtenir des résultats analysables. En effet, l'efficacité des oligodésoxynucléotides dépend de la composition générale en ARN et en ADNc des échantillons. Par conséquent, il est possible que l'absence dans la fraction nucléaire des ARN ribosomiques entraîne une altération de l'efficacité des oligodésoxynucléotides qui ne permettent plus de réaliser une quantification dans des conditions optimisées. Néanmoins nous avons commencé l'analyse par Western blot de la quantité des protéines SNRNP40, U1A et Sart1 qui sont associées respectivement aux snRNP U5, U1 et à la tri-snRNP U4/U6.U5. Les premiers résultats obtenus sont présentés dans la figure V.21. L'analyse de ces données semble suggérer des variations d'expressions des protéines Sart1, U1A et SNRNP40 au cours de la réponse précoce au stress oxydant. De plus il semble apparaître des changements de localisation cellulaire des protéines SNRNP40 et U1A au temps 4 h. En effet, on observe que la protéine est dans les temps 1 h et 2 h majoritairement dans le noyau puis se relocalise dans le cytoplasme 4 h après le début du

stress oxydant. A l'inverse, on remarque que la quantité de protéine U1A diminue fortement dans le cytoplasme et augmente faiblement dans le noyau. Ces résultats restent à être confirmés par l'analyse des autres séries biologiques mais semblent néanmoins montrer que le stress oxydant a un impact sur l'expression et la localisation cellulaire de certaines protéines des UsnRNP.

Afin de compléter cette étude, nous avons également entrepris d'étudier l'effet du stress oxydant sur d'autres protéines associées au UsnRNP et en particulier des protéines spécifiques de la di-snRNP U11/U12 telles que la protéine ZCRB1 pour mieux comprendre l'impact d'une augmentation de la snRNA U11 sur l'épissage des introns U12-dépendant lors la réponse précoce au stress oxydant. Une analyse par Western blot de la quantité et de la localisation des facteurs d'épissage appartenant à la famille des protéines SR et hnRNP est également en cours de réalisation. Cette étude permettra d'évaluer si certaines modulations d'épissage observées au cours du stress oxydant peuvent être expliquées par des variations de la quantité de protéines SR et hnRNP ou par une altération de leur localisation cellulaire.

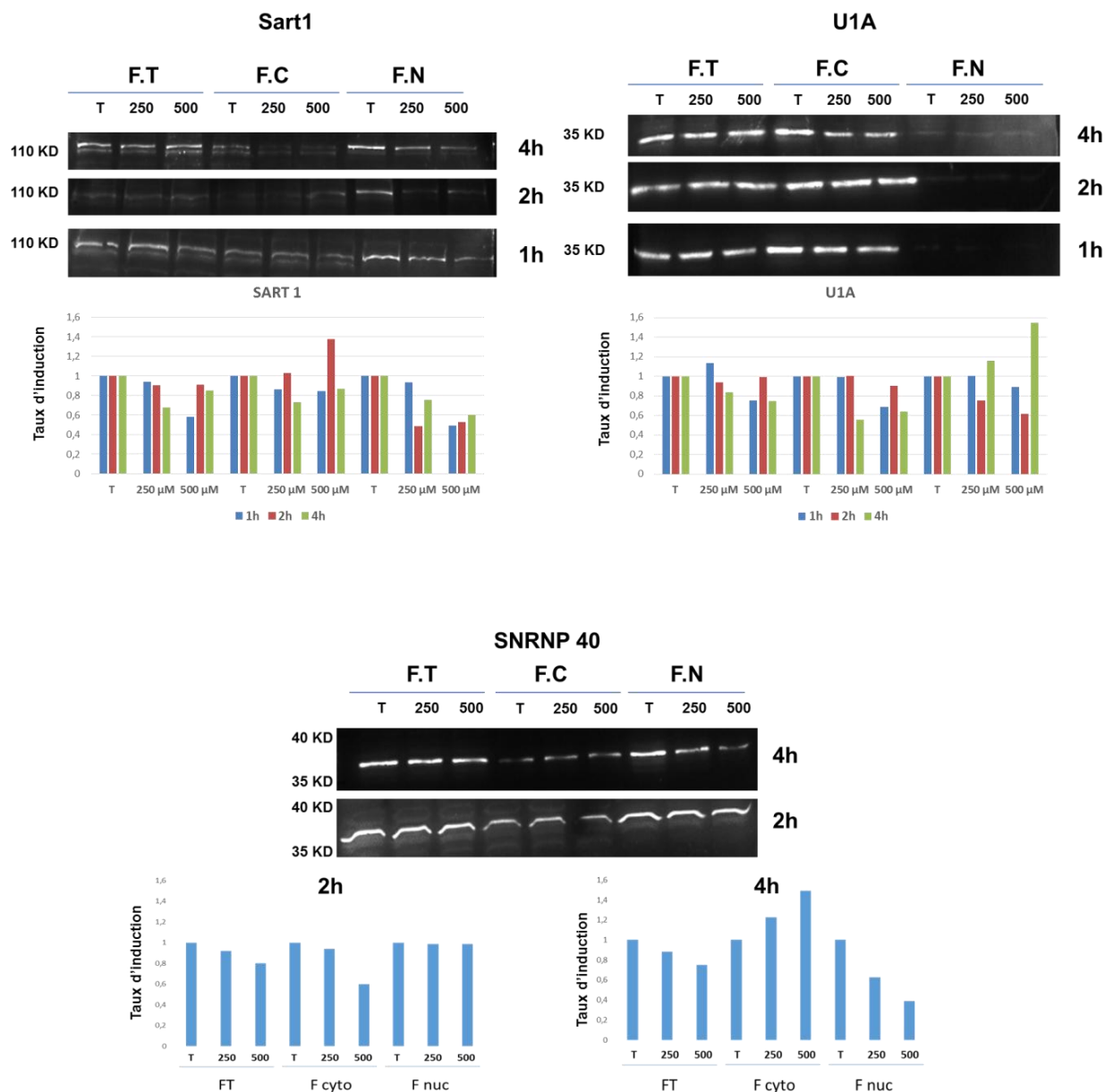


Figure V.21: Analyse préliminaire de l'effet du stress oxydant sur l'expression et la localisation des protéines SART1, U1A et SNRNP40 aux temps 1 h, 2 h et 4 h. Les différentes fractions totales, cytoplasmiques et nucléaires ont été fractionnées sur un gel SDS-PAGE 10% puis transférées sur une membrane de nitrocellulose. Des anticorps (Ac) dirigés spécifiquement contre les protéines SART1, U1A et SNRNP40 ont été utilisés. Une quantification des différents signaux obtenus a été réalisée en utilisant comme rapporteur le signal obtenu pour la GAPDH. Les signaux obtenus dans les fractions nucléaires ont été rapportés en utilisant les signaux obtenus lors de la coloration au rouge ponceau. Différentes doses de peroxyde d'hydrogène ont été testées : 250 μM et 500 μM. Un contrôle négatif a été réalisé en absence de stress oxydant (T).

V.6. CONCLUSION ET PERSPECTIVES.

Cette seconde partie de mon travail de thèse avait pour objectifs initiaux i) d'évaluer le degré de modulation de l'épissage alternatif lors de la réponse des cellules de vertébrés au stress oxydant, ii) de rechercher si les ARNm dont l'épissage était altéré au cours de la réponse au stress correspondaient à des fonctions cellulaires bien définies et iii) d'étudier les mécanismes cellulaires à l'origine des variations d'épissage observées après un stress oxydant au travers, en particulier, l'étude de l'impact du stress oxydant sur l'expression et la localisation des composants du spliceosome et des facteurs de régulation de l'épissage.

Afin de répondre à ces objectifs j'ai dû, dans un premier temps, établir des conditions de stress oxydant (choix de l'agent oxydant, concentration de cet agent, lot de cellules à utiliser, nombre de passages de ces cellules..) permettant d'induire un stress oxydant significatif, c'est-à-dire suffisamment fort pour pouvoir observer de façon reproductible une réponse cellulaire, mais modéré afin de ne pas décimer ma population de cellules et pouvoir observer sur une période de temps relativement courte (24 h) la récupération des cellules. Notre choix s'est porté sur le traitement de cellules HeLa S3 par 250 ou 500 μ M de peroxyde d'hydrogène.

J'ai ensuite entrepris d'étudier et de caractériser la réponse des cellules HeLa S3 à ce stress oxydant afin de choisir au mieux la cinétique de notre analyse transcriptomique. Cela m'a permis d'établir que la réponse des cellules HeLa au stress oxydant impliquait deux phases de réponse : une phase précoce (1h-4h) caractérisée par un fort taux de mortalité associé à une forte augmentation du taux d'entités oxygénées réactives (ROS) intracellulaires et une phase tardive (16h-24h) caractérisée par une augmentation de la viabilité cellulaire, une diminution du taux de ROS intracellulaires et une forte surexpression des ARN satellites III. Concernant l'augmentation de la viabilité cellulaire, la nature du test de viabilité effectué (test MTT, qui permet d'estimer le nombre de cellules vivantes et métaboliquement actives), ne permet pas de conclure quant à l'origine de cette augmentation, à savoir une reprise de l'activité métabolique ou de la prolifération cellulaire. Pour répondre à cette question des expériences complémentaires, basées sur l'analyse en temps réel de l'impédance des cellules par le système xCELLigence et permettant d'obtenir des informations quantitatives sur l'état physiologique de la cellule (adhésion, prolifération, viabilité, morphologie), sont

envisagées à court terme. Concernant la surexpression des ARN Sat III nos données montrent qu'elle est corrélée à la formation de granules de stress nucléaires contenant les protéines HSF1 et hnRNP HAP (SafB) et que la transcription des ARN Sat III est HSF1-dépendante. Nos données révèlent également la formation transitoire de foyers nucléaires contenant la protéine HSF1, mais ne contenant ni SafB, ni les ARN Sat III lors de la phase précoce de réponse au stress oxydant et dont la fonction et la contribution à l'adaptation des cellules HeLa au stress oxydant restent à déterminer.

Sur la base de ces premières observations, une analyse globale du transcriptome par emploi de puces à exons (Affymetrix) a été réalisée à partir d'ARN totaux isolés 1 h, 2 h, 4 h et 24 h après le début d'un stress oxydant généré par ajout de 500 μ M d' H_2O_2 . Nous avons ainsi identifié des modulations d'expression et d'épissage spécifiques de chacune des deux phases. Bien que la validation expérimentale des variations d'expression de gènes détectées sur les puces soit toujours en cours, une analyse préliminaire des données par des outils bio-informatiques a été réalisée. Cette analyse a permis de mettre en évidence des fonctions cellulaires bien définies qui semblent être affectées séquentiellement lors de la réponse au stress oxydant. Ainsi, lors de la phase précoce de réponse, qui est associée à une forte diminution de la viabilité cellulaire, ce sont les groupes fonctionnels correspondant à la régulation de l'apoptose et à la survie cellulaire qui sont les plus enrichis, suivis de près par des gènes relatifs à l'organisation du cytosquelette et à la régulation de la transcription. Alors qu'en phase tardive, les groupes fonctionnels les plus significativement enrichis sont impliqués dans la migration des cellules, la réponse au stress oxydant, la réponse à la vitamine A, mais également à l'angiogenèse. Les résultats des analyses par le logiciel Ingenuity suggèrent un rôle prépondérant du facteur de transcription Nupr1 dans la réponse des cellules HeLa au stress oxydant, ce qui est en accord avec l'observation d'une déplétion de Nupr1 qui entraîne une forte augmentation du taux de ROS intracellulaires (Weis et al., 2015). Des expériences de stress oxydant seront réalisées dans des cellules préalablement déplétées en Nupr1 par ARN interférence afin de confirmer la contribution de Nupr1 dans la réponse des cellules HeLa au stress oxydant et tester son éventuel impact sur les modulations d'épissage alternatif détectées lors du stress.

Concernant la modulation de l'épissage alternatif, nos données révèlent qu'elle concerne de nombreux ARNm et la validation expérimentale des événements les plus significatifs sera poursuivie afin d'une part de mieux appréhender l'importance fonctionnelle de ces modulations et d'autre part de constituer des listes plus grandes d'événements validés. Ces listes seront essentielles pour notre étude à plus long terme de l'effet du stress oxydant sur la machinerie d'épissage et pour la compréhension des mécanismes de régulation de l'épissage au cours de la réponse au stress oxydant. A ce propos, la contribution des ARN Sat III aux modulations d'épissage observées 24 h après le début du stress sera testée en déplaçant ces ARN par emploi de LNA gapmers. Les conditions de déplétion ont d'ores et déjà été mises au point au laboratoire par Christelle Aigueperse et permettent, tout comme la déplétion d'HSF1, de quasiment totalement abolir l'expression des ARN Sat III, ainsi que la formation des granules de stress nucléaires. Le transcriptome de cellules HeLa, n'exprimant plus les ARN Sat III ou le facteur de transcription HSF1, et soumis à un stress oxydant, sera analysé par l'approche puces à exons afin d'évaluer plus précisément la contribution des ARN Sat III et d'HSF1 dans les modulations de l'expression des gènes et d'épissage. L'effet de la déplétion des ARN Sat III a déjà été testé sur l'occurrence de deux événements d'épissage spécifiquement observés au temps 24 h et il semblerait que l'absence de ces ARN interfère avec ces modulations d'épissage. Ces données constitueraient la 1^{ère} preuve expérimentale de l'hypothèse émise depuis quelques années de l'implication des ARN Sat III dans la régulation de l'épissage alternatif.

L'analyse du type d'événements d'épissage altérés suite à un stress oxydant est également informative. Elle montre, en effet, que 52,4 % des événements d'épissage alternatifs détectés sur les puces changent la nature du premier ou des derniers exons des ARNm. Pour comparaison, notre analyse globale de l'effet d'un stress thermique (2h à 42°C suivies de 1 h à 37°C) réalisé sur les mêmes cellules HeLa, révèle que seuls 10 % des 420 événements détectés touchent des exons terminaux et que plus de 75 % des événements concernent des exons cassette (résultats non publiés). Il s'avère donc que, pour un même type cellulaire, la nature des événements d'épissage régulés diffère selon le type de stress subi par les cellules. En lien avec cette observation, moins de 4 % des variations d'épissage détectées suite à un stress oxydant (tous temps confondus) sont retrouvées lors d'un stress thermique

et cette fraction commune tombe à moins de 2 % si la comparaison est effectuée avec les données obtenues lors d'un stress génotoxique induit par la camptothécine (Dutertre et al., 2010). On notera cependant que dans ce dernier cas, le stress a été effectué sur des cellules MCF-7 et non de cellules HeLa, et que seules les données relatives aux exons dont l'inclusion diminue suite au stress, soit 354 exons, étaient disponibles et ont été prises en compte lors de notre analyse comparative. De façon surprenante cependant, on constate que 10 à 25% des gènes dont l'épissage est affecté suite à un stress oxydant voient leur épissage alternatif également affecté lors d'un stress thermique ou d'un stress génotoxique. Ces données suggèrent donc que la réponse cellulaire à ces 3 stress pourrait faire intervenir la régulation de l'épissage alternatif d'un socle commun de gènes, mais que cette régulation toucherait des exons différents d'un même pré-ARNm selon la nature du stress subi. Se pose alors la question de l'impact fonctionnel des variations d'épissage détectées et de façon plus large de la possibilité d'attribuer à chacun des exons d'un gène une fonction particulière. De nombreuses bases de données regroupant les motifs et domaines protéiques en fonction de leurs fonctions existent, mais, à notre connaissance, aucune base de données de ce type n'existe au niveau exonique. Or il pourrait être intéressant de comparer les séquences des exons spécifiquement régulés lors d'un stress donné afin de rechercher si les modulations observées pourraient affecter une fonction bien précise des protéines codées par les pré-ARNm dont l'épissage est modulé. L'équipe de Didier Auboeuf (LBMC-ENS, Lyon) travaille au développement d'une base de données d'ontologie des exons qui permettrait de prédire les conséquences fonctionnelles des changements d'épissage alternatif. Un contact a été pris avec cette équipe afin de rechercher si les exons dont l'inclusion varie suite à un stress oxydant peuvent être associées à des fonctions particulières.

Enfin, nos données mettent en évidence un effet marqué du stress oxydant sur l'expression de certains snRNA appartenant au spliceosome mineur lors de la phase précoce de réponse. La recherche expérimentale d'éventuels défauts d'épissage des introns mineurs a été débutée et est actuellement poursuivie, mais à l'heure actuelle nos données ne permettent pas de conclure sur l'effet d'un stress oxydant sur l'épissage des introns de type U12. C'est pourquoi, en vue de mieux appréhender l'origine des variations des snRNA U4atac et U11, ainsi que leur impact sur la fonctionnalité du spliceosome mineur, une collaboration avec l'équipe de Mikko Frilander (Université d'Helsinki), spécialiste international du spliceosome

mineur, a été établie. Cette équipe a développé des approches permettant d'analyser le taux de formation et la stabilité des différentes di- et tri-snRNP spliceosomales mineures (Argente et al., 2014 ; Niemelä et al., 2014). Une étude de la localisation cellulaire des snRNA mineurs au cours de la réponse au stress oxydant a également été entreprise afin de tester l'hypothèse d'une agrégation cytoplasmique de certaines UsnRNP mineures. Nous espérons au travers de ces expériences pouvoir apporter des informations importantes sur l'impact du stress oxydant sur la composition et la fonction du spliceosome mineur.

ANNEXES

ANNEXES

ANNEXE 1 : RESULTATS DE L'ANALYSE FONCTIONNELLE PAR LE LOGICIEL DAVID DES TERMES GO ASSOCIES AUX GENES DONT L'EXPRESSION EST MODIFIEE LORS D'UN STRESS OXYDANT AU TEMPS 4 H.

L'analyse a été réalisée en utilisant les gènes dont l'expression est modifiée 4 h après le début d'un stress oxydant. Pour chaque GO identifié, un taux d'enrichissement est calculé et des tests statistiques sont effectués pour évaluer le niveau de significativité de l'enrichissement du terme d'intérêt. A partir de ces résultats, le logiciel regroupe les GO biologiquement reliées entre eux en cluster et associe un score d'enrichissement. Ce score d'enrichissement est directement lié à la p-value calculée pour le groupe fonctionnel selon la formule suivante : $\text{score} = -\log(\text{p-value})$, par conséquent un score d'enrichissement ≥ 2 correspond à une p-value ≤ 0.01 et plus le score d'enrichissement est élevé, plus l'enrichissement de ces termes est significatif. Les groupes fonctionnels indiqués sont ceux qui ont un score d'enrichissement supérieur ou égal à 2.

Annotation Cluster 1 Score d'enrichissement : 5,53						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0043065~positive regulation of apoptosis	71	4.6	2,0E+00	5,5E-08	2,0E-04
terme GO BP	GO:0043068~positive regulation of programmed cell death	71	4.6	1,9E+00	7,3E-08	1,3E-04
terme GO BP	GO:0010942~positive regulation of cell death	71	4.6	1,9E+00	8,7E-08	1,1E-04
terme GO BP	GO:0042981~regulation of apoptosis	110	7.1	1,6E+00	3,1E-07	1,2E-04
terme GO BP	GO:0043067~regulation of programmed cell death	110	7.1	1,6E+00	5,1E-07	1,6E-04
terme GO BP	GO:0010941~regulation of cell death	110	7.1	1,6E+00	6,1E-07	1,5E-04
terme GO BP	GO:0006917~induction of apoptosis	53	3.4	2,0E+00	3,0E-06	5,9E-04
terme GO BP	GO:0012502~induction of programmed cell death	53	3.4	2,0E+00	3,3E-06	6,1E-04
terme GO BP	GO:0043066~negative regulation of apoptosis	52	3.4	1,7E+00	1,0E-04	1,2E-02
terme GO BP	GO:0043069~negative regulation of programmed cell death	52	3.4	1,7E+00	1,5E-04	1,6E-02
terme GO BP	GO:0060548~negative regulation of cell death	52	3.4	1,7E+00	1,6E-04	1,6E-02
terme GO BP	GO:0006916~anti-apoptosis	33	2.1	1,9E+00	5,3E-04	3,2E-02

Annotation Cluster 2 Score d'enrichissement : 5,47						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mot-clé Pir Superfamily	actin-binding	45	2.9	2,4E+00	1,0E-07	7,3E-06
terme GO MF	GO:0008092~cytoskeletal protein binding	75	4.9	1,8E+00	1,6E-06	1,7E-03
terme GO MF	GO:0003779~actin binding	54	3.5	2,0E+00	2,5E-06	1,4E-03
terme GO CC	GO:0015629~actin cytoskeleton	38	2.4	1,9E+00	3,0E-04	1,3E-02

Annotation Cluster 3 Score d'enrichissement : 5,18						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mot-clé Pir Superfamily	nucleus	435	28.4	1,3E+00	7,4E-12	2,3E-09
terme GO BP	GO:0031328~positive regulation of cellular biosynthetic process	99	6.4	1,7E+00	1,0E-07	9,5E-05
terme GO BP	GO:0010557~positive regulation of macromolecule biosynthetic process	95	6.2	1,7E+00	1,5E-07	1,1E-04
terme GO BP	GO:0045893~positive regulation of transcription, DNA-dependent	75	4.9	1,9E+00	1,7E-07	1,1E-04
terme GO BP	GO:0009891~positive regulation of biosynthetic process	99	6.4	1,7E+00	2,1E-07	1,1E-04
terme GO BP	GO:0051254~positive regulation of RNA metabolic process	75	4.9	1,8E+00	2,4E-07	1,1E-04
terme GO BP	GO:0045941~positive regulation of transcription	84	5.4	1,8E+00	3,0E-07	1,3E-04
terme GO BP	GO:0006357~regulation of transcription from RNA polymerase II promoter	101	6.6	1,6E+00	5,1E-07	1,7E-04
terme GO BP	GO:0051173~positive regulation of nitrogen compound metabolic process	92	6.0	1,7E+00	5,2E-07	1,5E-04
terme GO BP	GO:0010628~positive regulation of gene expression	85	5.5	1,7E+00	5,5E-07	1,5E-04
terme GO BP	GO:0045935~positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	89	5.8	1,7E+00	8,7E-07	2,0E-04
terme GO BP	GO:0010604~positive regulation of macromolecule metabolic process	113	7.3	1,6E+00	1,4E-06	3,0E-04
Mot-clé Pir Superfamily	Transcription	215	14.0	1,4E+00	1,5E-06	7,8E-05
Mot-clé Pir Superfamily	transcription regulation	211	13.7	1,4E+00	1,6E-06	7,6E-05
terme GO BP	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	59	3.8	1,9E+00	2,9E-06	6,0E-04

Annotation Cluster 4 Score d'enrichissement : 4,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mot-clé Pir Superfamily	atp-binding	161	10,53	1,59	2,47E-09	2,58E-07
UP SEQ FEATURE	nucleotide phosphate-binding region:ATP	123	8,04	1,67	1,84E-08	3,82E-05
Mot-clé Pir Superfamily	nucleotide-binding	187	12,23	1,45	1,07E-07	6,72E-06
UP SEQ FEATURE	binding site:ATP	74	4,84	1,78	1,62E-06	2,24E-03
Mot-clé Pir Superfamily	serine/threonine-protein kinase	56	3,66	1,92	3,45E-06	1,55E-04
Mot-clé Pir Superfamily	kinase	87	5,69	1,66	3,58E-06	1,50E-04
UP SEQ FEATURE	domain:Protein kinase	65	4,25	1,81	4,63E-06	3,83E-03
INTERPRO	IPR000719:Protein kinase, core	68	4,45	1,77	5,22E-06	9,45E-03
terme GO MF	GO:0032559~adenyl ribonucleotide binding	174	11,38	1,38	5,98E-06	2,21E-03
terme GO MF	GO:0005524~ATP binding	172	11,25	1,38	6,12E-06	1,70E-03
terme GO MF	GO:0001882~nucleoside binding	184	12,03	1,35	9,37E-06	2,08E-03
terme GO MF	GO:0030554~adenyl nucleotide binding	179	11,71	1,34	1,78E-05	3,28E-03
terme GO MF	GO:0001883~purine nucleoside binding	181	11,84	1,34	1,97E-05	3,13E-03
INTERPRO	IPR017442:Serine/threonine protein kinase-related	53	3,47	1,83	2,66E-05	1,60E-02
INTERPRO	IPR017441:Protein kinase, ATP binding site	63	4,12	1,71	3,17E-05	1,43E-02
terme GO MF	GO:0004672~protein kinase activity	81	5,30	1,58	3,50E-05	4,85E-03
UP SEQ FEATURE	active site:Proton acceptor	80	5,23	1,58	4,47E-05	2,29E-02
terme GO BP	GO:0006468~protein amino acid phosphorylation	86	5,62	1,53	6,72E-05	8,87E-03
terme GO MF	GO:0004674~protein serine/threonine kinase activity	61	3,99	1,68	7,07E-05	8,69E-03
INTERPRO	IPR008271:Serine/threonine protein kinase, active site	51	3,34	1,78	7,27E-05	2,61E-02
terme GO MF	GO:0032555~purine ribonucleotide binding	199	13,02	1,28	9,26E-05	1,02E-02
terme GO MF	GO:0032553~ribonucleotide binding	199	13,02	1,28	9,26E-05	1,02E-02
terme GO BP	GO:0006793~phosphorus metabolic process	116	7,59	1,41	9,82E-05	1,17E-02
terme GO BP	GO:0006796~phosphate metabolic process	116	7,59	1,41	9,82E-05	1,17E-02
terme GO MF	GO:0017076~purine nucleotide binding	204	13,34	1,26	2,18E-04	2,00E-02
terme GO BP	GO:0016310~phosphorylation	97	6,34	1,44	2,25E-04	1,88E-02
INTERPRO	transferase	142	9,29	1,33	2,72E-04	7,74E-03

Annotation Cluster 5 : Score d'enrichissement : 4,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mot-clé Pir Superfamily	metal-binding	311	20,34	1,37	1,15E-09	1,45E-07
Mot-clé Pir Superfamily	zinc	230	15,04	1,38	2,24E-07	1,28E-05
Mot-clé Pir Superfamily	zinc-finger	179	11,71	1,36	1,09E-05	4,26E-04
terme GO MF	GO:0046914~transition metal ion binding	284	18,57	1,21	1,84E-04	1,84E-02
terme GO MF	GO:0008270~zinc ion binding	232	15,17	1,19	1,91E-03	1,01E-01
terme GO MF	GO:0043167~ion binding	398	26,03	1,11	5,02E-03	2,08E-01
terme GO MF	GO:0043169~cation binding	392	25,64	1,11	5,74E-03	2,18E-01
terme GO MF	GO:0046872~metal ion binding	388	25,38	1,11	6,16E-03	2,11E-01

Annotation Cluster 6 : Score d'enrichissement : 3,7						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
UP SEQ FEATURE	domain:Leucine-zipper	25	1,64	2,96	1,15E-09	2,50E-03
INTERPRO	IPR004827:Basic-leucine zipper (bZIP) transcription factor	16	1,05	3,73	2,24E-07	1,13E-02
SMART	SM00338:BRLZ	16	1,05	3,17	1,09E-05	2,86E-02
INTERPRO	IPR011616:bZIP transcription factor, bZIP-1	10	0,65	3,99	1,84E-04	1,57E-01
UP SEQ FEATURE	DNA-binding region:Basic motif	18	1,18	1,46	1,91E-03	9,84E-01

Annotation Cluster 7 : Score d'enrichissement : 3,5						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Superfamily	isopeptide bond	47	3,07	1,93	2,18E-05	7,59E-04
Mots-clé Pir Superfamily	ubl conjugation	67	4,38	1,49	1,00E-03	2,30E-02
UP SEQ FEATURE	cross-link:Glycyl lysine isopeptide (Lys-Gly) (interchain with G-Cter in ubiquitin)	29	1,90	1,90	1,32E-03	2,29E-01

Annotation Cluster 8 : Score d'enrichissement : 3,3						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0070161~anchoring junction	32	2,09	2,45	5,03E-06	3,47E-04
terme GO CC	GO:0005912~adherens junction	29	1,90	2,46	1,36E-05	8,35E-04
terme GO CC	GO:0005924~cell-substrate adherens junction	20	1,31	2,48	3,54E-04	1,39E-02
terme GO CC	GO:0005925~focal adhesion	19	1,24	2,45	6,08E-04	2,08E-02
terme GO CC	GO:0016323~basolateral plasma membrane	30	1,96	1,95	6,98E-04	2,24E-02
terme GO CC	GO:0030055~cell-substrate junction	20	1,31	2,35	7,20E-04	2,07E-02
terme GO CC	GO:0030054~cell junction	58	3,79	1,47	2,77E-03	4,40E-02
Mots-clé Pir Superfamily	cell junction	41	2,68	1,35	4,63E-02	3,35E-01

Annotation Cluster 9 : Score d'enrichissement : 3,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0007010~cytoskeleton organization	58	3,79	1,58	5,58E-04	3,24E-02
terme GO BP	GO:0030029~actin filament-based process	36	2,35	1,77	1,04E-03	5,07E-02
terme GO BP	GO:0030036~actin cytoskeleton organization	34	2,22	1,78	1,30E-03	5,92E-02

Annotation Cluster 10 Score d'enrichissement : 3,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0006913~nucleocytoplasmic transport	32	2,09	2,43	5,29E-06	9,34E-04
terme GO BP	GO:0006606~protein import into nucleus	22	1,44	3,03	6,46E-06	1,09E-03
terme GO BP	GO:0051169~nuclear transport	32	2,09	2,40	6,95E-06	1,12E-03
terme GO BP	GO:0051170~nuclear import	22	1,44	2,96	9,50E-06	1,47E-03
terme GO BP	GO:0034504~protein localization in nucleus	22	1,44	2,77	2,79E-05	3,97E-03
terme GO BP	GO:0033365~protein localization in organelle	27	1,77	2,19	2,00E-04	1,84E-02
terme GO BP	GO:0017038~protein import	25	1,64	2,26	2,22E-04	1,90E-02
terme GO BP	GO:0006605~protein targeting	34	2,22	1,87	5,40E-04	3,23E-02
terme GO BP	GO:0046907~intracellular transport	78	5,10	1,41	1,74E-03	7,07E-02
terme GO BP	GO:0000060~protein import into nucleus, translocation	9	0,59	3,14	6,21E-03	1,68E-01
Mots-clé Pir Superfamily	protein transport	51	3,34	1,38	1,81E-02	1,98E-01
terme GO BP	GO:0008104~protein localization	92	6,02	1,24	2,37E-02	3,51E-01
terme GO BP	GO:0045184~establishment of protein localization	81	5,30	1,25	2,74E-02	3,80E-01
terme GO BP	GO:0034613~cellular protein localization	47	3,07	1,36	2,85E-02	3,85E-01
terme GO BP	GO:0015031~protein transport	80	5,23	1,24	3,02E-02	3,92E-01
terme GO BP	GO:0070727~cellular macromolecule localization	47	3,07	1,35	3,21E-02	4,02E-01
terme GO BP	GO:0006886~intracellular protein transport	42	2,75	1,33	4,91E-02	4,92E-01

Annotation Cluster 11 Score d'enrichissement : 3,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0010605~negative regulation of macromolecule metabolic process	96	6,28	1,55	1,33E-05	1,97E-03
terme GO BP	GO:0010558~negative regulation of macromolecule biosynthetic process	71	4,64	1,54	2,61E-04	2,08E-02
terme GO BP	GO:0031327~negative regulation of cellular biosynthetic process	72	4,71	1,52	3,25E-04	2,48E-02
terme GO BP	GO:0010629~negative regulation of gene expression	66	4,32	1,55	3,44E-04	2,43E-02
terme GO MF	GO:0016564~transcription repressor activity	46	3,01	1,72	3,52E-04	2,96E-02
terme GO BP	GO:0009890~negative regulation of biosynthetic process	72	4,71	1,49	5,93E-04	3,28E-02
terme GO BP	GO:0051253~negative regulation of RNA metabolic process	49	3,20	1,60	1,08E-03	5,15E-02
terme GO BP	GO:0045892~negative regulation of transcription, DNA-dependent	48	3,14	1,60	1,34E-03	6,05E-02
terme GO BP	GO:0016481~negative regulation of transcription	57	3,73	1,47	3,08E-03	1,05E-01
terme GO BP	GO:0045934~negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	60	3,92	1,39	8,22E-03	1,94E-01
terme GO BP	GO:0051172~negative regulation of nitrogen compound metabolic process	60	3,92	1,37	1,09E-02	2,25E-01
terme GO BP	GO:0000122~negative regulation of transcription from RNA polymerase II promoter	33	2,16	1,47	2,57E-02	3,67E-01

Annotation Cluster 12 : Score d'enrichissement : 2,7						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0002237~response to molecule of bacterial origin	20	1,31	2,76	7,78E-05	9,58E-03
terme GO BP	GO:0032496~response to lipopolysaccharide	16	1,05	2,46	1,70E-03	7,08E-02
terme GO BP	GO:0009617~response to bacterium	25	1,64	1,54	3,48E-02	4,19E-01

Annotation Cluster 13 Score d'enrichissement : 2,7						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0042325~regulation of phosphorylation	64	4,19	1,63	1,12E-04	1,25E-02
terme GO BP	GO:0051174~regulation of phosphorus metabolic process	65	4,25	1,59	2,02E-04	1,81E-02
terme GO BP	GO:0019220~regulation of phosphate metabolic process	65	4,25	1,59	2,02E-04	1,81E-02
terme GO BP	GO:0043549~regulation of kinase activity	51	3,34	1,69	2,43E-04	1,98E-02
terme GO BP	GO:0051338~regulation of transferase activity	52	3,40	1,66	3,55E-04	2,41E-02
terme GO BP	GO:0045859~regulation of protein kinase activity	49	3,20	1,68	3,77E-04	2,46E-02
terme GO BP	GO:0051347~positive regulation of transferase activity	33	2,16	1,63	6,32E-03	1,66E-01
terme GO BP	GO:0033674~positive regulation of kinase activity	32	2,09	1,64	6,49E-03	1,66E-01
terme GO BP	GO:0045860~positive regulation of protein kinase activity	31	2,03	1,65	7,08E-03	1,75E-01
terme GO BP	GO:0043085~positive regulation of catalytic activity	58	3,79	1,32	2,33E-02	3,47E-01
terme GO BP	GO:0044093~positive regulation of molecular function	64	4,19	1,29	2,59E-02	3,65E-01
terme GO BP	GO:0032147~activation of protein kinase activity	17	1,11	1,77	2,94E-02	3,87E-01

Annotation Cluster 14 Score d'enrichissement : 2,7						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0048514~blood vessel morphogenesis	33	2,16	1,85	8,01E-04	4,22E-02
terme GO BP	GO:0001568~blood vessel development	36	2,35	1,74	1,39E-03	6,10E-02
terme GO BP	GO:0001944~vasculature development	36	2,35	1,70	2,11E-03	8,02E-02
terme GO BP	GO:0001525~angiogenesis	23	1,50	1,84	6,27E-03	1,66E-01

Annotation Cluster 15 Score d'enrichissement : 2,7						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0031674~I band	13	0,85	3,11	7,07E-04	2,15E-02
terme GO CC	GO:0030016~myofibril	19	1,24	2,25	1,67E-03	3,50E-02
terme GO CC	GO:0043292~contractile fiber	20	1,31	2,18	1,86E-03	3,60E-02
terme GO CC	GO:0030018~Z disc	11	0,72	3,15	1,99E-03	3,60E-02
terme GO CC	GO:0044449~contractile fiber part	19	1,24	2,21	2,06E-03	3,60E-02
terme GO CC	GO:0030017~sarcomere	16	1,05	2,15	6,82E-03	9,23E-02

Annotation Cluster 16 Score d'enrichissement : 2,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Superfamily	ubl conjugation pathway	64	4,19	1,65	9,44E-05	2,96E-03
terme GO BP	GO:0009057~macromolecule catabolic process	93	6,08	1,41	5,56E-04	3,27E-02
terme GO BP	GO:0044265~cellular macromolecule catabolic process	85	5,56	1,39	1,58E-03	6,76E-02
terme GO BP	GO:0030163~protein catabolic process	74	4,84	1,41	2,24E-03	8,30E-02
terme GO BP	GO:0051603~proteolysis involved in cellular protein catabolic process	71	4,64	1,40	3,16E-03	1,07E-01
terme GO BP	GO:0044257~cellular protein catabolic process	71	4,64	1,40	3,57E-03	1,13E-01
terme GO BP	GO:0043632~modification-dependent macromolecule catabolic process	68	4,45	1,40	3,79E-03	1,18E-01
terme GO BP	GO:0019941~modification-dependent protein catabolic process	68	4,45	1,40	3,79E-03	1,18E-01
terme GO BP	GO:0006508~proteolysis	102	6,67	1,15	9,09E-02	6,40E-01

Annotation Cluster 17 Score d'enrichissement : 2,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0006417~regulation of translation	25	1,64	2,16	4,44E-04	2,75E-02
terme GO BP	GO:0017148~negative regulation of translation	10	0,65	3,49	1,57E-03	6,80E-02
terme GO BP	GO:0010608~posttranscriptional regulation of gene expression	32	2,09	1,80	1,62E-03	6,84E-02
terme GO BP	GO:0032268~regulation of cellular protein metabolic process	57	3,73	1,43	6,15E-03	1,67E-01
Mots-clé Pir Superfamily	translation regulation	12	0,78	2,35	1,20E-02	1,52E-01

Annotation Cluster 18 Score d'enrichissement : 2,4						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0000267~cell fraction	110	7,19	1,34	1,13E-03	2,94E-02
terme GO CC	GO:0005624~membrane fraction	85	5,56	1,38	1,80E-03	3,62E-02
terme GO CC	GO:0005626~insoluble fraction	87	5,69	1,37	2,35E-03	3,98E-02
terme GO CC	GO:0005792~microsome	30	1,96	1,67	7,16E-03	9,22E-02
terme GO CC	GO:0042598~vesicular fraction	30	1,96	1,62	1,06E-02	1,20E-01

Annotation Cluster 19 Score d'enrichissement : 2,4						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO MF	GO:0046983~protein dimerization activity	70	4,58	1,53	3,53E-04	2,76E-02
terme GO MF	GO:0042803~protein homodimerization activity	42	2,75	1,49	9,73E-03	3,04E-01
terme GO MF	GO:0042802~identical protein binding	70	4,58	1,29	2,06E-02	4,23E-01

Annotation Cluster 20 Score d'enrichissement : 2,4						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
UP SEQ FEATURE	domain:DHR-1	6	0,39	7,11	8,23E-04	1,82E-01
UP SEQ FEATURE	domain:DHR-2	6	0,39	7,11	8,23E-04	1,82E-01
INTERPRO	IPR010703:Dedicator of cytokinesis	3	0,20	5,30	1,04E-01	9,21E-01

Annotation Cluster 21 Score d'enrichissement : 2,3						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Superfamily	motor protein	23	1,50	2,28	4,08E-04	1,06E-02
terme GO MF	GO:0003774~motor activity	25	1,64	2,08	7,74E-04	5,24E-02
terme GO MF	GO:0003777~microtubule motor activity	15	0,98	2,31	4,69E-03	2,03E-01
terme GO CC	GO:0005874~microtubule	34	2,22	1,63	5,49E-03	7,88E-02
Mots-clé Pir Superfamily	microtubule	30	1,96	1,70	5,56E-03	9,25E-02
terme GO BP	GO:0007017~microtubule-based process	34	2,22	1,59	7,80E-03	1,89E-01
terme GO BP	GO:0007018~microtubule-based movement	18	1,18	1,89	1,35E-02	2,63E-01
terme GO CC	GO:0005875~microtubule associated complex	15	0,98	1,94	2,17E-02	2,08E-01

Annotation Cluster 22 Score d'enrichissement : 2,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0001501~skeletal system development	44	2,88	1,64	1,38E-03	6,14E-02
terme GO BP	GO:0060348~bone development	20	1,31	1,93	7,06E-03	1,76E-01
terme GO BP	GO:0001503~ossification	18	1,18	1,86	1,59E-02	2,88E-01

Annotation Cluster 23 Score d'enrichissement : 2,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0051270~regulation of cell motion	33	2,16	2,03	1,58E-04	1,58E-02
terme GO BP	GO:0030334~regulation of cell migration	28	1,83	1,96	9,00E-04	4,60E-02
terme GO BP	GO:0040012~regulation of locomotion	29	1,90	1,79	2,96E-03	1,04E-01
terme GO BP	GO:0051272~positive regulation of cell motion	18	1,18	2,18	3,20E-03	1,06E-01
terme GO BP	GO:0051271~negative regulation of cell motion	12	0,78	2,26	1,53E-02	2,85E-01
terme GO BP	GO:0030336~negative regulation of cell migration	11	0,72	2,29	1,96E-02	3,23E-01
terme GO BP	GO:0040013~negative regulation of locomotion	11	0,72	2,14	3,02E-02	3,91E-01
terme GO BP	GO:0040017~positive regulation of locomotion	15	0,98	1,81	3,49E-02	4,19E-01
terme GO BP	GO:0030335~positive regulation of cell migration	14	0,92	1,87	3,51E-02	4,20E-01

Annotation Cluster 24 Score d'enrichissement : 2,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0042641~actomyosin	8	0,52	3,90	3,27E-03	5,04E-02
terme GO CC	GO:0001725~stress fiber	7	0,46	3,84	7,73E-03	9,48E-02
terme GO CC	GO:0032432~actin filament bundle	7	0,46	3,54	1,16E-02	1,28E-01

Annotation Cluster 25 Score d'enrichissement : 2,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
UP SEQ FEATURE	repeat:HEAT 7	10	0,65	5,43	4,06E-05	2,38E-02
UP SEQ FEATURE	repeat:HEAT 5	11	0,72	4,09	2,12E-04	9,30E-02
UP SEQ FEATURE	repeat:HEAT 6	10	0,65	4,34	2,92E-04	1,04E-01
UP SEQ FEATURE	repeat:HEAT 1	14	0,92	3,14	3,64E-04	1,02E-01
UP SEQ FEATURE	repeat:HEAT 2	14	0,92	3,14	3,64E-04	1,02E-01
INTERPRO	IPR000357:HEAT	13	0,85	2,92	1,23E-03	2,01E-01
UP SEQ FEATURE	repeat:HEAT 4	11	0,72	3,33	1,26E-03	2,30E-01
INTERPRO	IPR011989:Armadillo-like helical	22	1,44	2,13	1,39E-03	1,90E-01
UP SEQ FEATURE	repeat:HEAT 8	7	0,46	4,80	2,29E-03	3,38E-01
UP SEQ FEATURE	repeat:HEAT 10	6	0,39	5,21	4,13E-03	4,46E-01
UP SEQ FEATURE	repeat:HEAT 9	6	0,39	5,21	4,13E-03	4,46E-01
UP SEQ FEATURE	repeat:HEAT 3	11	0,72	2,87	4,14E-03	4,36E-01
UP SEQ FEATURE	repeat:HEAT 16	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 17	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 22	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 21	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 24	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 23	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 19	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 18	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 20	3	0,20	9,77	3,18E-02	8,81E-01
UP SEQ FEATURE	repeat:HEAT 12	4	0,26	5,21	3,59E-02	8,99E-01
UP SEQ FEATURE	repeat:HEAT 13	4	0,26	5,21	3,59E-02	8,99E-01
UP SEQ FEATURE	repeat:HEAT 11	4	0,26	4,74	4,66E-02	9,34E-01
UP SEQ FEATURE	repeat:HEAT 15	3	0,20	5,58	9,53E-02	9,79E-01
UP SEQ FEATURE	repeat:HEAT 14	3	0,20	4,34	1,48E-01	9,93E-01

Annotation Cluster 26 Score d'enrichissement : 2,1						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
INTERPRO	IPR018097:EGF-like calcium-binding, conserved site	19	1,24	2,45	6,09E-04	1,46E-01
INTERPRO	IPR013032:EGF-like region, conserved site	41	2,68	1,73	7,31E-04	1,53E-01
UP SEQ FEATURE	domain:EGF-like 2; calcium-binding	13	0,85	2,97	1,10E-03	2,13E-01
INTERPRO	IPR006209:EGF	22	1,44	2,14	1,25E-03	1,87E-01
INTERPRO	IPR001881:EGF-like calcium-binding	18	1,18	2,32	1,65E-03	2,06E-01
INTERPRO	IPR013091:EGF calcium-binding	15	0,98	2,58	1,65E-03	1,94E-01
Mots-clé Pir Superfamily	egf-like domain	31	2,03	1,76	2,76E-03	5,28E-02
INTERPRO	IPR006210:EGF-like	29	1,90	1,78	3,22E-03	3,07E-01
INTERPRO	IPR000152:EGF-type aspartate/asparagine hydroxylation conserved site	17	1,11	2,17	4,63E-03	3,75E-01
UP SEQ FEATURE	domain:EGF-like 6	10	0,65	3,03	4,66E-03	4,54E-01
UP SEQ FEATURE	domain:EGF-like 3; calcium-binding	9	0,59	3,17	6,10E-03	5,15E-01
SMART	SM00179:EGF_CA	18	1,18	1,97	8,56E-03	6,33E-01
INTERPRO	IPR000742:EGF-like, type 3	26	1,70	1,66	1,35E-02	5,39E-01
UP SEQ FEATURE	domain:EGF-like 7; calcium-binding	7	0,46	3,38	1,47E-02	7,45E-01
UP SEQ FEATURE	domain:EGF-like 1	17	1,11	1,85	2,10E-02	8,04E-01
SMART	SM00181:EGF	29	1,90	1,51	2,53E-02	5,93E-01
UP SEQ FEATURE	domain:EGF-like 4; calcium-binding	7	0,46	2,76	3,73E-02	9,01E-01
UP SEQ FEATURE	domain:EGF-like 5; calcium-binding	7	0,46	2,46	6,05E-02	9,52E-01
UP SEQ FEATURE	domain:EGF-like 3	9	0,59	1,56	2,16E-01	9,98E-01
UP SEQ FEATURE	domain:EGF-like 6; calcium-binding	4	0,26	2,08	3,00E-01	1,00E+00

Annotation Cluster 27 Score d'enrichissement : 2,11						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
UP SEQ FEATURE	domain:C2	17	1,11	2,77	3,25E-04	9,85E-02
INTERPRO	IPR000008:C2 calcium-dependent membrane targeting	21	1,37	1,90	6,97E-03	3,99E-01
SMART	SM00239:C2	21	1,37	1,61	3,46E-02	6,42E-01
INTERPRO	IPR018029:C2 membrane targeting protein	15	0,98	1,75	4,60E-02	7,83E-01

Annotation Cluster 28 Score d'enrichissement : 2,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0015630~microtubule cytoskeleton	67	4,38	1,61	1,16E-04	5,82E-03
terme GO CC	GO:0005815~microtubule organizing center	27	1,77	1,40	7,00E-02	4,27E-01
terme GO CC	GO:0005813~centrosome	24	1,57	1,41	8,49E-02	4,75E-01

Annotation Cluster 29 Score d'enrichissement : 2,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0042035~regulation of cytokine biosynthetic process	15	0,98	2,40	3,18E-03	1,06E-01
terme GO BP	GO:0001817~regulation of cytokine production	27	1,77	1,77	5,01E-03	1,45E-01
terme GO BP	GO:0042108~positive regulation of cytokine biosynthetic process	9	0,59	2,22	4,57E-02	4,76E-01

Annotation Cluster 30 Score d'enrichissement : 2,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0006984~ER-nuclear signaling pathway	10	0,65	3,39	1,96E-03	7,62E-02
terme GO BP	GO:0051789~response to protein stimulus	19	1,24	2,11	3,47E-03	1,12E-01
terme GO BP	GO:0034976~response to endoplasmic reticulum stress	9	0,59	3,14	6,21E-03	1,68E-01
terme GO BP	GO:0034620~cellular response to unfolded protein	7	0,46	3,95	6,36E-03	1,65E-01
terme GO BP	GO:0030968~endoplasmic reticulum unfolded protein response	7	0,46	3,95	6,36E-03	1,65E-01
terme GO BP	GO:0006983~ER overload response	4	0,26	4,74	4,57E-02	4,75E-01
terme GO BP	GO:0006986~response to unfolded protein	11	0,72	1,84	7,32E-02	5,94E-01

Annotation Cluster 31 Score d'enrichissement : 2,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0048870~cell motility	40	2,62	1,54	6,37E-03	1,64E-01
terme GO BP	GO:0051674~localization of cell	40	2,62	1,54	6,37E-03	1,64E-01
terme GO BP	GO:0016477~cell migration	36	2,35	1,55	9,62E-03	2,14E-01
terme GO BP	GO:0006928~cell motion	54	3,53	1,35	2,10E-02	3,32E-01

ANNEXE 2 : RESULTATS DE L'ANALYSE FONCTIONNELLE PAR LE LOGICIEL DAVID DES TERMES GO ASSOCIES AUX GENES DONT L'EXPRESSION EST MODIFIEE LORS D'UN STRESS OXYDANT AU TEMPS 24 H.

L'analyse a été réalisée en utilisant les gènes dont l'expression est modifiée 24 h après le début d'un stress oxydant. Pour chaque GO identifié, un taux d'enrichissement est calculé et des tests statistiques sont effectués pour évaluer le niveau de significativité de l'enrichissement du terme d'intérêt. A partir de ces résultats, le logiciel regroupe les GO biologiquement reliées entre eux en cluster et associe un score d'enrichissement. Ce score d'enrichissement est directement lié à la p-value calculée pour le groupe fonctionnel selon la formule suivante : $\text{score} = -\log(\text{p-value})$, par conséquent un score d'enrichissement ≥ 2 correspond à une p-value ≤ 0.01 et plus le score d'enrichissement est élevé, plus l'enrichissement de ces termes est significatif. Les groupes fonctionnels indiqués sont ceux qui ont un score d'enrichissement supérieur ou égal à 3.

Annotation Cluster 1 Score d'enrichissement : 6,7						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0030334~regulation of cell migration	35	2,8	2,9647162	1,59E-08	1,12E-05
terme GO BP	GO:0040012~regulation of locomotion	37	3,0	2,75868607	4,13E-08	1,32E-05
terme GO BP	GO:0051270~regulation of cell motion	37	3,0	2,74439236	4,75E-08	1,19E-05
terme GO BP	GO:0030335~positive regulation of cell migration	22	1,8	3,53862434	5,72E-07	9,58E-05
terme GO BP	GO:0040017~positive regulation of locomotion	23	1,9	3,35972357	7,75E-07	1,18E-04
terme GO BP	GO:0051272~positive regulation of cell motion	22	1,8	3,21364863	3,06E-06	3,71E-04

Annotation Cluster 2 Score d'enrichissement : 6,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0031012~extracellular matrix	56	4,5	2,30017674	7,97E-09	1,96E-06
terme GO CC	GO:0005578~proteinaceous extracellular matrix	52	4,2	2,3027439	2,85E-08	4,68E-06
Mots-clé Pir Family	extracellular matrix	40	3,2	2,63627673	4,94E-08	1,52E-05
terme GO CC	GO:0044420~extracellular matrix part	27	2,2	3,27016886	1,18E-07	1,16E-05
terme GO CC	GO:0005604~basement membrane	20	1,6	3,63352095	1,35E-06	8,33E-05
terme GO MF	GO:0005201~extracellular matrix structural constituent	18	1,5	2,97631116	8,41E-05	2,27E-02

Annotation Cluster 3 Score d'enrichissement : 6,4						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0000267~cell fraction	124	10,0	1,622502984	4,23E-08	5,22E-06
terme GO CC	GO:0005626~insoluble fraction	97	7,8	1,63833251	1,13E-06	9,26E-05
terme GO CC	GO:0005624~membrane fraction	94	7,6	1,64653743	1,34E-06	9,47E-05

Annotation Cluster 4 Score d'enrichissement : 6,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0033273~response to vitamin	21	1,7	4,554882155	1,15E-08	1,34E-05
terme GO BP	GO:0009991~response to extracellular stimulus	40	3,2	2,602789803	5,57E-08	1,15E-05
terme GO BP	GO:0007584~response to nutrient	29	2,3	2,96532124	3,32E-07	5,83E-05
terme GO BP	GO:0031667~response to nutrient levels	35	2,8	2,543335213	7,52E-07	1,20E-04
terme GO BP	GO:0033189~response to vitamin A	11	0,9	3,74925674	4,92E-04	1,87E-02

Annotation Cluster 5 Score d'enrichissement : 6,1						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0010035~response to inorganic substance	44	3,6	3,072561621	4,89E-11	8,60E-08
terme GO BP	GO:0006979~response to oxidative stress	33	2,7	2,88052652	8,90E-08	1,74E-05
terme GO BP	GO:0000302~response to reactive oxygen species	17	1,4	3,244811287	4,77E-05	3,22E-03
terme GO BP	GO:0042542~response to hydrogen peroxide	12	1,0	3,067573696	0,001468065	4,36E-02

Annotation Cluster 6 Score d'enrichissement : 5,7						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0001568~blood vessel development	43	3,5	2,51	4,66E-08	1,26E-05
terme GO BP	GO:0001944~vasculature development	43	3,5	2,45	9,42E-08	1,74E-05
terme GO BP	GO:0001525~angiogenesis	26	2,1	2,51	3,12E-05	2,15E-03
terme GO BP	GO:0048514~blood vessel morphogenesis	32	2,6	2,17	6,27E-05	3,86E-03

Annotation Cluster 7 Score d'enrichissement : 5,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	actin-binding	38	3,1	2,44	7,81E-07	9,62E-05
terme GO MF	GO:0008092~cytoskeletal protein binding	65	5,3	1,83	2,47E-06	2,69E-03
terme GO MF	GO:0003779~actin binding	46	3,7	2,01	9,48E-06	5,16E-03

Annotation Cluster 8 Score d'enrichissement : 5,5						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0007155~cell adhesion	83	6,7	1,70	1,96E-06	2,55E-04
terme GO BP	GO:0022610~biological adhesion	83	6,7	1,69	2,10E-06	2,64E-04
Mots-clé Pir Family	cell adhesion	52	4,2	1,96	5,76E-06	4,44E-04

Annotation Cluster 9 Score d'enrichissement : 5,4						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0010033~response to organic substance	100	8,1	1,99	3,30E-11	1,16E-07
terme GO BP	GO:0009719~response to endogenous stimulus	61	4,9	2,16	1,96E-08	1,15E-05
terme GO BP	GO:0009725~response to hormone stimulus	56	4,5	2,18	5,18E-08	1,14E-05
terme GO BP	GO:0043434~response to peptide hormone stimulus	27	2,2	2,51	2,20E-05	1,65E-03
terme GO BP	GO:0032868~response to insulin stimulus	19	1,5	2,72	1,71E-04	7,89E-03
terme GO BP	GO:0032869~cellular response to insulin stimulus	15	1,2	3,16	2,09E-04	9,13E-03
terme GO BP	GO:0032870~cellular response to hormone stimulus	22	1,8	2,37	3,44E-04	1,37E-02
terme GO BP	GO:0008286~insulin receptor signaling pathway	9	0,7	3,48	3,42E-03	7,98E-02

Annotation Cluster 10 Score d'enrichissement : 5,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0035295~tube development	37	3,0	2,41	1,34E-06	1,81E-04
terme GO BP	GO:0048754~branching morphogenesis of a tube	17	1,4	3,74	7,00E-06	7,03E-04
terme GO BP	GO:0001763~morphogenesis of a branching structure	18	1,5	3,48	9,90E-06	8,71E-04
terme GO BP	GO:0035239~tube morphogenesis	23	1,9	2,59	6,21E-05	3,89E-03

Annotation Cluster 11 Score d'enrichissement : 5,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0042981~regulation of apoptosis	99	8,0	1,76	2,62E-08	1,32E-05
terme GO BP	GO:0043067~regulation of programmed cell death	99	8,0	1,75	4,32E-08	1,27E-05
terme GO BP	GO:0010941~regulation of cell death	99	8,0	1,74	5,07E-08	1,19E-05
terme GO BP	GO:0043068~positive regulation of programmed cell death	57	4,6	1,88	5,00E-06	5,67E-04
terme GO BP	GO:0010942~positive regulation of cell death	57	4,6	1,88	5,70E-06	6,26E-04
terme GO BP	GO:0043065~positive regulation of apoptosis	56	4,5	1,86	8,45E-06	7,82E-04
terme GO BP	GO:0043066~negative regulation of apoptosis	46	3,7	1,86	6,37E-05	3,85E-03
terme GO BP	GO:0043069~negative regulation of programmed cell death	46	3,7	1,83	8,90E-05	4,73E-03
terme GO BP	GO:0060548~negative regulation of cell death	46	3,7	1,83	9,57E-05	5,01E-03
terme GO BP	GO:0012502~induction of programmed cell death	41	3,3	1,83	2,43E-04	1,03E-02
terme GO BP	GO:0006917~induction of apoptosis	40	3,2	1,79	4,57E-04	1,79E-02
terme GO BP	GO:0006916~anti-apoptosis	28	2,3	1,95	1,13E-03	3,68E-02

Annotation Cluster 12 Score d'enrichissement : 4,9						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0012501~programmed cell death	73	5,9	1,71	6,95E-06	7,19E-04
Mots-clé Pir Family	Apoptosis	48	3,9	2,00	7,67E-06	5,25E-04
terme GO BP	GO:0006915~apoptosis	72	5,8	1,71	7,76E-06	7,58E-04
terme GO BP	GO:0008219~cell death	81	6,6	1,61	1,82E-05	1,42E-03
terme GO BP	GO:0016265~death	81	6,6	1,60	2,31E-05	1,69E-03

Annotation Cluster 13 Score d'enrichissement : 4,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0009991~response to extracellular stimulus	40	3,2	2,60	5,57E-08	1,15E-05
terme GO BP	GO:0048545~response to steroid hormone stimulus	30	2,4	2,24	6,45E-05	3,84E-03
terme GO BP	GO:0031960~response to corticosteroid stimulus	17	1,4	2,86	2,27E-04	9,71E-03
terme GO BP	GO:0051384~response to glucocorticoid stimulus	16	1,3	2,94	2,77E-04	1,14E-02

Annotation Cluster 14 Score d'enrichissement : 4,5						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0044421~extracellular region part	116	9,4	1,71	6,63E-09	3,27E-06
terme GO CC	GO:0005615~extracellular space	73	5,9	1,51	3,71E-04	1,30E-02
terme GO CC	GO:0005576~extracellular region	169	13,7	1,19	8,10E-03	1,21E-01

Annotation Cluster 15 Score d'enrichissement : 4,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0009611~response to wounding	74	6,0	2,00	1,33E-08	1,17E-05
terme GO BP	GO:0006954~inflammatory response	42	3,4	1,85	1,58E-04	7,39E-03
terme GO BP	GO:0006952~defense response	53	4,3	1,23	8,44E-02	5,68E-01

Annotation Cluster 16 Score d'enrichissement : 3,9						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
INTERPRO	IPR013091:EGF calcium-binding	18	1,5	3,70	4,38E-06	7,02E-03
INTERPRO	IPR013032:EGF-like region, conserved site	41	3,3	2,07	1,63E-05	8,73E-03
INTERPRO	IPR001881:EGF-like calcium-binding	20	1,6	3,08	1,90E-05	7,62E-03
INTERPRO	IPR018097:EGF-like calcium-binding, conserved site	20	1,6	3,08	1,90E-05	7,62E-03
UP SEQ FEATURE	domain:EGF-like 2; calcium-binding	14	1,1	3,88	4,36E-05	3,87E-02
INTERPRO	IPR000742:EGF-like, type 3	30	2,4	2,29	4,47E-05	1,43E-02
INTERPRO	IPR006209:EGF	22	1,8	2,56	1,15E-04	3,04E-02
SMART	SM00179:EGF_CA	20	1,6	2,68	1,18E-04	1,84E-02
Mots-clé Pir Family	egf-like domain	31	2,5	2,14	1,20E-04	3,69E-03
INTERPRO	IPR000152:EGF-type aspartate/asparagine hydroxylation conserved site	18	1,5	2,75	2,46E-04	5,50E-02
INTERPRO	IPR006210:EGF-like	28	2,3	2,06	4,79E-04	8,21E-02
UP SEQ FEATURE	domain:EGF-like 1	18	1,5	2,37	1,45E-03	2,04E-01
SMART	SM00181:EGF	28	2,3	1,79	3,41E-03	1,64E-01
UP SEQ FEATURE	domain:EGF-like 3	12	1,0	2,53	7,29E-03	5,02E-01

Annotation Cluster 17 Score d'enrichissement : 3,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	ubl conjugation	64	5,2	1,73	2,45E-05	1,08E-03
Mots-clé Pir Family	isopeptide bond	38	3,1	1,89	2,44E-04	5,56E-03
UP SEQ FEATURE	cross-link:Glycyl lysine isopeptide (Lys-Gly) (interchain with G-Cter in ubiquitin)	25	2,0	1,98	1,79E-03	2,21E-01

Annotation Cluster 18 Score d'enrichissement : 3,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
INTERPRO	IPR004827:Basic-leucine zipper (bZIP) transcription factor	15	1,2	4,19	7,54E-06	6,05E-03
UP SEQ FEATURE	domain:Leucine-zipper	21	1,7	3,02	1,67E-05	2,00E-02
SMART	SM00338:BRLZ	15	1,2	3,64	3,48E-05	1,09E-02
UP SEQ FEATURE	DNA-binding region:Basic motif	22	1,8	2,16	1,27E-03	1,89E-01
INTERPRO	IPR011700:Basic leucine zipper	6	0,5	5,92	2,36E-03	2,24E-01
INTERPRO	IPR011616:bZIP transcription factor, bZIP-1	7	0,6	3,34	1,60E-02	5,35E-01

Annotation Cluster 19 Score d'enrichissement : 3,5						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	signal	267	21,6	1,30	1,35E-06	1,19E-04
UP SEQ FEATURE	signal peptide	267	21,6	1,30	2,09E-06	3,77E-03
Mots-clé Pir Family	glycoprotein	332	26,9	1,22	2,05E-05	9,69E-04
UP SEQ FEATURE	glycosylation site:N-linked (GlcNAc,,,)	315	25,5	1,21	1,03E-04	7,20E-02
Mots-clé Pir Family	Secreted	134	10,8	1,26	3,60E-03	4,26E-02
Mots-clé Pir Family	disulfide bond	215	17,4	1,17	8,01E-03	8,06E-02
terme GO CC	GO:0005576~extracellular region	169	13,7	1,19	8,10E-03	1,21E-01
UP SEQ FEATURE	disulfide bond	205	16,6	1,15	1,84E-02	7,13E-01

Annotation Cluster 20 Score d'enrichissement : 3,5						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0030016~myofibril	21	1,7	2,68	8,58E-05	4,69E-03
terme GO CC	GO:0043292~contractile fiber	22	1,8	2,58	1,01E-04	4,97E-03
terme GO CC	GO:0044449~contractile fiber part	20	1,6	2,51	3,25E-04	1,33E-02
terme GO CC	GO:0031674~I band	13	1,1	3,35	3,63E-04	1,37E-02
terme GO CC	GO:0030018~Z disc	11	0,9	3,39	1,14E-03	3,08E-02
terme GO CC	GO:0030017~sarcomere	17	1,4	2,46	1,30E-03	3,15E-02

Annotation Cluster 21 Score d'enrichissement : 3,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0009792~embryonic development ending in birth or egg hatching	43	3,5	1,84	1,42E-04	6,91E-03
terme GO BP	GO:0043009~chordate embryonic development	41	3,3	1,77	4,63E-04	1,79E-02
terme GO BP	GO:0001701~in utero embryonic development	24	1,9	1,95	2,64E-03	6,75E-02

Annotation Cluster 22 Score d'enrichissement : 3,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0035295~tube development	37	3,0	2,41	1,34E-06	1,81E-04
terme GO BP	GO:0060541~respiratory system development	18	1,5	2,39	1,26E-03	3,97E-02
terme GO BP	GO:0030324~lung development	17	1,4	2,46	1,31E-03	3,99E-02
terme GO BP	GO:0030323~respiratory tube development	17	1,4	2,39	1,80E-03	5,03E-02
terme GO BP	GO:0048286~lung alveolus development	6	0,5	3,90	1,59E-02	2,29E-01

Annotation Cluster 23 Score d'enrichissement : 3,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0030036~actin cytoskeleton organization	31	2,5	1,96	4,98E-04	1,87E-02
terme GO BP	GO:0007010~cytoskeleton organization	50	4,0	1,64	6,13E-04	2,18E-02
terme GO BP	GO:0030029~actin filament-based process	32	2,6	1,90	6,96E-04	2,44E-02

Annotation Cluster 24 Score d'enrichissement : 3,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0019220~regulation of phosphate metabolic process	64	5,2	1,89	1,10E-06	1,61E-04
terme GO BP	GO:0051174~regulation of phosphorus metabolic process	64	5,2	1,89	1,10E-06	1,61E-04
terme GO BP	GO:0042325~regulation of phosphorylation	62	5,0	1,90	1,26E-06	1,78E-04
terme GO BP	GO:0043549~regulation of kinase activity	45	3,6	1,80	1,57E-04	7,54E-03
terme GO BP	GO:0051338~regulation of transferase activity	46	3,7	1,77	2,06E-04	9,12E-03
terme GO BP	GO:0045859~regulation of protein kinase activity	42	3,4	1,74	5,55E-04	1,99E-02
terme GO BP	GO:0044093~positive regulation of molecular function	58	4,7	1,42	6,46E-03	1,24E-01
terme GO BP	GO:0051347~positive regulation of transferase activity	28	2,3	1,67	9,46E-03	1,64E-01
terme GO BP	GO:0033674~positive regulation of kinase activity	27	2,2	1,67	1,07E-02	1,79E-01
terme GO BP	GO:0043085~positive regulation of catalytic activity	51	4,1	1,40	1,27E-02	2,00E-01
terme GO BP	GO:0045860~positive regulation of protein kinase activity	25	2,0	1,60	2,26E-02	2,77E-01
terme GO BP	GO:0032147~activation of protein kinase activity	10	0,8	1,26	4,01E-01	9,46E-01

Annotation Cluster 25 Score d'enrichissement : 3,1						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO MF	GO:0001871~pattern binding	26	2,1	2,40	6,83E-05	2,45E-02
terme GO MF	GO:0030247~polysaccharide binding	26	2,1	2,40	6,83E-05	2,45E-02
Mots-clé Pir Family	heparin-binding	13	1,1	3,18	6,56E-04	1,15E-02
terme GO MF	GO:0005539~glycosaminoglycan binding	22	1,8	2,23	7,53E-04	6,13E-02
terme GO MF	GO:0030246~carbohydrate binding	40	3,2	1,61	3,36E-03	1,68E-01
terme GO MF	GO:0008201~heparin binding	15	1,2	2,07	1,28E-02	3,73E-01

Annotation Cluster 26 Score d'enrichissement : 3,1						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0051130~positive regulation of cellular component organization	28	2,3	2,21	1,41E-04	6,97E-03
terme GO BP	GO:0033043~regulation of organelle organization	30	2,4	1,98	5,49E-04	1,99E-02
terme GO BP	GO:0010638~positive regulation of organelle organization	14	1,1	2,41	4,73E-03	1,01E-01

Annotation Cluster 27 Score d'enrichissement : 3,1						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0031099~regeneration	17	1,4	3,53	1,59E-05	1,33E-03
terme GO BP	GO:0042246~tissue regeneration	9	0,7	4,03	1,27E-03	3,95E-02
terme GO BP	GO:0040007~growth	25	2,0	1,96	2,06E-03	5,58E-02
terme GO BP	GO:0048589~developmental growth	14	1,1	2,30	7,08E-03	1,32E-01

Annotation Cluster 28 Score d'enrichissement : 3,1						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0032496~response to lipopolysaccharide	17	1,4	3,16	6,69E-05	3,91E-03
terme GO BP	GO:0002237~response to molecule of bacterial origin	18	1,5	3,00	7,77E-05	4,33E-03
terme GO BP	GO:0009617~response to bacterium	20	1,6	1,48	8,09E-02	5,59E-01

ANNEXE 3 : RESULTATS DES ANALYSES ET PREDICTIONS EN AMONT DES DONNEES D'EXPRESSION DES PUCES PAR LE LOGICIEL INGENUITY AU TEMPS 2 H, 4 H ET 24 H.

2 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
PDGF BB		Activated	3,276	6,23E-21
NUPR1	1,137	Activated	2,594	2,09E-12
HIF1A	-1,829	Activated	2,918	2,10E-11
TREM1	-1,189	Activated	3,12	4,14E-11
FSH			1,377	1,04E-10
TGFB1	1,08		-0,452	1,33E-09
FOXO3	-1,114		1,268	2,34E-09
Hdac		Inhibited	-2,956	9,20E-09
TP63	-1,06		0,056	1,16E-08
SMAD4	-1,225		0,535	1,17E-08
FOXO4	1,178		0,933	1,28E-08
ERK			0,968	5,45E-08
NR3C1	-1,159		0,153	8,22E-08
Lh			1,71	1,31E-07
TP53	2,235		0,097	1,97E-07
TNF	1,265	Activated	2,112	2,01E-07
P38 MAPK			1,198	2,27E-07
STAT3	1,111		-0,415	6,34E-07
CD40LG	1,009	Activated	2,262	1,10E-06
FOXL2	-1,084	Activated	3,138	1,52E-06
Pkc(s)			1,549	1,70E-06
F7	-1,012		1,005	2,25E-06
F2	1,144		0,789	2,51E-06
ITGB1	1,302		-0,051	3,32E-06
SUPT16H	-1,08			3,74E-06
SSRP1	1,009			3,74E-06
FOXO1	1,051		1,437	3,90E-06
EGR1	3,531		-0,335	4,79E-06
TCR			1,689	1,45E-05
RBPJ	1,044		0,928	1,92E-05
ERK1/2			0,441	2,11E-05
MAP2K1/2			0,543	2,31E-05
BCR (complex)			0,315	2,31E-05
SRF	-1,629			2,31E-05
NOX4	-1,051		0,128	3,19E-05
NDRG1	1,527		-1,231	3,19E-05
CCL5	1,037		1,238	3,92E-05
TP73	1,16		1,16	4,54E-05

2 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
Akt			0,564	5,04E-05
IGF1	1,004		0,823	5,21E-05
SP1	-1,058		0,477	5,75E-05
ERBB2	-1,043		0,63	7,48E-05
S100A6	1,185		-1,633	7,77E-05
Cyclin E				8,16E-05
MAP3K3	1,027			8,16E-05
GAST	1,121			8,16E-05
IgG		Inhibited	-2,879	9,12E-05
ELK1	-1,003		0,152	9,33E-05
ESR1	-1,197		-1,473	9,68E-05
CD40	-1,025		-0,664	1,11E-04
MNT	-1,006	Inhibited	-2	1,11E-04
PI3K (family)			0,486	1,13E-04
NEDD9	1,012	Activated	2,646	1,17E-04
KIAA1524	1,138		-1,093	1,90E-04
MYC	-1,006		-0,335	2,00E-04
MUC1	-1,155		0,271	2,14E-04
KLF6	1,161		0,204	2,14E-04
PPP2R5C	1,069	Inhibited	-2,236	2,19E-04
ATF4	1,135			2,19E-04
CREBBP	-1,183			2,39E-04
NFkB (complex)		Activated	3,104	2,56E-04
TLR7	1,005		1,695	2,71E-04
SMARCA4	-1,089		-0,132	2,82E-04
CD3		Inhibited	-2,753	2,90E-04
COL18A1	-1,099		-1	3,07E-04
Cdk				3,16E-04
A2M	-1,143		1,342	3,17E-04
CXCL12	-1,097		0,293	3,24E-04
Igm				3,51E-04
EGF	-1,146		1,538	3,85E-04
BRAF	1,035		-1,969	4,45E-04
PPRC1	-1,144	Activated	2,611	4,48E-04
CAMP	-1,14		0,138	4,51E-04
GSK3B	1,019		-0,672	4,51E-04
FAS	-1,08		-1,331	4,51E-04
Cg		Activated	2,051	4,66E-04
Jnk			1,221	5,36E-04
EGFR	-1,194		1,055	5,36E-04
SBDS	1,146			5,87E-04
SIN3B	-1,077			6,09E-04
WT1	-1,062			6,09E-04

2 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
EPAS1	-1,272			6,31E-04
KRAS	1,374		0	6,54E-04
SIRT6	1,051			7,64E-04
ITGA5	-1,245			7,64E-04
EIF3E	1,268			7,64E-04
mir-7				7,64E-04
AKT3	-1,131			7,64E-04
SMAD2	1,187		1,987	8,13E-04
IL13	-1,059		-0,322	8,28E-04
MAX	1,17	Inhibited	-2	8,87E-04
HGF	-1,113		1,637	9,36E-04
PI3K (complex)			0,055	9,80E-04
Growth hormone			1,89	1,29E-03
IL12B	-1,018	Activated	2	1,32E-03
NCR2	-1,193		1,067	1,32E-03
BMP2	-1,069		1,03	1,32E-03
AR	-1,126		-0,832	1,32E-03
VPRBP	-1,171	Inhibited	-2	1,32E-03
TRIB3	1,034	Inhibited	-2	1,32E-03
CDK9	1,049			1,32E-03
KDM8	1,148			1,48E-03
SYK	-1,068			1,48E-03
ATG7	1,01			1,48E-03
ZBTB17	-1,082			1,48E-03
TGM2	-1,155		1,01	1,49E-03
ELAVL1	-1,15		1,974	1,50E-03
PTGS2	1,939		0,672	1,50E-03
IL1A	1,277		-0,07	1,68E-03
SIN3A	1,115	Inhibited	-2	1,73E-03
IL1B	-1,143		1,639	1,78E-03
FFAR4	-1,044			1,88E-03
CDCA2	1,26			1,88E-03
ITCH	-1,057			1,88E-03
BGN	-1,016			1,88E-03
CCL21	-1,062			1,88E-03
HSPB3	1,085			1,88E-03
HS6ST1	-1,063			1,88E-03
HSPB6	1,075			1,88E-03
CKS1B	1,329			1,88E-03
LPAR2	1,034			1,88E-03
HSPB8	1,088			1,88E-03
BECN1	1,021			1,88E-03
HS6ST2	-1,028			1,88E-03

2 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
CREB1	-1,012		1,974	1,90E-03
estrogen receptor			0,535	2,09E-03
SELPLG	1,307		1,342	2,16E-03
IL17A			1,06	2,26E-03
MAPK9	-1,247		-1,89	2,32E-03
GDF2	-1,029			2,50E-03
ERBB4	1,042			2,50E-03
PCGF2	1,025			2,50E-03
IL5	-1,159		1,664	2,66E-03
CSF2	-1,026		1,913	2,72E-03
RELA	1,071	Activated	2,066	2,93E-03
IL4	-1,026		1,673	3,05E-03
CD28	-1,34		-1,701	3,05E-03
PAF1		Activated	2,449	3,11E-03
HNRNPA2B1	1,232			3,15E-03
FGF2	-1,468		0,374	3,25E-03
UPF2	1,012	Inhibited	-2,213	3,25E-03
APP	-1,03		1,091	3,28E-03
PRDM5	1,067			3,79E-03
IL1				3,88E-03
IL31	1,061			3,88E-03
KDM5A	1,163			3,88E-03
TLR9	-1,069			3,88E-03
HDAC4	-1,114			3,88E-03
Immunoglobulin			0,447	3,92E-03
MGEA5	-1,026	Inhibited	-2	4,00E-03
JUN	1,91		0,933	4,21E-03
PRKCE	-1,055		1,067	4,22E-03
IGF1R	-1,391			4,22E-03
JUNB	1,718			4,22E-03
PRKCA	-1,407		1,387	4,82E-03
WBSCR22	1,003			5,49E-03
DGKH	1,077			5,49E-03
IRS1	-1,328			5,49E-03
NOD1	-1,314			5,49E-03
HSPB2				5,49E-03
PYCARD	1,057			5,49E-03
HOXB9	1,114			5,49E-03
TNFRSF25				5,49E-03
LIPE	-1,025			5,49E-03
PLA2G4A	1,206			5,49E-03
ERBB3	1,055			5,49E-03

2 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
N4BP1	-1,161			5,49E-03
TNFRSF12A	-1,446			5,49E-03
RGCC	1,076			5,49E-03
GH1	1,023			5,63E-03
VTN	-1,041			5,63E-03
EZH2	1,092	Activated	2,562	6,27E-03
SMAD3	-1,076			6,29E-03
mir-10	1,19		-0,332	6,51E-03
MYB	-1,143			6,63E-03
Mek			-0,023	6,88E-03
FANCC	1,028			7,13E-03
TLR4	-1,234		1,129	7,59E-03
IFNG	-1,12	Activated	2,056	7,73E-03
TRAF2	-1,107			7,78E-03
CXCR4	1,789			7,78E-03
PPARGC1A	1,275		1,067	8,11E-03
HDAC6	1,063		0	8,11E-03
CHI3L1	-1,067			8,11E-03
ATM	1,252			8,11E-03
MYCN	-1,004			8,11E-03
SOX2				8,79E-03
Notch			0,747	9,80E-03
YAP1	-1,272		0,283	9,80E-03
SATB1	-1,124	Inhibited	-2,228	9,81E-03
HSF1	-1,115		-1,809	3,63E-02

4 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
PDGF BB		Activated	4,566	2,79E-20
NUPR1	1,259	Activated	5,696	1,45E-14
TREM1	1,077	Activated	4,105	7,14E-12
P38 MAPK			1,433	3,08E-10
TNF	1,582	Activated	4,05	1,69E-09
F7	1,04	Activated	2,738	1,13E-08
TGFB1	1,137		-0,385	2,53E-08
FOXO3	1,229	Activated	2,146	8,04E-08
PI3K (complex)			0,445	1,74E-07
ITGB1	1,027		0,429	2,71E-07
SMAD4	-1,502		-0,068	3,20E-07
FOXO4	1,236		1,416	3,22E-07
RELA	1,015	Activated	3,293	6,65E-07
FSH		Activated	2,049	9,62E-07
Hdac		Inhibited	-3,276	1,08E-06
estrogen receptor			-0,284	1,14E-06
ERK1/2		Activated	2,25	1,40E-06
SCD	-1,187	Inhibited	-2,378	4,82E-06
TLR7	-1,1	Activated	3,263	1,04E-05
TGM2	-1,273	Activated	3,311	1,19E-05
ATF4	1,21	Activated	2,2	1,20E-05
TP63	1,146		0,165	1,25E-05
NR3C1	1,244		-0,849	1,29E-05
miR-155-5p		Inhibited	-3,123	1,34E-05
CREB1	-1,096	Activated	2,416	1,37E-05
S100A6	1,165		-1,667	1,40E-05
F2	1,106		0,46	1,66E-05
EGF	1,022	Activated	2,465	1,98E-05
IL17A		Activated	3,341	2,00E-05
HIF1A	-1,1	Activated	2,25	2,35E-05
SRF	-1,423			3,08E-05
EGR1	3,397		1,319	3,41E-05
FAS	-1,321		-0,431	3,41E-05
NFKB1	2,293	Activated	2,129	3,59E-05
JUN	1,842	Activated	2,131	3,82E-05
MAP2K1/2		Activated	2,066	3,93E-05
Lh		Activated	2,915	5,31E-05
FOXL2	-1,013	Activated	3,441	5,65E-05
CREBBP	-1,102			5,68E-05
ERK			1,907	5,89E-05
ERBB2	-1,126		1,427	7,01E-05
PTAFR	1,399		1,954	8,27E-05

4 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
TEAD			-1,067	8,27E-05
Mek			0,088	9,04E-05
TCR			1,591	1,22E-04
ATF2	-1,085		1,671	1,34E-04
PPRC1	-1,053	Activated	2,722	1,54E-04
Jnk			1,334	1,59E-04
FOXO1	1,092		1,293	1,73E-04
ESR1	1		-0,891	1,95E-04
CD24		Inhibited	-3	2,06E-04
PAF1		Activated	2,111	2,08E-04
SMARCA4	-1,242			2,12E-04
BCR (complex)			0	2,23E-04
IL27	1,016		1,517	2,77E-04
CD40LG	1,039	Activated	2,665	3,23E-04
Pka			0	3,44E-04
SIRT6	-1,001		1,969	3,82E-04
miR-130a-3p			1,331	3,82E-04
mir-7			0,039	3,82E-04
FASN	-1,304		-1,732	3,82E-04
SP1	-1,262		1,004	3,88E-04
KIAA1524	-1,418		-1,551	4,12E-04
NCR2	1,082		1,709	4,17E-04
CCL2	-1,292		-1,342	4,17E-04
IgG		Inhibited	-2,399	4,61E-04
MYC	-1,037		-1,457	5,94E-04
let-7	1,079		1,8	6,12E-04
EGFR	1,901		0,231	6,36E-04
ERG	1,05	Inhibited	-2	6,94E-04
COL18A1	-1,059		-1,387	7,76E-04
Growth hormone			1,508	7,80E-04
SP3	-1,107		0,882	8,21E-04
NFkB (complex)		Activated	3,742	8,23E-04
Pkc(s)			1,756	8,27E-04
mir-10	1,164		1,615	8,27E-04
IRS1	-1,074			8,68E-04
MAP3K3	-1,074			8,68E-04
STAT3	1,085		0,722	9,46E-04
PTGS2	2,833	Activated	2,061	9,84E-04
IL1RN	1,111	Inhibited	-2,714	9,84E-04
CD3		Inhibited	-2,529	1,21E-03
IGF1	-1,147		1,686	1,23E-03
NDRG1	1,605		-1,432	1,32E-03

4 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
IL1B	1,014	Activated	3,749	1,49E-03
CD40	1,018		-0,216	1,65E-03
Immunoglobulin		Activated	2,138	1,66E-03
YAP1	-1,285		0,221	1,81E-03
TP53	1,63		-0,37	1,90E-03
PRKCA	1,002	Activated	2,382	1,91E-03
MNT	1,09	Inhibited	-2	2,28E-03
CXCL8	3,818		0,675	2,49E-03
RBPJ	1,213		0,186	2,49E-03
TP73	-1,27		0,698	2,75E-03
PI3K (family)			0,468	2,77E-03
MAX	1,21	Inhibited	-2	2,87E-03
SMAD3	1,278		-0,135	2,93E-03
Calcineurin protein(s)				3,23E-03
AMH				3,23E-03
FANCA	-1,268			3,23E-03
ZNF148	-1,227			3,23E-03
miR-7a-5p				3,23E-03
NFIL3	2,016			3,23E-03
IL18RAP	1,15			3,23E-03
TAF4B	1,099			3,23E-03
S1PR3				3,23E-03
S100A7	1,242			3,23E-03
BRD4	-1,065		-1,667	3,34E-03
SELPLG	2,448		1,89	3,35E-03
IL12B	1,096	Activated	2,219	3,51E-03
CDK9	-1,154		1,969	3,51E-03
TRIB3	1,082	Inhibited	-2,186	3,51E-03
NFKBIA	1,576			3,51E-03
EIF2AK2	-1,401		-0,24	3,58E-03
SIN3B	-1,228			3,67E-03
WT1	-1,006			3,67E-03
TFAP2A	-1,038			3,72E-03
DDX58	1,971		1,067	4,23E-03
CAV1	-1,056		-0,159	4,23E-03
KLF6	1,209			4,23E-03
TLR3	-1,133		1,939	4,24E-03
Igm			-1,925	4,43E-03
KRAS	-1,28		1,067	4,59E-03
TLR4	-1,296		1,978	4,67E-03
SPDEF	-1,004		1,3	4,74E-03
SMAD2	-1,058		1,133	5,07E-03

4 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
TBP	-1,124			5,07E-03
PPARD	1,11		1,41	5,27E-03
FOS	3,242	Activated	2,236	5,59E-03
Cg		Activated	2,749	5,61E-03
EPAS1	1,302		1,11	5,78E-03
WISP2	-1,166		-0,707	5,84E-03
IFNG	-1,161	Activated	2,946	6,27E-03
EZH2	1,172	Activated	3,03	6,68E-03
IRF3	1,111		1,751	6,82E-03
MYCN	1,097		0,849	6,82E-03
HDAC6	-1,061		0	6,82E-03
OSM	1,069		-0,425	6,82E-03
CHI3L1	-1,097			6,82E-03
ATM	-1,286			6,82E-03
TGFA			1,987	7,03E-03
PF4	1,148		1,987	7,03E-03
IL32	1,005		1,957	7,03E-03
SMAD7	-1,001		1	7,03E-03
CCL5	1,066		1,414	7,23E-03
ITGA5	-1,204			7,49E-03
AKT3	1,059			7,49E-03
TLR5	-1,064			7,49E-03
XIAP	-1,105			7,49E-03
TAZ	1,01			7,49E-03
TLR6	1,019			7,49E-03
Fcer1		Activated	2,202	7,56E-03
PPP2R5C	-1,041	Inhibited	-2,236	7,56E-03
Akt			1,998	8,85E-03
IL1A	1,596	Activated	2,71	8,91E-03
MAP3K14	-1,088		1,231	8,97E-03
CEBPA	1,101	Activated	2,514	9,05E-03
FFAR4	1,039			9,11E-03
TNIP3	-1,408			9,11E-03
ZC3HAV1	-1,554			9,11E-03
Eotaxin				9,11E-03
FLII	-1,254			9,11E-03
H6PD	-1,032			9,11E-03
TRPM2	1,048			9,11E-03
ITCH	-1,038			9,11E-03
GJA1	1,029			9,11E-03
MAPKAPK2	1,016			9,11E-03

4 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
DNAJA3	-1,303			9,11E-03
LPAR2	1,014			9,11E-03
DSG1	-1,005			9,11E-03
NTRK1	1,117			9,11E-03
GNA13	-1,275			9,11E-03
PLA2G2A	1,116			9,11E-03
SLC37A4	-1,121			9,11E-03
LPAR3	1,057			9,11E-03
INHBC	1,962			9,11E-03
CD300C	1,054			9,11E-03
PGR	-1,007		0,99	9,54E-03

24 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
NUPR1	1,459	Activated	7,988	2,06E-16
TGFB1	-1,021		-0,03	8,08E-16
PDGF BB		Activated	2,987	9,34E-15
TNF	1,092	Activated	3,22	8,98E-14
ERBB2	-1,132		0,981	1,00E-13
TREM1	1,093	Activated	3,767	6,78E-13
TP53	-1,282	Activated	2,364	2,40E-12
estrogen receptor		Inhibited	-2,287	3,89E-12
FOXO1	1,349		-0,397	4,04E-12
MYC	1,009		0,371	8,86E-12
ERK1/2			1,932	2,28E-11
Cg		Activated	2,965	3,20E-11
ERK			1,045	7,15E-11
KRAS	-1,511		1,369	1,02E-10
JUN	1,999		1,446	3,44E-10
P38 MAPK			-0,659	1,17E-09
TP63	-1,086		1,675	2,57E-09
PPRC1	1,184	Activated	3,347	4,19E-09
ESR1	-1,009		-0,126	6,72E-09
S100A6	1,203	Inhibited	-2,714	9,04E-09
FSH			1,346	1,43E-08
TRIB3	2,172	Inhibited	-2,339	1,47E-08
HIF1A	-2,233		0,275	1,90E-08
Akt			0,158	2,55E-08
SMARCA4	-1,014			2,92E-08
KIAA1524	-1,275		1,042	3,70E-08
Lh			1,635	4,36E-08
FOS	1,583		0,82	5,27E-08
NR3C1	-1,113		0,51	8,20E-08
MGEA5	-1,154		1,521	9,86E-08
ATF4	1,279		1,706	1,16E-07
F7	1,05	Activated	3,248	1,51E-07
Mek		Activated	2,392	1,98E-07
SBDS	1,166			3,36E-07
MYOC	1,161			4,77E-07
MAP2K1/2			1,994	5,41E-07
CD24		Inhibited	-2,714	5,41E-07
COL18A1	1,052		-0,5	6,29E-07
NFkB (complex)		Activated	2,441	6,90E-07
IL1A	1,775		1,349	8,45E-07
STAT3	-1,276		-0,177	1,03E-06
EGF	-1,21	Activated	2,88	1,39E-06

24 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
NDRG1	1,49		-1,254	1,41E-06
PI3K (complex)			-0,059	1,49E-06
FOXO3	1,253		1,288	1,68E-06
SP1	-1,045		0,92	1,85E-06
EGR1	1,59	Activated	2,067	4,53E-06
SMAD4	-1,934		1,213	4,73E-06
SYVN1	1,172		-0,626	5,94E-06
IgG			-1,235	6,11E-06
BRAF	-1,202		0,505	8,07E-06
SRC	1,006		0,299	8,13E-06
CCL5	-1,052		0,777	9,55E-06
IL13	1,125		0,009	1,01E-05
ESR2			0,651	1,04E-05
GH1	1,113		1,516	1,31E-05
TAB1	-1,221		-0,956	1,65E-05
MTOR	-1,058		-1,348	1,65E-05
PI3K (family)			0,036	1,94E-05
AR	-1,102		0,314	2,26E-05
MAP2K1	-1,031		1,463	2,53E-05
ITGB1	-1,393		1,39	2,53E-05
TFAP2A	-1,204			2,61E-05
NOX4	-1,615		-0,035	3,07E-05
TRAF2	1,056		-0,97	3,07E-05
AMH				3,33E-05
IGF1	-1,869		0,329	3,50E-05
KDM5B	-1,083	Activated	3,626	4,36E-05
SCD	-2,037	Inhibited	-2,183	4,65E-05
MAPK9	-1,501		-1,772	5,90E-05
PRKCA	1,057		0,896	6,31E-05
F2	1,102		0,721	6,31E-05
CTNNB1	-1,363		0,156	7,16E-05
E2F4	1,221			7,17E-05
EGFR	2,473	Activated	2,458	7,55E-05
PPARG	1,168		0,526	8,28E-05
PRKCE	1,271		1,941	8,95E-05
Ap1		Activated	2,236	9,04E-05
RELA	1,219	Activated	2,07	9,45E-05
Growth hormone			0,389	1,07E-04
Creb				1,07E-04
INS			0,851	1,10E-04
PRKCD	1,01		0,777	1,10E-04
IKBKE	-1	Activated	2,429	1,18E-04

24 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
VPRBP	-1,226		0	1,18E-04
ACSL5	1,248		0	1,18E-04
Gsk3			-1,098	1,33E-04
HIC1	1,115		-0,142	1,38E-04
PTEN	-1,166		0,314	1,42E-04
HLX	1,058		-1,89	1,42E-04
SND1	-1,024			1,42E-04
Hdac			-1,17	1,50E-04
Pkc(s)			1,977	1,51E-04
LGALS3	1,283			1,51E-04
FAT1	-1,153		1,969	1,56E-04
mir-7			0,144	1,56E-04
FOXL2	1,021	Activated	2,506	1,96E-04
FOXM1	-1,064		-1,755	2,04E-04
FOSL1	2,393		-0,44	2,05E-04
TCR			1,007	2,29E-04
HNRNPA2B1	-1,175			2,40E-04
GSK3B	-1,109		-0,109	2,66E-04
FAS	1,106		-1,803	2,66E-04
WISP2	-2,053		-1	2,72E-04
HMOX1	1,315		-1,108	2,72E-04
ATM	-1,59		0	3,22E-04
CHI3L1	-1,103			3,22E-04
ATF2	-1,804		1,091	3,36E-04
PGR	1,075		0,596	3,38E-04
NFKB1	1,006			3,46E-04
E2F1	-1,08			3,66E-04
IL1B	1,107	Activated	2,777	3,80E-04
PDLIM2	1,041		1,604	4,29E-04
SYK/ZAP				4,39E-04
IRS1	1,34			4,39E-04
KLF9	-1,442			4,39E-04
MAP3K3	-1,117			4,39E-04
LIPE	1,039			4,39E-04
SF1	1,076			4,39E-04
PPP2R5B	1,117			4,39E-04
MIF	-1,011		0,663	4,40E-04
EDNRB	1,037		0,141	4,40E-04
SYK	1,058		-0,689	4,40E-04
PTTG1	-1,213			4,59E-04
IL17A		Activated	2,265	4,65E-04
MITF	-1,475		-1,912	4,93E-04

24 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
ZNF217	-1,049		0,378	5,05E-04
IL6	1,626		-0,823	5,17E-04
TP73	1,03		-0,401	5,77E-04
DICER1	1,024			6,36E-04
SPDEF	1,176		1,808	7,70E-04
NRG1	1,036		0,667	7,89E-04
HGF	1,084	Activated	2,316	7,94E-04
TCF				8,56E-04
AHR	1,749		1,941	8,90E-04
MAP3K7	1,05		-0,64	8,90E-04
miR-155-5p		Inhibited	-2,621	8,90E-04
SP3	-1,127		0,958	9,03E-04
SUPT16H	-1,293			1,14E-03
WT1	1,055			1,14E-03
Histone h4				1,18E-03
PPARA	1,032		-1,109	1,20E-03
CHUK	1,051	Activated	2,2	1,27E-03
CDK9	-1,031		0,692	1,27E-03
NAMPT	1,55		0,664	1,27E-03
ELK1	1,104		0,555	1,27E-03
JUND	-1,029			1,27E-03
CREBBP	1,094			1,41E-03
Jnk			1,772	1,42E-03
NEUROG1	1,008		0	1,42E-03
GLI1	-1,004		-0,374	1,59E-03
MAPK8	1,04		1,633	1,60E-03
MYB	-1,048			1,60E-03
IFNG	-1,048	Activated	2,882	1,65E-03
Calcineurin protein(s)				1,65E-03
E2F8	-1,14			1,65E-03
Cdk				1,65E-03
GNE	-1,124			1,65E-03
PTAFR	1,288			1,65E-03
PPP2R2A	1,069			1,65E-03
NFIL3	-1,037			1,65E-03
EREG	3,033			1,65E-03
DDX17	-1,271			1,65E-03
IL18RAP	1,033			1,65E-03
RIPK1	1,076			1,65E-03
PDCD4	-2,371			1,65E-03
CREB1	-1,067			1,77E-03
CDKN1A	1,387		1,316	1,79E-03

24 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
KLF6	1,137		0,204	1,82E-03
RPS6KB1			-1	1,82E-03
FOSL2	1,121			1,82E-03
YBX1	1,034	Activated	2	1,93E-03
FOXO4	-1,474		1,298	1,93E-03
Hsp27			1,154	1,93E-03
ATF3	2,241		0,669	1,93E-03
PSMD10	-1,176		0,447	1,93E-03
SPP1	1,107		-1,103	1,93E-03
AZGP1	1,063		-1,446	1,93E-03
Mapk				1,93E-03
HDAC1	-1,131			2,03E-03
CD44	1,503		0,266	2,18E-03
SREBF1				2,21E-03
TLR7	-1,178	Activated	2,46	2,61E-03
PTGS2	1,919		0,111	2,61E-03
UXT	-1,012		0,873	2,62E-03
ETV5	2,328			2,62E-03
NRXN1	1,071			2,62E-03
TWIST2	-1,186		0,314	2,82E-03
CTGF	-1,073		0,169	2,82E-03
PPP2R5C	-1,166	Inhibited	-2,236	2,82E-03
POU2F1	-1,115			2,82E-03
CEBPD	-1,34			2,82E-03
CEBPA	1,17		1,929	2,96E-03
TGFA			1,987	3,07E-03
PF4	1,092		0,927	3,07E-03
ITGB3	1,126		-0,278	3,07E-03
miR-296-5p			-1	3,07E-03
RUVBL1	1,005		0,378	3,31E-03
let-7	-2,436		0,169	3,31E-03
MAX	1,035			3,31E-03
PDX1	-1,055			3,31E-03
CDK4				3,48E-03
SRF	-1,094			3,74E-03
MAPK1	-1,003		-0,883	3,76E-03
IRF1	1,357		0,152	3,83E-03
HAND2	1,094		-0,115	3,83E-03
TBX5	1,07		-0,115	3,83E-03
DNMT3B	1,042	Activated	2,538	3,87E-03
SIRT6	1,23			3,90E-03
PBRM1	1,002			3,90E-03

24 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
PROK2	1,171			3,90E-03
EIF3E	-1,403			3,90E-03
AKT3	1,028			3,90E-03
PRKRA	-1,075			3,90E-03
FASN	-1,528			3,90E-03
WAC	-1,164			3,90E-03
TARBP2	1,121			3,90E-03
DDX5				3,90E-03
A2M	-3,083	Activated	2,236	3,97E-03
ADIPOQ	1,054		1,987	3,97E-03
SMARCD3	-1,016		-1,342	3,97E-03
T	-1,287		-1,342	3,97E-03
NFYA				3,97E-03
CCND1	2,325			4,20E-03
CD40LG	-1,015	Activated	2,05	4,43E-03
MMP9	1,029		1,154	4,81E-03
DEF6	1,065		-0,117	4,81E-03
WNT5A	1,01		-0,283	4,81E-03
CBX7	1,077	Inhibited	-2	4,81E-03
NCOR2	-1,058			4,81E-03
SP4	-1,016			4,81E-03
SOCS1	1,457		-1,432	4,93E-03
ERN1	1,447			5,10E-03
JUNB	1,079		-0,277	5,42E-03
ANXA2	-1,141		-0,391	5,42E-03
Rap1				5,78E-03
FFAR4	-1,052			5,78E-03
LTB4R2				5,78E-03
KLHL2	-1,058			5,78E-03
GPBAR1	1,093			5,78E-03
PRKDC	-1,669			5,78E-03
PECAM1	1,011			5,78E-03
AGR2	1,022			5,78E-03
FBLN2	1,079			5,78E-03
miR-183-5p				5,78E-03
miR-208a-3p				5,78E-03
BGN	-1,319			5,78E-03
CCL21	1,066			5,78E-03
GJA1	-1,207			5,78E-03
TLE1	1,054			5,78E-03
HSPB3	-1,944			5,78E-03
KCNMA1	1,136			5,78E-03

24 h				
Régulateur	taux d'induction	prédiction	z-score	p-value
SMAD2	-1,749		1,432	9,35E-03
KLF4	2,048			9,35E-03
FANCC	-1,309			9,40E-03
Immunoglobulin			1,633	9,63E-03

ANNEXE 4 : RESULTATS DE L'ANALYSE FONCTIONNELLE PAR LE LOGICIEL DAVID DES TERMES GO ASSOCIES AUX GENES DONT L'ÉPISSAGE EST MODULE LORS D'UN STRESS OXYDANT AU TEMPS 2H, 4 H ET 24 H.

L'analyse a été réalisée en utilisant les gènes dont l'épissage est modulé 2 h, 4 h ou 24 h après le début d'un stress oxydant. Pour chaque GO identifié, un taux d'enrichissement est calculé et des tests statistiques sont effectués pour évaluer le niveau de significativité de l'enrichissement du terme d'intérêt. A partir de ces résultats, le logiciel regroupe les GO biologiquement reliées entre eux en cluster et associe un score d'enrichissement. Ce score d'enrichissement est directement lié à la p-value calculée pour le groupe fonctionnel selon la formule suivante : $\text{score} = -\log(\text{p-value})$, par conséquent un score d'enrichissement ≥ 2 correspond à une p-value ≤ 0.01 et plus le score d'enrichissement est élevé, plus l'enrichissement de ces termes est significatif. Les groupes fonctionnels indiqués sont ceux qui ont un score d'enrichissement supérieur ou égal à 1,3.

RESULTATS OBTENUS POUR LE TEMPS 2 H

Annotation Cluster 1 Score d'enrichissement : 5,8						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0031974~membrane-enclosed lumen	129	15,28	1,59	2,99E-08	1,37E-05
terme GO CC	GO:0043233~organelle lumen	126	14,93	1,59	5,94E-08	1,36E-05
terme GO CC	GO:0070013~intracellular organelle lumen	123	14,57	1,59	1,00E-07	1,52E-05
terme GO CC	GO:0031981~nuclear lumen	101	11,97	1,60	1,49E-06	1,36E-04
terme GO CC	GO:0005654~nucleoplasm	66	7,82	1,72	1,76E-05	1,34E-03
terme GO CC	GO:0044451~nucleoplasm part	40	4,74	1,65	2,12E-03	9,23E-02

Annotation Cluster 2 Score d'enrichissement : 3,6						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0006366~transcription from RNA polymerase II promoter	27	3,20	2,39	6,20E-05	2,93E-02
terme GO BP	GO:0006351~transcription, DNA-dependent	29	3,44	2,06	4,09E-04	9,34E-02
terme GO BP	GO:0032774~RNA biosynthetic process	29	3,44	2,03	5,08E-04	8,74E-02

Annotation Cluster 3 Score d'enrichissement : 3,4						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	nucleotide-binding	112	13,27	1,55	3,27E-06	2,98E-04
terme GO MF	GO:0017076~purine nucleotide binding	129	15,28	1,38	6,91E-05	5,75E-02
Mots-clé Pir Family	atp-binding	87	10,31	1,53	7,60E-05	4,15E-03
terme GO MF	GO:0032555~purine ribonucleotide binding	121	14,34	1,36	2,98E-04	6,20E-02
terme GO MF	GO:0032553~ribonucleotide binding	121	14,34	1,36	2,98E-04	6,20E-02
terme GO MF	GO:0000166~nucleotide binding	143	16,94	1,31	3,25E-04	5,42E-02
UP SEQ FEATURE	nucleotide phosphate-binding region:ATP	65	7,70	1,57	3,77E-04	1,62E-01
terme GO MF	GO:0001883~purine nucleoside binding	107	12,68	1,38	4,42E-04	6,13E-02
terme GO MF	GO:0001882~nucleoside binding	107	12,68	1,37	5,75E-04	6,80E-02
terme GO MF	GO:0030554~adenyl nucleotide binding	105	12,44	1,37	6,01E-04	6,25E-02
Mots-clé Pir Family	kinase	48	5,69	1,62	1,20E-03	4,29E-02
terme GO MF	GO:0032559~adenyl ribonucleotide binding	97	11,49	1,33	2,29E-03	1,51E-01
terme GO MF	GO:0005524~ATP binding	95	11,26	1,32	3,34E-03	1,74E-01
UP SEQ FEATURE	binding site:ATP	37	4,38	1,58	6,89E-03	7,71E-01

Annotation Cluster 4 Score d'enrichissement : 3,2						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0010033~response to organic substance	68	8,06	1,95	1,49E-07	4,27E-04
terme GO BP	GO:0009719~response to endogenous stimulus	41	4,86	2,10	1,29E-05	1,84E-02
terme GO BP	GO:0009725~response to hormone stimulus	38	4,50	2,15	1,73E-05	1,65E-02
terme GO BP	GO:0032870~cellular response to hormone stimulus	18	2,13	2,80	2,20E-04	7,59E-02
terme GO BP	GO:0043434~response to peptide hormone stimulus	19	2,25	2,56	4,44E-04	8,16E-02
terme GO BP	GO:0032868~response to insulin stimulus	12	1,42	2,49	8,63E-03	3,00E-01
terme GO BP	GO:0032869~cellular response to insulin stimulus	9	1,07	2,74	1,61E-02	3,92E-01
terme GO BP	GO:0008286~insulin receptor signaling pathway	5	0,59	2,80	1,01E-01	7,50E-01
terme GO BP	GO:0046626~regulation of insulin receptor signaling pathway	3	0,36	3,27	2,32E-01	9,02E-01

Annotation Cluster 5 Score d'enrichissement : 3,1						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	nucleus	244	28,91	1,33	5,98E-07	8,17E-05
Mots-clé Pir Family	transcription regulation	125	14,81	1,44	2,78E-05	1,90E-03
Mots-clé Pir Family	Transcription	125	14,81	1,40	7,37E-05	4,47E-03
terme GO MF	GO:0030528~transcription regulator activity	104	12,32	1,42	1,94E-04	5,39E-02
terme GO BP	GO:0051252~regulation of RNA metabolic process	116	13,74	1,33	9,95E-04	1,08E-01
Mots-clé Pir Family	dna-binding	108	12,80	1,34	1,18E-03	4,51E-02
terme GO BP	GO:0006355~regulation of transcription, DNA-dependent	113	13,39	1,32	1,38E-03	1,20E-01
terme GO BP	GO:0045449~regulation of transcription	155	18,36	1,23	2,53E-03	1,79E-01
terme GO BP	GO:0006350~transcription	127	15,05	1,25	4,28E-03	2,27E-01
terme GO MF	GO:0003677~DNA binding	135	16,00	1,19	1,65E-02	4,35E-01
terme GO MF	GO:0003700~transcription factor activity	56	6,64	1,18	1,31E-01	8,67E-01

Annotation Cluster 6 Score d'enrichissement : 3,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0043232~intracellular non-membrane-bounded organelle	148	17,54	1,31	2,24E-04	1,45E-02
terme GO CC	GO:0043228~non-membrane-bounded organelle	148	17,54	1,31	2,24E-04	1,45E-02
terme GO CC	GO:0005856~cytoskeleton	78	9,24	1,30	1,27E-02	2,24E-01

Annotation Cluster 7 Score d'enrichissement : 3,0						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0006357~regulation of transcription from RNA polymerase II promoter	59	6,99	1,68	1,02E-04	4,09E-02
terme GO BP	GO:0051173~positive regulation of nitrogen compound metabolic process	52	6,16	1,67	3,13E-04	9,52E-02
terme GO BP	GO:0045941~positive regulation of transcription	47	5,57	1,73	3,22E-04	8,83E-02
terme GO BP	GO:0031328~positive regulation of cellular biosynthetic process	54	6,40	1,63	4,31E-04	9,09E-02
terme GO BP	GO:0009891~positive regulation of biosynthetic process	54	6,40	1,61	5,98E-04	9,12E-02
terme GO BP	GO:0010628~positive regulation of gene expression	47	5,57	1,68	6,13E-04	8,86E-02
terme GO BP	GO:0045935~positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	48	5,69	1,59	1,49E-03	1,22E-01
terme GO BP	GO:0010557~positive regulation of macromolecule biosynthetic process	49	5,81	1,55	2,36E-03	1,72E-01
terme GO BP	GO:0045893~positive regulation of transcription, DNA-dependent	38	4,50	1,65	2,91E-03	1,89E-01
terme GO BP	GO:0051254~positive regulation of RNA metabolic process	38	4,50	1,64	3,28E-03	2,06E-01
terme GO BP	GO:0010604~positive regulation of macromolecule metabolic process	59	6,99	1,43	5,32E-03	2,56E-01
terme GO BP	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	30	3,55	1,68	7,05E-03	2,80E-01

Annotation Cluster 8 Score d'enrichissement : 2,9						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
INTERPRO	IPR004827:Basic-leucine zipper (bZIP) transcription factor	12	1,42	5,00	2,03E-05	2,55E-02
SMART	SM00338:BRLZ	12	1,42	4,77	2,97E-05	7,49E-03
UP SEQ FEATURE	DNA-binding region:Basic motif	19	2,25	2,73	2,07E-04	1,49E-01
UP SEQ FEATURE	domain:Leucine-zipper	14	1,66	2,95	9,03E-04	2,98E-01
PIR_SUPERFAMILY	PIRSF001719:fos transforming protein	4	0,47	16,30	1,08E-03	3,59E-01
INTERPRO	IPR011616:bZIP transcription factor, bZIP-1	7	0,83	4,98	2,35E-03	6,32E-01
INTERPRO	IPR000837:Fos transforming protein	4	0,47	11,03	4,36E-03	6,72E-01
INTERPRO	IPR011700:Basic leucine zipper	4	0,47	5,88	2,80E-02	9,25E-01
terme GO MF	GO:0043565~sequence-specific DNA binding	34	4,03	1,15	2,66E-01	9,44E-01

Annotation Cluster 9 Score d'enrichissement : 2,8						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	isopeptide bond	29	3,44	2,11	2,89E-04	1,43E-02
Mots-clé Pir Family	ubl conjugation	42	4,98	1,66	1,64E-03	5,14E-02
UP SEQ FEATURE	cross-link:Glycyl lysine isopeptide (Lys-Gly) (interchain with G-Cter in ubiquitin)	18	2,13	2,10	5,75E-03	8,16E-01

Annotation Cluster 10 Score d'enrichissement : 2,5						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO MF	GO:0003712~transcription cofactor activity	32	3,79	1,81	1,60E-03	1,17E-01
terme GO MF	GO:0016563~transcription activator activity	34	4,03	1,71	2,98E-03	1,79E-01
terme GO MF	GO:0008134~transcription factor binding	40	4,74	1,60	3,58E-03	1,75E-01
terme GO MF	GO:0003713~transcription coactivator activity	20	2,37	1,92	8,14E-03	2,84E-01

Annotation Cluster 11 Score d'enrichissement : 2,4						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0051341~regulation of oxidoreductase activity	9	1,07	4,66	5,45E-04	8,80E-02
terme GO BP	GO:0032768~regulation of monooxygenase activity	7	0,83	5,80	9,87E-04	1,12E-01
terme GO BP	GO:0045428~regulation of nitric oxide biosynthetic process	7	0,83	5,37	1,52E-03	1,21E-01
terme GO BP	GO:0050999~regulation of nitric-oxide synthase activity	5	0,59	6,47	6,12E-03	2,74E-01
terme GO BP	GO:0032769~negative regulation of monooxygenase activity	4	0,47	9,21	7,53E-03	2,88E-01
terme GO BP	GO:0051001~negative regulation of nitric-oxide synthase activity	3	0,36	12,43	2,10E-02	4,38E-01
terme GO BP	GO:0051354~negative regulation of oxidoreductase activity	4	0,47	6,37	2,22E-02	4,38E-01

Annotation Cluster 12 Score d'enrichissement : 2,18						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0022402~cell cycle process	51	6,04	1,87	2,25E-05	1,61E-02
terme GO BP	GO:0007049~cell cycle	63	7,46	1,68	5,47E-05	3,10E-02
terme GO BP	GO:0022403~cell cycle phase	37	4,38	1,85	4,42E-04	8,69E-02
terme GO BP	GO:0000278~mitotic cell cycle	33	3,91	1,85	9,85E-04	1,16E-01
Mots-clé Pir Family	cell cycle	34	4,03	1,72	2,99E-03	7,18E-02
terme GO BP	GO:0000279~M phase	28	3,32	1,76	4,82E-03	2,39E-01
terme GO BP	GO:0048285~organelle fission	19	2,25	1,72	2,85E-02	4,91E-01
terme GO BP	GO:0000087~M phase of mitotic cell cycle	18	2,13	1,66	4,32E-02	5,78E-01
terme GO BP	GO:0007067~mitosis	17	2,01	1,60	6,64E-02	6,68E-01
terme GO BP	GO:0000280~nuclear division	17	2,01	1,60	6,64E-02	6,68E-01
Mots-clé Pir Family	mitosis	13	1,54	1,65	9,62E-02	5,46E-01
Mots-clé Pir Family	cell division	17	2,01	1,50	1,06E-01	5,57E-01
terme GO BP	GO:0051301~cell division	20	2,37	1,40	1,20E-01	7,78E-01

Annotation Cluster 13 Score d'enrichissement : 1,99						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0009057~macromolecule catabolic process	58	6,87	1,54	1,08E-03	1,13E-01
Mots-clé Pir Family	ubl conjugation pathway	37	4,38	1,69	2,46E-03	6,21E-02
terme GO BP	GO:0044265~cellular macromolecule catabolic process	52	6,16	1,49	4,29E-03	2,23E-01
terme GO BP	GO:0030163~protein catabolic process	45	5,33	1,50	6,79E-03	2,91E-01
terme GO BP	GO:0051603~proteolysis involved in cellular protein catabolic process	43	5,09	1,48	9,41E-03	3,08E-01
terme GO BP	GO:0044257~cellular protein catabolic process	43	5,09	1,48	1,03E-02	3,14E-01
terme GO BP	GO:0019941~modification-dependent protein catabolic process	40	4,74	1,44	1,86E-02	4,15E-01
terme GO BP	GO:0043632~modification-dependent macromolecule catabolic process	40	4,74	1,44	1,86E-02	4,15E-01
terme GO BP	GO:0006508~proteolysis	53	6,28	1,04	4,49E-01	9,77E-01

Annotation Cluster 14 Score d'enrichissement : 1,99						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0006917~induction of apoptosis	28	3,32	1,81	3,41E-03	2,04E-01
terme GO BP	GO:0012502~induction of programmed cell death	28	3,32	1,81	3,51E-03	2,05E-01
terme GO BP	GO:0010941~regulation of cell death	56	6,64	1,42	6,90E-03	2,86E-01
terme GO BP	GO:0042981~regulation of apoptosis	55	6,52	1,42	8,26E-03	2,96E-01
terme GO BP	GO:0043065~positive regulation of apoptosis	33	3,91	1,59	9,55E-03	3,08E-01
terme GO BP	GO:0043067~regulation of programmed cell death	55	6,52	1,40	9,66E-03	3,07E-01
terme GO BP	GO:0043068~positive regulation of programmed cell death	33	3,91	1,58	1,05E-02	3,15E-01
terme GO BP	GO:0010942~positive regulation of cell death	33	3,91	1,57	1,12E-02	3,27E-01
terme GO BP	GO:0060548~negative regulation of cell death	28	3,32	1,61	1,52E-02	3,81E-01
terme GO BP	GO:0043066~negative regulation of apoptosis	27	3,20	1,58	2,15E-02	4,36E-01
terme GO BP	GO:0006916~anti-apoptosis	18	2,13	1,81	2,16E-02	4,35E-01
terme GO BP	GO:0043069~negative regulation of programmed cell death	27	3,20	1,56	2,49E-02	4,62E-01

Annotation Cluster 15 Score d'enrichissement : 1,87						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO MF	GO:0016564~transcription repressor activity	29	3,44	1,89	1,54E-03	1,24E-01
terme GO BP	GO:0010629~negative regulation of gene expression	38	4,50	1,56	7,00E-03	2,82E-01
terme GO BP	GO:0010605~negative regulation of macromolecule metabolic process	51	6,04	1,44	8,24E-03	2,99E-01
terme GO BP	GO:0016481~negative regulation of transcription	35	4,15	1,58	8,32E-03	2,94E-01
terme GO BP	GO:0051172~negative regulation of nitrogen compound metabolic process	38	4,50	1,52	1,10E-02	3,25E-01
terme GO BP	GO:0009890~negative regulation of biosynthetic process	41	4,86	1,48	1,14E-02	3,28E-01
terme GO BP	GO:0031327~negative regulation of cellular biosynthetic process	40	4,74	1,48	1,34E-02	3,59E-01
terme GO BP	GO:0045934~negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	37	4,38	1,50	1,45E-02	3,76E-01
terme GO BP	GO:0010558~negative regulation of macromolecule biosynthetic process	39	4,62	1,48	1,46E-02	3,76E-01
terme GO BP	GO:0000122~negative regulation of transcription from RNA polymerase II promoter	21	2,49	1,64	3,26E-02	5,14E-01
terme GO BP	GO:0045892~negative regulation of transcription, DNA-dependent	25	2,96	1,45	6,02E-02	6,42E-01
terme GO BP	GO:0051253~negative regulation of RNA metabolic process	25	2,96	1,43	6,99E-02	6,78E-01

Annotation Cluster 16 Score d'enrichissement : 1,84						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	Cholesterol biosynthesis	6	0,71	7,34	1,01E-03	4,18E-02
terme GO BP	GO:0006695~cholesterol biosynthetic process	7	0,83	5,58	1,23E-03	1,15E-01
Mots-clé Pir Family	sterol biosynthesis	6	0,71	5,58	3,74E-03	8,19E-02
terme GO BP	GO:0016126~sterol biosynthetic process	7	0,83	4,14	6,01E-03	2,75E-01
KEGG_PATHWAY	hsa00900:Terpenoid backbone biosynthesis	5	0,59	5,97	7,87E-03	4,83E-01
Mots-clé Pir Family	lipid synthesis	11	1,30	2,61	9,38E-03	1,74E-01
terme GO BP	GO:0008203~cholesterol metabolic process	11	1,30	2,48	1,30E-02	3,57E-01
terme GO BP	GO:0008299~isoprenoid biosynthetic process	5	0,59	5,18	1,40E-02	3,69E-01
terme GO BP	GO:0016125~sterol metabolic process	11	1,30	2,26	2,36E-02	4,49E-01
Mots-clé Pir Family	Steroid biosynthesis	6	0,71	3,49	2,74E-02	2,81E-01
terme GO BP	GO:0008202~steroid metabolic process	17	2,01	1,74	3,52E-02	5,34E-01
terme GO BP	GO:0006694~steroid biosynthetic process	9	1,07	2,19	5,20E-02	6,12E-01
terme GO BP	GO:0006720~isoprenoid metabolic process	6	0,71	2,83	5,92E-02	6,37E-01
terme GO BP	GO:0008610~lipid biosynthetic process	15	1,78	0,96	6,97E-01	9,98E-01

Annotation Cluster 17 Score d'enrichissement : 1,81						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0005912~adherens junction	16	1,90	2,37	3,10E-03	1,21E-01
terme GO CC	GO:0016323~basolateral plasma membrane	19	2,25	2,15	3,30E-03	1,18E-01
terme GO CC	GO:0070161~anchoring junction	17	2,01	2,27	3,41E-03	1,13E-01
terme GO CC	GO:0005925~focal adhesion	12	1,42	2,70	4,73E-03	1,27E-01
terme GO CC	GO:0005924~cell-substrate adherens junction	12	1,42	2,60	6,32E-03	1,42E-01
terme GO CC	GO:0030055~cell-substrate junction	12	1,42	2,46	9,46E-03	1,95E-01
terme GO CC	GO:0030054~cell junction	33	3,91	1,46	2,82E-02	3,44E-01
Mots-clé Pir Family	cell junction	25	2,96	1,46	6,18E-02	4,36E-01
Mots-clé Pir Family	synapse	14	1,66	1,53	1,29E-01	5,98E-01
terme GO CC	GO:0045202~synapse	18	2,13	1,16	3,76E-01	8,51E-01

Annotation Cluster 18 Score d'enrichissement : 1,59						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0005856~cytoskeleton	78	9,24	1,30	1,27E-02	2,24E-01
terme GO CC	GO:0015630~microtubule cytoskeleton	34	4,03	1,42	3,65E-02	3,93E-01
terme GO CC	GO:0044430~cytoskeletal part	54	6,40	1,30	3,70E-02	3,80E-01

Annotation Cluster 19 Score d'enrichissement : 1,44						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO MF	GO:0008092~cytoskeletal protein binding	41	4,86	1,67	1,47E-03	1,31E-01
terme GO CC	GO:0015629~actin cytoskeleton	20	2,37	1,71	2,59E-02	3,29E-01
terme GO MF	GO:0003779~actin binding	21	2,49	1,33	1,65E-01	8,83E-01
Mots-clé Pir Family	actin-binding	14	1,66	1,32	2,66E-01	7,58E-01

RESULTATS OBTENUS POUR LE TEMPS 4 H

Annotation Cluster 1 Score d'enrichissement : 2,29						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0000267~cell fraction	28	10,57	1,88	1,65E-03	3,68E-01
terme GO CC	GO:0005626~insoluble fraction	22	8,30	1,90	5,31E-03	5,22E-01
terme GO CC	GO:0005624~membrane fraction	20	7,55	1,80	1,49E-02	7,50E-01

Annotation Cluster 2 Score d'enrichissement : 2,23						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	PIRSF018208:SSX1 protein	3	1,13	30,77	3,78E-03	2,02E-01
UP SEQ FEATURE	domain:KRAB-related	3	1,13	24,41	6,23E-03	6,00E-01
INTERPRO	IPR003655:Krueppel-associated box-related	3	1,13	24,36	6,25E-03	6,56E-01
INTERPRO	IPR019041:SSXRD motif	3	1,13	21,92	7,74E-03	6,29E-01

Annotation Cluster 3 Score d'enrichissement : 2,07						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
UP SEQ FEATURE	binding site:ATP	19	7,17	2,57	4,74E-04	2,42E-01
Mots-clé Pir Family	nucleotide-binding	40	15,09	1,75	6,47E-04	1,01E-01
UP SEQ FEATURE	active site:Proton acceptor	21	7,92	2,34	7,06E-04	2,41E-01
Mots-clé Pir Family	atp-binding	33	12,45	1,83	1,04E-03	8,17E-02
terme GO MF	GO:0001882~nucleoside binding	39	14,72	1,64	2,05E-03	5,68E-01
UP SEQ FEATURE	nucleotide phosphate-binding region:ATP	25	9,43	1,90	3,13E-03	5,20E-01
terme GO MF	GO:0001883~purine nucleoside binding	38	14,34	1,60	3,33E-03	3,65E-01
terme GO MF	GO:0032559~adenyl ribonucleotide binding	36	13,58	1,63	3,61E-03	3,09E-01
terme GO MF	GO:0004672~protein kinase activity	19	7,17	2,12	3,62E-03	2,57E-01
terme GO BP	GO:0006468~protein amino acid phosphorylation	19	7,17	2,11	3,85E-03	8,73E-01
terme GO MF	GO:0030554~adenyl nucleotide binding	37	13,96	1,59	4,65E-03	2,72E-01
UP SEQ FEATURE	domain:Protein kinase	15	5,66	2,34	5,14E-03	5,78E-01
INTERPRO	IPR000719:Protein kinase, core	15	5,66	2,30	5,82E-03	7,75E-01
terme GO MF	GO:0032555~purine ribonucleotide binding	41	15,47	1,51	6,27E-03	3,07E-01
terme GO MF	GO:0032553~ribonucleotide binding	41	15,47	1,51	6,27E-03	3,07E-01
terme GO MF	GO:0017076~purine nucleotide binding	42	15,85	1,48	7,77E-03	2,99E-01
terme GO MF	GO:0005524~ATP binding	34	12,83	1,56	9,30E-03	2,93E-01
terme GO MF	GO:0004713~protein tyrosine kinase activity	8	3,02	3,26	1,14E-02	3,02E-01
terme GO BP	GO:0006793~phosphorus metabolic process	23	8,68	1,75	1,14E-02	9,53E-01
terme GO BP	GO:0006796~phosphate metabolic process	23	8,68	1,75	1,14E-02	9,53E-01
Mots-clé Pir Family	kinase	18	6,79	1,93	1,28E-02	5,06E-01
terme GO MF	GO:0000166~nucleotide binding	46	17,36	1,39	1,60E-02	3,56E-01
INTERPRO	IPR017441:Protein kinase, ATP binding site	13	4,91	2,09	2,23E-02	8,53E-01
terme GO BP	GO:0016310~phosphorylation	19	7,17	1,76	2,23E-02	9,63E-01
terme GO MF	GO:0004674~protein serine/threonine kinase activity	13	4,91	2,04	2,53E-02	4,60E-01
INTERPRO	IPR017442:Serine/threonine protein kinase-related	11	4,15	2,24	2,56E-02	8,50E-01
Mots-clé Pir Family	serine/threonine-protein kinase	11	4,15	2,13	3,50E-02	5,94E-01
INTERPRO	IPR002290:Serine/threonine protein kinase	8	3,02	2,26	6,42E-02	8,48E-01
INTERPRO	IPR008271:Serine/threonine protein kinase, active site	9	3,40	1,86	1,12E-01	9,37E-01
SMART	SM00220:S_TKc	8	3,02	1,80	1,54E-01	9,59E-01
Mots-clé Pir Family	transferase	24	9,06	1,27	1,87E-01	8,10E-01

Annotation Cluster 4 Score d'enrichissement : 1,63						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
UP SEQ FEATURE	domain:ABC transporter 2	4	1,51	9,15	9,22E-03	6,62E-01
UP SEQ FEATURE	domain:ABC transporter 1	4	1,51	9,15	9,22E-03	6,62E-01
UP SEQ FEATURE	nucleotide phosphate-binding region:ATP 1	4	1,51	8,37	1,18E-02	7,18E-01
UP SEQ FEATURE	nucleotide phosphate-binding region:ATP 2	4	1,51	8,37	1,18E-02	7,18E-01
INTERPRO	IPR003593:ATPase, AAA+ type, core	7	2,64	3,55	1,41E-02	7,66E-01
KEGG_PATHWAY	hsa02010:ABC transporters	4	1,51	6,00	2,74E-02	5,41E-01
INTERPRO	IPR003439:ABC transporter-like	4	1,51	6,09	2,76E-02	7,95E-01
SMART	SM00382:AAA	7	2,64	2,83	3,66E-02	9,93E-01
INTERPRO	IPR017871:ABC transporter, conserved site	4	1,51	5,13	4,27E-02	8,44E-01
COG_ONTOLOGY	Defense mechanisms	3	1,13	7,96	5,04E-02	4,34E-01
terme GO MF	GO:0016887~ATPase activity	9	3,40	1,82	1,21E-01	8,68E-01

Annotation Cluster 5 Score d'enrichissement : 1,54						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO MF	GO:0003779~actin binding	13	4,91	2,70	3,22E-03	4,83E-01
terme GO MF	GO:0008092~cytoskeletal protein binding	15	5,66	2,01	1,69E-02	3,54E-01
terme GO CC	GO:0015629~actin cytoskeleton	9	3,40	2,43	3,19E-02	8,94E-01
terme GO BP	GO:0051493~regulation of cytoskeleton organization	6	2,26	3,26	3,67E-02	8,91E-01
Mots-clé Pir Family	actin-binding	8	3,02	2,39	5,12E-02	6,17E-01
terme GO BP	GO:0033043~regulation of organelle organization	6	2,26	2,04	1,68E-01	9,33E-01

Annotation Cluster 6 Score d'enrichissement : 1,45						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
BIOCARTA	h_gleevecpathway:Inhibition of Cellular Proliferation by Gleevec	4	1,51	10,08	5,77E-03	5,58E-01
BIOCARTA	h_fcer1Pathway:Fc Epsilon Receptor I Signaling in Mast Cells	4	1,51	6,39	2,08E-02	7,73E-01
BIOCARTA	h_pdgfPathway:PDGF Signaling Pathway	3	1,13	5,75	8,86E-02	7,30E-01
BIOCARTA	h_tcrPathway:T Cell Receptor Signaling Pathway	3	1,13	4,23	1,48E-01	8,02E-01

Annotation Cluster 7 Score d'enrichissement : 1,43						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
UP SEQ FEATURE	metal ion-binding site:Calcium 1; catalytic	3	1,13	27,46	4,89E-03	6,16E-01
UP SEQ FEATURE	metal ion-binding site:Calcium 2; via carbonyl oxygen	4	1,51	7,51	1,59E-02	7,90E-01
UP SEQ FEATURE	metal ion-binding site:Calcium 3; via carbonyl oxygen	3	1,13	10,46	3,27E-02	9,25E-01
UP SEQ FEATURE	metal ion-binding site:Calcium 2	4	1,51	5,33	3,89E-02	9,45E-01
INTERPRO	IPR000008:C2 calcium-dependent membrane targeting	6	2,26	3,20	3,96E-02	8,47E-01
INTERPRO	IPR018029:C2 membrane targeting protein	5	1,89	3,45	5,70E-02	8,29E-01
UP SEQ FEATURE	metal ion-binding site:Calcium 3	3	1,13	7,32	6,25E-02	9,81E-01
SMART	SM00239:C2	6	2,26	2,55	8,51E-02	9,81E-01
UP SEQ FEATURE	domain:C2	4	1,51	3,66	9,56E-02	9,91E-01

Annotation Cluster 8 Score d'enrichissement : 1,43						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0006006~glucose metabolic process	8	3,02	3,87	4,64E-03	8,45E-01
terme GO BP	GO:0019318~hexose metabolic process	8	3,02	3,08	1,52E-02	9,54E-01
terme GO BP	GO:0016052~carbohydrate catabolic process	6	2,26	4,07	1,58E-02	9,42E-01
Mots-clé Pir Family	glycolysis	4	1,51	6,70	2,15E-02	6,40E-01
terme GO BP	GO:0046164~alcohol catabolic process	5	1,89	4,56	2,36E-02	9,35E-01
terme GO BP	GO:0006096~glycolysis	4	1,51	6,29	2,52E-02	9,10E-01
terme GO BP	GO:0044275~cellular carbohydrate catabolic process	5	1,89	4,35	2,76E-02	8,94E-01
terme GO BP	GO:0005996~monosaccharide metabolic process	8	3,02	2,66	3,07E-02	8,97E-01
terme GO BP	GO:0006007~glucose catabolic process	4	1,51	5,10	4,31E-02	8,67E-01
terme GO BP	GO:0019320~hexose catabolic process	4	1,51	4,29	6,59E-02	8,63E-01
terme GO BP	GO:0046365~monosaccharide catabolic process	4	1,51	4,16	7,05E-02	8,63E-01
KEGG_PATHWAY	hsa00010:Glycolysis / Gluconeogenesis	3	1,13	3,30	2,26E-01	6,97E-01
terme GO BP	GO:0006091~generation of precursor metabolites and energy	6	2,26	1,42	4,19E-01	9,86E-01

Annotation Cluster 9 Score d'enrichissement : 1,4						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO MF	GO:0004620~phospholipase activity	6	2,26	4,89	7,52E-03	3,20E-01
terme GO MF	GO:0004629~phospholipase C activity	4	1,51	8,73	1,04E-02	3,00E-01
terme GO MF	GO:0016298~lipase activity	6	2,26	4,06	1,60E-02	3,76E-01
UP SEQ FEATURE	domain:PI-PLC X-box	3	1,13	12,20	2,44E-02	8,74E-01
INTERPRO	IPR000909:Phospholipase C, phosphatidylinositol-specific, X region	3	1,13	11,54	2,71E-02	8,28E-01
Mots-clé Pir Family	lipid degradation	5	1,89	4,28	2,93E-02	6,62E-01
terme GO MF	GO:0008081~phosphoric diester hydrolase activity	5	1,89	4,02	3,52E-02	5,57E-01
SMART	SM00148:PLCXc	3	1,13	9,19	4,10E-02	9,38E-01
INTERPRO	IPR017946:PLC-like phosphodiesterase, TIM beta/alpha- barrel domain	3	1,13	8,77	4,52E-02	8,15E-01
terme GO MF	GO:0004435~phosphoinositide phospholipase C activity	3	1,13	8,11	5,18E-02	6,63E-01
terme GO BP	GO:0016042~lipid catabolic process	5	1,89	2,14	2,04E-01	9,47E-01
Mots-clé Pir Family	transducer	8	3,02	0,67	9,54E-01	1,00E+00

RESULTATS OBTENUS POUR LE TEMPS 24 H

Annotation Cluster 1 Score d'enrichissement : 3,53						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0030030~cell projection organization	21	8,08	3,88	4,62E-07	8,15E-04
terme GO BP	GO:0030182~neuron differentiation	19	7,31	2,95	7,82E-05	3,39E-02
terme GO BP	GO:0031175~neuron projection development	14	5,38	3,72	1,02E-04	3,52E-02
terme GO BP	GO:0048666~neuron development	16	6,15	3,21	1,41E-04	4,05E-02
terme GO BP	GO:0000902~cell morphogenesis	16	6,15	3,06	2,40E-04	5,86E-02
terme GO BP	GO:0032989~cellular component morphogenesis	17	6,54	2,91	2,46E-04	5,28E-02
terme GO BP	GO:0007409~axonogenesis	11	4,23	3,87	5,52E-04	1,03E-01
terme GO BP	GO:0000904~cell morphogenesis involved in differentiation	12	4,62	3,34	9,32E-04	1,52E-01
terme GO BP	GO:0048858~cell projection morphogenesis	12	4,62	3,33	9,66E-04	1,44E-01
terme GO BP	GO:0048667~cell morphogenesis involved in neuron differentiation	11	4,23	3,58	1,02E-03	1,39E-01
terme GO BP	GO:0048812~neuron projection morphogenesis	11	4,23	3,51	1,17E-03	1,47E-01
terme GO BP	GO:0032990~cell part morphogenesis	12	4,62	3,19	1,38E-03	1,60E-01
terme GO BP	GO:0007411~axon guidance	7	2,69	4,45	4,85E-03	3,23E-01

Annotation Cluster 2 Score d'enrichissement : 2,96						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0005626~insoluble fraction	26	10,00	2,16	3,50E-04	3,24E-02
terme GO CC	GO:0000267~cell fraction	29	11,15	1,87	1,43E-03	7,76E-02
terme GO CC	GO:0005624~membrane fraction	23	8,85	1,99	2,59E-03	8,75E-02

Annotation Cluster 3 Score d'enrichissement : 2,18						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0003012~muscle system process	9	3,46	3,64	3,32E-03	2,66E-01
terme GO BP	GO:0006936~muscle contraction	8	3,08	3,55	7,29E-03	3,91E-01
terme GO BP	GO:0006939~smooth muscle contraction	4	1,54	8,24	1,22E-02	4,81E-01

Annotation Cluster 4 Score d'enrichissement : 2,02						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0006928~cell motion	20	7,69	2,86	7,06E-05	4,07E-02
terme GO BP	GO:0001667~ameboidal cell migration	4	1,54	7,35	1,67E-02	5,51E-01
terme GO BP	GO:0051674~localization of cell	10	3,85	2,21	3,70E-02	7,08E-01
terme GO BP	GO:0048870~cell motility	10	3,85	2,21	3,70E-02	7,08E-01
terme GO BP	GO:0016477~cell migration	9	3,46	2,22	5,01E-02	7,68E-01

Annotation Cluster 5 Score d'enrichissement : 1,94						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0051270~regulation of cell motion	11	4,23	3,87	5,52E-04	1,03E-01
terme GO BP	GO:0040012~regulation of locomotion	10	3,85	3,54	2,09E-03	2,06E-01
terme GO BP	GO:0051272~positive regulation of cell motion	7	2,69	4,86	3,15E-03	2,66E-01
terme GO BP	GO:0030334~regulation of cell migration	9	3,46	3,62	3,43E-03	2,62E-01
terme GO BP	GO:0030155~regulation of cell adhesion	8	3,08	3,97	4,02E-03	2,87E-01
terme GO BP	GO:0030335~positive regulation of cell migration	6	2,31	4,58	9,83E-03	4,30E-01
terme GO BP	GO:0040017~positive regulation of locomotion	6	2,31	4,16	1,45E-02	5,21E-01
terme GO BP	GO:0048661~positive regulation of smooth muscle cell proliferation	3	1,15	6,58	7,52E-02	8,25E-01
terme GO BP	GO:0048660~regulation of smooth muscle cell proliferation	3	1,15	4,43	1,46E-01	8,88E-01
terme GO BP	GO:0007265~Ras protein signal transduction	3	1,15	1,94	4,56E-01	9,84E-01

Annotation Cluster 6 Score d'enrichissement : 1,90						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0014031~mesenchymal cell development	5	1,92	6,66	6,51E-03	3,94E-01
terme GO BP	GO:0048762~mesenchymal cell differentiation	5	1,92	6,66	6,51E-03	3,94E-01
terme GO BP	GO:0060485~mesenchyme development	5	1,92	6,54	6,97E-03	4,02E-01
terme GO BP	GO:0014033~neural crest cell differentiation	4	1,54	8,24	1,22E-02	4,81E-01
terme GO BP	GO:0014032~neural crest cell development	4	1,54	8,24	1,22E-02	4,81E-01
terme GO BP	GO:0007507~heart development	7	2,69	2,21	9,65E-02	8,48E-01

Annotation Cluster 7 Score d'enrichissement : 1,83						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0045202~synapse	14	5,38	2,75	1,72E-03	7,75E-02
terme GO CC	GO:0044456~synapse part	11	4,23	3,12	2,76E-03	8,29E-02
terme GO CC	GO:0030054~cell junction	15	5,77	2,02	1,62E-02	2,37E-01
Mots-clé Pir Family	synapse	8	3,08	2,82	2,38E-02	4,33E-01
Mots-clé Pir Family	cell junction	11	4,23	2,07	4,11E-02	5,00E-01
terme GO BP	GO:0007268~synaptic transmission	8	3,08	1,82	1,48E-01	8,83E-01

Annotation Cluster 8 Score d'enrichissement : 1,81						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
UP SEQ FEATURE	domain:Laminin G-like 3	4	1,54	14,93	2,26E-03	6,88E-01
UP SEQ FEATURE	domain:Laminin G-like 5	3	1,15	24,89	6,00E-03	6,44E-01
UP SEQ FEATURE	domain:Laminin G-like 1	4	1,54	10,30	6,64E-03	6,25E-01
UP SEQ FEATURE	domain:Laminin G-like 2	4	1,54	10,30	6,64E-03	6,25E-01
UP SEQ FEATURE	domain:Laminin G-like 4	3	1,15	14,93	1,66E-02	7,91E-01
INTERPRO	IPRO12680:Laminin G, subdomain 2	4	1,54	6,56	2,27E-02	8,35E-01
INTERPRO	IPRO13032:EGF-like region, conserved site	10	3,85	2,35	2,72E-02	8,49E-01
INTERPRO	IPRO01791:Laminin G	4	1,54	6,12	2,72E-02	8,14E-01
SMART	SM00282:LamG	4	1,54	4,95	4,59E-02	7,34E-01
INTERPRO	IPRO13320:Concanavalin A-like lectin/glucanase, subgroup	4	1,54	3,58	1,01E-01	9,30E-01

Annotation Cluster 9 Score d'enrichissement : 1,68						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
KEGG_PATHWAY	hsa04510:Focal adhesion	12	4,62	3,61	3,71E-04	4,11E-02
KEGG_PATHWAY	hsa04512:ECM-receptor interaction	7	2,69	5,04	2,32E-03	1,23E-01
UP SEQ FEATURE	region of interest:Domain II and I	3	1,15	28,00	4,71E-03	7,03E-01
terme GO CC	GO:0043256~laminin complex	3	1,15	23,28	6,80E-03	1,60E-01
KEGG_PATHWAY	hsa05222:Small cell lung cancer	6	2,31	4,32	1,16E-02	2,80E-01
Mots-clé Pir Family	basement membrane	4	1,54	7,71	1,48E-02	4,21E-01
KEGG_PATHWAY	hsa05200:Pathways in cancer	12	4,62	2,21	1,67E-02	3,17E-01
terme GO CC	GO:0005605~basal lamina	3	1,15	12,33	2,38E-02	2,77E-01
UP SEQ FEATURE	domain:Laminin EGF-like 3	3	1,15	10,18	3,44E-02	8,80E-01
UP SEQ FEATURE	domain:Laminin EGF-like 1	3	1,15	8,61	4,67E-02	9,25E-01
UP SEQ FEATURE	domain:Laminin EGF-like 2	3	1,15	8,30	5,00E-02	9,29E-01
Mots-clé Pir Family	laminin egf-like domain	3	1,15	7,51	5,97E-02	5,56E-01
terme GO CC	GO:0044420~extracellular matrix part	5	1,92	2,98	8,59E-02	4,23E-01
terme GO CC	GO:0005604~basement membrane	4	1,54	3,58	9,99E-02	4,68E-01
INTERPRO	IPRO02049:EGF-like, laminin	3	1,15	5,04	1,18E-01	9,44E-01
SMART	SM00180:EGF_Lam	3	1,15	4,08	1,66E-01	9,42E-01

Annotation Cluster 10 Score d'enrichissement : 1,66						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO CC	GO:0045202~synapse	14	5,38	2,75	1,72E-03	7,75E-02
terme GO CC	GO:0044463~cell projection part	11	4,23	3,28	1,92E-03	7,44E-02
terme GO CC	GO:0044456~synapse part	11	4,23	3,12	2,76E-03	8,29E-02
terme GO CC	GO:0043197~dendritic spine	4	1,54	7,35	1,66E-02	2,11E-01
terme GO CC	GO:0030425~dendrite	7	2,69	3,00	2,92E-02	3,04E-01
terme GO CC	GO:0014069~postsynaptic density	4	1,54	3,94	8,05E-02	4,31E-01
terme GO CC	GO:0045211~postsynaptic membrane	4	1,54	2,07	3,02E-01	7,60E-01
Mots-clé Pir Family	postsynaptic cell membrane	3	1,15	2,05	4,30E-01	9,26E-01

Annotation Cluster 11 Score d'enrichissement : 1,63						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	cytoskeleton	22	8,46	2,60	1,22E-04	1,34E-02
terme GO CC	GO:0005856~cytoskeleton	32	12,31	1,62	6,85E-03	1,49E-01
terme GO CC	GO:0005815~microtubule organizing center	9	3,46	2,48	2,85E-02	3,10E-01
terme GO CC	GO:0005813~centrosome	8	3,08	2,49	4,16E-02	3,48E-01
terme GO CC	GO:0043228~non-membrane-bounded organelle	46	17,69	1,24	8,32E-02	4,27E-01
terme GO CC	GO:0043232~intracellular non-membrane-bounded organelle	46	17,69	1,24	8,32E-02	4,27E-01
terme GO CC	GO:0044430~cytoskeletal part	20	7,69	1,47	8,45E-02	4,25E-01
terme GO CC	GO:0015630~microtubule cytoskeleton	12	4,62	1,53	1,60E-01	6,12E-01

Annotation Cluster 12 Score d'enrichissement : 1,61						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
Mots-clé Pir Family	collagen	6	2,31	4,75	8,67E-03	3,81E-01
UP SEQ FEATURE	domain:Collagen-like 1	3	1,15	13,18	2,11E-02	8,16E-01
UP SEQ FEATURE	domain:Collagen-like 2	3	1,15	13,18	2,11E-02	8,16E-01
INTERPRO	IPR008160:Collagen triple helix repeat	5	1,92	4,10	3,36E-02	8,19E-01
Mots-clé Pir Family	hydroxylation	4	1,54	4,23	6,83E-02	5,79E-01

Annotation Cluster 13 Score d'enrichissement : 1,59						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
terme GO BP	GO:0007610~behavior	20	7,69	2,90	5,96E-05	5,12E-02
terme GO BP	GO:0007626~locomotory behavior	12	4,62	2,98	2,37E-03	2,18E-01
terme GO BP	GO:0042330~taxis	7	2,69	2,97	3,03E-02	7,00E-01
terme GO BP	GO:0006935~chemotaxis	7	2,69	2,97	3,03E-02	7,00E-01
Mots-clé Pir Family	chemotaxis	3	1,15	3,13	2,47E-01	8,24E-01
Mots-clé Pir Family	cytokine	4	1,54	1,66	4,31E-01	9,24E-01
terme GO MF	GO:0005125~cytokine activity	4	1,54	1,36	5,52E-01	9,93E-01

Annotation Cluster 14 Score d'enrichissement : 1,58						
Catégorie	Terme	Nombre	%	Enrichissement	PValue	Benjamini
KEGG_PATHW AY	hsa04512:ECM-receptor interaction	7	2,69	5,04	2,32E-03	1,23E-01
terme GO CC	GO:0031012~extracellular matrix	12	4,62	2,43	1,05E-02	1,80E-01
Mots-clé Pir Family	extracellular matrix	9	3,46	2,81	1,53E-02	4,00E-01
terme GO CC	GO:0005578~proteinaceous extracellular matrix	11	4,23	2,40	1,63E-02	2,28E-01
Mots-clé Pir Family	cell binding	3	1,15	15,03	1,64E-02	3,91E-01
Mots-clé Pir Family	cell adhesion	12	4,62	2,14	2,58E-02	3,98E-01
UP SEQ FEATURE	short sequence motif:Cell attachment site	5	1,92	4,24	3,02E-02	8,78E-01
terme GO CC	GO:0044420~extracellular matrix part	5	1,92	2,98	8,59E-02	4,23E-01
terme GO BP	GO:0007155~cell adhesion	15	5,77	1,46	1,47E-01	8,84E-01
terme GO BP	GO:0022610~biological adhesion	15	5,77	1,45	1,48E-01	8,84E-01

Matériel et Méthodes

Matériel et Méthodes

Matériel

I. Souches bactériennes, cellules humaines et plasmides.

I.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 a été 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 M15Tn10 (TetR))* (Stratagène). Cette souche a été utilisée pour la production de plasmides.

Souche *E. coli* XL1-Codon Plus (DE3)-RIL : F- *ompt hsdS (rB-mB-) dcm +TetR gal λ (DE3) endAhte [argU ileY leuW CamR]* (Stratagène). Cette souche a été utilisée pour la production la production des protéines recombinantes

I.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, Henrietta Lacks.

Cellules HEK 293T : N°ATCC® : CRL-1573™ (Graham et al., 1977) (Human Embryonic Kidney fibroblast-like) cellules adhérentes issues de tissus rénaux embryonnaires humains. Les lignées utilisées dans ce travail sont des HEK-293T qui expriment l'antigène T de SV40 et permettent donc d'atteindre des niveaux élevés de surexpression de protéines.

I.3. Plasmides utilisés.

Le plasmide pUC18-T7 : Le plasmide pUC18 a servi à réaliser les constructions permettant de purifier les complexes RNP se formant entre différents fragments de 50 nucléotides A, B, C, D, E et F avec un extrait nucléaire, cela par la technique de chromatographie d'affinité MBP-MS2. Les vecteurs produits sont appelés pUC18-T7-hcTNT-MS2 et pUC18-T7-hcTNT (A, B, C, D, E ou F)-MS2.

Le plasmide pSVIRB-hcTNT : Ce plasmide nous a été fourni par l'équipe de N. Sergeant. Il provient du Dr A. Andreadis (Université du Massachusetts, Waltham, Massachusetts, USA). Ce minigène a été décrit dans Andreadis *et al.*, (1993 et 1995). Il contient trois exons du gène de l'insuline de rat et une portion de leurs séquences introniques flanquantes. Différentes tailles de séquences du gène hcTNT ont été insérées au sein d'un site *EcoRI*, situé dans la région intronique entre les exons 2 et 3 de l'insuline. La séquence hcTNT insérée contient l'exon 5 du pré-ARNm hcTNT, avec ces régions introniques flanquant l'exon 5 variant de 50 entre 250 nt. Des mutants ponctuels ont également été réalisés à partir de ces constructions.

Le plasmide RTB300-hcTNT : Ce plasmide a été fourni par l'équipe de T. Cooper (Houston USA). Ce minigène contient l'exon 5 hcTNT, flanqué de séquences introniques de 300 nts en amont et 372 nts en aval, respectivement en amont et en aval de l'exon 5 bordé par des exons du gène *STNI*.

Plasmide pMBP-MS2 : Ce plasmide a été donné par l'équipe de Lührmann. 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 MBPMS2.

pcDNA3.1-GFP-MBNL1 : Ce vecteur permet l'expression du facteur d'épissage MBNL1 dans les lignées cellulaires. Il provient du Dr. N. Charlet (IGBMC, Strasbourg).

Plasmide pXJ41-vide, pXJ41-hnRNP H : Le plasmide pXJ41-hnRNP H a été donné par l'équipe de J. Stevenin (Strasbourg). Ils portent la résistance à l'ampicilline. Il permet la production de la protéine hnRNP H dans les lignées cellulaires. Le pXJ41-vide, qui est le plasmide contrôle a été produit au laboratoire en retirant la séquence codante pour hnRNP H.

II. Anticorps.

Les anticorps utilisés au cours de ce travail sont répertoriés dans le Tableau VI.1.

Anticorps	Provenance	Dilution	Origine	Cible
Anticorps Primaires	Abcam	1/10000	souris	hnRNP H/F
	Abcam	1/1000	lapin	hnRNP H
	Abcam	1/1000	souris	hnRNP U
	Abcam	1/5000	lapin	GAPDH
	ImmuQuest	1/20000	souris	hnRNP K
	Sigma	1/7000	souris	MBNL1
	Sigma	1/20000	souris	β -Tubuline
	ThermoFisher	1/1000	lapin	SNRNP40
	ThermoFisher	1/1000	lapin	U1A
	ThermoFisher	1/1000	lapin	SART1
	ThermoFisher	1/500 (pour IF)	rat	HSF1
	Santa cruz	1/100 (pour IF)	souris	SAF-B/hnRNP HAP
Anticorps Secondaires	Jackson Immuno Research	1/20000	mouton	IgG de souris
	Alexa 555	1/100	chèvre	IgG de souris
	Alexa 482	1/100	chèvre	IgG de lapin
	Pierce	1/20000	chèvre	IgG de lapin

Tableau VI.1 : Caractéristiques et conditions d'utilisation des anticorps. Pour chaque anticorps est référencé une protéine cible, un fournisseur, l'organisme de production ainsi que le facteur de dilution utilisé pour les expériences de Western blot ou d'immunofluorescence (IF).

III. Oligodésoxyribonucléotides et siRNA.

Les oligodésoxyribonucléotides et les siRNA utilisés lors de ce travail sont décrits dans les tableaux VI.2 1-4 et tableaux VI.3.

Tableau VI.2-1 : Désignations, séquences et usages des oligodésoxynucleotides employés lors de ce travail.

Numéro Identifiant	Séquence 5'-3'	Description
67	GTTTCCAGTCACGAC	Oligo universel, appelé "primer -40" ou "-40 M13 forward" (permet un séquençage à l'opposé du 122).
122	CAGGAAACAGCTATGAC	Oligo universel, appelé "primer -21" ou primer: "-21 M13 reverse" (permet un séquençage à l'opposé des oligos 67 ou 967).
5991	CCTGGTGTGTGGGGAGCGTGG	oligo 5' séquençage plasmide SVIRB
7254	TAATACGACTCACTATAGGGCCTTTGTGGTCTCACTTGG	oligo 5' T7 interne exon 2 insuln pour amplification hcTNT pour transcription/épissage in vitro
7255	GCGCTCTGCCGGGCCACT	Oligo 3' interne exon 3 insuln pour amplification hcTNT pour transcription/épissage in vitro
7392	GATACCATGATCACGAAGGTGGT	oligo sens pour analyser le snRNA U1 humain et de souris par qRT-PCR
7393	CACAAAATTATGCAAGTCGAGTTTCC	oligo antisens pour analyser le snRNA U1 humain et de souris par qRT-PCR
7394	GTGCGGAAATCGACATCAAGAG	oligo sens pour analyser le snRNA U11 humain et de souris par qRT-PCR
7395	CGCCGGGACCAACGAT	oligo antisens pour analyser le snRNA U11 humain et de souris par qRT-PCR
7396	AACCTATGAGTAAGGAAAATAACGATTCCG	oligo sens pour analyser le snRNA U12 humain et de souris par qRT-PCR
7397	CGACCTTACCCGCTCAAAA	oligo antisens pour analyser le snRNA U12 humain par qRT-PCR
7399	CGGCCATACCCTGAAC	oligo antisens pour analyser l'ARN 5S humain et de souris par qRT-PCR
7400	GCGGTCTCCCATCCAAGTAC	oligo sens pour analyser l'ARN 5S humain et de souris par qRT-PCR
7566	TCGCTTCGGCAGCACATATACT	oligo sens de RT-PCRq analyse du snRNA U6 de souris et humain
7567	CGCTTACGAATTTGCGTGTCA	oligo anti-sens de RT-PCRq analyse du snRNA U6 de souris et humain
7568	TTTCTTGGGGTTGCGTACTGT	oligo sens de RT-PCRq analyse du snRNA U4 atac de souris et humain
7569	GCACAAAATAAAGCAAAAGCTCTA	oligo anti-sens de RT-PCRq analyse du snRNA U4 atac humain
7574	ATAAAGGATCCAGTGGCTTGGGAATTGG	oligo 5' Bam HI hcTNT intron 4 150nt (clonage MS2)
7575	ATAAAGGATCCAAACCACTGGCTGGTGGCT	oligo 5' Bam HI hcTNT intron 4 100nt (clonage MS2)
7576	ATAAAGGATCCCTAACTGTCTCGCTTTTCC	oligo 5' Bam HI hcTNT intron 4 50nt (clonage MS2)
7577	ATAAAGGATCCAAGAGGAGGACTGGAGAGAGG	oligo 5' Bam HI hcTNT exon 5 1nt (clonage MS2)
7578	ATAAAGGATCCGTAGTACAGCTTTCTTCTG	oligo 5' Bam HI hcTNT intron 5 1nt (clonage MS2)
7579	ATAAAGGATCCGACGCTCCTTGTCCAGG	oligo 5' Bam HI hcTNT intron 5 50nt (clonage MS2)
7580	ATAAAGGATCCAGCGCCAGGGACGGAGTGG	oligo 5' Bam HI hcTNT intron 5 100nt (clonage MS2)
7581	ATAAACTCGAGTAAGAGTGTGGGCTTTGG	oligo 3' Xho I hcTNT intron 4 100nt (clonage MS2)
7582	ATAAACTCGAGTCTGGGAGCAGAGAGC	oligo 3' Xho I hcTNT intron 4 50nt (clonage MS2)
7583	ATAAACTCGAGCTGAGGTTACGGGAGTG	oligo 3' Xho I hcTNT intron 4 1nt (clonage MS2)
7584	ATAAACTCGAGCGTCTTCTGCCTCTCTCC	oligo 3' Xho I hcTNT intron 5 1nt (clonage MS2)
7585	ATAAACTCGAGCTTGGGACAGCAGGCAGCAG	oligo 3' Xho I hcTNT intron 5 50nt (clonage MS2)
7586	ATAAACTCGAGACTCACACCCCCAGCCCCCA	oligo 3' Xho I hcTNT intron 5 100nt (clonage MS2)
7587	ATAAACTCGAGCCCTGGGAGGAGGGAGGAAACG	oligo 3' Xho I hcTNT intron 5 150nt (clonage MS2)
7796	CCGTAAAACGACGGCCAGTCCCTTCCATTCCATTATTATCC	RT-PCR pour la détection d'ARN sat III dans les extraits ARN de cellules HeLa soumises à un Effet des stress sur l'épissage alternatif
7797	TTCCATTCCAATCCTGACTCG	RT-PCR pour la détection d'ARN sat III dans les extraits ARN de cellules HeLa soumises à un Effet des stress sur l'épissage alternatif
7799	CTCTAAAACCACTGCACTGGGTGGCTGCTCTG	Oligo 5' mutagenese YACU du site 94-91 en amont hcTNT
7800	CAGGAGCAGCCACCAAGTGCAGTGGTTTTAAGAG	Oligo 3' mutagenese YACU du site 94-91 en amont hcTNT
7801	ACTGCGTGGTGGTACTCTGCGCGGGCTC	Oligo 5' mutagenese YACU du site 83-80 en amont hcTNT
7802	GAGCCCGGGCAGGAGTAGCCACCCAGCGCAGT	Oligo 3' mutagenese YACU du site 83-80 en amont hcTNT
7803	CCTGCCGGGGCTCTACTCCAGACTAACTG	Oligo 5' mutagenese YACU du site 60-63 en amont hcTNT
7804	CAGGTTAGTCTGGGAGTAGAGCCCGCGGCAGG	Oligo 3' mutagenese YACU du site 60-63 en amont hcTNT
7805	CTCTAAAACCACTGCACTGGGTGGTACTCTGCGCGGGCTC	Oligo 5' mutagenese YACU des sites 94-91 et 83-80 en amont hcTNT
7806	GAGCCCGGGCAGGAGTAGCCACCCAGTGCAGTGGTTTTAAGAG	Oligo 3' mutagenese YACU des sites 94-91 et 83-80 en amont hcTNT
7807	CCTTCTCTGTGGTACTTTCTGCTGCCTGCT	Oligo 5' mutagenese YACU du site 25-30 en aval hcTNT
7808	AGCAGGCAGCAGAAAGTACCACAGAAAGGAAAGG	Oligo 3' mutagenese YACU du site 25-30 en aval hcTNT
7809	TTCTGTGGTCTTCTACTGCTGCTGTCCCAA	Oligo 5' mutagenese YACU du site 32-37 en aval hcTNT

Tableau VI.2-2 : Désignations, séquences et usages des oligodésoxynucléotides employés lors de ce travail.

Numéro Identifiant	Séquence 5'- 3'	Description
7810	TTGGGACAGCAGGCAGTAGAAAAGCACCCACAGAA	Oligo 3' mutagenese YACU du site 32-37 en aval hcTNT
7811	GGTGTCTTCTGCTGCCTACTGTCCCAAGTGCAGC	Oligo 5' mutagenese YACU du site 39-44 en aval hcTNT
7812	GCTGCACCTGGGACAGTAGGCAGCAGAAAAGCACCC	Oligo 3' mutagenese YACU du site 39-44 en aval hcTNT
7813	CCTTCTCTCTGTGTACTTTCTACTGCTGCTGTCCCAA	Oligo 5' mutagenese YACU des sites 25-30 et 32-37 en aval hcTNT
7814	TTGGGACAGCAGGCAGTAGAAAAGTACCACAGAAGGAAAGG	Oligo 3' mutagenese YACU des sites 25-30 et 32-37 en aval hcTNT
7815	CCTTCTCTCTGTGTACTTTCTGCTGCCTACTGTCCCAAGTGCAGC	Oligo 5' mutagenese YACU des sites 25-30 et 39-44 en aval hcTNT
7816	GCTGCACCTGGGACAGTAGGCAGCAGAAAAGTACCACAGAAGGAAAGG	Oligo 3' mutagenese YACU des sites 25-30 et 39-44 en aval hcTNT
7817	CTTCTGTGTGCTTTCTACTGCCTACTGTCCCAAGTGCAGC	Oligo 5' mutagenese YACU des sites 32-37 et 39-44 en aval hcTNT
7818	GCTGCACCTGGGACAGTAGGCAGTAGAAAAGCACCCACAGAAG	Oligo 3' mutagenese YACU des sites 32-37 et 39-44 en aval hcTNT
7862	GCTACAGTCGGAAACCATCAGCAAGCAG	Oligo SENS dans exon 1 Insuline rat, pour analyse RT-PCR SVIRB
7925	CCTTCTCTCTGTGTACTTTCTGCTGCCT	Oligo 5' mutagenese YACU du site 25-30 en aval hcTNT + DG24
7926	AGGCAGCAGAAAAGTACACAGAAGGAAAGG	Oligo 3' mutagenese YACU du site 25-30 en aval hcTNT + DG24
7927	CCTTCTCTCTGTGTACTTTCTACTGCCTGCTGTCCCAA	Oligo 5' mutagenese YACU du site 32-37 en aval hcTNT + DG24
7928	TTGGGACAGCAGGCAGTAGAAAAGCACACAGAAGGAAAGG	Oligo 3' mutagenese YACU du site 32-37 en aval hcTNT + DG24
7929	CCTTCTCTCTGTGTACTTTCTACTGCCT	Oligo 5' mutagenese YACU des sites 25-30 et 32-37 en aval hcTNT + DG24
7930	AGGCAGTAGAAAAGTACACAGAAGGAAAGG	Oligo 3' mutagenese YACU des sites 25-30 et 32-37 en aval hcTNT + DG24
7931	CTTCTGTGTGCTTTCTACTGCCTACTGTCCCAAGTGCAGC	Oligo 5' mutagenese YACU des sites 32-37 et 39-44 en aval hcTNT + DG24
7932	GCTGCACCTGGGACAGTAGGCAGTAGAAAAGCACACAGAAG	Oligo 3' mutagenese YACU des sites 32-37 et 39-44 en aval hcTNT + DG24
7933	GTACAGCCTTCTCTGTGTGCTTTCTGCTGCCT	Oligo 5' DG24 en aval de l'exon 5 hcTNT (mut sur WT ou sur mut 3'3)
7934	AGGCAGCAGAAAAGCACACAGAAGGAAAGGCTGTAC	Oligo 3' DG24 en aval de l'exon 5 hcTNT (mut sur WT ou sur mut 3'3)
8078	TCACCTGCCCCCGGGCCTCC	oligo 3' RT-PCR épissage hcTNT dans SVIRB-utiliser avec le n 5991 mutant 6479
8079	TTGTTCCACATGGCCCTGTGG	Oligo 5' exon 2 IR rat pour RT-PCR
8080	CTTGTGGTCTCCACTTCGCG	Oligo 3' exon 2 IR rat pour RT-PCR
8081	GGGATTTGAGGGACGCTGTGGG	Oligo 5' intron 1 IR rat pour RT-PCR
8082	CCCCCGCACCATATGCACAT	Oligo 3' 3'-UTR IR rat pour RT-PCR
8146	ATAAAGGATCCATGATGTTGGGCACGGAAGGT	oligo 5' clonage hnRNP H pXJ-41 (BamHI)
8147	CTATGCAATGTTTGATTGAAATCA	oligo 3' clonage hnRNP H pXJ-41 BOUT FRANC, STOP INCLUS)
8148	GTAGTAGGGCGCAGAGACGGAGTGGGTACAG	Oligo 5' mutagenese site 1 fragment F HnRNP H
8149	CTGACCCACTCCGTCTCTGCCGCTACTCAC	Oligo 3' mutagenese site 1 fragment F HnRNP H
8150	GTGAGTAGGGCGCAGAGACGGAGTGAGTCAGTCGTTTCCTC	Oligo 5' mutagenese sites 1 et 2 fragment F HnRNP H
8151	GAGGAAACGACTGACTCACTCCGTCTCTGCCGCTACTCAC	Oligo 3' mutagenese sites 1 et 2 fragment F HnRNP H
8152	GGCAGGGACGGAGTGAGTCAGTCGTTTCCTC	Oligo 5' mutagenese site 2 fragment F HnRNP H
8153	GAGGAAACGACTGACTCACTCCGTCCCTGCCGC	Oligo 3' mutagenese site 2 fragment F HnRNP H
8154	TCAGTCGTTTCTCTCTGCTCCAGGG	Oligo 5' mutagenese compensatoire structure site 1 hnRNP H
8155	CCCTGGGAGGCAGAGGAGAAACGACTGA	Oligo 3' mutagenese compensatoire structure site 1 hnRNP H
8208	TAATACGACTCACTATAGGGAGGCGCAGAGACGGAGTG	Oligo T7 5' site F sans seq MS2 + mut site hnRNP H
8209	CCCTGGGAGGCAGAGGAGG	Oligo 3' fragment F mut comp site 1 hnRNP H
8210	TAATACGACTCACTATAGGGAGGCGCAGGACGGAGTG	Oligo T7 5' site F sans seq MS2
8211	CCCTGGGAGGCAGGGGAGG	Oligo 3' fragment F
8217	AGGTGCTGCCGCCGGGCGGTGGCTG	Oligo RT et reverse pour analyse hcTNT splicing RTB300
8218	CATTACCACATTGGTGTGC	Oligo forward pour analyse hcTNT splicing RTB300
8401	CGGCCTTTGGCTAAGATCAAGTG	Oligo sens snRNA U2 murin et humain pour qRT-PCR (publi Workman et al, 2010)
8402	TCCTCGGATAGAGGACGATACAGA	Oligo anti-sens snRNA U2 murin et humain pour qRT-PCR (publi Workman et al, 2010)
8425	TCTCCATCCCATTGACAGAA	oligo 5' amplification RTPCR gène Spag9

Tableau VI.2-3 : Désignations, séquences et usages des oligodésoxynucleotides employés lors de ce travail.

Numéro Identifiant	Séquence 5'- 3'	Description
9756	TGGTTGTTCACTGTGTGTTAGTG	oligo 3' oligo spécifique de l'exon 10 de l'ARNm codant ECM2 E10R
9757	ATCGCCTCCATCCAGATGA	oligo 5' oligo spécifique de l'exon 5 de l'ARNm codant ECM2 E5F
9758	TGGACCTATGCCTGAAGAAGTG	oligo 3' oligo spécifique de l'exon 5 de l'ARNm codant ECM2 E5R
9759	CCGGATCAACCTCAGCAAGA	oligo 5' oligo spécifique de l'exon 1 de l'ARNm codant FBXW10 E1F
9760	GCTGTTCCGAAACCAAGTACA	oligo 3' oligo spécifique de l'exon 1 de l'ARNm codant FBXW10 E1R
9761	CAGGTTACAGCCCAAGAGA	oligo 5' oligo spécifique de l'exon 12 de l'ARNm codant FBXW10 E12F
9762	GGAGTACTCGGGACTTGGGA	oligo 3' oligo spécifique de l'exon 12 de l'ARNm codant FBXW10 E12R
9763	GAGCACCAGGAAGCCAAGAT	oligo 5' oligo spécifique de l'exon 14 de l'ARNm codant FBXW10 E14F
9764	AGGCAGGCCTTTGATCTTCC	oligo 3' oligo spécifique de l'exon 14 de l'ARNm codant FBXW10 E14R
9765	GACACCTTGTCTGACACCC	oligo 5' oligo spécifique de l'exon 1 de l'isoforme NR_051988.1 pré-ARNm NON codant
9766	AGTTGTGACTTCTCTGTCAT	oligo 3' oligo spécifique de l'exon 1 de l'isoforme NR_051988.1 pré-ARNm NON codant
9862	TAATGGAACCCCTGGTAGTCAG	oligo 5' oligo spécifique de l'exon 7 de l'ARNm codant DCLK1 E7F
9863	TGCTCCGACGGCTTCTCT	oligo 3' oligo spécifique de l'exon 7 de l'ARNm codant DCLK1 E7R
9864	GGTTCACAGCCCAAGAGAGT	oligo 5' oligo spécifique de l'exon 12 de l'ARNm codant FBXW10 E12F
9865	GAGAACTGGCCGGCTTCA	oligo 3' oligo spécifique de l'exon 12 de l'ARNm codant FBXW10 E12R
9866	AATCCATCGCCTCCATCCC	oligo 5' oligo spécifique de l'exon 5 de l'ARNm codant ECM2 E5F
9867	TGCTTTTGGACCTATGCCTGA	oligo 3' oligo spécifique de l'exon 5 de l'ARNm codant ECM2 E5R
9935	CTCATCTCCAGATCACACCGCT	Oligo 5' oligo spécifique de l'exon 20 du pré-ARNm codant SERCA2
9936	CAGGCTGCACACTCTTTAC	Oligo 3' oligo spécifique de l'exon 20 du pré-ARNm codant SERCA2b
A051	GAGGTTTATCCGAGGCGGATTAT	Oligo antisens d'analyse par RT-PCR du snRNA U4 de souris et humain
A052	CATGGCGGGGATTGGGAAAAGTT	Oligo antisens d'analyse par RT-PCR du snRNA U4 de souris et humain
A120	CTAGAACTCGCTCTAAAGACGG	Oligo 5' spécifique de l'exon 1 du pré-ARNm NR4A3 RT-PCR
A121	CTGGAGGGGAAGGGCTATATTGG	Oligo 3' spécifique de l'exon 3 du pré-ARNm NR4A3 RT-PCR
A122	TGCTGCACCTGTCTATACACAT	Oligo 5' spécifique de l'exon 5B du pré-ARNm NR4A3 RT-qPCR
A123	GGCGTTCCATCCCAGTTCTC	Oligo 3' spécifique de l'exon 5B du pré-ARNm NR4A3 RT-qPCR
A124	CTCAAGCCTTCTCGCTGTGA	Oligo 5' spécifique de l'exon 3 du pré-ARNm NR4A3 RT-qPCR
A125	CAGTGCTCGTCCCATAACG	Oligo 3' spécifique de l'exon 3 du pré-ARNm NR4A3 RT-qPCR
A359	GCGTAGTGCTTTCTGATTCCAG	Oligo 3' spécifique de l'exon 2 du pré-ARNm codant tra2B4
A360	TACGGCGAGCGGGAATCC	Oligo 5' spécifique de l'exon 1/3 du pré-ARNm codant tra2B1
A361	TCTAGACTAGATTAGATCGGGAC	Oligo 3' spécifique de l'exon 3/4 du pré-ARNm codant tra2B1
A362	GGCGAGCGGGTCTAGATC	Oligo 5' spécifique de l'exon 1/4 du pré-ARNm codant tra2B3
A363	TTTGCCCGATTCCCAACATG	Oligo 3' spécifique de l'exon 4/5 du pré-ARNm codant tra2B3
A364	ATGATCAGGGTCTCTATTTATGGCT	Oligo 5' spécifique de l'exon 5 du pré-ARNm codant NR4A3
A365	GCACGTGTCAGGGTCACTTGT	Oligo 3' spécifique de l'exon 5 du pré-ARNm codant NR4A3
A366	AGGAAAATGCGGAAGTCGTC	Oligo 5' spécifique de l'exon 2 du pré-ARNm codant tra2B4
A392	TGGAATGGTAAAAGAAGTTGTCCG	Oligo 5' spécifique de la jonction exon 4/5 du pré-ARNm codant NR4A3
A655	ATGTAGAGAAGGGTGACAACC	Oligo 5' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 4 U12 dépendant au sein de l'ARNm GBP1 par PCR
A656	TCGGATTCTATGTGTGACGTC	Oligo 3' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 4 U12 dépendant au sein de l'ARNm GBP1 par PCR
A657	GTATCCACTTGAAGACTGACTG	Oligo 5' amplification PCR pour vérifier la présence de l'intron 4 U12 dépendant au sein de l'ARNm GBP1 par PCR
A658	GACTGGTACCTTCTCAGCT	Oligo 3' amplification PCR pour vérifier la présence de l'intron 4 U12 dépendant au sein de l'ARNm GBP1 par PCR
A659	TCAATGGCCTCGTCTAGAGAG	Oligo 5' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 7 U12 dépendant au sein de l'ARNm GBP2 par PCR
A660	AGTCATCTCGCCTTCTTCC	Oligo 3' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 7 U12 dépendant au sein de l'ARNm GBP2 par PCR
A661	GGGCCTTTGTCTCAGACCT	Oligo 5' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 4 U12 dépendant au sein de l'ARNm MORC3 par PCR
A662	GCATTTTGTAGCTTCTAAGATCC	Oligo 3' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 4 U12 dépendant au sein de l'ARNm MORC3 par PCR

Numéro Identifiant	Séquence 5'- 3'	Description
A663	ACACAGCTGGTTTCGAAGAGT	Oligo 5' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 7 U12 dépendant au sein de l'ARNm MORC3 par PCR
A664	CCATGTTGTTGCCCTAACTG	Oligo 3' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 7 U12 dépendant au sein de l'ARNm MORC3 par PCR
A665	GCAATCAGATCGCCTCTGC	Oligo 5' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 8 U12 dépendant au sein de l'ARNm RNF38 par PCR
A666	TACGTCCTGTAAGCCTCCGA	Oligo 3' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 8 U12 dépendant au sein de l'ARNm RNF38 par PCR
A667	TTGAAGAACACGTCCTCAAGC	Oligo 5' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 13 U12 dépendant au sein de l'ARNm MARCH6 par PCR
A668	TAGCAGACTCCAGTAAGCG	Oligo 3' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 13 U12 dépendant au sein de l'ARNm MARCH6 par PCR
A669	TGCGGTGATGACTTCGATGG	Oligo 5' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 10 U12 dépendant au sein de l'ARNm SBNO1 par PCR
A670	GGTTCAGAAGCACCAGTTGCAC	Oligo 3' amplification PCR pour vérifier l'inclusion/exclusion de l'intron 10 U12 dépendant au sein de l'ARNm SBNO1 par PCR
A750	AGCTGAAGAAAGGTACCAGT	Oligo 5' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 6 U2 dépendant au sein de l'ARNm GBP1 par qPCR
A751	TTGTTGCACAAATTCGGGGT	Oligo 3' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 6 U2 dépendant au sein de l'ARNm GBP1 par qPCR
A752	TCATTTGAAATCAAATGCCATGT	Oligo 5' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 4 U2 dépendant au sein de l'ARNm MORC3 par qPCR
A753	CTGACGTGCACCATGTGTAG	Oligo 5' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 5 U2 dépendant au sein de l'ARNm MORC3 par qPCR
A754	GCAATAAGCCCTCAGGGAATAG	Oligo 3' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 5 U2 dépendant au sein de l'ARNm MORC3 par qPCR
A755	ACTGAGTGAGGTGTTGTGTTCA	Oligo 5' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 7 U12 dépendant au sein de l'ARNm MORC3 par qPCR
A756	CATGTTGTTGGCCCTAACTGAC	Oligo 3' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 7 U12 dépendant au sein de l'ARNm MORC3 par qPCR
A757	TTTCCAAGTGAATGCTGTCGAG	Oligo 3' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 8 U12 dépendant au sein de l'ARNm RNF38 par qPCR
A758	GCCTCCTCAATGCTTCAG	Oligo 5' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 7 U2 dépendant au sein de l'ARNm RNF38 par qPCR
A759	TCGTCTGGTTCTTAAATCCACT	Oligo 3' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 7 U2 dépendant au sein de l'ARNm RNF38 par qPCR
A760	TCCACACTGACCCTTACTGC	Oligo 3' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 13 U12 dépendant au sein de l'ARNm MARCH6 par qPCR
A761	GCAATCAGAAGCCTCAGGG	Oligo 5' amplification qPCR pour vérifier l'inclusion/exclusion de l'intron 10 U12 dépendant au sein de l'ARNm SBNO1 par qPCR
B250	CCGTAACGACGGCCAGTTCCTCTTCGTC	Oligo 3' de RT PCR pour la rétrotranscription spécifique de l'ARN circulaire exon 5
B251	TCTTCGTCCTCTCCAG	Oligo 3' oligo spécifique de l'exon 5 du pré-ARNm codant hcTNT

Tableau VI.2.4 : Désignations, séquences et usages des oligodésoxynucléotides employés lors de ce travail.

nom cible	nom du siRNA	séquence (brin sens)
non-targeting pool (NT)	si non-target	5' UGGUUUACAUGUCGACUAA 3'
hnRNP H1	siRNP H1-4	5' GAUCCACCACGAAAGCUUA 3'
hnRNP H1	siRNP H1-5	5' CAGAGUAACUGGUGAAGCA 3'
hnRNP H1	siRNP H1-6	5' GUUCGCAACUCAUGAAGAU 3'
hnRNP H1	siRNP H1-7	5' UAACAUUGCCGGUGGACUU 3'
HSF1	si HSF1-1	5' CACAUCCAUGCCCAAGUA 3'
HSF1	si HSF1-2	5' CCACUUGGAUGCUAUGGAC 3'
HSF1	si HSF1-3	5' CCUGAAGAGUGAAGACAUA 3'
HSF1	si HSF1-4	5' GAACAGCAGCCCGGAUUCA 3'

Tableau VI.3 : Caractéristiques des siRNA utilisés. Pour chaque siRNA utilisé, il est indiqué le nom de la cible, le nom du siRNA et sa séquence en nucléotides.

IV. 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.

Milieux 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) : NaCl 137 mM, KCl 2,7 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-mercaptoéthanol 100 mM ; bleu de bromophénol 0,01 % (p/v) ; Tris-HCl 63,65 mM pH 6,8.

Tampon D : MgCl_2 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 cellules pour l'extraction de 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 $\mu\text{L}/\text{mL}$)

Tampon TfbI (préparation de bactéries compétentes) : KAc 30 mM; KCl 100 mM; CaCl_2 10 mM ; Glycérol 15% (v/v) ; MnCl_2 50 mM.

Tampon TfbII (préparation de bactéries compétentes) : MOPS Na pH7 10 mM ; KCl 10 mM ; CaCl_2 75 mM ; Glycérol 15% (v/v).

Tampon de transcription (5x) : DTT 200 mM ; spermidine 10 mM ; MgCl_2 120 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 PBS-Tween 0,1 % : PBS ; Tween 20 0,1 % (v/v).

Tampon de transcription US (5x) : MgCl_2 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.

Solution de polyacrylamide dénaturante : Acrylamide/Bisacrylamide (23,75/1,25) 25 % (p/v) ; TBE; urée 8 M.

Solution de polyacrylamide native : Acrylamide/Bisacrylamide (38/2) 40 % (p/v) ; TBE.

HBS (2X) : HEPES-KOH 50 mM, pH 7,05 ; KCl 10 mM ; NaCl 280 mM ; Na_2HPO_4 1,5 mM.

Méthodes

V. Méthodes utilisées pour la culture cellulaire.

V.1. Culture cellulaire.

V.1.1. les Lignées HEK293 et HeLa S3 et réalisation d'un stress oxydant.

Les lignées cellulaires HEK293 et HeLa S3 sont cultivées dans le milieu DMEM (Dulbecco's Modified Eagle Medium) supplémenté avec 10% de sérum de veau foetal (FBS), 100 U/ml de pénicilline, 100 µg/mL streptomycine et 4 mM de glutamine (Gibco). Les cellules sont cultivées en flasques ou en boîtes de différents diamètres (en fonction de l'utilisation) à 37°C, sous une atmosphère humide et à 5% de CO₂. La confluence des cellules est vérifiée par microscopie. Une fois à 70% ou 90%, les cellules sont rincées avec du PBS puis incubées avec 1 mL de trypsine (Invitrogen) à 37°C environ 2 min pour décrocher les cellules. L'activité enzymatique de la trypsine est arrêtée par ajout de 1 mL de milieu. Pour la réalisation d'un stress oxydant efficace, une nouvelle boîte contenant 10 mL de milieu estensemencée avec 2x10⁶ cellules 24 h avant le début du stress. Pour le stress du milieu de culture DMEM contenant 250 µM ou 500 µM d'H₂O₂ est préparé à partir d'une solution mère à 35%, puis dans les boites contenant les cellules.

V.1.2. Transfection de cellules HEK293 par un plasmide au réactif JetPEI.

Les cellules HEK293 (300 000 cellules par puits en plaque 6 puits dans 3 mL de milieu) sont transfectées 1,5 µg de plasmide pSVIRB-hcTNT ou ses différents mutants et 1,5 µg de plasmide pcDNA3.1-(CTG)960 ou 0,1 µg de plasmide pcDNA3.1-MBNL1 ou complété par le plasmide pcDNA3.1 vide en utilisant le réactif JetPEIPM (PolyPlus Transfection). Après 5 h d'incubation avec l'ADN, les cellules sont lavées avec 2 mL de PBS et remises en culture pendant 48 h.

V.1.3. Transfection de cellules humaines par un siRNA au phosphate de calcium.

La veille de la transcription, les cellules HEK293 sontensemées à 200 000 cellules par puits en plaque 6 puits dans 2 mL de milieu DMEM sans antibiotique. Les siRNA (50 pmoles) sont dilués à 25 mM dans 36,8 µl de tampon TE 0,1X, 5,2 µl de CaCl₂ 2 M, 42 µl de HBS (2X) pH 7,5. Le mélange est incubé 20 min à température ambiante, avant d'être distribué sur les cellules en culture. Après 16h de culture, le milieu de culture est remplacé par 2 ml de milieu DMEM sans antibiotique frais. Après 48 h, les cellules sont co-transfectées par un mélange siRNA (50-100 pmoles) et de plasmide pSVIRB-hcTNT (0,6 µg) en utilisant le réactif lipofectamine 2000 (Invitrogen). Après 24 h, les cellules sont récoltées dans 1 ml de PBS (70% sert à la préparation des ARN totaux et 30 % sert à la préparation d'extraits protéiques).

V.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 10.10⁶ 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 µL de H₂O, dont 1 µL sert au dosage par absorption optique à 260 nm grâce à un appareil nanodrop.

V.3. Dosages de la viabilité cellulaire et des ROS intracellulaires.

V.3.1. Dosage de la viabilité cellulaire :

Le dosage de la viabilité cellulaire a été réalisé par utilisation du kit Vybrant MTT cell proliferation Assay Kit (Invitrogen), dont le principe repose sur un test colorimétrique par ajout d'un sel de tétrazolium (MTT) (jaune) qui est réduit en formazan (violet) par la succinate déshydrogénase mitochondriale qui est active seulement dans les cellules viables. Suite aux différents stress, les cellules sont extraites et mises en solution dans 2 mL DMEM. 100 µL de cette solution (représentant 5% des cellules totales) sont additionnées à 10 µL d'une solution de MTT à 5 mg/mL (jaune), puis incubée pendant 3 h à 37°C. Afin de lyser les cellules, 100 µL d'une solution de SDS à 200 mg/mL sont ensuite ajoutés au mélange. Après une nouvelle incubation de 3 h à 37°C, l'absorbance de la solution devenue violette suite à la réduction du MTT par la succinate déshydrogénase mitochondriale est mesurée à 570 nm par un spectrophotomètre. Un contrôle négatif est réalisé en parallèle avec 100 µL de DMEM (Dulbecco's Modified Eagle Medium) sans cellule. Pour chaque condition, l'absorbance mesurée est rapportée à l'absorbance de l'échantillon issus des cellules non stressés (100% de viabilité cellulaire).

V.3.2. Dosage des ROS intracellulaires :

Le dosage de la quantité de ROS intracellulaire a été réalisé par utilisation du kit OxySelect ROS Assay kit (Euromedex) basé sur l'utilisation d'un agent réducteur, le H2DCFDA (dichlorodihydrofluorescein diacetate), qui s'oxyde au contact d'un dérivé réactif de l'oxygène et devient alors fluorescent. Suite aux différents stress, les cellules sont extraites et mises en solution dans 2 mL de DMEM (Dulbecco's Modified Eagle Medium). 100 µL de cette solution (représentant 5% des cellules totales) sont additionnées à 5 µL de H2DCFDA 20X. Le mélange est incubé à 37°C pendant 1 h, puis centrifugé à 200g pendant 5 min pour culotter les cellules. Le culot est repris dans une solution de lyse fournie dans le kit et après 5 min d'incubation à 37°C, la fluorescence est mesurée par spectrofluorimétrie (longueur d'onde d'excitation 480 nm ; longueur d'onde d'émission 530 nm).

VI. Méthodes d'analyse in vitro.

VI.1. Introduction des mutations par mutagenèse dirigée.

Les mutations des sites MBNL1 ou hnRNP H dans les minigènes IR-hcTNT-150-E5-150 et RTB300 ont été réalisées par mutagenèse dirigée en utilisant les couples d'oligodésoxynucléotides décrits dans les Tableaux VI.2 1-4. Le protocole de PCR utilisé est décrit dans les tableaux VI.4 et VI.5. Une fraction (10 μL) des produits PCR est ensuite digérée par 10 UE d'enzyme DpnI (fastDigest MBI Fermentas), qui hydrolyse l'ADN parental méthylé, ceci dans un volume réactionnel de 30 μL .

Réactifs	Volume (μL)
tampon Pfu 5x (Fermentas)	10
Pfu DNA polymérase (Fermentas; 2,5U/ μL)	1
mélange de dNTP (1,25 mM chaque)	8
Matrice plasmidique (200 ng/ μL)	1
Oligodésoxyribonucléotide sens (10 pmol/ μL)	1
Oligodésoxyribonucléotide antisens (10 pmol/ μL)	1
H ₂ O	qps 50

Tableau VI.4 : Mélange réactionnel pour la réalisation d'une mutagenèse dirigée.

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	68°C	8 min
1	Elongation finale	72°C	15 min
1	Stockage	20°C	∞

Tableau VI.5 : Les conditions expérimentales des cycles de la PCR.

Les produits de digestion sont utilisés pour transformer des bactéries *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 incorporés 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).

VI.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 un oligodésoxyribonucléotide sens contenant le promoteur T7, combiné à un oligodésoxyribonucléotide antisens dont l'identité dépend de la séquence à amplifier. Le protocole de PCR utilisé est indiqué dans le tableau VI.6. Les produits PCR obtenus sont fractionnés en gel d'agarose à 1 ou 2 % 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'élution est réalisée dans un volume de 20 μL .

Réactifs	Volume (μl)
Tampon DreamTaq 5x (Fermentas)	5
Dream Taq ADN polymérase (Fermentase; 5 U/ μL)	1
mélange de dNTP (1,25 mM chaque)	8
Matrice plasmidique (10 ng/ μL)	1
Oligodésoxyribonucléotide sens (100 pmol/ μL)	1
Oligodésoxyribonucléotide antisens (100 pmol/ μL)	1
H ₂ O	qps 50

Tableau VI.6 : Réaction de PCR pour la synthèse de matrice de transcription.

VI.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 VI.7. Le mélange réactionnel est incubé pendant 4 h à 37°C. Puis la matrice ADN est dégradée par incubation pendant 30 min à 37°C en présence de 10 UE de l'enzyme DNase I (Fermentas). 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 mesurée par spectrophotométrie à 260 nm.

Réactifs	Volume (µl)
Tampon de transcription 5x	10
NTP (25 mM chaque)	15
Produit PCR purifié	20
ARN polymérase T7 (produite au laboratoire)	5

Tableau VI.7 : Réaction de la transcription *in vitro*.

VI.4. Etude de l'interaction des protéines avec les ARN.

VI.4.1. Purification par chromatographie d'affinité MBP-MS2.

VI.4.1.1. Production et purification de la protéine de fusion MBP-MS2.

Des bactéries *E. coli* BL21 (λDE3) 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 UA600nm/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 2000 g 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.

VI.4.1.2. 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 µL de tampon D, dénaturés et renaturés comme décrit précédemment en présence de 3 mM de MgCl₂. Un aliquot de 5 µL (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 alors 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. Les ARN sont révélés par exposition aux UV en présence de révélateur GelRed (FluoProbes). Le mélange ARN-MS2/MBP-MS2 est ensuite mis en contact avec 40 µl d'une suspension de résine amylose (New England Biolabs), préalablement lavée 3 fois avec 500 µl 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 200 g puis sont lavées 3 fois avec du tampon D avant d'ajouter 250 µl d'extrait nucléaire de cellules HeLa (CilBiotech), des ARNt en excès 50x molaire, en présence de 3 mM de MgCl₂ 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 200 g,

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 µl 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 µl 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 importante (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%.

Pour l'analyse du contenu en protéines par spectrométrie de masse, une première étape a consisté à séparer les protéines de la fraction d'élution (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).

VI.4.2. Estimation des quantités de protéines par « western blot ».

Un microlitre d'extrait nucléaire ou 1 µL 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 de protéines a lieu sur membrane de nitrocellulose 0.45 µm pendant 1 h à 100 V dans du tampon de transfert refroidi. La membrane est ensuite saturée avec 5 mL de PBS/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 la nuit 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.

VI.4.3 Expériences de retard sur gel.

VI.4.3.1 Marquage des ARN en 5'.

Deux microgrammes d'ARN à marquer radioactivement sont déphosphorylés par action de la phosphatase alcaline de veau (CIAP, MBI-Fermentas) à raison de 1 UE par μg d'ARN pendant 1 h à 37°C (voir Tableau VI.8). Les ARN sont ensuite extraits par un mélange de phénol-chloroforme, puis précipités à -80°C et lavés avec de l'éthanol 70%. Les culots sont repris dans 14 μL d'eau stérile. 7 μL de cette solution sont utilisés pour le marquage radioactif selon le protocole décrit dans le Tableau VI.9. Le mélange réactionnel est incubé pendant 1 h à 37°C.

Déphosphorylation de l'extrémité 5' des ARN	
Réactifs	Volume (μl)
ARN à déphosphoryler (2 μg)	
Tampon CIP 10X (MgCl ₂ 0,1 M; Tris/HCl 0,1 M, pH 7,5)	1
Calf Intestine Alkaline Phosphatase (fermentas 1 U/ μL)	2
H ₂ O	qsp 10

Tableau VI.8: Protocole pour la déphosphorylation en 5' des ARN.

Marquage en 5' des ARN	
Réactifs	Volume (μl)
ARN déphosphorylés repris dans eau stérile	7
Tampon PNK 10X	1
ATP [γ -32P] (Amersham, 10 mCi/ml, 3000 Ci/mmol)	1
T4 polynucléotide kinase (MBI Fermentas, 10 U/ μl)	1

Tableau VI.9: Protocole pour le marquage en 5' des ARN déphosphorylés. Le mélange est incubé 1 h à 37°C puis les transcrits sont purifiés sur gel de polyacrylamide dénaturant 8 %, élués, et précipités comme décrit précédemment. Le tampon PNK : MgCl₂ 0.1 M ; DTT 50 mM ; spermidine 1 mM ; EDTA 1 mM ; Tris-HCl 0,5 M, pH 7,6.

VI.4.3.2 Formation des complexes [ARN/protéine].

Ces expériences sont réalisées avec 20 cps d'ARN radiomarqué, en présence de $MgCl_2$ 2,8 mM. Les ARN sont dénaturés pendant 10 min à 65°C puis renaturés pendant 15 min par refroidissement progressif à température ambiante (Tableau VI.10). Quatre 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.

Réactifs	Volume (μ l)
ARN radioactif (20 cps)	
Tampon D	1,8
$MgCl_2$ (62,5 mM)	0,5
H ₂ O	qps 7

Tableau VI.10: Composition du mélange réactionnel pour la formation de complexes ARN-protéine avant ajout de protéine purifiée. Le mélange est dénaturé pendant 5 min à 65°C, puis renaturé lentement à 37°C.

VI.4.3.3 Fractionnement en gel natif.

Du bleu CSB (4 μ l) est ajouté au mélange réactionnel ayant servi à former le complexe. L'échantillon est ensuite déposé sur un gel de polyacrylamide natif 8% (acrylamide/bisacrylamide 38/2), en présence de glycérol 0,5 % et de TBE 0,5X. Le pourcentage de ce gel peut varier de 5 à 8 % en acrylamide/bisacrylamide suivant la taille des ARN utilisés lors de l'expérience. L'électrophorèse se fait à 100 volts et à température ambiante. Le gel est ensuite séché et exposé en autoradiographie durant une nuit. Ainsi, les complexes [ARN/protéine] formés sont retardés dans leur migration par rapport aux ARN non complexés.

VI.4.3.4 Détermination de la constante de dissociation (Kd) des complexes.

La détermination de la constante de dissociation d'un complexe [ARN/protéine] permet d'évaluer l'affinité de la protéine d'intérêt pour l'ARN. Cette constante correspond à la concentration en protéine pour laquelle 50% de l'ARN est lié à la protéine. Les quantités d'ARN libre et lié sont quantifiées par mesure de la radioactivité dans les bandes correspondantes, en utilisant l'appareil PhosphorImager à l'aide du logiciel « ImageQuant Software Version 3.3 » (Molecular Dynamics). La sigmoïde obtenue par représentation graphique de la concentration en protéine (en abscisse) par rapport au pourcentage d'ARN lié est tracée à l'aide du logiciel « SigmaPlot 2000 ». Le point d'inflexion de la sigmoïde indique la valeur de la constante de dissociation (concentration en protéine pour laquelle 50% des ARN sont liés à la protéine d'intérêt) est donnée par le logiciel.

VI.4.4. Etude de la structure des ARN et identification des sites de fixation de protéines par la méthode d'empreintes en solution.

Les protocoles relatifs à ces expériences ont fait l'objet d'une publication méthodologique, dont je suis premier auteur, dans *Methods in Molecular Biology* en 2015. Cette publication est présentée dans la partie Publications.

VI.5. Transcription et épissage in vitro d'un ARN pré-messager.

Le protocole expérimental utilisé pour la transcription des ARN d'intérêt est décrit dans le Tableau VI.11.

Réactifs	Volume (µl)
Matrice PCR (1,5 pmol)	
Tp 5X (MgCl ₂ 30 mM ; Spermidine 10 mM ; NaCl 50 mM ; Tris-HCl 200 mM, pH 7,5)	8
NTP (A, C, U, G ; 5, 5, 5, 2 mM)	4
Cap 5m7GpppG (10 mM)	2
Inhibiteur de RNases (RNasin, 20 U/µl, Amersham Biosciences)	2
BSA (2 mg/ml)	2
DTT (200 mM)	2
ARN polymérase T7	2
H ₂ O	qsp 40

Tableau VI.11 : Mélange réactionnel pour la transcription d'ARN pré-messagers.

Le mélange réactionnel est incubé pendant 2 h à 37°C. Après une étape de dégradation de la matrice ADN, les ARN sont purifiés sur colonne Sephadex G-50 (mini Quick Spin Columns, Roche).

Les expériences d'épissage *in vitro* sont réalisées avec des extraits nucléaires de cellules HeLa provenant de la société CilBiotech. Le mélange réactionnel est décrit dans le Tableau VI.12.

Réactifs	Volume (µl)
ARN pré-messager (10 fmol/µl)	2
MgCl ₂ (62,5 mM)	1
ATP (25 mM)	1
Créatine phosphate (625 mM)	1
Inhibiteur de RNases (RNasin, 20 U/µl, Amersham Biosciences)	0,5
Alcool polyvinylique 13 %	6
Extrait nucléaire (10 mg/ml, traité au préalable 10 min à 30°C)	8

Tableau VI.12 : Mélange réactionnel pour la réalisation d'un épissage *in vitro* d'ARN pré-messagers.

Le mélange est complété par 2,5 µl de Tampon D ou de concentration croissante de protéine hnRNP H-Nter. Une incubation de 5 min à 0°C puis 2h à 30°C est réalisée. La réaction d'épissage est ensuite arrêtée par ajout de 20 µg de protéinase K (Roche) fonctionnant dans le tampon suivant : EDTA 12,5 mM ; NaCl 150 mM ; SDS 1 % ; Tris-HCl 100 mM, pH 7,4. Après une incubation de 30 minutes à 30°C, les ARN sont extraits au phénol/chloroforme (1/1), puis lavé au chloroforme/alcool isoamylique (24/1) et les ARN sont précipités en présence d'éthanol absolu sans ajout de sels. Le précipité est lavé puis séché et repris dans 10 µl d'eau stérile.

VII. Méthode d'analyse RT-PCR et RT-qPCR.

VII.1. Digestion des ADN par la DNase RQ1 dans des extraits d'ARN.

Avant chaque réaction de transcription inverse, 10 µg d'ARN totaux sont soumis à digestion par 10 U 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.

VII.2. La 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 oligodT(16), soit l'oligonucléotide spécifique de l'ARN sat III (7796), soit un oligonucléotide hexamérique aléatoirement dégénéré (N6). 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).

VII.3. Détection des produits rétrotranscrits par PCR point final.

Deux microlitres de la solution d'ADNc produits par RT-PCR sont utilisés comme matrice pour une réaction de PCR. Le milieu réactionnel utilisé est décrit dans le Tableau VI.13 et les conditions de réalisation de la PCR dans le Tableau VI.14.

Réactifs pour PCR	Volume (μ l)
Tampon DreamTaq buffer green 10x (Fermentas)	5
Dream Taq ADN polymérase (Fermentase; 5 U/ μ L)	0,5
mélange de dNTP (1,25 mM chaque)	8
ADNc	2
Oligodésoxyribonucléotide sens (20 μ M)	1
Oligodésoxyribonucléotide antisens (20 μ M)	1
H ₂ O	qsp 50

Tableau VI.13 : Mélange réactionnel de PCR pour la détection de produits spécifiques à partir d'ARN totaux rétrotranscrits.

Nombre de cycles	Etape	Température	Durée
1	Dénaturation initiale	95°C	5 min
30	Dénaturation	95°C	30 sec
	Hybridation	58°C	30 sec
	Elongation	72°C	30 sec
1	Elongation finale	72°C	10 min
1	Stockage	20°C	∞

Tableau VI.14 : Conditions expérimentales des cycles de PCR pour la détection de produits spécifiques à partir d'ARN totaux rétrotranscrits.

Les produits PCR obtenus sont fractionnés en gel d'agarose à bas point de fusion à 2, 3 ou 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).

VII.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 le produit à détecter, l'ADNc matrice peut être dilué de 1 à 300 fois, afin que la détection des produits d'intérêt apparaissent entre 20 et 32 Ct pour optimiser 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 (Tableau VI.15). Le programme utilisé pour la qPCR est détaillé dans le Tableau VI.16.

Réactifs pour qPCR	Volume (μ l)
ADNc dilué ou non	1
Master MIX 2x	5
Oligodésoxyribonucléotide sens (5 μ M)	1
Oligodésoxyribinucléotide antisens (5 μ M)	1
H ₂ O	2

Tableau VI.15 : Mélange réactionnel pour la réalisation d'une réaction de qPCR.

Nombre de cycles	Etape	Température	Durée	Vitesse d'élongation
1	Dénaturation initiale	95°C	10 min	-
45	Dénaturation	95°C	10 sec	20°C/sec
	Hybridation	55 à 58°C	10 sec	20°C/sec
	Elongation	72°C	10 sec	5°C/sec
	1	Courbe de fusion	95°C	10 sec
60°C			5 sec	20°C/sec
95°C			-	0,1°C/sec

Tableau VI.16 : Conditions expérimentales des cycles pour la réaction de qPCR pour la détection de produits spécifique à partie d'ARN totaux rétrotranscrits. La température d'hybration varie en fonction des oligodésoxyribonucléotides utilisés afin d'obtenir une efficacité optimale.

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 ARN invariant, celui de la protéine GAPDH ou de l'ARN 5S. La quantification relative est réalisée en prenant en considération l'efficacité de chaque couple d'oligonucléotide utilisé.

VIII. Microscopie de fluorescence.

VIII.1. Visualisation des ARN satellite III après un stress oxydant par RNA-FISH.

VIII.1.1. Préparation des sondes LNA biotinylées.

Les sondes LNA fournies par la société Exiqon correspondent à 4 répétitions de séquences GGAAU canoniques et sont diluées à 30 nM dans une solution [SSC2x ; 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.

VIII.1.2. Fixation des cellules et incubation avec les sondes.

Les cellules sont mises en culture comme décrit dans le paragraphe V.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 stress oxydant, 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. Suite à cela les cellules sont déshydratées par 3 bains successifs d'éthanol à 70, 90 et 100%. Les cellules sont ensuite séchées et incubées pendant environ 30 min à 45°C avec la solution contenant les sondes biotinylées. Ces lames sont ensuite lavées 3x5 min à 65°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 1 à 2 h environ. 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 VIII.2. Dans le cas contraire, des lamelles sont directement montées dessus pour la visualisation au microscope à épifluorescence ou au microscope confocal (voir paragraphe VIII.3).

VIII.2. Visualisation des protéines par immunofluorescence.

Les cellules sont mises en culture comme décrit dans le paragraphe V.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 stress oxydant, 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, §II) 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.

VIII.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.

PUBLICATIONS

Tau exon 2 responsive elements deregulated in myotonic dystrophy type I are proximal to exon 2 and synergistically regulated by MBNL1 and MBNL2

C. Carpentier, D.Ghanem, F.J. Fernandez-Gomez, F.Jumeau, **J.V.Philippe**, F. Freyermuth, A. Labudeck, S. Eddarkaoui, C.M.Dhaenens, I.Holt, I. Behm-Ansmant, N. Marmier-Gourrier, C. Branlant, N. Charlet-Berguerand, J.Marie, S.Schraen-Maschke, L.Buée, N. Sergeant, M.L.Caillet-Boudin

Probing Small Non-Coding RNAs Structures

Jean-Vincent Philippe, Lilia Ayadi, Christiane Branlant,
and Isabelle Behm-Ansmant.

Bibliographie

Bibliographie

- Achsel, T., Brahms, 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.
- Achsel, T., Stark, H., and Lührmann, R. (2001). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 3685–3689.
- Ahn, S.-G., and Thiele, D.J. (2003). Redox regulation of mammalian heat shock factor 1 is essential for Hsp gene activation and protection from stress. *Genes Dev.* *17*, 516–528.
- Akaike, Y., Masuda, K., Kuwano, Y., Nishida, K., Kajita, K., Kurokawa, K., Satake, Y., Shoda, K., Imoto, I., and Rokutan, K. (2014). HuR Regulates Alternative Splicing of the TRA2 β Gene in Human Colon Cancer Cells under Oxidative Stress. *Mol. Cell. Biol.* *34*, 2857–2873.
- Alani, B., Salehi, R., Sadeghi, P., Khodaghali, F., Digaleh, H., Jabbarzadeh-Tabrizi, S., Zare, M., and Korbekandi, H. (2015). Silencing of Hsp70 intensifies 6-OHDA-induced apoptosis and Hsp90 upregulation in PC12 cells. *J. Mol. Neurosci.* *MN 55*, 174–183.
- Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., Morozova, N., Fenouil, R., Descostes, N., Andrau, J.-C., et al. (2012). Argonaute proteins couple chromatin silencing to alternative splicing. *Nat. Struct. Mol. Biol.* *19*, 998–1004.
- Amit, M., Donyo, M., Hollander, D., Goren, A., Kim, E., Gelfman, S., Lev-Maor, G., Burstein, D., Schwartz, S., Postolsky, B., et al. (2012). Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. *Cell Rep.* *1*, 543–556.
- Argente, J., Flores, R., Gutiérrez-Arumí, A., Verma, B., Martos-Moreno, G.Á., Cuscó, I., Oghabian, A., Chowen, J.A., Frilander, M.J., and Pérez-Jurado, L.A. (2014). Defective minor spliceosome mRNA processing results in isolated familial growth hormone deficiency. *EMBO Mol. Med.* *6*, 299–306.
- Arimoto, K., Fukuda, H., Imajoh-Ohmi, S., Saito, H., and Takekawa, M. (2008). Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. *Nat. Cell Biol.* *10*, 1324–1332.
- 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.
- Ashwal-Fluss, R., Meyer, M., Pamudurti, N.R., Ivanov, A., Bartok, O., Hanan, M., Evtal, N., Memczak, S., Rajewsky, N., and Kadener, S. (2014). circRNA Biogenesis Competes with Pre-mRNA Splicing. *Mol. Cell* *56*, 55–66.
- Aubol, B.E., Jamros, M.A., McGlone, M.L., and Adams, J.A. (2013). Splicing kinase SRPK1 conforms to the landscape of its SR protein substrate. *Biochemistry (Mosc.)* *52*, 7595–7605.
- Auweter, S.D., Fasan, R., Reymond, L., Underwood, J.G., Black, D.L., Pitsch, S., and Allain, F.H.-T. (2006). Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. *EMBO J.* *25*, 163–173.

- Ayala, A., Muñoz, M.F., and Argüelles, S. (2014). Lipid Peroxidation: Production, Metabolism, and Signaling Mechanisms of Malondialdehyde and 4-Hydroxy-2-Nonenal. *Oxid. Med. Cell. Longev.* 2014.
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124.
- Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., et al. (2012). The evolutionary landscape of alternative splicing in vertebrate species. *Science* 338, 1587–1593.
- Barreau, C., Paillard, L., Méreau, A., and Osborne, H.B. (2006). Mammalian CELF/Bruno-like RNA-binding proteins: molecular characteristics and biological functions. *Biochimie* 88, 515–525.
- Batra, R., Charizanis, K., Manchanda, M., Mohan, A., Li, M., Finn, D.J., Goodwin, M., Zhang, C., Sobczak, K., Thornton, C.A., et al. (2014). Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell* 56, 311–322.
- 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.
- Battle, D.J., Kasim, M., Yong, J., Lotti, F., Lau, C.-K., Mouaikel, J., Zhang, Z., Han, K., Wan, L., and Dreyfuss, G. (2006). The SMN Complex: An Assembly Machine for RNPs. *Cold Spring Harb. Symp. Quant. Biol.* 71, 313–320.
- Beard, J.A., Tenga, A., and Chen, T. (2015). The interplay of NR4A receptors and the oncogene–tumor suppressor networks in cancer. *Cell. Signal.* 27, 257–266.
- Beckman, J.S., Beckman, T.W., Chen, J., Marshall, P.A., and Freeman, B.A. (1990). Apparent hydroxyl radical production by peroxynitrite: implications for endothelial injury from nitric oxide and superoxide. *Proc. Natl. Acad. Sci. U. S. A.* 87, 1620–1624.
- Bernardi, R., and Pandolfi, P.P. (2007). Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nat. Rev. Mol. Cell Biol.* 8, 1006–1016.
- 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. (2004). Nuclear stress bodies: a heterochromatin affair? *Nat. Rev. Mol. Cell Biol.* 5, 493–498.
- Biamonti, G., and Caceres, J.F. (2009). Cellular stress and RNA splicing. *Trends Biochem. Sci.* 34, 146–153.
- Biamonti, G., and Vourc’h, C. (2010). Nuclear stress bodies. *Cold Spring Harb. Perspect. Biol.* 2, a000695.
- Bienert, G.P., Heinen, R.B., Berny, M.C., and Chaumont, F. (2014). Maize plasma membrane aquaporin ZmPIP2;5, but not ZmPIP1;2, facilitates transmembrane diffusion of hydrogen peroxide. *Biochim. Biophys. Acta* 1838, 216–222.

- 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.
- Blechingberg, J., Lykke-Andersen, S., Jensen, T.H., Jørgensen, A.L., and Nielsen, A.L. (2007). Regulatory mechanisms for 3'-end alternative splicing and polyadenylation of the Glial Fibrillary Acidic Protein, GFAP, transcript. *Nucleic Acids Res.* *35*, 7636–7650.
- Blencowe, B.J. (2006). Alternative splicing: new insights from global analyses. *Cell* *126*, 37–47.
- Bley, N., Lederer, M., Pfalz, B., Reinke, C., Fuchs, T., Glaß, M., Möller, B., and Hüttelmaier, S. (2015). Stress granules are dispensable for mRNA stabilization during cellular stress. *Nucleic Acids Res.* *43*, e26.
- Bono, F., and Gehring, N.H. (2011). Assembly, disassembly and recycling. *RNA Biol.* *8*, 24–29.
- Braberg, H., Jin, H., Moehle, E.A., Chan, Y.A., Wang, S., Shales, M., Benschop, J.J., Morris, J.H., Qiu, C., Hu, F., et al. (2013). From Structure to Systems: High-Resolution, Quantitative Genetic Analysis of RNA Polymerase II. *Cell* *154*, 775–788.
- Bradley, T., Cook, M.E., and Blanchette, M. (2015a). SR proteins control a complex network of RNA-processing events. *RNA N. Y. N* *21*, 75–92.
- Bradley, T., Cook, M.E., and Blanchette, M. (2015b). SR proteins control a complex network of RNA-processing events. *RNA N. Y. N* *21*, 75–92.
- 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. (2013a). Dynamic integration of splicing within gene regulatory pathways. *Cell* *152*, 1252–1269.
- Braunschweig, U., Gueroussov, S., Plocik, A.M., Graveley, B.R., and Blencowe, B.J. (2013b). Dynamic integration of splicing within gene regulatory pathways. *Cell* *152*, 1252–1269.
- Braunschweig, U., Gueroussov, S., Plocik, A., Graveley, B.R., and Blencowe, B.J. (2013c). Dynamic integration of splicing within gene regulatory pathways. *Cell* *152*, 1252–1269.
- Brigelius-Flohé, R. (2006). Glutathione peroxidases and redox-regulated transcription factors. *Biol. Chem.* *387*, 1329–1335.
- Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., and Hudson, T. (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* *68*, 799–808.

- Brookes, P.S., Yoon, Y., Robotham, J.L., Anders, M.W., and Sheu, S.-S. (2004). Calcium, ATP, and ROS: a mitochondrial love-hate triangle. *Am. J. Physiol. Cell Physiol.* *287*, C817–C833.
- Brunner, H.G., Brüggewirth, H.T., Nillesen, W., Jansen, G., Hamel, B.C., Hoppe, R.L., de Die, C.E., Höweler, C.J., van Oost, B.A., and Wieringa, B. (1993). Influence of sex of the transmitting parent as well as of parental allele size on the CTG expansion in myotonic dystrophy (DM). *Am. J. Hum. Genet.* *53*, 1016–1023.
- Buchan, J.R., and Parker, R. (2009). Eukaryotic Stress Granules: The Ins and Outs of Translation. *Mol. Cell* *36*, 932–941.
- Buchan, J.R., Muhlrads, D., and Parker, R. (2008). P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. *J. Cell Biol.* *183*, 441–455.
- 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.
- Buée, L., Bussièrre, T., Buée-Scherrer, V., Delacourte, A., and Hof, P.R. (2000). Tau protein isoforms, phosphorylation and role in neurodegenerative disorders. *Brain Res. Brain Res. Rev.* *33*, 95–130.
- Buratti, E., Dörk, T., Zuccato, E., Pagani, F., Romano, M., and Baralle, F.E. (2001). Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J.* *20*, 1774–1784.
- Burge, C.B., Padgett, R.A., and Sharp, P.A. (1998). Evolutionary fates and origins of U12-type introns. *Mol. Cell* *2*, 773–785.
- Busà, R., Geremia, R., and Sette, C. (2010). Genotoxic stress causes the accumulation of the splicing regulator Sam68 in nuclear foci of transcriptionally active chromatin. *Nucleic Acids Res.* *38*, 3005–3018.
- Busch, A., and Hertel, K.J. (2012). Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip. Rev. RNA* *3*, 1–12.
- Cáceres, 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.
- Cadenas, E., and Sies, H. (1998). The lag phase. *Free Radic. Res.* *28*, 601–609.
- Cammas, A., Pileur, F., Bonnal, S., Lewis, S.M., Lévêque, N., Holcik, M., and Vagner, S. (2007). Cytoplasmic relocalization of heterogeneous nuclear ribonucleoprotein A1 controls translation initiation of specific mRNAs. *Mol. Biol. Cell* *18*, 5048–5059.
- Cao, W., Jamison, S.F., and Garcia-Blanco, M.A. (1997). Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *RNA N. Y. N* *3*, 1456–1467.
- Carpentier, C., Ghanem, D., Fernandez-Gomez, F.J., Jumeau, F., Philippe, J.V., Freyermuth, F., Labudeck, A., Eddarkaoui, S., Dhaenens, C.M., Holt, I., et al. (2014). Tau exon 2 responsive elements deregulated in myotonic dystrophy type I are proximal to exon 2 and synergistically regulated by MBNL1 and MBNL2. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* *1842*, 654–664.

- Casas-Finet, J.R., Smith, J.D., 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.
- Caspersen, C., Pedersen, P.S., and Treiman, M. (2000). The sarco/endoplasmic reticulum calcium-ATPase 2b is an endoplasmic reticulum stress-inducible protein. *J. Biol. Chem.* **275**, 22363–22372.
- Catalá, Á. (2014). Lipid peroxidation modifies the assembly of biological membranes “The Lipid Whisker Model.” *Front. Physiol.* **5**.
- Chaturvedi, P., and Parnaik, V.K. (2010). Lamin A Rod Domain Mutants Target Heterochromatin Protein 1 α and β for Proteasomal Degradation by Activation of F-Box Protein, FBXW10. *PLoS ONE* **5**.
- Chen, K. (2015). Alternative splicing: An important mechanism in stem cell biology. *World J. Stem Cells* **7**, 1.
- Chen, C.Y., and Sarnow, P. (1995). Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* **268**, 415–417.
- Chen, B., Lu, Y., Chen, Y., and Cheng, J. (2015a). The role of Nrf2 in oxidative stress-induced endothelial injuries. *J. Endocrinol.* **225**, R83–R99.
- Chen, K., Dai, X., and Wu, J. (2015b). Alternative splicing: An important mechanism in stem cell biology. *World J. Stem Cells* **7**, 1–10.
- Cho, S., Hoang, A., Sinha, R., Zhong, X.-Y., Fu, X.-D., Krainer, A.R., and Ghosh, G. (2011). Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 8233–8238.
- Choi, N.-R., Lee, H.-J., Jung, S.-H., Hong, C.Y., Vo, M.-C., Hoang, M.-D., Kim, H.-J., and Lee, J.-J. (2015). Generation of potent dendritic cells with improved migration ability through p-cofilin and sarco/endoplasmic reticulum Ca²⁺ transport ATPase 2 regulation. *Cytotherapy* **17**, 1421–1433.
- Claeskens, A., Ongenae, N., Neefs, J.M., Cheyins, P., Kaijen, P., Cools, M., and Kutoh, E. (2000). Hevin is down-regulated in many cancers and is a negative regulator of cell growth and proliferation. *Br. J. Cancer* **82**, 1123–1130.
- 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.
- Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J.L., Bell, J.C., and Duncan, P.I. (1996a). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J.* **15**, 265–275.
- Colwill, K., Feng, L.L., Yeakley, J.M., Gish, G.D., Cáceres, J.F., Pawson, T., and Fu, X.D. (1996b). SRPK1 and Clk/Sty protein kinases show distinct substrate specificities for serine/arginine-rich splicing factors. *J. Biol. Chem.* **271**, 24569–24575.
- Cooke, M.S., Evans, M.D., Dizdaroglu, M., and Lunec, J. (2003a). Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J.* **17**, 1195–1214.

- Cooke, M.S., Evans, M.D., Dizdaroglu, M., and Lunec, J. (2003b). Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *17*, 1195–1214.
- Cooper, D.R., Carter, G., Li, P., Patel, R., Watson, J.E., and Patel, N.A. (2014). Long Non-Coding RNA NEAT1 Associates with SRp40 to Temporally Regulate PPAR γ 2 Splicing during Adipogenesis in 3T3-L1 Cells. *Genes* *5*, 1050–1063.
- Cooper, T.A., Wan, L., and Dreyfuss, G. (2009). RNA and Disease. *Cell* *136*, 777–793.
- Corcus, L., and Solier, S. (2005). Épissage alternatif, pathologie et thérapeutique moléculaire. *Médecine/sciences* *21*, 253–260.
- Cosentino-Gomes, D., Rocco-Machado, N., and Meyer-Fernandes, J.R. (2012). Cell Signaling through Protein Kinase C Oxidation and Activation. *Int. J. Mol. Sci.* *13*, 10697–10721.
- Cote, G.J., Zhu, W., Thomas, A., Martin, E., Murad, F., and Sharina, I.G. (2012). Hydrogen Peroxide Alters Splicing of Soluble Guanylyl Cyclase and Selectively Modulates Expression of Splicing Regulators in Human Cancer Cells. *PLoS ONE* *7*.
- Crawford, J.B., and Patton, J.G. (2006). Activation of α -Tropomyosin Exon 2 Is Regulated by the SR Protein 9G8 and Heterogeneous Nuclear Ribonucleoproteins H and F. *Mol. Cell. Biol.* *26*, 8791–8802.
- Cuppens, H., Lin, W., Jaspers, M., Costes, B., Teng, H., Vankeerberghen, A., Jorissen, M., Droogmans, G., Reynaert, I., Goossens, M., et al. (1998). Polyvariant mutant cystic fibrosis transmembrane conductance regulator genes. The polymorphic (Tg)m locus explains the partial penetrance of the T5 polymorphism as a disease mutation. *J. Clin. Invest.* *101*, 487–496.
- Dai, B., Hu, Z., Huang, H., Zhu, G., Xiao, Z., Wan, W., Zhang, P., Jia, W., and Zhang, L. (2014). Overexpressed KDM5B is associated with the progression of glioma and promotes glioma cell growth via downregulating p21. *Biochem. Biophys. Res. Commun.* *454*, 221–227.
- Damiano, F., Rochira, A., Tocci, R., Alemanno, S., Gnani, A., and Siculella, L. (2013). hnRNP A1 mediates the activation of the IRES-dependent SREBP-1a mRNA translation in response to endoplasmic reticulum stress. *Biochem. J.* *449*, 543–553.
- Damianov, A., and Black, D.L. (2010). Autoregulation of Fox protein expression to produce dominant negative splicing factors. *RNA N. Y. N* *16*, 405–416.
- Dansithong, W., Paul, S., Comai, L., and Reddy, S. (2005). MBNL1 is the primary determinant of focus formation and aberrant insulin receptor splicing in DM1. *J. Biol. Chem.* *280*, 5773–5780.
- Darley-Usmar, V.M., Hogg, N., O’Leary, V.J., Wilson, M.T., and Moncada, S. (1992). The simultaneous generation of superoxide and nitric oxide can initiate lipid peroxidation in human low density lipoprotein. *Free Radic. Res. Commun.* *17*, 9–20.
- Darnell, R.B. (2013). RNA Protein Interaction in Neurons. *Annu. Rev. Neurosci.* *36*, 243–270.
- Das, S., and Krainer, A.R. (2014). Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Mol. Cancer Res. MCR* *12*, 1195–1204.
- 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.

- Dasgupta, T., and Ladd, A.N. (2012). The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA binding proteins. *Wiley Interdiscip. Rev. RNA* 3, 104–121.
- Daubner, G.M., Cléry, A., and Allain, F.H.-T. (2013). RRM-RNA recognition: NMR or crystallography...and new findings. *Curr. Opin. Struct. Biol.* 23, 100–108.
- 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.
- Dayalan Naidu, S., Kostov, R.V., and Dinkova-Kostova, A.T. (2015). Transcription factors Hsf1 and Nrf2 engage in crosstalk for cytoprotection. *Trends Pharmacol. Sci.* 36, 6–14.
- 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.
- De Temmerman, N., Sermon, K., Seneca, S., De Rycke, M., Hilven, P., Lissens, W., Van Steirteghem, A., and Liebaers, I. (2004). Intergenerational instability of the expanded CTG repeat in the DMPK gene: studies in human gametes and preimplantation embryos. *Am. J. Hum. Genet.* 75, 325–329.
- Disher, K., and Skandalis, A. (2007). Evidence of the modulation of mRNA splicing fidelity in humans by oxidative stress and p53. *Genome Natl. Res. Counc. Can. Génome Cons. Natl. Rech. Can.* 50, 946–953.
- Dizdaroglu, M., Jaruga, P., Birincioglu, M., and Rodriguez, H. (2002). Free radical-induced damage to DNA: mechanisms and measurement. *Free Radic. Biol. Med.* 32, 1102–1115.
- Dominguez, C., and Allain, F.H.-T. (2005). Resonance assignments of the two N-terminal RNA recognition motifs (RRM) of the human heterogeneous nuclear ribonucleoprotein F (hnRNP F). *J. Biomol. NMR* 33, 282.
- Dominguez, C., and Allain, F.H.-T. (2006). NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. *Nucleic Acids Res.* 34, 3634–3645.
- Dostie, J., and Dreyfuss, G. (2002). Translation is required to remove Y14 from mRNAs in the cytoplasm. *Curr. Biol. CB* 12, 1060–1067.
- 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.
- Du, H., Cline, M.S., Osborne, R.J., Tuttle, D.L., Clark, T.A., Donohue, J.P., Hall, M.P., Shiue, L., Swanson, M.S., Thornton, C.A., et al. (2010). Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat. Struct. Mol. Biol.* 17, 187–193.
- Dujardin, G., Buratti, E., Charlet-Berguerand, N., Martins de Araujo, M., Mbopda, A., Le Jossic-Corcos, C., Pagani, F., Ferec, C., and Corcos, L. (2010). CELF proteins regulate CFTR pre-mRNA splicing: essential role of the divergent domain of ETR-3. *Nucleic Acids Res.* 38, 7273–7285.
- Dujardin, G., Lafaille, C., de la Mata, M., Marasco, L.E., Muñoz, M.J., Le Jossic-Corcos, C., Corcos, L., and Kornblihtt, A.R. (2014). How slow RNA polymerase II elongation favors alternative exon skipping. *Mol. Cell* 54, 683–690.

- Dundr, M., and Misteli, T. (2010). Biogenesis of nuclear bodies. *Cold Spring Harb. Perspect. Biol.* 2, a000711.
- Dutertre, M., Sanchez, G., De Cian, M.-C., Barbier, J., Dardenne, E., Gratadou, L., Dujardin, G., Le Jossic-Corcoss, C., Corcos, L., and Auboeuf, D. (2010). Cotranscriptional exon skipping in the genotoxic stress response. *Nat. Struct. Mol. Biol.* 17, 1358–1366.
- Echeverria, G.V., and Cooper, T.A. (2014). Muscleblind-like 1 activates insulin receptor exon 11 inclusion by enhancing U2AF65 binding and splicing of the upstream intron. *Nucleic Acids Res.* 42, 1893–1903.
- Edey, P., Marcaillou, C., Sahbatou, M., Labalme, A., Chastang, J., Touraine, R., Tubacher, E., Senni, F., Bober, M.B., Nampoothiri, S., et al. (2011). Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science* 332, 240–243.
- Edge, C., Gooding, C., and Smith, C.W.J. (2013). Dissecting domains necessary for activation and repression of splicing by Muscleblind-like protein 1. *BMC Mol. Biol.* 14, 29.
- Eggert, C., Chari, A., Lagerbauer, B., and Fischer, U. (2006). Spinal muscular atrophy: the RNP connection. *Trends Mol. Med.* 12, 113–121.
- Egloff, S., Szczepaniak, S.A., Dienstbier, M., Taylor, A., Knight, S., and Murphy, S. (2010). The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J. Biol. Chem.* 285, 20564–20569.
- 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.
- Eskiw, C.H., Dellaire, G., Mymryk, J.S., and Bazett-Jones, D.P. (2003). Size, position and dynamic behavior of PML nuclear bodies following cell stress as a paradigm for supramolecular trafficking and assembly. *J. Cell Sci.* 116, 4455–4466.
- Espinosa-Diez, C., Miguel, V., Mennerich, D., Kietzmann, T., Sánchez-Pérez, P., Cadenas, S., and Lamas, S. (2015). Antioxidant responses and cellular adjustments to oxidative stress. *Redox Biol.* 6, 183–197.
- 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.
- Fan, L., Li, A., Li, W., Cai, P., Yang, B., Zhang, M., Gu, Y., Shu, Y., Sun, Y., Shen, Y., et al. (2014). Novel role of Sarco/endoplasmic reticulum calcium ATPase 2 in development of colorectal cancer and its regulation by F36, a curcumin analog. *Biomed. Pharmacother.* 68, 1141–1148.
- Fardaei, M., Larkin, K., Brook, J.D., and Hamshere, M.G. (2001). In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res.* 29, 2766–2771.
- Faustino, N.A., and Cooper, T.A. (2005). Identification of putative new splicing targets for ETR-3 using sequences identified by systematic evolution of ligands by exponential enrichment. *Mol. Cell. Biol.* 25, 879–887.

- 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.
- Fischer, U., Englbrecht, C., and Chari, A. (2011). Biogenesis of spliceosomal small nuclear ribonucleoproteins. *Wiley Interdiscip. Rev. RNA* *2*, 718–731.
- Fisette, J.-F., Toutant, J., Dugré-Brisson, S., Desgroseillers, L., and Chabot, B. (2010). hnRNP A1 and hnRNP H can collaborate to modulate 5' splice site selection. *RNA N. Y. N* *16*, 228–238.
- Fransen, M., Nordgren, M., Wang, B., and Apanasets, O. (2012). Role of peroxisomes in ROS/RNS-metabolism: implications for human disease. *Biochim. Biophys. Acta* *1822*, 1363–1373.
- Friguet, B., and Szveda, L.I. (1997). Inhibition of the multicatalytic proteinase (proteasome) by 4-hydroxy-2-nonenal cross-linked protein. *FEBS Lett.* *405*, 21–25.
- Frilander, M.J., and Steitz, J.A. (1999). Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/branchpoint interactions. *Genes Dev.* *13*, 851–863.
- Fujimura, K., Kano, F., and Murata, M. (2008). Dual localization of the RNA binding protein CUGBP-1 to stress granule and perinucleolar compartment. *Exp. Cell Res.* *314*, 543–553.
- Fukumura, K., Kato, A., Jin, Y., Ideue, T., Hirose, T., Kataoka, N., Fujiwara, T., Sakamoto, H., and Inoue, K. (2007). Tissue-specific splicing regulator Fox-1 induces exon skipping by interfering E complex formation on the downstream intron of human F1 γ gene. *Nucleic Acids Res.* *35*, 5303–5311.
- Fulda, S., Gorman, A.M., Hori, O., and Samali, A. (2010). Cellular Stress Responses: Cell Survival and Cell Death. *Int. J. Cell Biol.* *2010*.
- Furling, D., Lemieux, D., Taneja, K., and Puymirat, J. (2001a). Decreased levels of myotonic dystrophy protein kinase (DMPK) and delayed differentiation in human myotonic dystrophy myoblasts. *Neuromuscul. Disord. NMD* *11*, 728–735.
- Furling, D., Coiffier, L., Mouly, V., Barbet, J.P., St Guily, J.L., Taneja, K., Gourdon, G., Junien, C., and Butler-Browne, G.S. (2001b). Defective satellite cells in congenital myotonic dystrophy. *Hum. Mol. Genet.* *10*, 2079–2087.
- Gama-Carvalho, M., and Carmo-Fonseca, M. (2001). The rules and roles of nucleocytoplasmic shuttling proteins. *FEBS Lett.* *498*, 157–163.
- Gates, D.P., Coonrod, L.A., and Berglund, J.A. (2011). Autoregulated splicing of muscleblind-like 1 (MBNL1) Pre-mRNA. *J. Biol. Chem.* *286*, 34224–34233.
- Ge, Y., and Porse, B.T. (2014). The functional consequences of intron retention: alternative splicing coupled to NMD as a regulator of gene expression. *BioEssays News Rev. Mol. Cell. Dev. Biol.* *36*, 236–243.
- Gelfman, S., Cohen, N., Yearim, A., and Ast, G. (2013). DNA-methylation effect on cotranscriptional splicing is dependent on GC architecture of the exon–intron structure. *Genome Res.* *23*, 789–799.
- Genestra, M. (2007). Oxyl radicals, redox-sensitive signalling cascades and antioxidants. *Cell. Signal.* *19*, 1807–1819.

- Gerstberger, S., Hafner, M., Ascano, M., and Tuschl, T. (2014). Evolutionary Conservation and Expression of Human RNA-Binding Proteins and Their Role in Human Genetic Disease. *Adv. Exp. Med. Biol.* 825, 1–55.
- Giorgi, C., Agnoletto, C., Baldini, C., Bononi, A., Bonora, M., Marchi, S., Missiroli, S., Patergnani, S., Poletti, F., Rimessi, A., et al. (2010). Redox control of protein kinase C: cell- and disease-specific aspects. *Antioxid. Redox Signal.* 13, 1051–1085.
- 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.
- Girotti, A.W. (1998). Lipid hydroperoxide generation, turnover, and effector action in biological systems. *J. Lipid Res.* 39, 1529–1542.
- Goers, E.S., Purcell, J., Voelker, R.B., Gates, D.P., and Berglund, J.A. (2010). MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res.* 38, 2467–2484.
- Gomes, A.V., Potter, J.D., and Szczesna-Cordary, D. (2002). The role of troponins in muscle contraction. *IUBMB Life* 54, 323–333.
- Gomes, A.V., Venkatraman, G., Davis, J.P., Tikunova, S.B., Engel, P., Solaro, R.J., and Potter, J.D. (2004). Cardiac Troponin T Isoforms Affect the Ca²⁺ Sensitivity of Force Development in the Presence of Slow Skeletal Troponin I INSIGHTS INTO THE ROLE OF TROPONIN T ISOFORMS IN THE FETAL HEART. *J. Biol. Chem.* 279, 49579–49587.
- Goo, Y.-H., and Cooper, T.A. (2009). CUGBP2 directly interacts with U2 17S snRNP components and promotes U2 snRNA binding to cardiac troponin T pre-mRNA. *Nucleic Acids Res.* 37, 4275–4286.
- Gooding, C., Edge, C., Lorenz, M., Coelho, M.B., Winters, M., Kaminski, C.F., Cherny, D., Eperon, I.C., and Smith, C.W.J. (2013). MBNL1 and PTB cooperate to repress splicing of Tpm1 exon 3. *Nucleic Acids Res.* 41, 4765–4782.
- Grammatikakis, I., Goo, Y.-H., Echeverria, G.V., and Cooper, T.A. (2011). Identification of MBNL1 and MBNL3 domains required for splicing activation and repression. *Nucleic Acids Res.* 39, 2769–2780.
- Graveley, B.R. (2000). Sorting out the complexity of SR protein functions. *RNA N. Y. N* 6, 1197–1211.
- 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.
- Hall, L.L., Smith, K.P., Byron, M., and Lawrence, J.B. (2006). Molecular anatomy of a speckle. *Anat. Rec. A. Discov. Mol. Cell. Evol. Biol.* 288, 664–675.
- 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, J., and Cooper, T.A. (2005). Identification of CELF splicing activation and repression domains in vivo. *Nucleic Acids Res.* 33, 2769–2780.

- Han, J., Ding, J.-H., Byeon, C.W., Kim, J.H., Hertel, K.J., Jeong, S., and Fu, X.-D. (2011). SR proteins induce alternative exon skipping through their activities on the flanking constitutive exons. *Mol. Cell Biol.* *31*, 793–802.
- Han, S.P., Tang, Y.H., and Smith, R. (2010). Functional diversity of the hnRNPs: past, present and perspectives. *Biochem. J.* *430*, 379–392.
- Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* *495*, 384–388.
- 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.
- Hayes, J.D., and McMahon, M. (2001). Molecular basis for the contribution of the antioxidant responsive element to cancer chemoprevention. *Cancer Lett.* *174*, 103–113.
- He, Y., and Smith, R. (2009). Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B. *Cell. Mol. Life Sci. CMLS* *66*, 1239–1256.
- He, H., Liyanarachchi, S., Akagi, K., Nagy, R., Li, J., Dietrich, R.C., Li, W., Sebastian, N., Wen, B., Xin, B., et al. (2011a). Mutations in U4atac snRNA, a Component of the Minor Spliceosome, in the Developmental Disorder MOPD I. *Science* *332*, 238–240.
- He, H., Liyanarachchi, S., Akagi, K., Nagy, R., Li, J., Dietrich, R.C., Li, W., Sebastian, N., Wen, B., Xin, B., et al. (2011b). Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science* *332*, 238–240.
- Henry, R.W., Ford, E., Mital, R., Mittal, V., and Hernandez, N. (1998). Crossing the line between RNA polymerases: transcription of human snRNA genes by RNA polymerases II and III. *Cold Spring Harb. Symp. Quant. Biol.* *63*, 111–120.
- Hensen, S.M.M., Heldens, L., van Enkevort, C.M.W., van Genesen, S.T., Pruijn, G.J.M., and Lubsen, N.H. (2013). Activation of the antioxidant response in methionine deprived human cells results in an HSF1-independent increase in HSPA1A mRNA levels. *Biochimie* *95*, 1245–1251.
- Hernandez, N., and Weiner, A.M. (1986). Formation of the 3' end of U1 snRNA requires compatible snRNA promoter elements. *Cell* *47*, 249–258.
- Hnilicová, J., Hozeifi, S., Dušková, E., Icha, J., Tománková, T., and Staněk, D. (2011). Histone Deacetylase Activity Modulates Alternative Splicing. *PLoS ONE* *6*.
- Ho, T.H., Charlet-B, N., Poulos, M.G., Singh, G., Swanson, M.S., and Cooper, T.A. (2004a). Muscleblind proteins regulate alternative splicing. *EMBO J.* *23*, 3103–3112.
- Ho, T.H., Charlet-B, N., Poulos, M.G., Singh, G., Swanson, M.S., and Cooper, T.A. (2004b). Muscleblind proteins regulate alternative splicing. *EMBO J.* *23*, 3103–3112.
- Ho, T.H., Bundman, D., Armstrong, D.L., and Cooper, T.A. (2005). Transgenic mice expressing CUG-BP1 reproduce splicing mis-regulation observed in myotonic dystrophy. *Hum. Mol. Genet.* *14*, 1539–1547.

- Hogg, N., and Kalyanaraman, B. (1999). Nitric oxide and lipid peroxidation. *Biochim. Biophys. Acta* *1411*, 378–384.
- Holmström, K.M., and Finkel, T. (2014). Cellular mechanisms and physiological consequences of redox-dependent signalling. *Nat. Rev. Mol. Cell Biol.* *15*, 411–421.
- Hoskins, A.A., and Moore, M.J. (2012a). The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.* *37*, 179–188.
- Hoskins, A.A., and Moore, M.J. (2012b). The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.* *37*, 179–188.
- Howard, J.M., and Sanford, J.R. (2015). The RNAissance family: SR proteins as multifaceted regulators of gene expression. *Wiley Interdiscip. Rev. RNA* *6*, 93–110.
- Höweler, C.J., Busch, H.F., Geraedts, J.P., Niermeijer, M.F., and Staal, A. (1989). Anticipation in myotonic dystrophy: fact or fiction? *Brain J. Neurol.* *112 (Pt 3)*, 779–797.
- Huang, Y., and Steitz, J.A. (2005). SRprises along a messenger’s journey. *Mol. Cell* *17*, 613–615.
- Huang, H.-C., Nguyen, T., and Pickett, C.B. (2002). Phosphorylation of Nrf2 at Ser-40 by protein kinase C regulates antioxidant response element-mediated transcription. *J. Biol. Chem.* *277*, 42769–42774.
- Huang, J., Jiao, J., Xu, W., Zhao, H., Zhang, C., Shi, Y., and Xiao, Z. (2015a). miR-155 is upregulated in patients with active tuberculosis and inhibits apoptosis of monocytes by targeting FOXO3. *Mol. Med. Rep.*
- 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.
- Huang, Y., Li, W., Su, Z.-Y., and Kong, A.-N.T. (2015b). The complexity of the Nrf2 pathway: beyond the antioxidant response. *J. Nutr. Biochem.*
- Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B.Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S., et al. (2012). Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep.* *1*, 167–178.
- Huichalaf, C., Sakai, K., Jin, B., Jones, K., Wang, G.-L., Schoser, B., Schneider-Gold, C., Sarkar, P., Pereira-Smith, O.M., Timchenko, N., et al. (2010). Expansion of CUG RNA repeats causes stress and inhibition of translation in myotonic dystrophy 1 (DM1) cells. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *24*, 3706–3719.
- Hurley, P.J., Marchionni, L., Simons, B.W., Ross, A.E., Peskoe, S.B., Miller, R.M., Erho, N., Vergara, I.A., Ghadessi, M., Huang, Z., et al. (2012). Secreted protein, acidic and rich in cysteine-like 1 (SPARCL1) is down regulated in aggressive prostate cancers and is prognostic for poor clinical outcome. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 14977–14982.
- Hurley, P.J., Hughes, R.M., Simons, B., Huang, J., Miller, R.M., Shinder, B., Haffner, M.C., Esopi, D., Kinura, Y., Jabbari, J., et al. (2015). Androgen Regulated SPARCL1 in the Tumor Microenvironment Inhibits Metastatic Progression. *Cancer Res.*

- Iakova, P., Wang, G.-L., Timchenko, L., Michalak, M., Pereira-Smith, O.M., Smith, J.R., and Timchenko, N.A. (2004). Competition of CUGBP1 and calreticulin for the regulation of p21 translation determines cell fate. *EMBO J.* *23*, 406–417.
- 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.
- Ideue, T., Sasaki, Y.T.F., Hagiwara, M., and Hirose, T. (2007). Introns play an essential role in splicing-dependent formation of the exon junction complex. *Genes Dev.* *21*, 1993–1998.
- Indo, H.P., Yen, H.-C., Nakanishi, I., Matsumoto, K.-I., Tamura, M., Nagano, Y., Matsui, H., Gusev, O., Cornette, R., Okuda, T., et al. (2015). A mitochondrial superoxide theory for oxidative stress diseases and aging. *J. Clin. Biochem. Nutr.* *56*, 1–7.
- Inukai, A., Doyu, M., Kato, T., Liang, Y., Kuru, S., Yamamoto, M., Kobayashi, Y., and Sobue, G. (2000). Reduced expression of DMAHP/SIX5 gene in myotonic dystrophy muscle. *Muscle Nerve* *23*, 1421–1426.
- Ip, J.Y., Schmidt, D., Pan, Q., Ramani, A.K., Fraser, A.G., Odom, D.T., and Blencowe, B.J. (2011). Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res.* *21*, 390–401.
- 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.
- Izquierdo, J.M., Majós, N., Bonnal, S., Martínez, C., Castelo, R., Guigó, R., Bilbao, D., and Valcárcel, J. (2005). Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Mol. Cell* *19*, 475–484.
- Jangi, M., Boutz, P.L., Paul, P., and Sharp, P.A. (2014). Rbfox2 controls autoregulation in RNA-binding protein networks. *Genes Dev.* *28*, 637–651.
- Janshoff, A., and Steinem, C. (2015). Mechanics of lipid bilayers: What do we learn from pore-spanning membranes? *Biochim. Biophys. Acta BBA - Mol. Cell Res.*
- Jeck, W.R., Sorrentino, J.A., Wang, K., Slevin, M.K., Burd, C.E., Liu, J., Marzluff, W.F., and Sharpless, N.E. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA N. Y. N* *19*, 141–157.
- Jensen, K.B., Musunuru, K., Lewis, H.A., Burley, S.K., and Darnell, R.B. (2000). The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 5740–5745.
- Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C.Y., Xiao, R., Burge, C.B., and Fu, X.-D. (2013). SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* *153*, 855–868.
- 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., 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.
- Kalsotra, A., Xiao, X., Ward, A.J., Castle, J.C., Johnson, J.M., Burge, C.B., and Cooper, T.A. (2008). A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 20333–20338.
- Kalsotra, A., Wang, K., Li, P.-F., and Cooper, T.A. (2010). MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development. *Genes Dev.* *24*, 653–658.
- Kambach, C., Walke, S., and Nagai, K. (1999). Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. *Curr. Opin. Struct. Biol.* *9*, 222–230.
- Kanadia, R.N., Urbinati, C.R., Crusselle, V.J., Luo, D., Lee, Y.-J., Harrison, J.K., Oh, S.P., and Swanson, M.S. (2003a). Developmental expression of mouse muscleblind genes Mbnl1, Mbnl2 and Mbnl3. *Gene Expr. Patterns GEP* *3*, 459–462.
- Kanadia, R.N., Johnstone, K.A., Mankodi, A., Lungu, C., Thornton, C.A., Esson, D., Timmers, A.M., Hauswirth, W.W., and Swanson, M.S. (2003b). A muscleblind knockout model for myotonic dystrophy. *Science* *302*, 1978–1980.
- Kanadia, R.N., Shin, J., Yuan, Y., Beattie, S.G., Wheeler, T.M., Thornton, C.A., and Swanson, M.S. (2006). Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 11748–11753.
- Kanai, Y., Dohmae, N., and Hirokawa, N. (2004). Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* *43*, 513–525.
- Kano, S., Nishida, K., Kurebe, H., Nishiyama, C., Kita, K., Akaike, Y., Kajita, K., Kurokawa, K., Masuda, K., Kuwano, Y., et al. (2014). Oxidative stress-inducible truncated serine/arginine-rich splicing factor 3 regulates interleukin-8 production in human colon cancer cells. *Am. J. Physiol. - Cell Physiol.* *306*, C250–C262.
- Kato, T., Shimono, Y., Hasegawa, M., Jijiwa, M., Enomoto, A., Asai, N., Murakumo, Y., and Takahashi, M. (2009). Characterization of the HDAC1 complex that regulates the sensitivity of cancer cells to oxidative stress. *Cancer Res.* *69*, 3597–3604.
- Katrakha, I.A. (2013). Human cardiac troponin complex. Structure and functions. *Biochem. Biokhimiia* *78*, 1447–1465.
- 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.
- Khodor, Y.L., Menet, J.S., Tolan, M., and Rosbash, M. (2012). Cotranscriptional splicing efficiency differs dramatically between *Drosophila* and mouse. *RNA* *18*, 2174–2186.
- 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.

- Kiliszek, A., and Rypniewski, W. (2014). Structural studies of CNG repeats. *Nucleic Acids Res.* *42*, 8189–8199.
- Kiliszek, A., Kierzek, R., Krzyzosiak, W.J., and Rypniewski, W. (2009). Structural insights into CUG repeats containing the “stretched U-U wobble”: implications for myotonic dystrophy. *Nucleic Acids Res.* *37*, 4149–4156.
- Kim, J.H., and Richter, J.D. (2006). Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. *Mol. Cell* *24*, 173–183.
- Kim, D.-H., Langlois, M.-A., Lee, K.-B., Riggs, A.D., Puymirat, J., and Rossi, J.J. (2005). HnRNP H inhibits nuclear export of mRNA containing expanded CUG repeats and a distal branch point sequence. *Nucleic Acids Res.* *33*, 3866–3874.
- Kimura, T., Nakamori, M., Lueck, J.D., Pouliquin, P., Aoike, F., Fujimura, H., Dirksen, R.T., Takahashi, M.P., Dulhunty, A.F., and Sakoda, S. (2005). Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca²⁺-ATPase in myotonic dystrophy type 1. *Hum. Mol. Genet.* *14*, 2189–2200.
- Kino, Y., Mori, D., Oma, Y., Takeshita, Y., Sasagawa, N., and Ishiura, S. (2004). Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum. Mol. Genet.* *13*, 495–507.
- Kino, Y., Washizu, C., Oma, Y., Onishi, H., Nezu, Y., Sasagawa, N., Nukina, N., and Ishiura, S. (2009). MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1. *Nucleic Acids Res.* *37*, 6477–6490.
- Kino, Y., Washizu, C., Kurosawa, M., Oma, Y., Hattori, N., Ishiura, S., and Nukina, N. (2015). Nuclear localization of MBNL1: splicing-mediated autoregulation and repression of repeat-derived aberrant proteins. *Hum. Mol. Genet.* *24*, 740–756.
- Klesert, T.R., Cho, D.H., Clark, J.I., Maylie, J., Adelman, J., Snider, L., Yuen, E.C., Soriano, P., and Tapscott, S.J. (2000). Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy. *Nat. Genet.* *25*, 105–109.
- Kobayashi, M., and Yamamoto, M. (2006). Nrf2-Keap1 regulation of cellular defense mechanisms against electrophiles and reactive oxygen species. *Adv. Enzyme Regul.* *46*, 113–140.
- Konarska, M.M., Vilardell, J., and Query, C.C. (2006). Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol. Cell* *21*, 543–553.
- Konieczny, P., Stepniak-Konieczna, E., and Sobczak, K. (2014). MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Res.* *42*, 10873–10887.
- KORNBLIHTT, A.R., DE LA MATA, M., FEDEDA, J.P., MUÑOZ, M.J., and NOGUÉS, G. (2004). Multiple links between transcription and splicing. *RNA* *10*, 1489–1498.
- Kumar, A., Agarwal, S., Agarwal, D., and Phadke, S.R. (2013). Myotonic dystrophy type 1 (DM1): A triplet repeat expansion disorder. *Gene* *522*, 226–230.
- Kuroyanagi, H. (2009). Fox-1 family of RNA-binding proteins. *Cell. Mol. Life Sci. CMLS* *66*, 3895–3907.

- Kuyumcu-Martinez, N.M., Wang, G.-S., and Cooper, T.A. (2007). Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. *Mol. Cell* 28, 68–78.
- Kwak, M.-K., Wakabayashi, N., Itoh, K., Motohashi, H., Yamamoto, M., and Kensler, T.W. (2003). Modulation of Gene Expression by Cancer Chemopreventive Dithiolethiones through the Keap1-Nrf2 Pathway IDENTIFICATION OF NOVEL GENE CLUSTERS FOR CELL SURVIVAL. *J. Biol. Chem.* 278, 8135–8145.
- Kyburz, A., Friedlein, A., Langen, H., and Keller, W. (2006). Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. *Mol. Cell* 23, 195–205.
- Ladd, A.N. (2013). CUG-BP, Elav-like family (CELF)-mediated alternative splicing regulation in the brain during health and disease. *Mol. Cell. Neurosci.* 56, 456–464.
- Ladd, A.N., and Cooper, T.A. (2004). Multiple domains control the subcellular localization and activity of ETR-3, a regulator of nuclear and cytoplasmic RNA processing events. *J. Cell Sci.* 117, 3519–3529.
- Ladd, A.N., Charlet, N., and Cooper, T.A. (2001a). The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol. Cell. Biol.* 21, 1285–1296.
- Ladd, A.N., Charlet-B, N., and Cooper, T.A. (2001b). The CELF Family of RNA Binding Proteins Is Implicated in Cell-Specific and Developmentally Regulated Alternative Splicing. *Mol. Cell. Biol.* 21, 1285–1296.
- Ladd, A.N., Nguyen, N.H., Malhotra, K., and Cooper, T.A. (2004). CELF6, a Member of the CELF Family of RNA-binding Proteins, Regulates Muscle-specific Splicing Enhancer-dependent Alternative Splicing. *J. Biol. Chem.* 279, 17756–17764.
- Ladd, A.N., Stenberg, M.G., Swanson, M.S., and Cooper, T.A. (2005). Dynamic balance between activation and repression regulates pre-mRNA alternative splicing during heart development. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* 233, 783–793.
- Lallemant-Breitenbach, V., and de Thé, H. (2010). PML nuclear bodies. *Cold Spring Harb. Perspect. Biol.* 2, a000661.
- Lallemant-Breitenbach, V., Zhu, J., Puvion, F., Koken, M., Honoré, N., Doubeikovsky, A., Duprez, E., Pandolfi, P.P., Puvion, E., Freemont, P., et al. (2001). Role of promyelocytic leukemia (PML) sumolation in nuclear body formation, 11S proteasome recruitment, and As2O3-induced PML or PML/retinoic acid receptor alpha degradation. *J. Exp. Med.* 193, 1361–1371.
- Lallemant-Breitenbach, V., Jeanne, M., Benhenda, S., Nasr, R., Lei, M., Peres, L., Zhou, J., Zhu, J., Raught, B., and de Thé, H. (2008). Arsenic degrades PML or PML-RARalpha through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nat. Cell Biol.* 10, 547–555.
- Lee, K.-M., and Tarn, W.-Y. (2013). Coupling pre-mRNA processing to transcription on the RNA factory assembly line. *RNA Biol.* 10, 380–390.
- Lee, J.-A., Tang, Z.-Z., and Black, D.L. (2009). An inducible change in Fox-1/A2BP1 splicing modulates the alternative splicing of downstream neuronal target exons. *Genes Dev.* 23, 2284–2293.

- Lee, J.E., Lee, J.Y., Wilusz, J., Tian, B., and Wilusz, C.J. (2010). Systematic analysis of cis-elements in unstable mRNAs demonstrates that CUGBP1 is a key regulator of mRNA decay in muscle cells. *PLoS One* 5, e11201.
- Lee, K.-Y., Li, M., Manchanda, M., Batra, R., Charizanis, K., Mohan, A., Warren, S.A., Chamberlain, C.M., Finn, D., Hong, H., et al. (2013). Compound loss of muscleblind-like function in myotonic dystrophy. *EMBO Mol. Med.* 5, 1887–1900.
- Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* 19, 6860–6869.
- Leonard, S.S., Harris, G.K., and Shi, X. (2004). Metal-induced oxidative stress and signal transduction. *Free Radic. Biol. Med.* 37, 1921–1942.
- Levine, R.L., Mosoni, L., Berlett, B.S., and Stadtman, E.R. (1996). Methionine residues as endogenous antioxidants in proteins. *Proc. Natl. Acad. Sci. U. S. A.* 93, 15036–15040.
- Lev Maor, G., Yearim, A., and Ast, G. (2015). The alternative role of DNA methylation in splicing regulation. *Trends Genet.* 31, 274–280.
- 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, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707–719.
- Li, W., Thakor, N., Xu, E.Y., Huang, Y., Chen, C., Yu, R., Holcik, M., and Kong, A.-N. (2010). An internal ribosomal entry site mediates redox-sensitive translation of Nrf2. *Nucleic Acids Res.* 38, 778–788.
- Li, X., Su, Y., Pan, J., Zhou, Z., Song, B., Xiong, E., and Chen, Z. (2013). Connexin 26 is down-regulated by KDM5B in the progression of bladder cancer. *Int. J. Mol. Sci.* 14, 7866–7879.
- Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456, 464–469.
- 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, C.-C., Jiang, W., Mitra, R., Cheng, F., Yu, H., and Zhao, Z. (2015). Regulation rewiring analysis reveals mutual regulation between STAT1 and miR-155-5p in tumor immunosurveillance in seven major cancers. *Sci. Rep.* 5.
- Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X.-D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. *Nat. Struct. Mol. Biol.* 15, 819–826.
- Lin, X., Miller, J.W., Mankodi, A., Kanadia, R.N., Yuan, Y., Moxley, R.T., Swanson, M.S., and Thornton, C.A. (2006). Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.* 15, 2087–2097.

- 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, J.-L., and Gall, J.G. (2007). U bodies are cytoplasmic structures that contain uridine-rich small nuclear ribonucleoproteins and associate with P bodies. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 11655–11659.
- Liu, G., Chen, X., Gao, Y., Lewis, T., Barthelemy, J., and Leffak, M. (2012). Altered replication in human cells promotes DMPK (CTG)(n) · (CAG)(n) repeat instability. *Mol. Cell. Biol.* *32*, 1618–1632.
- 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.
- Loria, P.M., Duke, A., Rand, J.B., and Hobert, O. (2003). Two neuronal, nuclear-localized RNA binding proteins involved in synaptic transmission. *Curr. Biol. CB* *13*, 1317–1323.
- Loschi, M., Leishman, C.C., Berardone, N., and Boccaccio, G.L. (2009). Dynein and kinesin regulate stress-granule and P-body dynamics. *J. Cell Sci.* *122*, 3973–3982.
- Lu, Z.-X., Jiang, P., and Xing, Y. (2012). Genetic variation of pre-mRNA alternative splicing in human populations. *Wiley Interdiscip. Rev. RNA* *3*, 581–592.
- 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.
- Mailloux, R.J., and Harper, M.-E. (2012). Mitochondrial proticity and ROS signaling: lessons from the uncoupling proteins. *Trends Endocrinol. Metab. TEM* *23*, 451–458.
- Malhotra, D., Portales-Casamar, E., Singh, A., Srivastava, S., Arenillas, D., Happel, C., Shyr, C., Wakabayashi, N., Kensler, T.W., Wasserman, W.W., et al. (2010). Global mapping of binding sites for Nrf2 identifies novel targets in cell survival response through ChIP-Seq profiling and network analysis. *Nucleic Acids Res.* *38*, 5718–5734.
- Mangge, H., Becker, K., Fuchs, D., and Gostner, J.M. (2014). Antioxidants, inflammation and cardiovascular disease. *World J. Cardiol.* *6*, 462–477.
- Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C.A. (2000). Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* *289*, 1769–1773.
- Mankodi, A., Urbinati, C.R., Yuan, Q.P., Moxley, R.T., Sansone, V., Krym, M., Henderson, D., Schalling, M., Swanson, M.S., and Thornton, C.A. (2001). Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum. Mol. Genet.* *10*, 2165–2170.
- 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.
- Maracchioni, A., Totaro, A., Angelini, D.F., Di Penta, A., Bernardi, G., Carri, M.T., and Achsel, T. (2007). Mitochondrial damage modulates alternative splicing in neuronal cells: implications for neurodegeneration. *J. Neurochem.* *100*, 142–153.

- 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-Cianférani, 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.
- Maris, C., Dominguez, C., and Allain, F.H.-T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* *272*, 2118–2131.
- Marnett, L.J., Riggins, J.N., and West, J.D. (2003). Endogenous generation of reactive oxidants and electrophiles and their reactions with DNA and protein. *J. Clin. Invest.* *111*, 583–593.
- Marquis, J., Paillard, L., Audic, Y., Cosson, B., Danos, O., Le Bec, C., and Osborne, H.B. (2006). CUG-BP1/CELF1 requires UGU-rich sequences for high-affinity binding. *Biochem. J.* *400*, 291–301.
- Martinez-Contreras, R., Fiset, 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.
- Maslon, M.M., Heras, S.R., Bellora, N., Eyra, E., and Cáceres, J.F. (2014). The translational landscape of the splicing factor SRSF1 and its role in mitosis. *eLife* e02028.
- Masuda, A., Andersen, H.S., Doktor, T.K., Okamoto, T., Ito, M., Andresen, B.S., and Ohno, K. (2012). CUGBP1 and MBNL1 preferentially bind to 3' UTRs and facilitate mRNA decay. *Sci. Rep.* *2*, 209.
- de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell* *12*, 525–532.
- Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* *15*, 108–121.
- Matlin, A.J., Clark, F., and Smith, C.W.J. (2005). Understanding alternative splicing: towards a cellular code. *Nat. Rev. Mol. Cell Biol.* *6*, 386–398.
- McMahon, M., Thomas, N., Itoh, K., Yamamoto, M., and Hayes, J.D. (2006). Dimerization of substrate adaptors can facilitate cullin-mediated ubiquitylation of proteins by a “tethering” mechanism: a two-site interaction model for the Nrf2-Keap1 complex. *J. Biol. Chem.* *281*, 24756–24768.
- Meins, M., Schlickum, S., Wilhelm, C., Missbach, J., Yadav, S., Gläser, B., Grzmil, M., Burfeind, P., and Laccone, F. (2002). Identification and characterization of murine Brunol4, a new member of the elav/bruno family. *Cytogenet. Genome Res.* *97*, 254–260.
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* *495*, 333–338.
- Meng, T.-C., Fukada, T., and Tonks, N.K. (2002). Reversible oxidation and inactivation of protein tyrosine phosphatases in vivo. *Mol. Cell* *9*, 387–399.

- 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.
- Michalowski, S., Miller, J.W., Urbinati, C.R., Paliouras, M., Swanson, M.S., and Griffith, J. (1999). Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. *Nucleic Acids Res.* *27*, 3534–3542.
- Millhouse, S., and Manley, J.L. (2005). The C-terminal domain of RNA polymerase II functions as a phosphorylation-dependent splicing activator in a heterologous protein. *Mol. Cell. Biol.* *25*, 533–544.
- Misquitta, C.M., Ghosh, P., Mwanjewe, J., and Grover, A.K. (2005). Control of SERCA2a Ca²⁺ pump mRNA stability by nuclear proteins: role of domains in the 3'-untranslated region. *Cell Calcium* *37*, 17–24.
- 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.
- Moehle, E.A., Braberg, H., Krogan, N.J., and Guthrie, C. (2014). Adventures in time and space: splicing efficiency and RNA polymerase II elongation rate. *RNA Biol.* *11*, 313–319.
- 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.
- Moncada, S., and Bolaños, J.P. (2006). Nitric oxide, cell bioenergetics and neurodegeneration. *J. Neurochem.* *97*, 1676–1689.
- Mooers, B.H.M., Logue, J.S., and Berglund, J.A. (2005). The structural basis of myotonic dystrophy from the crystal structure of CUG repeats. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 16626–16631.
- 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.
- Moraes, K.C.M., Wilusz, C.J., and Wilusz, J. (2006). CUG-BP binds to RNA substrates and recruits PARN deadenylase. *RNA N. Y. N* *12*, 1084–1091.
- Mukhopadhyay, D., Houchen, C.W., Kennedy, S., Dieckgraefe, B.K., and Anant, S. (2003). Coupled mRNA stabilization and translational silencing of cyclooxygenase-2 by a novel RNA binding protein, CUGBP2. *Mol. Cell* *11*, 113–126.
- Muller, F.L., Liu, Y., and Van Remmen, H. (2004). Complex III releases superoxide to both sides of the inner mitochondrial membrane. *J. Biol. Chem.* *279*, 49064–49073.
- Muñoz, M.J., Santangelo, M.S.P., 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.
- Murphy, M.P. (2009). How mitochondria produce reactive oxygen species. *Biochem. J.* *417*, 1–13.

- 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.
- 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.
- Nayler, O., Cap, C., and Stamm, S. (1998). Human transformer-2-beta gene (SFRS10): complete nucleotide sequence, chromosomal localization, and generation of a tissue-specific isoform. *Genomics* *53*, 191–202.
- Niemela, E.H., Oghabian, A., Staals, R.H.J., Greco, D., Pruijn, G.J.M., and Frilander, M.J. (2014). Global analysis of the nuclear processing of transcripts with unspliced U12-type introns by the exosome. *Nucleic Acids Res.* *42*, 7358–7369.
- Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* *463*, 457–463.
- Ohno, M., and Shimura, Y. (1996). A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNA from the spliceosome. *Genes Dev.* *10*, 997–1007.
- Ohshima, D., Arimoto-Matsuzaki, K., Tomida, T., Takekawa, M., and Ichikawa, K. (2015). Spatio-temporal Dynamics and Mechanisms of Stress Granule Assembly. *PLoS Comput. Biol.* *11*, e1004326.
- Okunola, H.L., and Krainer, A.R. (2009). Cooperative-binding and splicing-repressive properties of hnRNP A1. *Mol. Cell. Biol.* *29*, 5620–5631.
- Ortuño-Pineda, C., Galindo-Rosales, J.M., Calderón-Salinas, J.V., Villegas-Sepúlveda, N., Saucedo-Cárdenas, O., De Nova-Ocampo, M., and Valdés, J. (2012). Binding of hnRNP H and U2AF65 to respective G-codes and a poly-uridine tract collaborate in the N50-5' splice site selection of the REST N exon in H69 cells. *PLoS One* *7*, e40315.
- Ott, M., Zhivotovsky, B., and Orrenius, S. (2007). Role of cardiolipin in cytochrome c release from mitochondria. *Cell Death Differ.* *14*, 1243–1247.
- Padgett, R.A. (2012). New connections between splicing and human disease. *Trends Genet.* *28*, 147–154.
- Paillard, L., Legagneux, V., and Beverley Osborne, H. (2003). A functional deadenylation assay identifies human CUG-BP as a deadenylation factor. *Biol. Cell Auspices Eur. Cell Biol. Organ.* *95*, 107–113.
- Pan, K., Lee, J.T.H., Huang, Z., and Wong, C.-M. (2015). Coupling and Coordination in Gene Expression Processes with Pre-mRNA Splicing. *Int. J. Mol. Sci.* *16*, 5682–5696.
- Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., Huang, J., Yeo, G.W., Ares, M., and Fu, X.-D. (2013). Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol. Cell* *50*, 223–235.
- Paradies, G., Petrosillo, G., Pistolese, M., and Ruggiero, F.M. (2001). Reactive oxygen species generated by the mitochondrial respiratory chain affect the complex III activity via cardiolipin peroxidation in beef-heart submitochondrial particles. *Mitochondrion* *1*, 151–159.

- Paradies, G., Petrosillo, G., Pistolese, M., and Ruggiero, F.M. (2002). Reactive oxygen species affect mitochondrial electron transport complex I activity through oxidative cardiolipin damage. *Gene* 286, 135–141.
- Paradies, G., Petrosillo, G., Paradies, V., and Ruggiero, F.M. (2009). Role of cardiolipin peroxidation and Ca²⁺ in mitochondrial dysfunction and disease. *Cell Calcium* 45, 643–650.
- Paradies, G., Petrosillo, G., Paradies, V., and Ruggiero, F.M. (2011). Mitochondrial dysfunction in brain aging: Role of oxidative stress and cardiolipin. *Neurochem. Int.* 58, 447–457.
- Paradies, G., Paradies, V., Ruggiero, F.M., and Petrosillo, G. (2014). Oxidative stress, cardiolipin and mitochondrial dysfunction in nonalcoholic fatty liver disease. *World J. Gastroenterol. WJG* 20, 14205–14218.
- 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.
- Parker, R., and Sheth, U. (2007). P bodies and the control of mRNA translation and degradation. *Mol. Cell* 25, 635–646.
- Pastore, A., Federici, G., Bertini, E., and Piemonte, F. (2003). Analysis of glutathione: implication in redox and detoxification. *Clin. Chim. Acta Int. J. Clin. Chem.* 333, 19–39.
- Paul, S., Dansithong, W., Kim, D., Rossi, J., Webster, N.J.G., Comai, L., and Reddy, S. (2006). Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing. *EMBO J.* 25, 4271–4283.
- Paul, S., Dansithong, W., Jog, S.P., Holt, I., Mittal, S., Brook, J.D., Morris, G.E., Comai, L., and Reddy, S. (2011). Expanded CUG Repeats Dysregulate RNA Splicing by Altering the Stoichiometry of the Muscleblind 1 Complex. *J. Biol. Chem.* 286, 38427–38438.
- Paushkin, S., Gubitz, A.K., Massenet, S., and Dreyfuss, G. (2002). The SMN complex, an assemblysome of ribonucleoproteins. *Curr. Opin. Cell Biol.* 14, 305–312.
- Pedrotti, S., and Cooper, T.A. (2014). In Brief: (mis)splicing in disease. *J. Pathol.* 233, 1–3.
- Petkovic, S., and Müller, S. (2015). RNA circularization strategies in vivo and in vitro. *Nucleic Acids Res.* 43, 2454–2465.
- Petrosillo, G., Casanova, G., Matera, M., Ruggiero, F.M., and Paradies, G. (2006). Interaction of peroxidized cardiolipin with rat-heart mitochondrial membranes: induction of permeability transition and cytochrome c release. *FEBS Lett.* 580, 6311–6316.
- Philippe, J.-V., Ayadi, L., Branlant, C., and Behm-Ansmant, I. (2015). Probing small non-coding RNAs structures. *Methods Mol. Biol. Clifton NJ* 1296, 119–136.
- Philips, A.V., Timchenko, L.T., and Cooper, T.A. (1998). Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* 280, 737–741.
- Piñol-Roma, S., and Dreyfuss, G. (1992). Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* 355, 730–732.

- Piñol-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.
- Pizzimenti, S., Briatore, F., Laurora, S., Toaldo, C., Maggio, M., De Grandi, M., Meaglia, L., Menegatti, E., Giglioni, B., Dianzani, M.U., et al. (2006). 4-Hydroxynonenal inhibits telomerase activity and hTERT expression in human leukemic cell lines. *Free Radic. Biol. Med.* *40*, 1578–1591.
- 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.
- Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R., and Bickmore, W.A. (2012). Psp1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet.* *8*, e1002717.
- Pratte, A., Prévost, C., Puymirat, J., and Mathieu, J. (2015). Anticipation in myotonic dystrophy type 1 parents with small CTG expansions. *Am. J. Med. Genet. A.* *167*, 708–714.
- Purcell, J., Oddo, J.C., Wang, E.T., and Berglund, J.A. (2012). Combinatorial Mutagenesis of MBNL1 Zinc Fingers Elucidates Distinct Classes of Regulatory Events. *Mol. Cell. Biol.* *32*, 4155–4167.
- Radi, R., Beckman, J.S., Bush, K.M., and Freeman, B.A. (1991). Peroxynitrite-induced membrane lipid peroxidation: the cytotoxic potential of superoxide and nitric oxide. *Arch. Biochem. Biophys.* *288*, 481–487.
- Radi, R., Peluffo, G., Alvarez, M.N., Naviliat, M., and Cayota, A. (2001). Unraveling peroxynitrite formation in biological systems. *Free Radic. Biol. Med.* *30*, 463–488.
- Raker, V.A., Plessel, G., and Lührmann, R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *EMBO J.* *15*, 2256–2269.
- Raponi, M., and Baralle, D. (2010). Alternative splicing: good and bad effects of translationally silent substitutions. *FEBS J.* *277*, 836–840.
- Rappsilber, J., Ryder, U., Lamond, A.I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. *Genome Res.* *12*, 1231–1245.
- Rau, F., Freyermuth, F., Fugier, C., Villemin, J.-P., Fischer, M.-C., Jost, B., Dembele, D., Gourdon, G., Nicole, A., Duboc, D., et al. (2011a). Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nat. Struct. Mol. Biol.* *18*, 840–845.
- Rau, F., Freyermuth, F., Fugier, C., Villemin, J.-P., Fischer, M.-C., Jost, B., Dembele, D., Gourdon, G., Nicole, A., Duboc, D., et al. (2011b). Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nat. Struct. Mol. Biol.* *18*, 840–845.
- Reczek, C.R., and Chandel, N.S. (2015). ROS-dependent signal transduction. *Curr. Opin. Cell Biol.* *33*, 8–13.
- Reddy, K., Schmidt, M.H.M., Geist, J.M., Thakkar, N.P., Panigrahi, G.B., Wang, Y.-H., and Pearson, C.E. (2014). Processing of double-R-loops in (CAG)_n·(CTG)_n and C9orf72 (GGGGCC)_n·(GGCCCC)_n repeats causes instability. *Nucleic Acids Res.* *42*, 10473–10487.

- Rehman, S., Gladman, J.T., Periasamy, A., Sun, Y., and Mahadevan, M.S. (2014). Development of an AP-FRET Based Analysis for Characterizing RNA-Protein Interactions in Myotonic Dystrophy (DM1). *PLoS ONE* 9.
- Rošić, S., Köhler, F., and Erhardt, S. (2014). Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J. Cell Biol.* 207, 335–349.
- Safe, S., Jin, U.-H., Morpurgo, B., Abudayyeh, A., Singh, M., and Tjalkens, R.B. Nuclear receptor 4A (NR4A) family – orphans no more. *J. Steroid Biochem. Mol. Biol.*
- Sahin, U., de Thé, H., and Lallemand-Breitenbach, V. (2014). PML nuclear bodies: Assembly and oxidative stress-sensitive sumoylation. *Nucl. 1949-1034* 5, 499–507.
- Saint-André, V., Batsché, E., Rachez, C., and Muchardt, C. (2011). Histone H3 lysine 9 trimethylation and HP1 γ favor inclusion of alternative exons. *Nat. Struct. Mol. Biol.* 18, 337–344.
- Salton, M., Voss, T.C., and Misteli, T. (2014). Identification by high-throughput imaging of the histone methyltransferase EHMT2 as an epigenetic regulator of VEGFA alternative splicing. *Nucleic Acids Res.* 42, 13662–13673.
- Salzman, J., Gawad, C., Wang, P.L., Lacayo, N., and Brown, P.O. (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PloS One* 7, e30733.
- 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.
- 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.
- Sarkar, P.S., Appukuttan, B., Han, J., Ito, Y., Ai, C., Tsai, W., Chai, Y., Stout, J.T., and Reddy, S. (2000). Heterozygous loss of Six5 in mice is sufficient to cause ocular cataracts. *Nat. Genet.* 25, 110–114.
- 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.
- Schafer, F.Q., and Buettner, G.R. (2001). Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radic. Biol. Med.* 30, 1191–1212.
- Schellenberg, M.J., Dul, E.L., and MacMillan, A.M. (2011). Structural model of the p14/SF3b155 · branch duplex complex. *RNA N. Y. N* 17, 155–165.
- Schultz, A., Nottrott, S., Hartmuth, K., and Lührmann, R. (2006). RNA structural requirements for the association of the spliceosomal hPrp31 protein with the U4 and U4atac small nuclear ribonucleoproteins. *J. Biol. Chem.* 281, 28278–28286.
- 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.
- Sergeant, N., Sablonnière, B., Schraen-Maschke, S., Ghestem, A., Maurage, C.A., Wattez, A., Vermersch, P., and Delacourte, A. (2001). Dysregulation of human brain microtubule-associated tau mRNA maturation in myotonic dystrophy type 1. *Hum. Mol. Genet.* 10, 2143–2155.

- Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* 76, 75–100.
- Shalgi, R., Hurt, J.A., Lindquist, S., and Burge, C.B. (2014). Widespread Inhibition of Posttranscriptional Splicing Shapes the Cellular Transcriptome following Heat Shock. *Cell Rep.* 7, 1362–1370.
- Sharma, S., Kohlstaedt, L.A., Damianov, A., Rio, D.C., and Black, D.L. (2008). Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. *Nat. Struct. Mol. Biol.* 15, 183–191.
- Shen, H., Zheng, X., Luecke, S., and Green, M.R. (2010). The U2AF35-related protein Urp contacts the 3' splice site to promote U12-type intron splicing and the second step of U2-type intron splicing. *Genes Dev.* 24, 2389–2394.
- Sherer, T.B., Betarbet, R., and Greenamyre, J.T. (2002). Environment, mitochondria, and Parkinson's disease. *Neurosci. Rev. J. Bringing Neurobiol. Neurol. Psychiatry* 8, 192–197.
- Shigenaga, M.K., Gimeno, C.J., and Ames, B.N. (1989). Urinary 8-hydroxy-2'-deoxyguanosine as a biological marker of in vivo oxidative DNA damage. *Proc. Natl. Acad. Sci. U. S. A.* 86, 9697–9701.
- Singh, R.K., and Cooper, T.A. (2012a). Pre-mRNA splicing in disease and therapeutics. *Trends Mol. Med.* 18, 472–482.
- Singh, R.K., and Cooper, T.A. (2012b). 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., Valcárcel, J., and Green, M.R. (1995). Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* 268, 1173–1176.
- 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.
- Skandalis, A., and Uribe, E. (2004). A survey of splice variants of the human hypoxanthine phosphoribosyl transferase and DNA polymerase beta genes: products of alternative or aberrant splicing? *Nucleic Acids Res.* 32, 6557–6564.
- Smith, K.J., and Lassmann, H. (2002). The role of nitric oxide in multiple sclerosis. *Lancet Neurol.* 1, 232–241.
- Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W.J. (2003). RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res.* 31, 5469–5482.
- Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* 3.
- Spickett, C.M., Pitt, A.R., Morrice, N., and Kolch, W. (2006). Proteomic analysis of phosphorylation, oxidation and nitrosylation in signal transduction. *Biochim. Biophys. Acta* 1764, 1823–1841.

- Spiller, M.P., Reijns, M.A.M., and Beggs, J.D. (2007a). Requirements for nuclear localization of the Lsm2-8p complex and competition between nuclear and cytoplasmic Lsm complexes. *J. Cell Sci.* *120*, 4310–4320.
- Spiller, M.P., Boon, K.-L., Reijns, M.A.M., and Beggs, J.D. (2007b). The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. *Nucleic Acids Res.* *35*, 923–929.
- 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.
- Stark, H., Dube, P., Lührmann, R., and Kastner, B. (2001). Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* *409*, 539–542.
- Starkov, A.A. (2008). The Role of Mitochondria in Reactive Oxygen Species Metabolism and Signaling. *Ann. N. Y. Acad. Sci.* *1147*, 37–52.
- Szweda, L.I., Uchida, K., Tsai, L., and Stadtman, E.R. (1993). Inactivation of glucose-6-phosphate dehydrogenase by 4-hydroxy-2-nonenal. Selective modification of an active-site lysine. *J. Biol. Chem.* *268*, 3342–3347.
- Tacke, R., Tohyama, M., Ogawa, S., and Manley, J.L. (1998). Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. *Cell* *93*, 139–148.
- Takahashi, A., Mori, E., Somakos, G.I., Ohnishi, K., and Ohnishi, T. (2008). Heat induces gammaH2AX foci formation in mammalian cells. *Mutat. Res.* *656*, 88–92.
- Takahashi, M., Higuchi, M., Matsuki, H., Yoshita, M., Ohsawa, T., Oie, M., and Fujii, M. (2013). Stress Granules Inhibit Apoptosis by Reducing Reactive Oxygen Species Production. *Mol. Cell. Biol.* *33*, 815–829.
- Takeo, K., Kawai, T., Nishida, K., Masuda, K., Teshima-Kondo, S., Tanahashi, T., and Rokutan, K. (2009). Oxidative stress-induced alternative splicing of transformer 2 β (SFRS10) and CD44 pre-mRNAs in gastric epithelial cells. *Am. J. Physiol. - Cell Physiol.* *297*, C330–C338.
- Tang, R., Qi, Q., Wu, R., Zhou, X., Wu, D., Zhou, H., Mao, Y., Li, R., Liu, C., Wang, L., et al. (2015). The polymorphic terminal-loop of pre-miR-1307 binding with MBNL1 contributes to colorectal carcinogenesis via interference with Dicer1 recruitment. *Carcinogenesis*.
- 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* *1792*, 14–26.
- Teplova, M., and Patel, D.J. (2008). Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat. Struct. Mol. Biol.* *15*, 1343–1351.
- Thimmulappa, R.K., Lee, H., Rangasamy, T., Reddy, S.P., Yamamoto, M., Kensler, T.W., and Biswal, S. (2006). Nrf2 is a critical regulator of the innate immune response and survival during experimental sepsis. *J. Clin. Invest.* *116*, 984–995.

- Thornton, C.A., Wymer, J.P., Simmons, Z., McClain, C., and Moxley, R.T. (1997). Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nat. Genet.* *16*, 407–409.
- Timchenko, L.T., Miller, J.W., Timchenko, N.A., DeVore, D.R., Datar, K.V., Lin, L., Roberts, R., Caskey, C.T., and Swanson, M.S. (1996). Identification of a (CUG)_n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res.* *24*, 4407–4414.
- Timchenko, N.A., Patel, R., Iakova, P., Cai, Z.-J., Quan, L., and Timchenko, L.T. (2004). Overexpression of CUG triplet repeat-binding protein, CUGBP1, in mice inhibits myogenesis. *J. Biol. Chem.* *279*, 13129–13139.
- Tong, K.I., Kobayashi, A., Katsuoka, F., and Yamamoto, M. (2006). Two-site substrate recognition model for the Keap1-Nrf2 system: a hinge and latch mechanism. *Biol. Chem.* *387*, 1311–1320.
- Trachootham, D., Lu, W., Ogasawara, M.A., Valle, N.R.-D., and Huang, P. (2008). Redox Regulation of Cell Survival. *Antioxid. Redox Signal.* *10*, 1343–1374.
- Tran, H., Gourrier, N., Lemercier-Neuillet, C., Dhaenens, C.-M., Vautrin, A., Fernandez-Gomez, F.J., Arandel, L., Carpentier, C., Obriot, H., Eddarkaoui, S., et al. (2011). Analysis of exonic regions involved in nuclear localization, splicing activity, and dimerization of Muscleblind-like-1 isoforms. *J. Biol. Chem.* *286*, 16435–16446.
- Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., et al. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* *39*, 925–938.
- Turunen, J.J., Will, C.L., Grote, M., Lührmann, R., and Frilander, M.J. (2008). The U11-48K protein contacts the 5' splice site of U12-type introns and the U11-59K protein. *Mol. Cell. Biol.* *28*, 3548–3560.
- Turunen, J.J., Niemelä, E.H., Verma, B., and Frilander, M.J. (2013). The significant other: splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* *4*, 61–76.
- Udd, B., and Krahe, R. (2012). The myotonic dystrophies: molecular, clinical, and therapeutic challenges. *Lancet Neurol.* *11*, 891–905.
- Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.-S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., et al. (2005). Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.* *37*, 844–852.
- Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). An RNA map predicting Nova-dependent splicing regulation. *Nature* *444*, 580–586.
- 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. (2005). 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. (2008). Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res.* *36*, 423–434.
- Valko, M., Rhodes, C.J., Moncol, J., Izakovic, M., and Mazur, M. (2006). Free radicals, metals and antioxidants in oxidative stress-induced cancer. *Chem. Biol. Interact.* *160*, 1–40.
- Valko, M., Leibfritz, D., Moncol, J., Cronin, M.T.D., Mazur, M., and Telser, J. (2007a). Free radicals and antioxidants in normal physiological functions and human disease. *Int. J. Biochem. Cell Biol.* *39*, 44–84.
- Valko, M., Leibfritz, D., Moncol, J., Cronin, M.T.D., Mazur, M., and Telser, J. (2007b). Free radicals and antioxidants in normal physiological functions and human disease. *Int. J. Biochem. Cell Biol.* *39*, 44–84.
- Valverde, R., Edwards, L., and Regan, L. (2008). Structure and function of KH domains. *FEBS J.* *275*, 2712–2726.
- Venables, J.P., Klinck, R., Koh, C., Gervais-Bird, J., Bramard, A., Inkel, L., Durand, M., Couture, S., Froehlich, U., Lapointe, E., et al. (2009). Cancer-associated regulation of alternative splicing. *Nat. Struct. Mol. Biol.* *16*, 670–676.
- Vidal, V.P., Verdone, L., Mayes, A.E., and Beggs, J.D. (1999). Characterization of U6 snRNA-protein interactions. *RNA N. Y. N* *5*, 1470–1481.
- Vinereanu, D., Bajaj, B.P.S., Fenton-May, J., Rogers, M.T., Mädler, C.F., and Fraser, A.G. (2004). Subclinical cardiac involvement in myotonic dystrophy manifesting as decreased myocardial Doppler velocities. *Neuromuscul. Disord. NMD* *14*, 188–194.
- Vitali, F., Henning, A., Oberstrass, F.C., Hargous, Y., Auweter, S.D., Erat, M., and Allain, F.H.-T. (2006). Structure of the two most C-terminal RNA recognition motifs of PTB using segmental isotope labeling. *EMBO J.* *25*, 150–162.
- Vlasova, I.A., Tahoe, N.M., Fan, D., Larsson, O., Rattenbacher, B., Sternjohn, J.R., Vasdewani, J., Karypis, G., Reilly, C.S., Bitterman, P.B., et al. (2008). Conserved GU-rich elements mediate mRNA decay by binding to CUG-binding protein 1. *Mol. Cell* *29*, 263–270.
- Vlasova-St Louis, I., and Bohjanen, P.R. (2011). Coordinate regulation of mRNA decay networks by GU-rich elements and CELF1. *Curr. Opin. Genet. Dev.* *21*, 444–451.
- Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* *136*, 701–718.
- Wan, L., Ottinger, E., Cho, S., and Dreyfuss, G. (2008). Inactivation of the SMN Complex by Oxidative Stress. *Mol. Cell* *31*, 244–254.
- Wang, Z., and Burge, C.B. (2008). Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA N. Y. N* *14*, 802–813.
- Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* *456*, 470–476.

- Wang, E.T., Cody, N.A.L., Jog, S., Biancolella, M., Wang, T.T., Treacy, D.J., Luo, S., Schroth, G.P., Housman, D.E., Reddy, S., et al. (2012). Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* *150*, 710–724.
- Wang, E.T., Ward, A.J., Cherone, J.M., Giudice, J., Wang, T.T., Treacy, D.J., Lambert, N.J., Freese, P., Saxena, T., Cooper, T.A., et al. (2015). Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res.* *25*, 858–871.
- Wang, Z., Xiao, X., Van Nostrand, E., and Burge, C.B. (2006). General and specific functions of exonic splicing silencers in splicing control. *Mol. Cell* *23*, 61–70.
- Ward, A.J., and Cooper, T.A. (2010). The pathobiology of splicing. *J. Pathol.* *220*, 152–163.
- Ward, A.J., Rimer, M., Killian, J.M., Dowling, J.J., and Cooper, T.A. (2010). CUGBP1 overexpression in mouse skeletal muscle reproduces features of myotonic dystrophy type 1. *Hum. Mol. Genet.* *19*, 3614–3622.
- Warf, M.B., and Berglund, J.A. (2007). MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA N. Y. N* *13*, 2238–2251.
- Warf, M.B., Diegel, J.V., von Hippel, P.H., and Berglund, J.A. (2009). The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 9203–9208.
- Warzecha, C.C., Sato, T.K., Nabet, B., Hogenesch, J.B., and Carstens, R.P. (2009). ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. *Mol. Cell* *33*, 591–601.
- Weis, S., Bielaw, T., Sommerer, I., Iovanna, J., Malicet, C., Mössner, J., and Hoffmeister, A. (2015). P8 deficiency increases cellular ROS and induces HO-1. *Arch. Biochem. Biophys.* *565*, 89–94.
- Weyn-Vanhentenryck, S.M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P.A., Zhang, M.Q., et al. (2014). HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep.* *6*, 1139–1152.
- Winczura, A., Zdzalik, D., and Tudek, B. (2012). Damage of DNA and proteins by major lipid peroxidation products in genome stability. *Free Radic. Res.* *46*, 442–459.
- Wong-ekkabut, J., Xu, Z., Triampo, W., Tang, I.-M., Peter Tieleman, D., and Monticelli, L. (2007). Effect of Lipid Peroxidation on the Properties of Lipid Bilayers: A Molecular Dynamics Study. *Biophys. J.* *93*, 4225–4236.
- 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.
- Wu, J.Y., and Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* *75*, 1061–1070.
- Xiang, Y., Qiu, Q., Jiang, M., Jin, R., Lehmann, B.D., Strand, D.W., Jovanovic, B., DeGraff, D.J., Zheng, Y., Yousif, D.A., et al. (2013). SPARCL1 suppresses metastasis in prostate cancer. *Mol. Oncol.* *7*, 1019–1030.
- 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.

- Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.-S., Zhang, C., Yeo, G., Black, D.L., Sun, H., et al. (2009). Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol. Cell* *36*, 996–1006.
- Yearim, A., Gelfman, S., Shayevitch, R., Melcer, S., Glaich, O., Mallm, J.-P., Nissim-Rafinia, M., Cohen, A.-H.S., Rippe, K., Meshorer, E., et al. (2015). HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Rep.* *10*, 1122–1134.
- Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.-D., and Gage, F.H. (2009). An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat. Struct. Mol. Biol.* *16*, 130–137.
- Yin, H., Xu, L., and Porter, N.A. (2011). Free radical lipid peroxidation: mechanisms and analysis. *Chem. Rev.* *111*, 5944–5972.
- Yoo, H.J., Im, C.-N., Youn, D.-Y., Yun, H.H., and Lee, J.-H. (2014). Bis is Induced by Oxidative Stress via Activation of HSF1. *Korean J. Physiol. Pharmacol. Off. J. Korean Physiol. Soc. Korean Soc. Pharmacol.* *18*, 403–409.
- Yuan, Y., Compton, S.A., Sobczak, K., Stenberg, M.G., Thornton, C.A., Griffith, J.D., and Swanson, M.S. (2007). Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res.* *35*, 5474–5486.
- 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.
- Zhang, Z., and Krainer, A.R. (2004). Involvement of SR proteins in mRNA surveillance. *Mol. Cell* *16*, 597–607.
- Zhang, C., Zhang, Z., Castle, J., Sun, S., Johnson, J., Krainer, A.R., and Zhang, M.Q. (2008a). Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. *Genes Dev.* *22*, 2550–2563.
- Zhang, G., Shi, H., Wang, L., Zhou, M., Wang, Z., Liu, X., Cheng, L., Li, W., and Li, X. (2015). MicroRNA and Transcription Factor Mediated Regulatory Network Analysis Reveals Critical Regulators and Regulatory Modules in Myocardial Infarction. *PLoS One* *10*, e0135339.
- Zhang, M., Mileykovskaya, E., and Dowhan, W. (2005). Cardiolipin Is Essential for Organization of Complexes III and IV into a Supercomplex in Intact Yeast Mitochondria. *J. Biol. Chem.* *280*, 29403–29408.
- Zhang, X., Zhou, J., Fernandes, A.F., Sparrow, J.R., Pereira, P., Taylor, A., and Shang, F. (2008b). THE PROTEASOME IS A TARGET OF OXIDATIVE DAMAGE IN HUMAN RETINA PIGMENT EPITHELIAL CELLS. *Invest. Ophthalmol. Vis. Sci.* *49*, 3622–3630.
- Zhao, Y., Nomiya, T., Findeisen, H.M., Qing, H., Aono, J., Jones, K.L., Heywood, E.B., and Bruemmer, D. (2014). Epigenetic regulation of the NR4A orphan nuclear receptor NOR1 by histone acetylation. *FEBS Lett.* *588*, 4825–4830.
- Zhong, H., and Yin, H. (2015). Role of lipid peroxidation derived 4-hydroxynonenal (4-HNE) in cancer: Focusing on mitochondria. *Redox Biol.* *4*, 193–199.

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., and Lou, H. (2008). Repression of Prespliceosome Complex Formation at Two Distinct Steps by Fox-1/Fox-2 Proteins. *Mol. Cell. Biol.* 28, 5507–5516.

Zhou, H.-L., Luo, G., Wise, J.A., and Lou, H. (2014). Regulation of alternative splicing by local histone modifications: potential roles for RNA-guided mechanisms. *Nucleic Acids Res.* 42, 701–713.

Zhou, X., Ferraris, J.D., and Burg, M.B. (2006a). Mitochondrial reactive oxygen species contribute to high NaCl-induced activation of the transcription factor TonEBP/OREBP. *Am. J. Physiol. Renal Physiol.* 290, F1169–F1176.

Zhou, X., Yang, W., and Li, J. (2006b). Ca²⁺- and Protein Kinase C-dependent Signaling Pathway for Nuclear Factor- κ B Activation, Inducible Nitric-oxide Synthase Expression, and Tumor Necrosis Factor- α Production in Lipopolysaccharide-stimulated Rat Peritoneal Macrophages. *J. Biol. Chem.* 281, 31337–31347.

Zhu, J., Koken, M.H., Quignon, F., Chelbi-Alix, M.K., Degos, L., Wang, Z.Y., Chen, Z., and de Thé, H. (1997). Arsenic-induced PML targeting onto nuclear bodies: implications for the treatment of acute promyelocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.* 94, 3978–3983.

Zhu, J., Mayeda, A., and Krainer, A.R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol. Cell* 8, 1351–1361.

Résumé

La dystrophie myotonique de type 1 (DM1) est une maladie génétique caractérisée par une dégénérescence des muscles squelettiques accompagnée d'une myotonie. Cette maladie est due à une expansion instable de triplets CTG dans la région 3' non traduite du gène *DMPK*. L'accumulation des ARNm *DMPK* mutés au sein de *foci* nucléaires conduit à la séquestration du facteur d'épissage MBNL1 et à des altérations de l'épissage alternatif de nombreux ARNm. En particulier, l'inclusion de l'exon 5 au sein du pré-ARNm de la troponine T cardiaque (hcTNT) est renforcée chez les patients DM1. Cette inclusion anormale participe aux anomalies cardiaques présentées par les patients. Les travaux de l'équipe, menés en collaboration avec l'équipe de Nicolas Sergeant à Lille, sur la régulation de l'épissage de l'ARNm hcTNT avaient établi l'existence de 8 sites MBNL1, dont 6 nouveaux, situés de part et d'autre de l'exon 5 et la présence de régions activatrices et inhibitrices de l'inclusion fixant des facteurs d'épissage dont l'identité n'était pas connue. L'un des objectifs de ma thèse était d'étudier l'importance fonctionnelle *in cellulo* de chacun des 8 sites MBNL1. J'ai ainsi pu montrer que chacun des 6 nouveaux sites participe à l'inhibition de l'inclusion de l'exon 5 par MBNL1. Les données obtenues nous ont amené à proposer un modèle dans lequel MBNL1 s'associe avec les triplets de sites MBNL1 situés de part et d'autre de l'exon 5 et entraîne la formation d'une structure à longue distance via des interactions protéiques MBNL1-MBNL1. Cette structure isolerait l'exon 5 dans une boucle et limiterait la fixation du spliceosome. Par ailleurs, j'ai mis en œuvre une approche de purification de RNP formées en extrait nucléaire pour identifier d'autres facteurs régulant l'inclusion de l'exon 5. La protéine hnRNP H a ainsi pu être identifiée. Sa capacité à activer l'inclusion de l'exon 5 *in cellulo* et à entrer en compétition avec MBNL1 pour la régulation de l'inclusion de l'exon 5 via sa fixation sur des sites localisés dans l'exon 5 et en aval de cet exon a pu être confirmée.

La seconde partie de ma thèse a porté sur l'étude de l'effet d'un stress oxydatif généré par 500 μM d' H_2O_2 sur le profil global d'épissage alternatif des pré-ARNm de cellules HeLa. Lors de ce travail, j'ai pu établir que la réponse des cellules HeLa au stress oxydatif implique deux phases de réponse : une phase précoce (1h-8h) caractérisée par un fort taux de mortalité associé à une forte augmentation du taux de d'entités oxygénées réactives (ROS) intracellulaire et une phase tardive (16h-24h) corrélée à une diminution du taux de ROS intracellulaires et une surexpression des ARN satellite III. Sur la base de ces données, une analyse globale du transcriptome par emploi de puces à exons (Affymetrix) a été réalisée à partir d'ARN totaux isolés 1h, 2h, 4h et 24h après le début du stress. Nous avons ainsi identifié des modulations d'expression et d'épissage spécifiques de chacune des deux phases. L'analyse des données par des outils bio-informatiques a permis de mettre en évidence des fonctions cellulaires bien définies qui sont plus particulièrement affectées lors d'un stress oxydant. Enfin, pour comprendre l'origine des variations d'épissage observées lors d'un stress oxydant, j'ai entrepris d'analyser les effets de ce stress sur le niveau d'expression et la localisation cellulaire des composants du spliceosome ou des facteurs qui s'associent pour réguler son activité.

Mots clés : épissage alternatif, dystrophie myotonique de type 1, stress oxydant, MBNL1, hnRNP H, micropuces

Abstract

Myotonic dystrophy of type 1 (DM1) is a genetic disease characterized by skeletal muscle degeneration associated to myotonia. DM1 results from the instable expansion of CTG repeats within the 3' untranslated region of the *DMPK* gene. The accumulation of mutated *DMPK* mRNAs within nuclear foci leads to the sequestration of the MBNL1 splicing factor and causes splicing misregulation of numerous pre-mRNAs. Among altered events the increase of the inclusion of exon 5 in the human cardiac troponin T (hcTNT) mRNA is of particular importance, since it contributes to the cardiac symptoms presented by the patients. Through collaborative work with N. Sergeant's team from Lille, the team has studied the molecular bases of hcTNT exon 5 inclusion regulation and mapped 8 MBNL1 binding sites, including 6 new ones, within intronic regions surrounding exon 5. They also identified positive and negative splicing regulatory elements of which protein partners remain unidentified. The first objective of my PhD thesis was to test the functional importance of each individual MBNL1 binding site. The obtained results established that the 6 newly identified MBNL1 binding sites are involved in splicing regulation by MBNL1 and lead us to propose a new regulation model in which MBNL1 binds on triplets of MBNL1 sites present on each side of exon 5 and form a long distance structure via MBNL1-MBNL1 protein interaction. The formation of this looping-structure is expected to isolate exon 5 and limit its recognition by the spliceosome. In addition I searched for protein partners of the identified regulatory elements by affinity chromatography. By this way, I identified hnRNP H as a positive regulator of exon 5 inclusion. Its capacity to compete with MBNL1 to regulate splicing *in cellulo* by binding on exonic and intronic binding sites was further confirmed.

The second part of my PhD work corresponds to the study of the global impact of oxidative stress, generated by exposition of HeLa cells to 500 μM of H_2O_2 , on alternative splicing. This allows us to establish that the response of HeLa cells to oxidative stress involve two distinct phases: an early one (1h-8h) characterized by poor survival rate and high intracellular ROS content and a late phase (16-24h), associated with a decrease of the intracellular ROS level and the overexpression of the long non coding sat III RNAs. Based on this observation, we performed a transcriptome global analysis by using exon arrays from Affymetrix on RNA samples isolated 1, 2, 4 or 24 hours after the induction of the oxidative stress. We identified changes of the gene expression level or mRNA splicing pattern specific of each of the response phases. Data computing by bio-informatic tools identified the most affected cellular processes and functions during the cell response to oxidative stress. In order to better understand the mechanisms underlying alternative splicing modulation during oxidative stress, I started to study the impact of oxidative stress on the expression level and the cellular localization of spliceosome components and most common splicing regulation factors.

Keywords : alternative splicing, myotonic dystrophy of type 1, oxidative stress, MBNL1, hnRNP H, microarray