

UNIVERSITE DU DROIT ET DE LA SANTE – LILLE II

THESE DE DOCTORAT D'UNIVERSITE

Pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DU DROIT ET DE LA SANTE – LILLE II

Discipline : Sciences de la Vie et de la Santé

Présentée par

Lucie COPPIN

**Mécanismes de régulation post-transcriptionnelle de
l'expression des mucines par la galectine-3**

Thèse dirigée par Monsieur le Professeur Pascal PIGNY

Soutenue le 12 juin 2017

Membres du jury :

Président	Madame le Professeur Nicole PORCHET
Rapporteurs	Madame le Docteur Stéfania MILLEVOI Monsieur le Professeur Pascal REBOUL
Examineurs	Madame le Docteur Delphine DELACOUR Monsieur le Professeur Hervé PUY
Directeur de Thèse	Monsieur le Professeur Pascal PIGNY

Ce travail de thèse a été effectué au sein de
l'équipe « Mucines, différenciation et cancérogenèse épithéliales »
de l'UMR-S 1172 Inserm - Université Lille - CHRU de Lille
Directrice d'équipe : Madame le **Docteur Isabelle Van Seuning**

sous la direction scientifique de
Monsieur le **Professeur Pascal Pigny**

Je remercie Madame le Professeur **Nicole Porchet** pour l'honneur qu'elle me fait de présider ce jury ;

Je tiens à remercier Madame le Docteur **Stéfania Millevoi** et Monsieur le Professeur **Pascal Reboul** d'avoir accepté d'être rapporteurs de cette thèse ;

J'exprime également toute ma reconnaissance à Madame le Docteur **Delphine Delacour**, à Monsieur le Professeur **Hervé Puy** et à Monsieur le Professeur **Pascal Pigny** pour avoir accepté de juger ce travail.

.

Je tiens à remercier :

Monsieur le Docteur **Luc Buée**, directeur de l'UMR-S 1172 et Monsieur le Professeur **Pierre Formstecher**, ancien directeur du Centre (ex UMR n°837) pour m'avoir accueillie au sein du Centre de Recherche Jean-Pierre Aubert où j'ai effectué mon Master 2 Recherche puis ma thèse d'Université ;

Madame le Docteur **Isabelle Van Seuning** qui m'a accueillie au sein de son équipe ;

Monsieur le Professeur **Pascal Pigny** pour m'avoir encadrée pendant cette thèse ;

Madame le Professeur **Nicole Porchet** pour son soutien, son aide et ses précieux conseils ;

Madame le Professeur **Marie-Pierre Buisine** pour son accueil dans le secteur d'oncologie au centre de Biologie-Pathologie-Génétique du CHRU de Lille ;

Mesdames les Docteurs **Audrey Vincent, Florence Renaud et Julie Leclerc** pour leur aide et leurs conseils scientifiques ;

Messieurs les Docteurs **Nicolas Jonckheere et Mickael Perrais** ainsi que Mesdames les Docteurs **Bernadette Neve et Guillemette Huet** pour les remarques pertinentes sur le sujet ;

Madame le Professeur **Emmanuelle Leteurtre** pour son expertise d'anatomopathologiste sur le sujet ;

Madame **Céline Schulz** pour ses conseils, sa présence et son amitié ;

Madame **Marie-Paule Ducourouble** pour ses conseils techniques et son soutien tout au long de ses années ;

Mesdames **Belinda Dûchene et Brigitte Hémon** pour leur aide technique ;

L'ensemble des collaborateurs extérieurs, les bioinformaticiens et ingénieurs du centre de Biologie Pathologie qui ont participé à ce travail ;

Les membres de l'équipe d'oncologie moléculaire du centre de Biologie-Pathologie-Génétique ainsi que Madame le Docteur **Marie-Françoise Odou** pour leur accueil au centre de Biologie Pathologie, leur gentillesse et leur aide ;

L'ensemble des membres actuels et anciens membres de l'équipe IVS que j'ai rencontré au cours de ces 7 années : **Nicolas, Johann, Emilie, Audrey, Romain, Matthieu, François,**

Fatima, Mehdi, Marjolaine, Nihad, Solange, Kelly, Mouloud et tous les autres pour les moments agréables partagés avec eux ;

Ma famille et plus spécialement mes parents pour leur soutien et confiance tout au long de ces années ;

Johan pour son soutien sans faille surtout ces derniers mois et ma petite Margaux pour ses jolis sourires.

Je remercie :

Le Centre Hospitalier Régional Universitaire de Lille,

La Région Hauts de France,

Le SIRIC ONCOLille,

La Ligue Nationale Contre le Cancer,

L'Université de Lille 2,

Pour leur soutien financier qui a permis la réalisation de ces travaux.

Les travaux exposés dans ce mémoire et auxquels j'ai participé ont fait l'objet des publications et communications suivantes :

Articles scientifiques publiés

“Galectin-3 is a non-classic RNA binding protein that stabilizes the mucin MUC4 mRNA in the cytoplasm of cancer cells”

Coppin L, Vincent A, Frénois F, Duchêne B, Lahdaoui F, Stechly L, Renaud F, Villenet C, van Seuning I, Leteurtre E, Dion J, Grandjean C, Poirier F, Figeac M, Delacour D, Porchet N, Pigny P.

Sci Rep. 2017;7:43927. doi: 10.1038/srep43927

“CA-125, but not galectin-3, complements CA 19-9 for discriminating ductal adenocarcinoma versus non-malignant pancreatic diseases”

Coppin L, Benomar K, Corfiotti F, Cattan S, Renaud F, Lapere C, Leteurtre E, Vantyghem MC, Truant S, Pigny P.

Pancreatology. 2016; 16:115-20

Articles scientifiques issus de collaborations pendant la thèse

“The serrated neoplasia pathway of colorectal tumors: Identification of MUC5AC hypomethylation as an early marker of polyps with malignant potential”

Renaud F, Mariette C, Vincent A, Wacrenier A, Maunoury V, Leclerc J, Coppin L, Crépin M, Van Seuning I, Leteurtre E, Buisine MP.

Int J Cancer. 2016; 138:1472-81

“MUC1 expression in papillary thyroid carcinoma is associated with BRAF mutation and lymph node metastasis; the latter is the most important risk factor of relapse”

Renaud F, Gnemmi V, Devos P, Aubert S, Crépin M, Coppin L, Ramdane N, Bouchindhomme B, Boidein M, Van Seuning I, Do Cao C, Pattou F, Carnaille B, Pigny P, Wemeau JL, Leteurtre E.

Thyroid. 2014; 24: 1375-84

Communications orales

“Galectin-3 interacts with large exon mucin transcripts and prevents MUC4 mRNA decay in pancreatic cancer cells”

Lucie Coppin, Frédéric Frénois, Fatima Lahdaoui, Laurence Stechly, Florence Renaud, Evelyne Crème, Claudine Grutzmacher, Emmanuelle Leteurtre, Isabelle Van Seuningen, Martin Figeac, Nicole Porchet & Pascal Pigny

Journée des doctorants, Inserm, Lille, janvier 2015

“Post-transcriptional regulation of mucins expression by galectin-3 in cancer”

Lucie Coppin, Frédéric Frénois, Fatima Lahdaoui, Laurence Stechly, Florence Renaud, Evelyne Crème, Claudine Grutzmacher, Emmanuelle Leteurtre, Isabelle Van seuningen, Martin Figeac, Nicole Porchet & Pascal Pigny

Journée lilloise des doctorants en cancérologie, Lille, Juin 2014

Communications par affiche

“Identification of a new function for galectin-3 : control of mucin mRNAs metabolism in cancer »

Lucie Coppin, Audrey Vincent, Frédéric Frénois, Belinda Duchêne, Fatima Lahdaoui, Laurence Stechly, Florence Renaud, Céline Villenet, Isabelle van Seuningen, Emmanuelle Leteurtre, Johann Dion, Cyrille Grandjean, Françoise Poirier, Martin Figeac, Delphine Delacour, Nicole Porchet & Pascal Pigny,

Biochemical society Londres, 4-7 septembre 2016

“Galectin-3 interacts with large exons mucin transcripts and prevents MUC4 mRNA decay”

Lucie Coppin, Frédéric Frénois, Fatima Lahdaoui, Laurence Stechly, Florence Renaud, Evelyne Crème, Claudine Grutzmacher, Emmanuelle Leteurtre, Isabelle Van seuningen, Martin Figeac, Nicole Porchet & Pascal Pigny

Journée André Verbert, 11 septembre 2014

“Galectin-3 regulates MUC1 and MUC4 expression : evidence for a new role of galectin-3 in the control of mRNA stability »

Lucie Coppin, Frédéric Frénois, Laurence Stechly, Fatima Lahdaoui, Martin Figeac, Evelyne Crème, Isabelle Van Seuning, Nicole Porchet & Pascal Pigny,

Journée Jeunes chercheurs Arc, 23-24 Octobre 2013

“Galectin-3 regulates MUC1 and MUC4 expression: Evidence for a new role of galectin-3 in the control of mRNA stability”

Lucie Coppin, Frédéric Frénois, Laurence Stechly, Fatima Lahdaoui, Martin Figeac, Evelyne Crème, Isabelle Van Seuning, Nicole Porchet & Pascal Pigny

EMBO symposium, Strasbourg, avril 2013

“Galectin-3 regulates MUC1 and MUC4 expression: Evidence for a new role of galectin-3 in the control of mRNA stability”

Lucie Coppin, Frédéric Frénois, Laurence Stechly, Fatima Lahdaoui, Martin Figeac, Evelyne Crème, Isabelle Van Seuning, Nicole Porchet & Pascal Pigny

Journée des doctorants, Mars 2013

“Galectin-3 regulates MUC1 and MUC4 expression: Evidence for a new role of galectin-3 in the control of mRNA stability”

Lucie Coppin, Frédéric Frénois, Laurence Stechly, Fatima Lahdaoui, Evelyne Crème, Isabelle Van Seuning, Nicole Porchet & Pascal Pigny

EMBO meeting, Nice, septembre 2012

“Rôle de la galectine-3 dans la régulation de l'expression des mucines membranaires MUC1 et MUC4”

Lucie Coppin, Evelyne Crème, Isabelle Van Seuning, Nicole Porchet & Pascal Pigny

Cancéropôle Nord-Ouest, Deauville, mai 2012

Résumé/Summary

L'adénocarcinome pancréatique canalaire s'accompagne d'une néoexpression de la mucine membranaire MUC4 et d'une surexpression des mucines membranaires MUC1 et MUC16. Ces O-glycoprotéines de haut poids moléculaire sont codées par des ARNm possédant des particularités inhabituelles par rapport aux autres transcrits humains, comme une longue demi-vie et une très grande taille. La galectine-3, une lectine endogène également surexprimée au cours du cancer pancréatique, exerce de très nombreuses fonctions biologiques, en particulier dans le domaine du trafic intracellulaire des glycoprotéines et de l'épissage des pré-ARNm. Cependant, l'implication de cette galectine à un autre niveau du cycle de vie des ARNm n'avait pas été explorée jusque-là dans la littérature. De précédents travaux du laboratoire ont démontré que la suppression de l'expression de la galectine-3 dans la lignée cellulaire cancéreuse pancréatique humaine CAPAN-1 s'accompagne d'une diminution de l'expression des transcrits de certaines mucines membranaires. L'objectif de ce travail a donc été d'étudier les mécanismes de régulation de l'expression des mucines membranaires, et plus particulièrement *MUC4*, par la galectine-3.

Nous avons démontré que la galectine-3, *in vitro*, régule l'expression de *MUC4* au niveau post-transcriptionnel en stabilisant les transcrits de cette mucine. Ceci passe par la potentialisation de la fixation de la *RNA Binding Protein* hnRNP-L sur l'élément *cis*-régulateur *CA repeat* présent dans le 3'UTR de *MUC4*. Nos résultats indiquent que cette régulation est présente *in vivo* au niveau physiologique dans des tissus épithéliaux digestifs murins. Par ailleurs, nous avons mis en évidence que la galectine-3 interagit avec hnRNP-L dans le cytoplasme mais qu'elle interagit faiblement avec des marqueurs de P-Bodies ou de granules de stress. Concernant le rôle de la galectine-3 dans le cycle de vie des ARNm, nos données révèlent que celle-ci se lie aux transcrits matures de *MUC4* au niveau périnucléaire, probablement dans des granules de stockage qui ne sont ni des granules de stress ni des *P-bodies* et dont le type reste à déterminer. Nous avons également élargi nos résultats en étudiant l'implication de cette lectine dans le métabolisme d'autres ARNm et nos analyses indiquent que la galectine-3 serait impliquée dans la régulation post-transcriptionnelle positive ou négative d'un ensemble de transcrits dont les fonctions convergent vers les voies UPR (*Unfolded protein response*) et ERAD (*Endoplasmic-reticulum-associated protein degradation*) mais également plus généralement vers le *processing* des protéines en réponse au stress du réticulum endoplasmique.

En conclusion, nos travaux mettent en évidence un nouveau rôle de la galectine-3 en tant que *RNA binding protein* dans la stabilisation des ARNm de *MUC4* mais aussi un

nouveau rôle dans la coordination de l'expression de répertoires de transcrits matures ayant des rôles biologiques communs (*RNA regulon*) permettant à la cellule de s'adapter au plan morphologique, métabolique et biologique à des changements physiopathologiques. Ceci renforce les interconnexions largement décrites dans la littérature entre mucines, galectine-3 et les grandes fonctions cellulaires qui sont perturbées en situation cancéreuse.

Pancreatic ductal adenocarcinoma is characterized by a neoexpression of the membrane-bound mucin MUC4 and an overexpression of membrane-bound mucins MUC1 and MUC16. These high molecular weight O-glycoproteins are encoded by mRNA sharing unusual features among human transcripts, such as a long half-life and a very large size. Galectin-3, an endogenous lectin frequently over-expressed in pancreatic cancer, has many biological functions, especially in intracellular glycoprotein trafficking and pre-mRNA splicing. However, the involvement of this lectin in another step of mRNA life cycle has not been explored in literature yet. Previous works performed in the laboratory have demonstrated that *LGALS3* gene knock-down in a human cancerous pancreatic cancer cell line is followed by a decrease of the expression of several membrane-bound mucin mRNAs. The aim of this present work was to study the mechanism of the regulation of mucins expression, especially for *MUC4*, by galectin-3.

We have demonstrated that galectin-3, *in vitro*, regulates *MUC4* expression at the post-transcriptional level through the stabilization of the transcripts of this mucin. Galectin-3 potentiates the binding of hnRNP-L, a RNA-Binding protein, on the CA repeat region present in *MUC4* 3'UTR. Our results show that this regulation occurs physiologically *in vivo* in mice digestive epithelial tissues. Moreover, we have demonstrated that galectin-3 interacts with hnRNP-L in cell cytoplasm but scarcely with protein markers of P-Bodies or stress granules markers. Regarding the influence of galectin-3 in mRNA life cycle, our results suggest that it binds to mature *MUC4* transcripts in the perinuclear area, probably in storage granules whose type should to be determined. We have also broadened our results by studying this lectin's involvement in the metabolism of other mRNA. Our analyzes suggest that galectin-3 could be involved in the positive or negative post-transcriptional regulation of a mRNA subset whose functions are linked to unfolded protein response (UPR) and Endoplasmic-reticulum-associated protein degradation (ERAD) pathways, but also more generally towards protein processing in response to endoplasmic reticulum stress.

In conclusion, our work highlights a new function for galectin-3 as a RNA binding protein in the stabilization of *MUC4* mRNA, but also a new function in the coordination of the expression of repertoires of mature transcripts with shared functions or (RNA regulon) allowing morphological, biological and metabolic cell adaptation to physiopathological changes. These results strengthen the interplay between mucins, galectin-3 and cellular functions which are disturbed in cancer.

Sommaire

Liste des abréviations.....	1
Préambule	2
Introduction	5
I. Adénocarcinome pancréatique canalaire.....	6
I.1. Carcinogénèse pancréatique	7
I.1.1. Lésions précurseurs de l'adénocarcinome pancréatique canalaire	7
I.1.2. Altérations géniques.....	8
I.2. Dépistage	9
II. Les mucines	11
II.1. Généralités.....	11
II.2. Structure et classification	14
II.2.1. Apomucine et chaînes glycaniques	14
a. Apomucine.....	14
b. Chaînes glycaniques	15
II.2.2. Classification des mucines et relation structure-fonction	17
a. Mucines sécrétées	18
b. Mucines membranaires	19
II.3. Territoires d'expression physiologique	21
II.4. Mucines et cancérogénèse	23
II.4.1. Modifications qualitatives.....	24
a. Altération de la glycosylation.....	24
b. Délocalisation des mucines	25
c. Variants d'épissage.....	26
II.4.2. Modifications quantitatives.....	27
II.4.3. Rôles des mucines dans la cancérogenèse	29
a. Propriétés d'adhérence et d'anti-adhérence des mucines	30
b. Modulation de la réponse immunitaire	32
c. Signalisation cellulaire	32
d. Chimiorésistance	33
III. Les mucines MUC4, MUC1 et MUC16.....	35
III.1. MUC4.....	35

III.1.1.	Structure	35
III.1.2.	Gène et ARNm.....	36
III.1.3.	Caractérisation de la région 5'-flanquante de <i>MUC4</i> et régulation de son expression	38
III.2.	MUC1.....	40
III.2.1.	Structure	40
III.2.2.	Gène, ARNm et régulation de l'expression de MUC1	40
III.3.	MUC16.....	43
III.3.1.	Structure protéique	43
III.3.2.	Régulation de l'expression de MUC16.....	44
III.3.3.	Expression de MUC16 en situations physiologiques et pathologiques	45
IV.	Galectines	46
IV.1.	Généralités.....	46
IV.1.1.	Classification et structure.....	46
IV.1.2.	Expression tissulaire et cellulaire des galectines	48
IV.1.3.	Fonctions des galectines	50
IV.1.4.	Galectines et cancer	52
IV.2.	Galectine-3	52
IV.2.1.	Gène	52
IV.2.2.	Protéine	54
IV.2.3.	Sites tissulaires et cellulaires d'expression	56
IV.2.4.	Ligands.....	56
IV.2.5.	Fonctions.....	58
IV.2.6.	Galectine-3 et cancer.....	62
IV.2.7.	Les inhibiteurs de galectine-3	65
IV.2.8.	Galectine-3 et autres pathologies	66
IV.2.9.	Modèles animaux	69
a.	Souris KO et anomalies du développement.....	69
b.	Souris KO et anomalies des cellules épithéliales	70
c.	Souris KO et cancer	70
d.	Souris KO et fibrose	71
e.	Souris KO et métabolisme	71
f.	Souris KO et inflammation.....	71
V.	Le métabolisme et le devenir des ARN pré-messagers.....	73

V.1.	Les acteurs du métabolisme des ARN : les RNA-binding proteins (RBP).....	75
V.1.1.	Les RBP	75
V.1.2.	Les mRBP	78
V.2.	Le cycle de vie de l'ARNm.....	82
V.2.1.	Maturation.....	82
a.	Epissage	82
b.	Ajout de la coiffe et de la queue poly(A)	87
V.2.2.	Export nucléaire	88
V.2.3.	Mécanisme de dégradation intra-cytoplasmique des ARNm.....	90
V.2.4.	Facteurs influençant la stabilité des ARNm.....	93
a.	Queue poly(A) et PABP	94
b.	Micro-ARN.....	97
c.	ARE	99
V.2.5.	Stockages dans les granules	102
a.	P-bodies	104
b.	Granules de stress	105
c.	Autres granules	108
Problématique et objectifs du travail		109
Résultats 1^{ère} partie : Galectine-3 et ARNm de mucines		112
I.	Etude de l'expression des gènes et protéines d'intérêt dans les lignées cellulaires Sh1 et Sc.....	113
II.	Etude du rôle de la galectine-3 dans la régulation de l'expression des mucines membranaires MUC1 et MUC4	115
II.1.	Etude de l'activité du promoteur de MUC1 et MUC4 sous l'influence de la galectine-3	115
II.2.	Etude de l'influence de la galectine-3 sur la stabilité des transcrits de MUC1 et de MUC4.....	117
II.2.1.	Etude de la demi-vie des transcrits de MUC1 et de MUC4.....	117
II.2.2.	Confirmation de la régulation de l'expression de MUC4 par la galectine-3 au niveau post-traductionnel	119
II.3.	Relevance des résultats dans des tissus d'adénocarcinomes pancréatiques canaux humains	121

III. Mécanisme de stabilisation des transcrits de MUC4 par la galectine-3.....	123
III.1. Rôle du 3'UTR de MUC4 dans l'effet stabilisateur de la galectine-3	123
III.2. Relevance des résultats in vivo étudiée sur des tissus murins.....	136
IV. Etude de la liaison de la galectine-3 aux ARNm matures dans le cytoplasme.	137
Résultats 2^{ème} partie : Galectine-3 et autres ARNm.....	148
I. Y a-t-il d'autres ARNm régulés de manière post-transcriptionnelle par la galectine-3 ?	149
II. Y-a-t-il, parmi les ARNm régulés de manière post-transcriptionnelle par la galectine-3 certains ARNm-cibles qui soient aussi liés par la galectine-3 ?.....	151
III. Analyses des cibles stabilisées et déstabilisées par la galectine-3	153
IV. Analyse individuelles des 15 transcrits à la fois stabilisés et liés par la galectine-3..	158
V. La voie UPR et les voies qu'elle active secondairement peuvent-elles rendre compte de toutes les catégories fonctionnelles préalablement caractérisées ?	165
VI. Etude détaillée des transcrits-cibles rattachables directement ou indirectement à la voie UPR.....	170
VI.1. Transcrits directement rattachés à la voie UPR	170
VI.1.1. Voie UPR et transcrits stabilisés par la galectine-3	171
a. Analyse orientée à l'aide de l'outil GSEA	171
b. Analyse orientée par les bases de données	172
c. Analyse orientée par la bibliographique	173
VI.1.2. Voie UPR et transcrits déstabilisés par la galectine-3	174
VI.1. Transcrits ayant des fonctions dans le retour à l'homéostasie du réticulum endoplasmique mais non répertoriés comme appartenant à la voie UPR	175
VI.1.1. Cibles à fonctions de chaperonnes stabilisées par la galectine-3.....	175
VI.1.2. Cibles stabilisées par la galectine-3 et impliquées dans les voies secondairement activées à l'issue de la réponse UPR.....	176
VI.1.3. Cibles déstabilisées par la galectine-3	177
VII. Qu'en est-il des transcrits liés à la fois par la galectine-3 et HuR ?.....	180
Discussion	185
I. Etude de l'influence de la galectine-3 sur l'expression des mucines membranaires .	186
II. Etude des mécanismes de stabilisation des transcrits de <i>MUC4</i> par la galectine-3 ...	188
III. Liaison de la galectine-3 à des ARNm matures de mucines dans le cytoplasme.....	192

IV. Liens entre mucines et galectine-3 dans le cancer du pancréas	193
V. Rôle de la galectine-3 dans le métabolisme des ARNm	195
Conclusions et perspectives	203
Appendice technique	208
I. Modèles d'étude	209
I.1. Lignées cellulaires humaines	209
I.1.1. Lignée cancéreuse humaine pancréatique CAPAN-1	209
I.1.2. Lignées cellulaires Sh1 et Sc	209
I.2. Tissus humains	209
I.3. Tissus murins.....	210
II. Biologie cellulaire	210
II.1. Culture cellulaire	210
II.1.1. Milieux de culture	210
II.1.2. Conditions de culture	210
II.2. Restauration de phénotype : traitement des cellules par la galectine-3 recombinante	211
II.3. Inhibiteurs de la galectine-3	212
II.4. Etude des cellules en phase G2/M et cytométrie en flux	213
II.5. Transfections transitoires	213
II.5.1. Transfections transitoires à l'Effectene®	213
II.5.2. Transfections transitoires par des siARN	214
II.5.3. Transfections transitoires par des oligonucléotides antisens	215
II.6. Imagerie.....	216
II.6.1. Hybridation in situ couplée à de l'immunofluorescence (FISH)	216
II.6.2. Immunofluorescence ou Co-immunofluorescence	217
II.6.3. Proximity Ligation Assay	218
II.7. Immunohistochimie.....	219
III. Biologie moléculaire	221
III.1. Extraction des ARN totaux	221
III.1.1. Extraction des ARN totaux à partir de culture cellulaire	221
III.1.2. Extraction des ARN totaux à partir de tissus	222
III.2. Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)	222

III.2.1.	Reverse Transcription : synthèse de l'ADNc.....	222
III.2.2.	Amplification de l'ADNc par PCR.....	222
III.3.	Extraction des ARN et Reverse transcription spécifique lors de la réversion de phénotype	223
III.3.1.	Récupération des cellules.....	224
III.3.2.	Extraction des ARN	224
III.3.3.	Synthèse de l'ADNc	224
III.4.	PCR quantitative en temps réel (qPCR).....	225
III.4.1.	Quantification avec la technique TaqMan®	225
a.	Quantification absolue	225
b.	Quantification relative	227
III.4.2.	Quantification avec la technique EvaGreen®	229
III.5.	Etude de la demi-vie des transcrits.....	230
III.6.	Analyse par séquençage à haut débit.....	230
III.6.1.	Analyse de la demi-vie des transcrits.....	230
III.6.2.	Etude des ARNm immunoprécipités lors de la RNA-IP.....	232
III.7.	Sous-clonage de produits PCR en vecteur plasmidique pCR® 2.1	233
III.7.1.	Ligation	233
III.7.2.	Transformation bactérienne par choc thermique	233
III.7.3.	Sélection des clones	234
III.8.	Mutagenèse dirigée	234
III.9.	Clonage du motif exon2-exon3 de MUC4	236
III.10.	Préparation d'ADN plasmidique en petit volume (Mini-Prep).....	239
III.11.	Préparation d'ADN plasmidique en grande quantité (Maxi-Prep).....	240
III.11.1.	Transformation des bactéries E coli JM109.....	240
III.11.2.	Extraction plasmidique	240
IV.	Biochimie	241
IV.1.	Mesure de l'activité luciférase	241
IV.2.	Dosage de l'activité β -galactosidase	242
IV.3.	Préparation d'extraits cellulaires totaux	243
IV.4.	Dosage protéique.....	244
IV.5.	Western-blot	244
IV.6.	Western-blot MUC4.....	246
IV.7.	Co-immunoprécipitation	246

IV.8.	Immunoprécipitation des complexes ARNm-protéines (RNA- IP)	248
IV.9.	RNA-pull down	250
IV.9.1.	Synthèse de la sonde ARN et marquage avec la biotine	250
IV.9.2.	Interaction ARN-protéines	252
IV.9.3.	Purification des complexes ARN-protéines	252
V.	Statistiques	252
Bibliographie.....		253

Liste des abréviations

aa	Acide aminé	DDX6	DEAD (Asp-Glu-Ala-Asp)
Act. D	Actinomycine D		box helicase 6
ADN	Acide DésoxyriboNucléique	DNase	DésoxyriboNucléase
ADNc	Acide DésoxyriboNucléique complémentaire	dNTP	désoxyNucléotide TriPhosphate
AGO	Argonaute	DO	Densité Optique
Akt	v-akt murine thymoma viral oncogene	D-PBS	Dubelcco-Phosphate-Buffered Saline
AMOP	Adhesion-associated domain present in MUC4 and other proteins	DPC4	Deleted in pancreatic cancer locus 4
AP-1	Activator Protein-1	DTT	Dithiothréitol
ARE	AU-rich Element	EDTA	Acide Ethylène Diamine Tétracétique
ARE-BP	ARE Binding Protein	EGTA	Acide Ethylène Glycol
ARN	Acide RiboNucléique		Tétraacétique
ARNm	Acide RiboNucléique messenger	EGF	Epidermal Growth Factor
ARNt	ARN de transfert	EGFR	Epidermal Growth Factor Receptor
ATF 4/6	Activating transcription factor 4/6	eIF	Eukaryotic Initiation Factor
Bcl-2	B-Cell lymphoma 2	EJC	Exon Junction Complex
BSA	SérumAlbumine Bovine	ERAD	Endoplasmic-Reticulum- Associated protein
CA 19-9/125	Carbohydrate Antigen 19- 9/125		Degradation
CARE	CA Repeat Element	ErbB2	Erythroblastic leukemia viral oncogene homolog 2
CBC	Cap Binding Complex	ERF	Eukaryotic Release Factor
CDK	Cyclin-Dependent Kinase	Erk	Extracellular signal-regulated kinase
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator	FISH	Fluorescence In Situ Hybridization
CO2	Dioxyde de carbone	G3BP	RasGAP SH3 domain Binding Protein
CRD	Carbohydrate Recognition Domain	Gal	Galactose
Ct	Cycle seuil	GalNAc	N-acétylgalactosamine
DAPI	4',6-diamidino-2-phénylindole	GAPDH	Glycéraldéhyde 3-phosphate déshydrogénase
DBP5	DEAD-box protein		
DCP 1-2	Decapping Protein 1-2	GDP	Guanosine diphosphate

GlcNAc	N-acétylglucosamine	NK	Natural Killer
GM-CSF	Granulocyte-macrophage colony-stimulating factor	NMD	Non sens Mediated Decay
GTP	Guanosine triphosphate	NP-40	Nonidet P40
GTPase	Hydrolase de GTP	NS	Non Significatif
GW182	Glycine-tryptophan protein of 182 kDa	nt	nucléotide
HIF	Hypoxia Inducible Factor	NXF1	nuclear export factor 1
hnRNP	Heterogeneous Nuclear Ribonucleoprotein	NXT1	NTF-2-related protein 1
HSP	Heat Shock Protein	p	p-value
HuR	Human antigen R	PABP	Poly(A) binding protein
ICAM	Intercellular Adhesion Molecule	PanIN	Pancreatic Intraepithelial Neoplasm
IF	Immunofluorescence	pb	paire de base
IFN	Interféron	PB	Processing Bodies
IL	Interleukine	PBS	Phosphate Buffered Saline
INF	interferon	PCR	Polymerase Chain Reaction
IP	Immunoprecipitation	PDAC	Pancreatic Ductal Adeno Carcinoma
IRE1	inositol-requiring enzyme 1	PERK	PRKR-Like Endoplasmic Reticulum Kinase
JNK	Jun N-terminal Kinase	PFA	Paraformaldehyde
kb	kilobase	pH	Potentiel Hydrogène
kDa	kiloDalton	PI3K	Phosphatidylinositol 3-Kinase
KO	Knock-Out	PKA/B/C	Protéine Kinase A/B/C
K-Ras	Kirsten rat 2 viral oncogene	PLA	Proximity Ligation Assay
LB	Lysogeny Broth	PMSF	Phenyl Methyl Sulfonyl Fluoride
m/v	masse/volume	PPAR	Pperoxisome Proliferator-Activated Receptor
MAPK	Mitogen-Activated Protein Kinase	PTC	Premature Terminaison Codon
miR	microARN	qPCR	PCR quantitative
MMP	Matrix Metalloproteinases	RBP	RNA Binding Protein
mTOR	mammalian Target of Rapamycin	RBPDB	RNA Binding protein database
MUC	Mucine	RE	Réticulum Endoplasmique
Myc	Myelocytomatosis oncogene	RISC	RNA-Induced Silencing Complex
NF-κB	Nuclear Factor Kappa B	RLB	Reporter Lysis Buffer
NIDO	Nidogène	RNase	RiboNucléase

RNA-IP	Immunoprécipitation des complexes ARN-protéines	TBS	Tampon Tris-HCl Salin
RNA-Seq	RNA sequencing	TF	Thomsen-Friedenreich
RNP	Ribonucléoprotéine	TGF α/β	Transforming Growth Factor alpha/beta
RPMI	Roswell Park Memorial Institute	TIA-1	T-cell intracytoplasmic antigen 1
RT-PCR	Reverse transcriptase-Polymerase chain reaction	TIAR	TIA-1-related
RT-qPCR	quantitative Reverse Transcriptase-Polymerase Chain Reaction	Tm	Melting Temperature
SSC	Saline-Sodium Citrate	TNF	Tumor Necrosis Factor
SDS	Sodium Dodecyl Sulphate	TP53	Tumor Protein p53
SEA	Sea urchin sperm protein Enterokinase and Agrin	TR	Tandem Repeat
SEM	Standard Error Mean	Tris	Tris-(hydroxyméthyl)aminoéthane
Ser	Sérine	TTF-1	Thyroid Transcription Factor-1
SG	Granules de stress	TTP	Tristetrapolin
sh-RNA	short hairpain RNA	Tyr	Tyrosine
si-RNA	short interfering RNA	UPF	Up-frameshift suppressor homolog
SMAD	Mothers against decapentaplegic homolog	UPR	Unfolded Protein Response
SMG	Smaug	UTR	Untranslated Region
snRNP	small nuclear ribonucleoproteins	v/v	volume/volume
SOC	Super Optimal broth with Catabolite repression	VEGF	Vascular Endothelial Growth Factor
Sp1	Specific protein 1	VNTR	Variable Number of Tandem Repeat
SR	serine/arginine-rich	vWF	von Willebrand Factor
STAT	Signal Transduction and Activator of Transcription	Wnt	Wingless
SVF	Sérum de veau foetal	WT	Wild Type
T°amb.	Température ambiante	XBP1	X-Box Binding Protein 1
		X-Gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
		XRN1	XRN1 (5'-3' Exoribonuclease 1)

Préambule

La cancérogenèse épithéliale s'accompagne de la modification d'expression de nombreux gènes, entraînant ainsi des anomalies des protéines d'ordre quantitatif (augmentation ou diminution de leur niveau d'expression) ou qualitatif (activation constitutive, isoformes d'épissage alternatif, variations des modifications post-traductionnelles...). Parmi les gènes affectés figurent ceux codant les **mucines**, une famille de glycoprotéines typiquement exprimées par tous les types de cellules épithéliales, dont les ARNm possèdent des propriétés inhabituelles par rapport aux autres transcrits humains et dont l'implication dans le processus tumoral est maintenant bien établie. La problématique générale de notre travail s'inscrit donc dans la compréhension des mécanismes contribuant à la modification de l'expression des mucines en situation cancéreuse et ceci au sein de l'équipe « Mucines, différenciation, cancérogenèse épithéliales ».

Pour mener à bien cette étude, nous avons choisi comme modèle le **cancer pancréatique**. En effet, au cours du développement embryonnaire, la mucine membranaire MUC4 est exprimée transitoirement par le pancréas pour être ensuite réprimée dans le tissu pancréatique sain adulte, mais est néo-exprimée dès les stades précoces de la cancérogenèse pancréatique. À l'heure actuelle, les mécanismes à l'origine de cette réexpression ne sont pas entièrement connus. Dans les cancers de la sphère pancréato-bilaire, l'augmentation de l'expression des mucines membranaires est associée à une élévation parallèle de celle de la **galectine-3**, lectine endogène, dont le rôle dans la cancérogenèse est largement documenté. Le modèle du cancer pancréatique est donc un modèle de choix pour étudier en situation cancéreuse les rôles de ces deux protéines. Le premier objectif de cette thèse était donc d'étudier plus précisément l'influence de la galectine-3 sur la régulation de l'expression de la mucine membranaire MUC4. Pour répondre à cette question, nous avons posé nos hypothèses de travail à partir des données de la littérature concernant les mucines et la galectine-3 et à partir des mécanismes généraux gouvernant l'expression des gènes chez les eucaryotes, eux-mêmes en lien avec le métabolisme général des ARNm. Dans un second temps, l'ensemble des résultats obtenus nous ont permis d'élargir l'étude du rôle de la galectine-3 à d'autres ARNm que ceux codant les mucines et de tenter d'inscrire ce rôle dans un contexte physiopathologique en lien avec la cancérogenèse.

Le manuscrit est ainsi organisé :

Dans la première partie de l'introduction nous ferons une brève présentation de la pathologie qui a servi de modèle à notre étude : **l'adénocarcinome pancréatique canalaire**. Puis nous présenterons les **mucines** et plus particulièrement celles dont les ARN messagers ont fait l'objet de notre étude : **MUC4, MUC1 et MUC16**. Nous nous intéresserons ensuite à la **galectine-3**, et pour comprendre son rôle sur la **régulation post-transcriptionnelle** des mucines, nous terminerons par une synthèse sur le **cycle de vie des ARNm**. Puis, l'ensemble des résultats obtenus seront présentés et discutés en deux parties distinctes, la première correspondant aux travaux publiés, la seconde aux travaux en cours qui feront l'objet d'une publication en cours de préparation. Enfin, la dernière partie de ce manuscrit sera consacrée à la description des techniques qui ont été mises en œuvre pour la réalisation de ce travail.

Introduction

I. Adénocarcinome pancréatique canalaire

Le pancréas est une glande amphicrine, c'est-à-dire formée de tissus à fonctions à la fois exocrine et endocrine. Chez l'adulte, les cellules exocrines correspondent à environ 95 % du volume du pancréas (10 % de cellules canalaire et 85 % de cellules acineuses), alors que les cellules endocrines ne représentent qu'1 à 2 % du volume pancréatique (Pour *et al.*, 2003; Coppin, 2012). Cette glande mixte exerce ainsi deux fonctions physiologiques majeures :

- grâce à ses sécrétions exocrines formant le suc pancréatique, elle intervient dans la digestion,
- *via* ses sécrétions endocrines, elle exerce un rôle important dans la régulation du métabolisme glucidique.

Le cancer du pancréas exocrine ou adénocarcinome pancréatique canalaire (PDAC) représente 90 % des tumeurs pancréatiques. Les 10 % restants des cancers pancréatiques sont formés de types histologiques beaucoup plus rares nécessitant des prises en charge spécifiques tels que les cystadénomes séreux, les tumeurs endocrines, les carcinomes à cellules acineuses, les pancréatoblastomes et les néoplasies pseudopapillaires (Hezel *et al.*, 2006). Notre étude ne portera que sur **l'adénocarcinome pancréatique canalaire**. Ce cancer est dit de type épithélial, c'est-à-dire se développant aux dépens des cellules épithéliales qui forment des structures spécialisées où les cellules, appelées épithéliums, sont jointives.

L'adénocarcinome pancréatique canalaire représente la quatrième cause de décès par cancer dans les pays occidentaux (Ferlay *et al.*, 2016). Le plus souvent découvert entre 60 et 70 ans, son diagnostic est très fréquemment réalisé à un stade avancé de la maladie du fait de l'expression clinique tardive, et seuls 20 % des patients sont diagnostiqués à un stade où la tumeur est encore résécable. Tous stades confondus, la survie à 5 ans est inférieure à 5 %. En cas de chirurgie suivie de chimiothérapie, celle-ci passe de 10 à 30 % à 5 ans (Bardeesy and DePinho, 2002).

I.1. Carcinogénèse pancréatique

I.1.1. Lésions précurseurs de l'adénocarcinome pancréatique canalaire

Les lésions précurseurs de l'adénocarcinome pancréatique peuvent être classées en trois catégories (figure 1) : les PanIN (*pancreatic intraepithelial neoplasia* ou néoplasie intraépithéliale pancréatique), les MCN (*mucinous cystic neoplasm* ou néoplasie mucineuse kystique) et les IPMN (*intraductal papillary mucinous neoplasm* ou néoplasie mucineuse papillaire intracanaulaire) (Hezel *et al.*, 2006).

Parmi celles-ci, les PanIN sont les plus étudiées. Elles sont divisées en plusieurs stades progressifs: PanIN-1A, -1B, -2 et -3 qui correspondent à l'évolution d'une prolifération intracanaulaire vers un carcinome invasif (figure 1). Les stades PanIN-1A et -1B sont associés à des lésions de type hyperplasie sans dysplasie, PanIN-2 correspond à une dysplasie variable et PanIN-3 à un carcinome *in situ* (Bardeesy and DePinho, 2002).

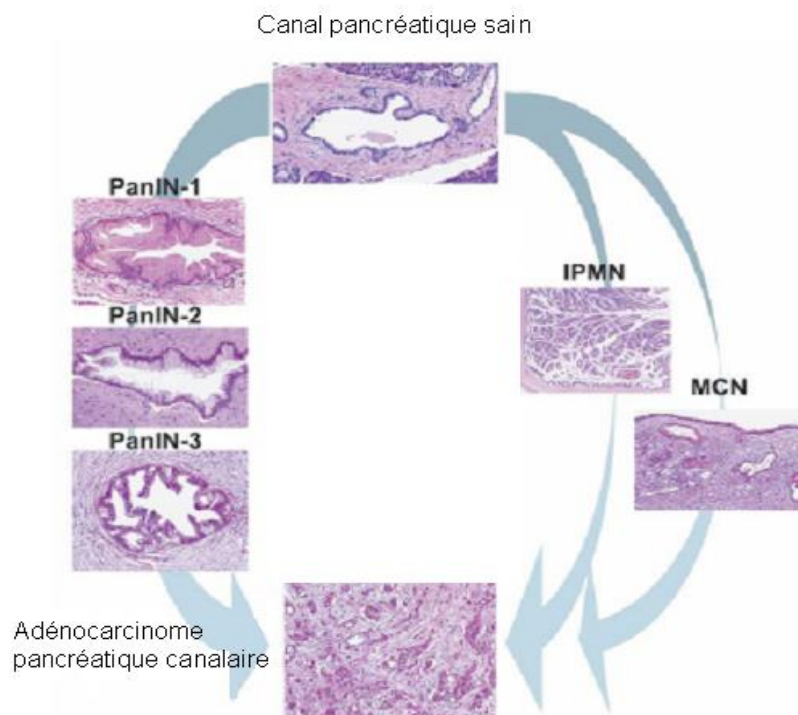


Figure 1 : Les différentes lésions précurseurs de l'adénocarcinome pancréatique canalaire : PanIN, IPMN et MCN (d'après Hezel *et al.*, 2006).

Le PDAC, qui est la forme la plus fréquente du cancer pancréatique, se développe aux dépens des canaux pancréatiques.

1.1.2. Altérations géniques

La croissance non contrôlée et autonome des cellules cancéreuses contribuant à la tumorigenèse est généralement liée à l'activation d'oncogènes et à la perte de fonction de gènes suppresseurs de tumeurs.

La progression tumorale pancréatique est associée à l'accumulation d'altérations génétiques et à des modifications cytologiques (figure 2). Dans le modèle de progression moléculaire décrit par Hruban *et al.*, des anomalies génétiques s'accumulent dès les stades PanIN-1 (activation de *KRAS*, surexpression d'*ERBB2*) et PanIN-2 (inactivation de *INK4A*) jusqu'au stade PanIN-3 (inactivation de *TP53*, *BRCA2*, *DPC4*) (Hruban *et al.*, 2000). En plus des phénomènes de mutations ou de pertes alléliques, l'inactivation de gènes suppresseurs de tumeurs par hyperméthylation des îlots CpG situés dans leurs régions promotrices est fréquemment observée dès les stades précoces de la cancérogenèse (tableau I) (Distler *et al.*, 2014).

Récemment, l'étude du transcriptome de 96 tumeurs pancréatiques a permis d'identifier 4 sous-types de tumeurs: squameux, progéniteur, ADEX (*aberrantly differentiated endocrine exocrine*) et immunogénique. Cette classification a pour objectif d'améliorer la prise en charge thérapeutique en proposant un traitement ciblé en fonction des caractéristiques de chaque sous-type tumoral (Bailey *et al.*, 2016).

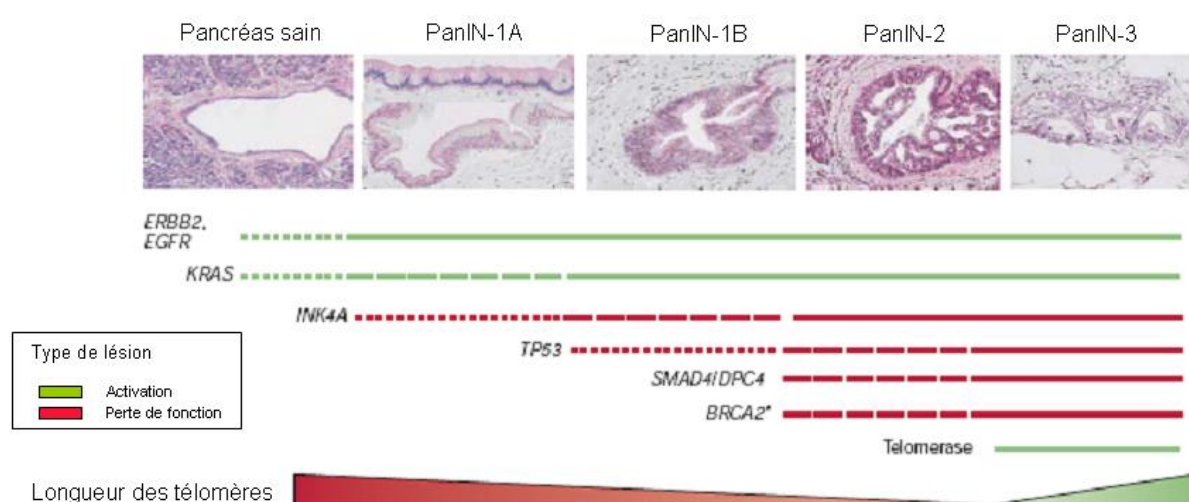


Figure 2 : Modèle de progression de l'adénocarcinome pancréatique (PanIN-1A à PanIN-3) et évolution des anomalies géniques observées au cours de la carcinogenèse pancréatique (d'après Bardeesy and DePinho, 2002).

Tableau I : Anomalies géniques observées dans le PDAC (Stewart and Kleihues 2005).

Gène	Chromosome	Mécanisme d'altération	% des cancers
Oncogenes			
<i>KRAS</i>	12p	Mutation ponctuelle	> 90
<i>MYB, AKT2, AIB 1</i>	6q, 19q, 20q	Amplification ¹	10-20
<i>ERBB2 (HER/2-neu)</i>	17q	Surexpression	70
Gènes suppresseurs de tumeur			
<i>p16^{INK4A}</i>	9p	Délétion homozygote	40
		Perte d'hétérozygotie et mutation intragénique	40
		Hyperméthylation du promoteur	15
<i>p53</i>	17p	Perte d'hétérozygotie et mutation intragénique	50-70
<i>DPC4</i>	18q	Délétion homozygote	35
		Perte d'hétérozygotie et mutation intragénique	20
<i>BRCA2</i>	13q	Mutation intragénique héritée et perte d'hétérozygotie	7
<i>MKK4</i>	17p	Délétion homozygote, perte d'hétérozygotie et mutation intragénique	4
<i>LKB1/STK11</i>	19p	Perte d'hétérozygotie et mutation intragénique	5
		délétion homozygote	
<i>ALK5</i> et <i>TGF βR2</i>	9q, 3p	Délétion homozygote	4
Réparation de mésappariements de l'ADN			
<i>MSH2, MLH1</i> , autres	2p, 3p, autres	Inconnu	< 5

¹En cas d'amplification, il n'est généralement pas possible d'identifier sans ambiguïté l'oncogène clé en raison des nombreux gènes impliqués.

1.2. Dépistage

Un des enjeux importants pour diminuer la mortalité du PDAC est d'en faire le diagnostic à un stade le plus précoce possible. Cependant, à la différence du cancer du côlon, il n'existe pas de dépistage généralisé du cancer du pancréas dans la population. Ceci est lié à l'absence de définition de groupes à risque et au manque d'examens peu coûteux, facilement réalisables et ayant une sensibilité et spécificité suffisantes (Lévy, 2009 ; Coppin, 2012).

A l'heure actuelle, le dosage de l'antigène CA19-9 est peu contributif puisqu'il manque de sensibilité et de spécificité. Parmi les techniques d'imagerie, l'écho-endoscopie représente la technique de choix car elle permet de détecter des lésions pancréatiques de petites tailles (Hammel, 2006; Rosty, 2004) mais celle-ci est invasive. Il est donc nécessaire

de **découvrir de nouveaux biomarqueurs** afin de diagnostiquer précocement ce cancer. De nombreuses études ont vu le jour ces dernières années afin de découvrir des cibles permettant un diagnostic à la fois précoce et non invasif (tableau II). Cependant, parmi toutes celles récemment décrites dans la littérature, il n'existe pas encore de nouveaux marqueurs validés et utilisables en routine. Parmi ces nouvelles approches, nous pouvons citer la recherche des exosomes circulants exprimant la protéine GPC1 (*glypican 1*) (Robinson *et al.*, 2016), l'étude de l'ADN tumoral circulant et la détection de cellules tumorales circulantes (Majumder *et al.*, 2015), la détermination du profil d'expression de panels de microARN (comme les miR-21,-155 et -216) (Le *et al.*, 2016) ainsi que l'étude au niveau sanguin de l'expression de différentes protéines telles que MIC1 (*Macrophage Inhibitory Cytokine-1*) (Le *et al.*, 2016). De nombreux travaux ont montré que le profil d'expression d'une famille de glycoprotéines, les mucines, est modifié au cours des différentes étapes de la cancérogénèse pancréatique (Kaur *et al.*, 2013). Ces études se sont intéressées aux changements observés lors de la progression tumorale et proposent d'étudier l'**expression de panels de mucines** non seulement pour la détection mais aussi pour l'évaluation du pronostic associé à ce cancer. Cette famille de protéines, que sont les mucines, fait l'objet de la partie suivante de l'introduction.

Tableau II : Biomarqueurs potentiels pour la détection du PDAC (Majumder *et al.*, 2015).

Bio specimen	Type of biomarker	Examples of molecular markers
Blood	Conventional protein markers	Carbohydrate antigen 19-9 (CA 19-9) Carcinoembryonic antigen (CEA)
	Novel proteins	Intercellular adhesion molecule-1 (ICAM-1) Osteoprotegerin (OPG) Macrophage inhibitory cytokine-1 (MIC-1) Tissue inhibitor of metalloproteinases-1 (TIMP-1) S100 calcium-binding protein P (S100P)
	Mutated genes	KRAS, TP53, SMAD4, CDKN2A, KDM6A, PREX2
	Aberrantly methylated genes	p16, ppENK, cyclin D2, SPARC/osteonectin SOCS-1, TSLC1
	Micro-RNAs	miR-1290, miR-145, miR-150, miR-223, miR-636, miR-26b, miR-34a, miR-122, miR-126, miR-145, miR-150, miR-223, miR-505, miR-636, miR-885.5p.
	Circulating tumor cells	(molecular markers not yet interrogated)
Cyst fluid	Mutated genes	KRAS, GNAS
	Aberrantly methylated genes	BNIP3, PTCHD2, SOX17, NXPH1 and EBF3
	Micro-RNAs	miR-138, miR-195, miR-204, miR-216a, miR-217, miR-218, miR-802, miR-155, miR-214, miR-26a, miR-30b, miR-31, and miR-125
Tumor tissue	Novel proteins	Gelsolin, Lumican, Galectin-1 and Laminin
Pancreatic juice	Mutated genes	KRAS, TP53
	Aberrantly methylated genes	ADCY1, CD1D, BMP3
Stool	Mutated genes	KRAS, BMP3

II. Les mucines

II.1. Généralités

Les mucines sont connues depuis plusieurs décennies pour être les constituants caractéristiques du gel de mucus, substance visqueuse et très hydrophile située à l'interface entre le milieu extérieur et l'épithélium sous-jacent (figure 3). Elles assurent la protection des cellules épithéliales contre les agressions extérieures qui peuvent être variées selon les épithéliums considérés, en particulier digestifs, respiratoires, urogénitaux ou des organes neurosensoriels. Il peut s'agir d'une protection contre des agents pathogènes bactériens ou viraux, enzymes digestives, sels biliaires, pH acide ou alcalin, ou encore de polluants chimiques de l'atmosphère. Initialement, les mucines ont été caractérisées sur le plan biochimique comme des glycoprotéines sécrétées de haut poids moléculaire conférant au gel de mucus ses propriétés rhéologiques (viscosité, fluidité, adhérence, élasticité, rôle de lubrifiant). En effet, grâce à l'établissement de ponts disulfures, elles forment des assemblages multimériques responsables de l'organisation d'un réseau tridimensionnel. Elles sont constituées d'une partie protéique appelée apomucine sur laquelle viennent se lier de façon covalente de très nombreuses chaînes glycaniques. Dans les années 1990, grâce au clonage moléculaire, cette famille s'est élargie au groupe des mucines membranaires. Bien que dépourvues de la capacité de former un réseau tridimensionnel, cette catégorie de mucines contribue aux propriétés biologiques et physicochimiques du glycocalyx, mince film hydrophile interposé entre le pôle apical de la cellule épithéliale et le gel de mucus. Elles sont aussi impliquées dans les interactions cellules-cellules et cellules-matrice extracellulaire ainsi que dans la modulation de la transduction de signaux (Andrianifahanana *et al.*, 2006; Hattrup and Gendler, 2008; Porchet and Aubert, 2004).

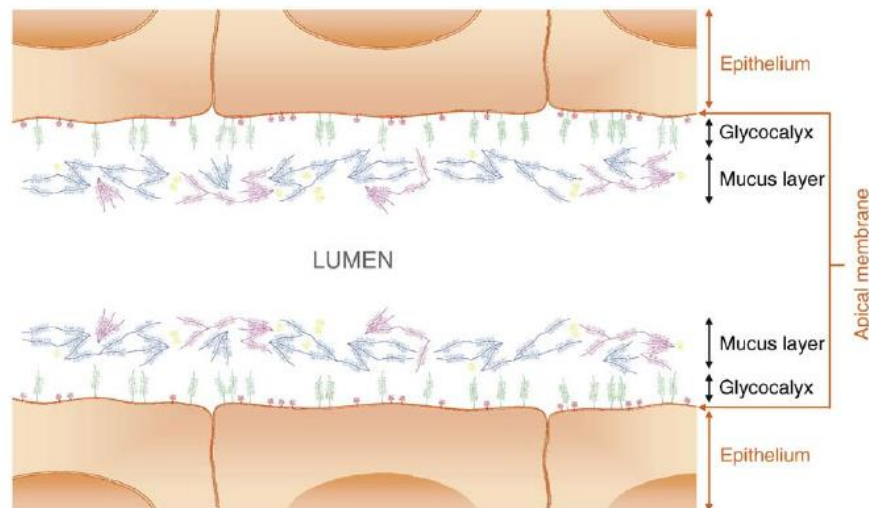


Figure 3 : Rôles des mucines dans la composition et les fonctions du mucus (Andrianifahanana *et al.*, 2006).

Les mucines secrétées, qui contiennent dans leur squelette peptidique des modules riches en cystéines, forment des ponts disulfures entre les monomères de mucines et sont responsables de l'organisation du gel de mucus. Le gel de mucus varie en épaisseur entre 50 et 450 nm selon les épithéliums. Les mucines transmembranaires participent à la constitution du glycocalyx, mince film hydrophile d'une épaisseur d'environ 10 nm, qui environne la membrane plasmique et duquel elles débordent largement pour pénétrer dans le mucus, en raison de leur taille importante (0,5 à 10 μm de longueur). Très riche en eau, le mucus contient bien d'autres constituants nécessaires à la physiologie du mucus (minéraux, immunoglobulines, cytokines, peptides à fonction antibactérienne, protéases...).

La famille des gènes humains codant les apomucines (*MUC family*, HGNC), comporte 21 gènes (tableau III) (dont certains sont organisés en complexes, tableau IV), qui possèdent des caractéristiques structurales communes : des séquences codantes répétées en tandem, des exons de grande, voire de très grande taille, un polymorphisme de répétition (petites variations de séquences entre les différentes unités élémentaires répétées), et des messages codants modulaires c'est-à-dire que les protéines codées résultent de la juxtaposition de différents domaines fonctionnels indépendants appelés modules (Porchet and Aubert, 2004).

Tableau III : La famille des gènes *MUC* (données extraites de la base de données HGNC (Human Genome Nomenclature Committee) et rédigées par D Swallow). Celle-ci comporte 21 gènes répartis sur 8 chromosomes. Certains sont proches et organisés en complexes (11p15, 7q22 et 6p21, 3q29).

Approved Symbol	Approved Name	Previous Symbols	Synonyms	Chromosome
MUC1	mucin 1, cell surface associated	PUM, MCKD1	CD227, PEM, ADMCKD, ADMCKD1, MCKD, MCD	1q22
MUC2	mucin 2, oligomeric mucus/gel-forming			11p15.5
MUC3A	mucin 3A, cell surface associated	MUC3		7q22.1
MUC3B	mucin 3B, cell surface associated			7q22 not on reference assembly
MUC4	mucin 4, cell surface associated			3q29
MUC5AC	mucin 5AC, oligomeric mucus/gel-forming		MUC5	11p15.5
MUC5B	mucin 5B, oligomeric mucus/gel-forming	MUC5	MG1	11p15.5
MUC6	mucin 6, oligomeric mucus/gel-forming			11p15.5
MUC7	mucin 7, secreted		FLJ27047, MG2	4q13.3
MUC8	mucin 8			12q24.33
OVGP1	oviductal glycoprotein 1	MUC9	CHIT5	1p13.2
MUC12	mucin 12, cell surface associated	MUC11		7q22.1
MUC13	mucin 13, cell surface associated	DRCC1		3q21.2
EMCN	endomucin		MUC14	4q24
MUC15	mucin 15, cell surface associated			11p14.3
MUC16	mucin 16, cell surface associated		CA125, FLJ14303	19p13.2
MUC17	mucin 17, cell surface associated			7q22.1
MUC19	mucin 19, oligomeric		FLJ35746	12q12
MUC20	mucin 20, cell surface associated		FLJ14408, KIAA1359	3q29
MUC21	mucin 21, cell surface associated	C6orf205	bCX31G15.2	6p21.33
MUC22	mucin 22		PBMUCL1	6p21.33

Tableau IV : Localisation chromosomique des gènes de mucines organisés en complexes. Le complexe localisé en 11p15 contient 4 gènes qui codent les mucines sécrétées responsables de l'organisation des gels de mucus (*gel-forming mucins*). Ces mucines sont fortement exprimées au niveau du tube digestif et de l'appareil respiratoire.

Position chromosomique	MUC
7q22	MUC3A/3B, MUC12, MUC17
3q29	MUC4, MUC20
11p15	MUC6, MUC2, MUC5AC, MUC5B
6p21	MUC21, MUC22

II.2. Structure et classification

II.2.1. Apomucine et chaines glycaniques

a. Apomucine

Le squelette protéique des mucines est appelé apomucine. Il est caractérisé par la juxtaposition de domaines fonctionnels et notamment par la présence de séquences d'acides aminés (aa) répétées en tandem, riches en résidus sérines (S), thréonines (T) et prolines (P) (20 à 55 % de la composition en aa), appelées séquences PTS et constituant le domaine élémentaire *tandem repeat* (TR). Sur les résidus sérines et thréonines de nombreuses chaînes O-glycaniques viennent se lier de façon covalente (60 à 80 % du poids sec des mucines), les résidus prolines donnant à la protéine la conformation nécessaire à la glycosylation et au maintien de l'apomucine sous forme étirée. La composition moyenne de ces motifs TR élémentaires, leur longueur et le nombre de leurs répétitions en tandem varient d'une mucine à l'autre (Porchet and Aubert, 2004). Le tableau V rassemble ces motifs pour différentes mucines, MUC1 ayant par exemple un motif TR composé de 20 acides aminés et MUC4 de 16 acides aminés. Il existe, pour une même mucine, un degré de diversité supplémentaire engendré par la variation du nombre de répétitions du motif TR entre les individus, à l'origine d'un polymorphisme de type VNTR (*Variable Number of tandem repeat*). Ainsi pour MUC1, le motif TR est répété de 25 à 125 fois et pour MUC4, le nombre de répétitions varie entre 145 et 395 fois (Porchet and Aubert, 2004; Rachagani *et al.*, 2009). Les autres domaines fonctionnels des mucines, sur lesquels s'est exercée la pression de sélection au cours de l'Evolution, sont conservés entre les espèces alors que les régions TR ont évoluées indépendamment et divergent entre elles (Dekker *et al.*, 2002).

Tableau V : Organisation du motif TR de quelques mucines (d'après Chaturvedi *et al.*, 2008; Rachagani *et al.*, 2009).

Mucine	Nombre d'AA	Nombre de répétitions	Séquence peptidique du motif
MUC1	20	25-125	PDTRPAPGSTAPPAHGV TSA
MUC2	23	50-100	PTTTPITTTTTVTPTPTPTGTQT
MUC4	16	145-395	TSSASTGHATPLPVT D
MUC5AC	8	TR1 :124 TR2 :17 TR 3 :34 TR4 :66	TTSTTSAP
MUC5B	29	11	AT(G/S)STATPSS(T/S)PGT(T/A)(H/W)T(P/L) (P/T)VL(T/S)(T/S)T(A/T)TT(P/T)T
MUC7	23	6	TTAAPPTPSATTPAPPSSSAPPG
MUC16	156	60	FNPWSSVPTTSTPGTSTPGTSTVHLATSGTP SSLPGHTAPVPLLIPFTLNFTITNLHYEENM QHPGSRKFNTTERVLQGLLKPLFKSTSVGP LYSGCRLTLRPEKHGAATGVDAICTLR LD PTGPGLDRERLYWELSQLTNSVTELGPYTL DRDSL YVNG

b. Chaînes glycaniques

Les mucines sont des protéines hautement glycosylées, portant de très nombreuses chaînes O-glycaniques et pour certaines d'entre elles également des chaînes N-glycaniques (22 sites consensuels de N-glycosylation au sein de MUC4, 150 au sein de MUC16). Les structures glycaniques, portées par une apomucine donnée, dépendent essentiellement de l'expression des enzymes permettant leur biosynthèse, les glycosyltransférases (Singh and Hollingsworth, 2006). Les glycosyltransférases constituent chez l'homme une superfamille comportant 214 gènes, répartis en 26 sous-familles et comptant de nombreuses isoformes possibles par gène. Ces motifs glycaniques varient en fonction du tissu, du type cellulaire, du stade de différenciation cellulaire et des situations physiopathologiques, comme l'inflammation ou le cancer (Andrianifahanana *et al.*, 2006). Leur assemblage variable conduit à une grande diversité de structures différant par leur composition, leur longueur (2 à

20 résidus de sucres), leur ramification, leur acidité et donc par l'hydrophilie et l'encombrement stérique qu'elles créent autour du squelette peptidique.

Les chaînes O-glycosylées sont formées de 3 domaines : le noyau, le squelette et la région périphérique (Lamblin *et al.*, 1991) (figure 4). Leur biosynthèse débute par l'addition d'un résidu N-acétylgalactosamine (GalNac) sur une sérine ou une thréonine pour former l'antigène Tn. Puis des glycosyltransférases viennent additionner des résidus de galactose ou de N-acétylglucosamine de façon linéaire ou branchée conduisant à la formation du noyau. Il en existe 4 types différents selon le ou les glycanes ajoutés. Par exemple, l'addition d'un résidu de galactose forme le motif Gal β 1-3GalNac-O-Ser/Thr (appelé l'antigène T ou core 1). Le squelette est quant à lui constitué de la répétition d'un motif disaccharidique composé de Galactose (Gal) et de N-acétylglucosamine (GlcNac). Enfin, la région périphérique est formée par des résidus de fucose, d'acide sialique (N acétylneuraminique ou NeuAc), de galactose ou de N-acétylgalactosamine, ces deux derniers pouvant être sulfatés. Ces terminaisons créent des motifs antigéniques de type groupe sanguin (Lewis A, B, H). La présence de résidus sulfates et d'acides neuraminiques sur la périphérie des chaînes oligossacharidiques confère aux mucines des propriétés polyanioniques. Ainsi, les mucines ont pu être classées en trois catégories selon leur charge : les mucines neutres pauvres en résidus acides (fucomucines), les mucines acides riches en résidus d'acides N-acétylneuraminiques (sialomucines) et les mucines très acides riches en résidus sulfates (sulfomucines) (Groux-Degroote *et al.*, 2008).

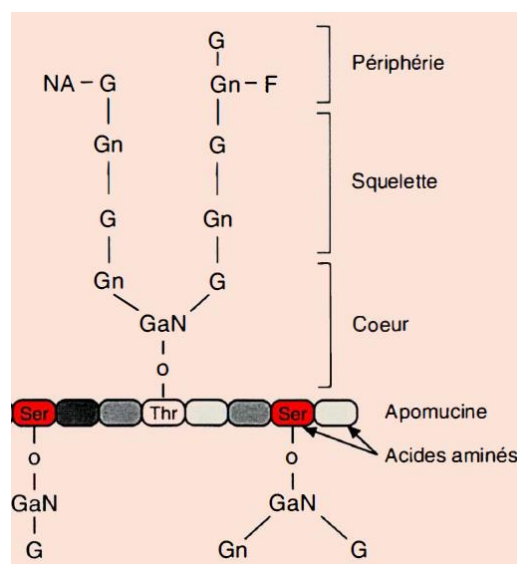


Figure 4 : Représentation schématique des chaînes O-glycaniques (d'après Lamblin *et al.*, 1991).

Celles-ci sont constituées de différentes parties : le cœur, le squelette et la région périphérique.

Les mucines comportent également à leurs extrémités C- et N- terminales des sites de N-glycosylation sur les résidus asparagines situés au sein de la séquence consensus Asn-X-Ser/Thr. Ces N-glycanes sont impliqués dans le transport, la maturation et la sécrétion des mucines (Parry *et al.*, 2006). Au sein de la cellule, la glycosylation de l'axe peptidique des mucines est amorcée au sein du réticulum endoplasmique lors d'étapes co-traductionnelles (étapes initiales de conformation de l'apomucine et de N-glycosylation) puis post-traductionnelles tout au long de l'appareil de Golgi.

II.2.2. Classification des mucines et relation structure-fonction

Les mucines sont classées en deux catégories principales en fonction leurs propriétés structurales : les mucines sécrétées et les mucines membranaires (tableau VI) (Porchet and Aubert, 2004).

Tableau VI : Classification des mucines (d'après Andrianifahanana *et al.*, 2006; Corfield, 2015).

Symbole	Localisation	Taille ARN kb	Symbole	Localisation	Taille ARN kb
<i>Soluble/ Gel-forming</i>			<i>Membranaires</i>		
MUC2	11p15.5	15.7	MUC1	1q21	4
MUC5AC	11p15.5	17.5	MUC3A/B	7q22	11.3/-
MUC5B	11p15.5	17.9	MUC4	3q29	26.5
MUC6	11p15.5	8	MUC12	7q22	16.3
MUC19	12q12	25.3	MUC13	3q21.2	2.8
			MUC15	11p14.3	3.4
<i>Soluble/ Non gel forming</i>			MUC16	19p13.2	43.8
MUC7	4q13-q21	2.5	MUC17	7q22	14.2
MUC8	12q24.3	-	MUC20	3q29	2.6
MUC9	1p13	2.3	MUC21	6p21	2.5
			MUC22	6p21	6

a. Mucines sécrétées

Les mucines sécrétées sont divisées en deux catégories : les mucines sécrétées formant le gel de mucus et les mucines solubles (Jonckheere *et al.*, 2013; Moniaux *et al.*, 2004; Porchet and Aubert, 2004).

Les principales mucines sécrétées formant le gel de mucus sont les mucines exprimées par les gènes du complexe 11p15 (Pigny *et al.*, 1996) : *MUC2*, *MUC5AC*, *MUC5B*, *MUC6* auxquelles vient s'ajouter *MUC19*. Elles sont synthétisées par des cellules épithéliales spécialisées dans le stockage et la sécrétion apicale des mucines, les cellules caliciformes ou cellules en gobelets (*goblet cells*). Elles possèdent toutes dans leur structure le domaine TR caractéristique ainsi que des régions riches en cystéines, appelées domaines CK (*Cystine-Knot*) et D, au niveau N- et C-terminal et présentant des homologies avec un module de même organisation présent dans le facteur de coagulation von Willebrand (vWF) et dans des facteurs de croissance constituant la superfamille des protéines à nœud cystine (figure 5). Par ailleurs, d'autres régions riches en cystéines sont localisées en alternance avec le domaine TR. Les monomères de cette catégorie de mucines sécrétées se dimérisent par l'intermédiaire de leur domaine CK et, grâce aux nombreux résidus cystéines de la partie N-terminale, forment des oligomères puis des multimères responsables de la formation du réseau tridimensionnel du gel de mucus.

Les mucines solubles (*MUC7* et *MUC9*) sont quant à elles plus petites, ne forment pas de gel, et sont principalement exprimées dans les sécrétions moins visqueuses comme les sécrétions salivaires et lacrymales.

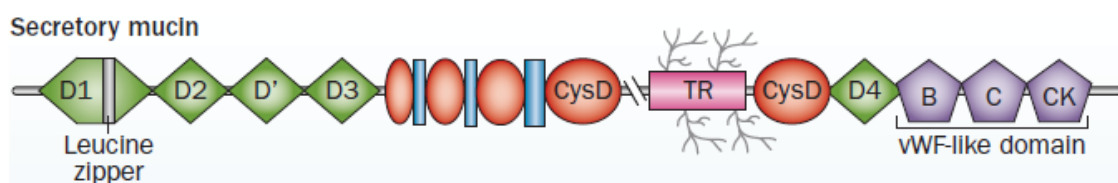


Figure 5 : Représentation schématique d'une mucine sécrétée avec ses différents domaines (Kaur *et al.*, 2013).

Les mucines sécrétées contribuent ainsi aux propriétés rhéologiques et biologiques du mucus qui recouvre les épithéliums bordant la lumière des tractus digestifs, respiratoires et urogénitaux. Elles jouent un rôle important dans leur protection (hydratation, lubrification, protection contre les pathogènes) et leur glycosylation crée un environnement hydrophile idéal pour ces fonctions (Porchet and Aubert, 2004). De plus, les mucines sécrétées interagissent avec des protéines présentes dans le milieu extracellulaire. Par exemple, l'interaction entre MUC5B et des protéines salivaires protégerait ces dernières de la protéolyse et contribuerait à la protection de la cavité buccale (Senapati *et al.*, 2010).

b. Mucines membranaires

Les mucines membranaires (MUC1, -3A, -3B, -4, -12, -13, -15, -16, -17, -20, -21 et -22) sont exprimées par de nombreux types de cellules épithéliales non spécialisées dans la mucosécrétion. Parmi les mucines de cette catégorie, MUC1 et MUC4 sont les mieux caractérisées. Ces mucines ont été longtemps appelées *mucin-like*, car non sécrétées et de ce fait considérées initialement comme ne faisant pas partie de la famille des mucines.

Ces mucines sont composées de 3 grandes parties (figure 6):

- un domaine extracellulaire, qui comporte la région TR caractéristique formant une structure rigide s'étendant au-delà du glycocalyx et pénétrant largement dans le gel de mucus ;
- un domaine hydrophobe transmembranaire unique ;
- une courte région intracytoplasmique pouvant comporter des sites de phosphorylations sur des résidus sérines, thréonines et tyrosines.

Le domaine extracellulaire est composé de différents modules : parmi eux, le domaine SEA (*Sea urchin sperm protein Enterokinase and Agrin*) est présent dans les mucines MUC1, -3A, -3B, -12, -13, -16, -17, et les domaines *Epidermal Growth Factor (EGF) like* sont présents dans MUC3, -4, -12, -13, -17. D'autres domaines sont spécifiques de la mucine MUC4, comme les domaines NIDO (*nidogen-like domain*), AMOP (*adhesion-associated domain in MUC4 and other proteins*) et vWFD (*von Willebrand factor D domain*) (figure 6). La mucine MUC15 est rattachée à ce groupe bien que dépourvue du domaine TR, le domaine le plus caractéristique des mucines (Jonckheere *et al.*, 2013).

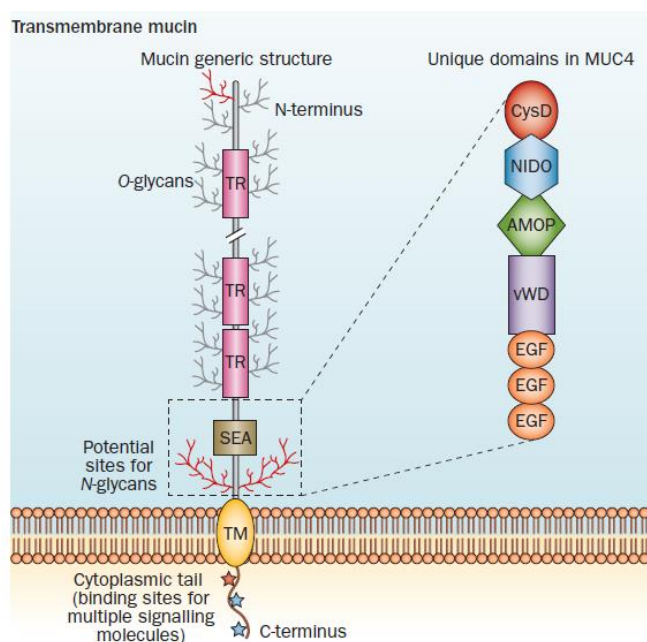


Figure 6 : Représentation schématique d’une mucine membranaire avec ses différents domaines (Kaur *et al.*, 2013).

Les domaines spécifiques à MUC4 sont représentés à droite. Abréviations : AMOP, adhesion-associated domain in MUC4 and other proteins; EGF, epidermal growth factor; NIDO, nidogen-like domain, SEA, sea urchin sperm protein– enterokinase–agrin; TM, transmembrane; TR, tandem repeat, vWD, von Willebrand factor D domain.

L’étude structurale et fonctionnelle des mucines membranaires a permis de montrer que par épissage alternatif certains gènes pouvaient inclure ou exclure le domaine transmembranaire et donc coder soit une isoforme membranaire, soit une isoforme sécrétée. C’est en particulier le cas pour MUC4. De plus, les mucines membranaires ont la possibilité de devenir soluble suite à un clivage protéolytique (Andrianifahanana *et al.*, 2006). Ceci a été démontré initialement pour MUC1.

Les mucines membranaires, exprimées à la surface cellulaire, typiquement au pôle apical de la cellule épithéliale, participent, en association avec les mucines sécrétées, à la fonction du gel de mucus et ainsi à la protection des épithéliums sous-jacents (figure 7), mais aussi à d’autres fonctions biologiques importantes telles que le filtrage ou la modulation des signaux qui parviennent au pôle apical de la cellule. Elles interagissent avec toutes sortes de ligands qui peuvent être extracellulaires, intracellulaires ou transmembranaires et ceci permet de comprendre l’origine de leurs nombreuses fonctions. Elles sont ainsi impliquées dans les interactions cellules-cellules (par exemple, via la liaison de MUC1 à ICAM1 (*Intercellular Adhesion Molecule 1*)) et les interactions cellules-matrice extracellulaire (par exemple,

interaction entre le domaine NIDO de MUC4 et la fibuline 2, (Senapati *et al.*, 2012)) ainsi que dans la signalisation cellulaire. Ainsi, la queue cytoplasmique de MUC1, qui comprend des sites de phosphorylation, sert de plateforme d'ancrage pour de nombreux acteurs des voies de signalisation comme par exemple la voie des MAPkinases (*mitogen-activated protein kinase*) (Carraway *et al.*, 2003; Singh and Hollingsworth, 2006). Les mucines sont aussi capables d'interagir avec les récepteurs de la famille ErbB. De cette façon, la mucine MUC4, par le biais de ses domaines de type EGF-like qui interagissent avec le récepteur ErbB2, exerce la fonction de co-récepteur (Jonckheere and Van Seuning, 2008; Singh and Hollingsworth, 2006).

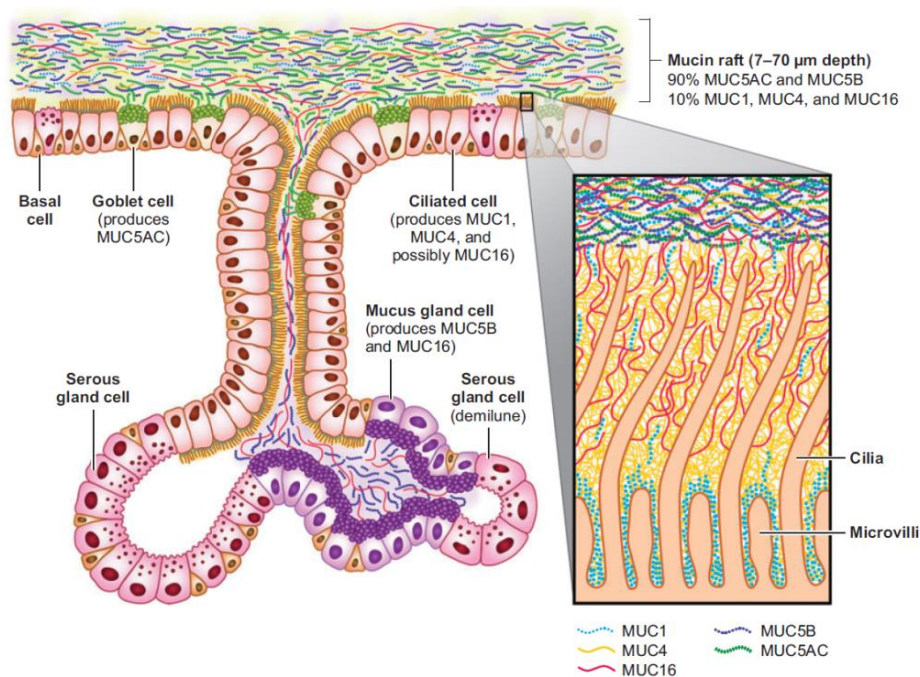


Figure 7 : Origine cellulaire et tissulaire des mucines sécrétées et membranaires entrant dans la composition du mucus bronchique (Hattrup and Gendler, 2008).

Cette composition peut varier selon les étages de l'arbre trachéobronchique et moduler la fluidité et la fonction du mucus.

II.3. Territoires d'expression physiologique

L'expression des mucines est régulée de façon spatio-temporelle et dépend du stade de développement des tissus (figure 8) (Audie *et al.*, 1993). Certaines mucines sont exprimées à un stade précédant la différenciation cellulaire et seraient impliquées dans la morphogenèse. Ainsi, MUC4 est exprimée dès 6 semaines et demie de gestation chez l'embryon et ses

ARNm sont détectés en premier au niveau de la trachée et des bronches (Buisine *et al.*, 1999). Pour MUC1, l'expression débute vers 8 semaines de gestation (Leroy *et al.*, 2006). Certaines mucines sont présentes de façon transitoire au cours du développement. Par exemple, l'expression MUC5AC au niveau du tractus intestinal décroît à partir de 12 semaines de gestation et elle n'est plus du tout détectée à partir de 18 semaines (Buisine *et al.*, 1998).

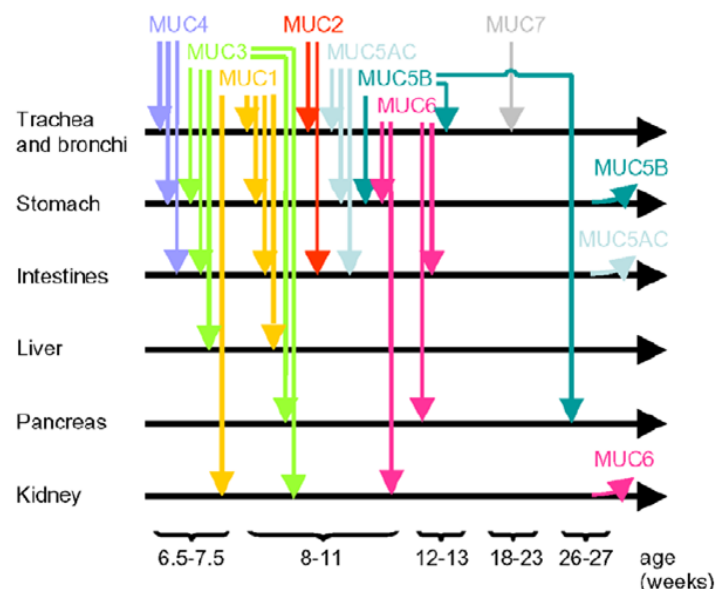


Figure 8 : Schéma temporel d'expression des mucines dans différents organes au cours du développement (Buisine *et al.*, 2008).

L'expression des mucines dépend également de l'état du tissu, sain, inflammatoire ou cancéreux. La figure 9 résume les sites d'expression des principales mucines en situation physiologique. Elles sont détectées principalement au niveau des épithéliums digestif, respiratoire et uro-génital (Andrianifahanana *et al.*, 2006). Par exemple, l'expression de MUC4 n'est pas restreinte aux cellules mucipares mais s'observe également au niveau des cellules épithéliales non spécialisées dans la mucosécrétion telles que les cellules ciliées bronchiques ou les entérocytes. Ainsi, cette mucine est exprimée par les épithéliums respiratoire, digestif et urogénital alors qu'elle n'est pas détectée dans le pancréas sain, la vésicule biliaire et le foie (Porchet and Aubert, 2004). L'expression de MUC4 est uniquement localisée au pôle apical des cellules épithéliales saines (Fauquette *et al.*, 2007). Elle est également présente sous forme sécrétée dans certaines sécrétions biologiques comme le lait, les larmes et la salive (Chaturvedi *et al.*, 2008). L'expression de la mucine MUC1 est très ubiquitaire. Elle est localisée à la surface apicale des cellules de nombreux épithéliums,

bronchique, mammaire, salivaire, pancréatique, prostatique et utérin. Son expression est plus faible dans les cellules épithéliales du côlon, de l'intestin grêle, de l'estomac et de la vésicule biliaire. Une forme soluble est aussi présente dans le lait maternel (Hatrup and Gendler, 2008).

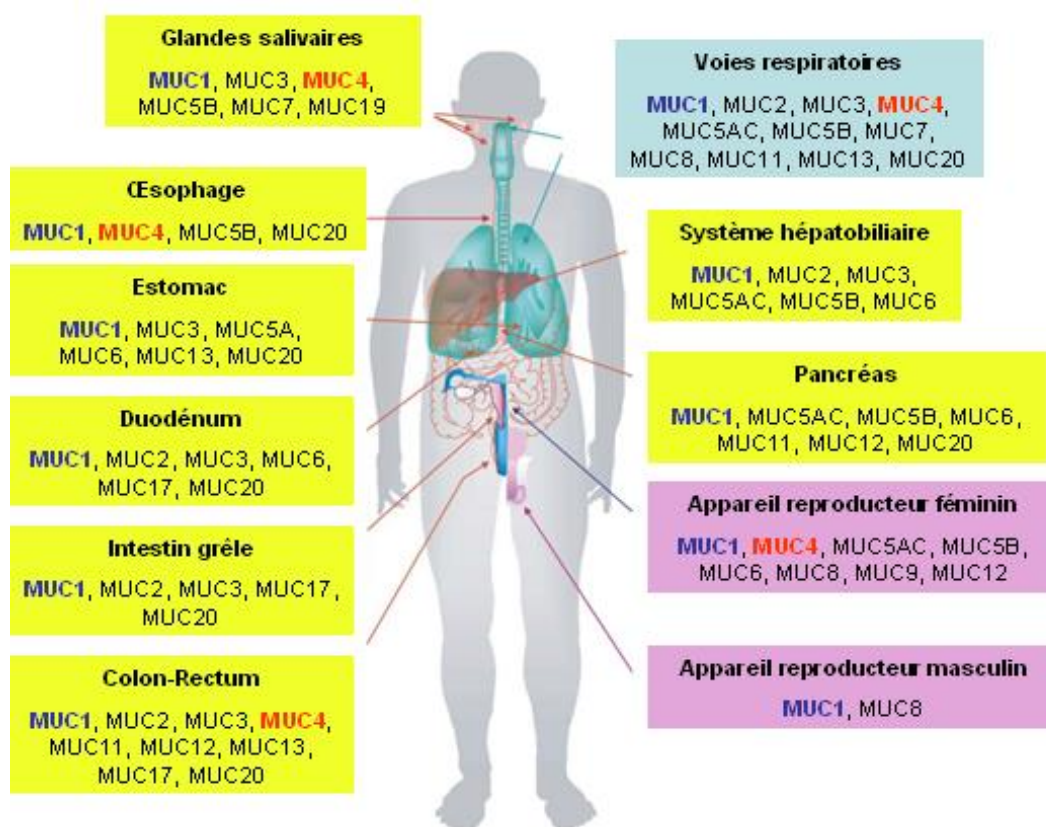


Figure 9 : Sites d'expression des mucines chez le sujet sain (adapté de Andrianifahanana *et al.*, 2006) : épithéliums digestif (jaune), respiratoire (bleu) et urogénital (rose).

II.4. Mucines et cancérogénèse

Les mucines sont impliquées dans la pathogenèse des cancers. En effet, des modifications qualitatives portant sur leur glycosylation, leur distribution cellulaire, leur épissage ainsi que des variations de leur niveau d'expression sont fréquemment observées au cours de la cancérogénèse. Ces changements favorisent la croissance ainsi que la survie des cellules cancéreuses dans des conditions souvent défavorables et modulent leur interaction avec le microenvironnement et leur reconnaissance par le système immunitaire (Hollingsworth and Swanson, 2004).

II.4.1. Modifications qualitatives

a. Altération de la glycosylation

La transformation maligne est caractérisée par une « simplification » de la glycosylation avec la synthèse de chaînes glycaniques tronquées, moins ramifiées (Ag Tn et T) associée à une augmentation de la synthèse des formes sialylées (figure 10). La néosynthèse des motifs sLe^a et sLe^x (Sialyl Lewis) est observée dans les stades plus tardifs de la cancérogénèse (Pinho and Reis, 2015). Ces modifications sont notamment liées à des changements d'expression des glycosyltransferases ou de protéines leur étant associées, comme la protéine chaperonne Cosm nécessaire à l'activité de la 1β3-galactosyltransferase (Hanson and Hollingsworth, 2016). Elles conduisent à la sous-expression ou à la sur-expression de motifs glycaniques normaux mais également à l'apparition de nouveaux motifs dont l'expression est normalement restreinte au tissu embryonnaire. Ces derniers correspondent aux *tumor associated antigen* (TAA) et peuvent être détectés dans la circulation et donc utilisés pour le diagnostic de certains cancers. Ainsi, le dosage du marqueur CA19-9 en cancérologie repose sur la détection du motif sLe^a (Glavey *et al.*, 2015; Pinho and Reis, 2015).

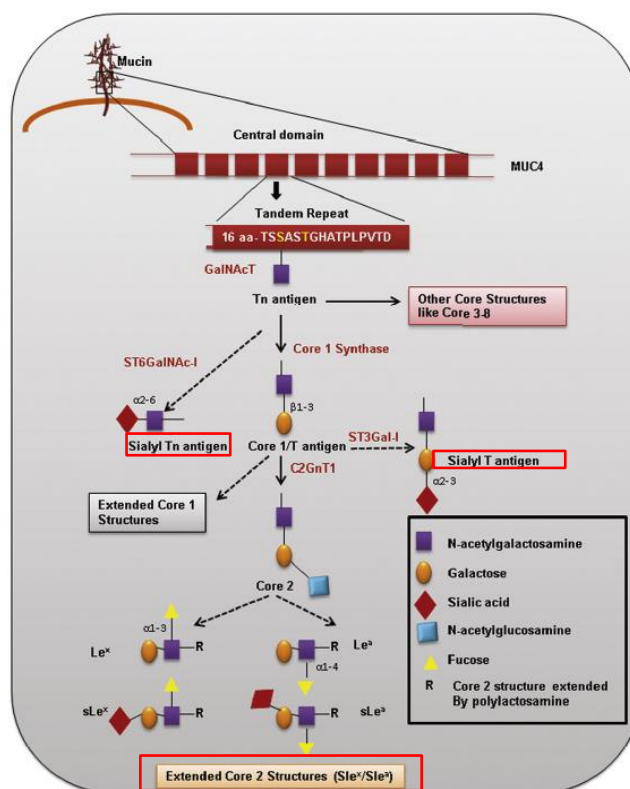


Figure 10 : Altérations de la synthèse des glycanes lors de la cancérogénèse avec notamment apparition des antigènes sT, sTn et sLe^{a/x} (Chugh *et al.*, 2015).

Ces anomalies de glycosylation sont mises à profit par les cellules tumorales de différentes façons. Ainsi, elles les protègent de la cytotoxicité médiée par les cellules *natural killer* (NK) et affectent la réponse immunitaire anti-tumorale. Par exemple les changements de glycosylation de MUC16 modifient la reconnaissance entre les cellules tumorales et les cellules NK. De plus, elles modifient les interactions des mucines avec certains de leurs ligands. Par exemple, le motif sLe^x favorise leurs interactions avec les sélectines (Madsen *et al.*, 2013; Remmers *et al.*, 2013). Ces modifications influencent également les propriétés adhésives des mucines, la charge négative conférée par les acides sialiques favorisant le détachement des cellules tumorales (figure 11).

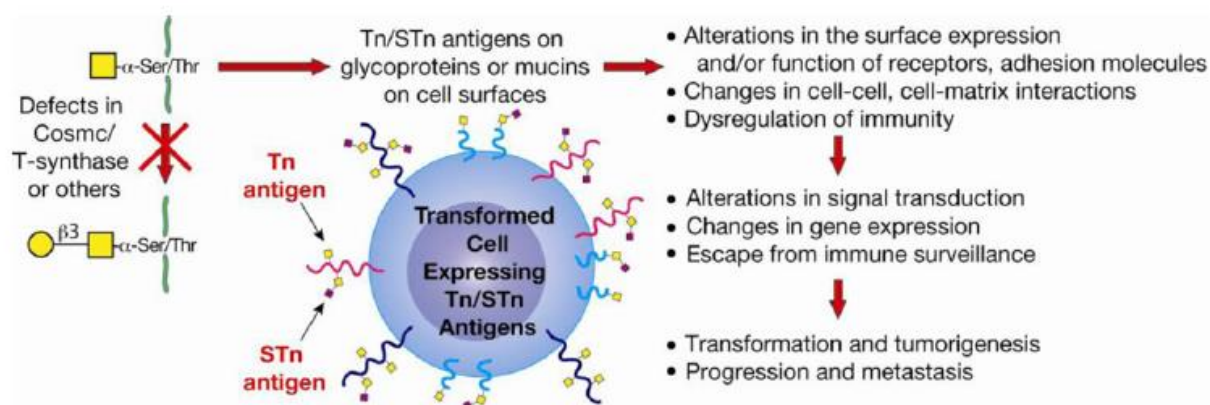


Figure 11 : Exemples de modifications des propriétés cellulaires induites par la présence des antigènes Tn et sTn sur les mucines exprimées par les cellules cancéreuses (Ju *et al.*, 2013).

b. Délocalisation des mucines

La cancérogenèse s'accompagne d'une perte de la polarité cellulaire qui conduit à une distribution circonférentielle des mucines, celles-ci étant situées normalement au pôle apical des cellules épithéliales saines (figure 12). Ces changements de localisation permettent leur interaction avec des molécules exprimées au pôle basolatéral, comme les récepteurs à activité tyrosine kinase, favorisant ainsi la prolifération et la survie cellulaire (Singh and Hollingsworth, 2006). Ainsi, la formation du complexe quaternaire MUC4-ErbB2-ErbB3-neuréguline, qui se forme uniquement dans les cellules ayant perdu leur polarité, active la voie PKB/Akt à l'origine d'une stimulation de la prolifération cellulaire (Jonckheere and Van Seuning, 2008). De plus, une séquestration cytoplasmique des mucines membranaires

MUC1 et MUC4 est observée dans les cellules cancéreuses (Fauquette *et al.*, 2007; Merlin *et al.*, 2011).

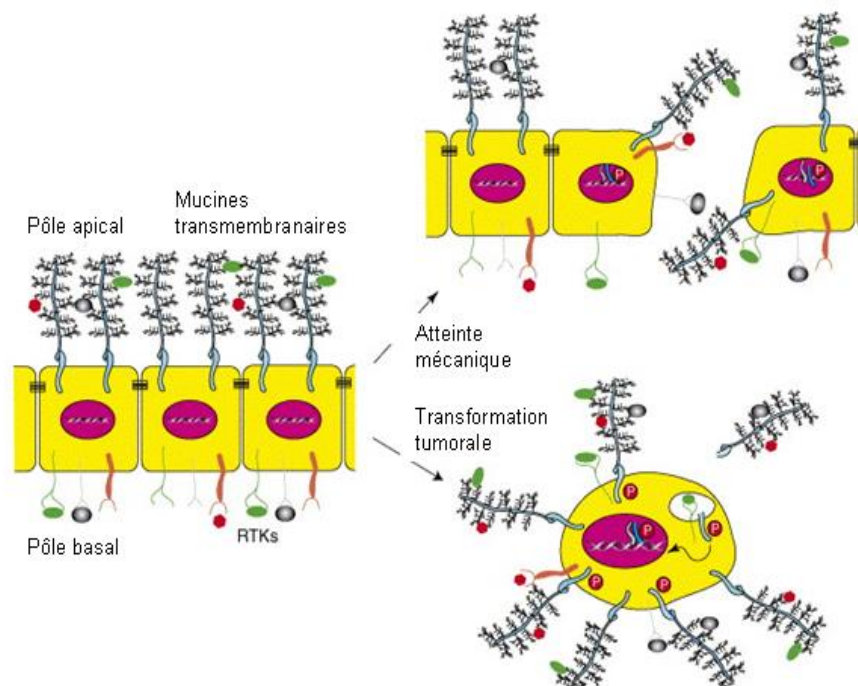


Figure 12 : Illustration de la perte de polarité des cellules cancéreuses et de la délocalisation circonférentielle des mucines permettant leur interaction avec des récepteurs basolatéraux (Singh and Hollingsworth, 2006).

c. Variants d'épissage

L'épissage alternatif des ARNm de mucine est à l'origine de variants correspondants à des formes sécrétées, membranaires et dépourvues du domaine central répété en tandem. MUC1 subit un épissage alternatif générant, tout comme pour MUC4, ces trois types de variants. Certaines de ces formes sont préférentiellement exprimées dans les tumeurs malignes (Apostolopoulos *et al.*, 2015). Ainsi, Obermair et al ont montré une modification du profil des variants d'épissage de MUC1 en comparant les tumeurs bénignes et malignes de l'ovaire (Obermair *et al.*, 2002).

II.4.2. Modifications quantitatives

A côté de ces anomalies qualitatives, des modifications du profil d'expression des mucines sont observées dans les maladies inflammatoires et le processus tumoral.

Par exemple, MUC4 est fréquemment surexprimée ou néo-exprimée dans différents cancers (figure 13) tels que l'adénocarcinome pulmonaire, la dysplasie de haut grade de l'œsophage et l'adénocarcinome œsophagien développé sur métaplasie de Barrett, les carcinomes ovariens, les cholangiocarcinomes intrahépatiques et les carcinomes développés à partir de canaux biliaires extrahépatiques. Dans toutes ces situations, cette surexpression est associée à un mauvais pronostic (Jonckheere and Van Seuning, 2010). Cependant, dans les carcinomes épidermoïdes des glandes salivaires, l'augmentation de l'expression de cette protéine est de bon pronostic (Weed *et al.*, 2004).

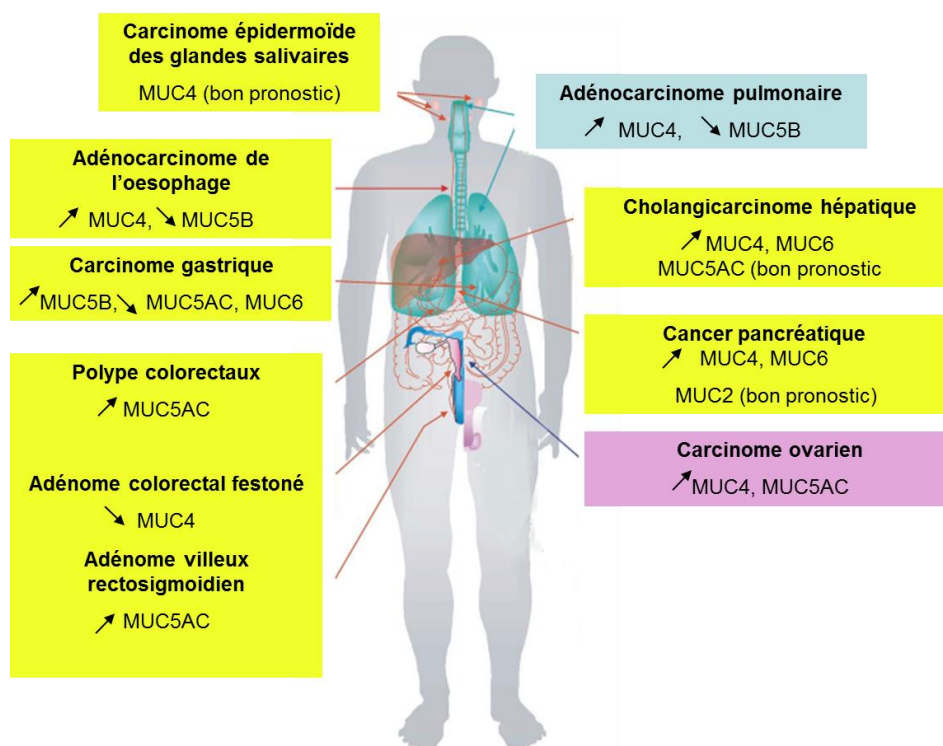


Figure 13 : Modifications de l'expression des mucines dans différents cancers (d'après A Vincent, thèse d'Université).

L'expression de MUC1 est également modifiée dans différents cancers. Ainsi, la surexpression de MUC1, souvent associée à un comportement tumoral agressif et à un mauvais pronostic, est présente dans plus de 90 % des cas de cancer du sein. Elle est

également retrouvée fréquemment dans d'autres cancers comme ceux des ovaires, du côlon, du poumon... (Hattrup and Gendler, 2008).

La cancérogenèse pancréatique a la particularité de s'accompagner d'une modification du profil d'expression des mucines (tableau VII). Les mucines MUC4, MUC16 et MUC5AC, qui sont absentes du tissu pancréatique sain, sont néo-exprimées dès les stades PanIN-1 et leurs niveaux d'expression augmentent avec le stade de la néoplasie pancréatique intraépithéliale (Kaur *et al.*, 2013). Les mucines MUC1 et MUC17, exprimées dans le pancréas sain, sont quant à elles surexprimées au cours de la progression tumorale. La mucine MUC2 reste absente du tissu tumoral pancréatique. Higashi *et al.* (2015) ont montré que l'étude du profil d'expression de MUC1, MUC2 et MUC4 sur le liquide de ponction pancréatique constituerait une stratégie efficace pour la détection précoce des lésions pancréatiques. La détermination du niveau d'expression de MUC5AC, MUC6 et MUC16 constituerait quant à elle un facteur prédictif de la survie des patients atteints de ce cancer.

Tableau VII : Modifications du profil d'expression des mucines observées dans différentes pathologies pancréatiques (Kaur *et al.*, 2013). Abbreviations: –, negative; +, low; ++, moderate; +++, intense; ±, negative or positive depending on study; IPMN, intraductal papillary mucinous neoplasms; MCN, mucinous cystic neoplasm; ND, not determined; PanIN, pancreatic intraepithelial neoplasia; PDAC, pancreatic ductal adenocarcinoma.

Mucin	Fetal pancreas	Adult pancreas	Pancreatitis	PanINs I/II/III	PDAC	Metastatic PDAC	MCN	IPMN
<i>Gel-forming mucins</i>								
MUC2	–	–	–	–	–	ND	++ (goblet cells)	++ (goblet cells)
MUC5B	±	±	++	ND	+++	ND	ND	++
MUC5AC	–	–	±	+ / ++ / +++	+++	+++	+++	++
MUC6	+	+ (interlobular and intralobular ducts; centroacinar cells)	++ to +++	++ / ++ / ++	+++	±	ND	++
<i>Membrane-bound mucins</i>								
MUC1	±	+ (ductal cells; centroacinar cells)	++	++ / ++ / +++	+++	+++	+++ (invasive cases)	+++
MUC3	±	± (interlobular and intralobular ducts; centroacinar cells)	±	+ / + / +	++	ND	ND	±
MUC4	–	–	–	+ / ++ / +++	+++	+++	+++	++
MUC7	–	–	+	ND	+	+++	ND	–
MUC13	ND	± (focal staining)	ND	ND	+++	ND	ND	ND
MUC16	ND	–	±	+ / ++ / ++	+++	+++	ND	ND
MUC17	–	++ (islets)	–	+++ / +++ / +++	++	ND	ND	ND
MUC20	+	+	ND	ND	ND	ND	ND	ND
MUC21	ND	+	ND	ND	ND	ND	ND	ND

II.4.3. Rôles des mucines dans la cancérogenèse

Les mucines participent à la cancérogenèse. Ainsi, Schroeder *et al.* (2004) ont mis en évidence le fait que des souris surexprimant MUC1 développent spontanément une tumeur mammaire, lui attribuant ainsi un rôle d'oncogène. Les mucines sont en effet mises à profit par les cellules tumorales pour contrôler le microenvironnement en modifiant d'une part leurs interactions avec les cellules et la matrice environnante et d'autre part en modulant les facteurs de croissance et cytokines présents dans cet environnement. Les mucines influencent également la croissance cellulaire, l'adhérence cellulaire et l'invasion ainsi que la surveillance immunitaire (figure 14) (Hollingsworth and Swanson, 2004).

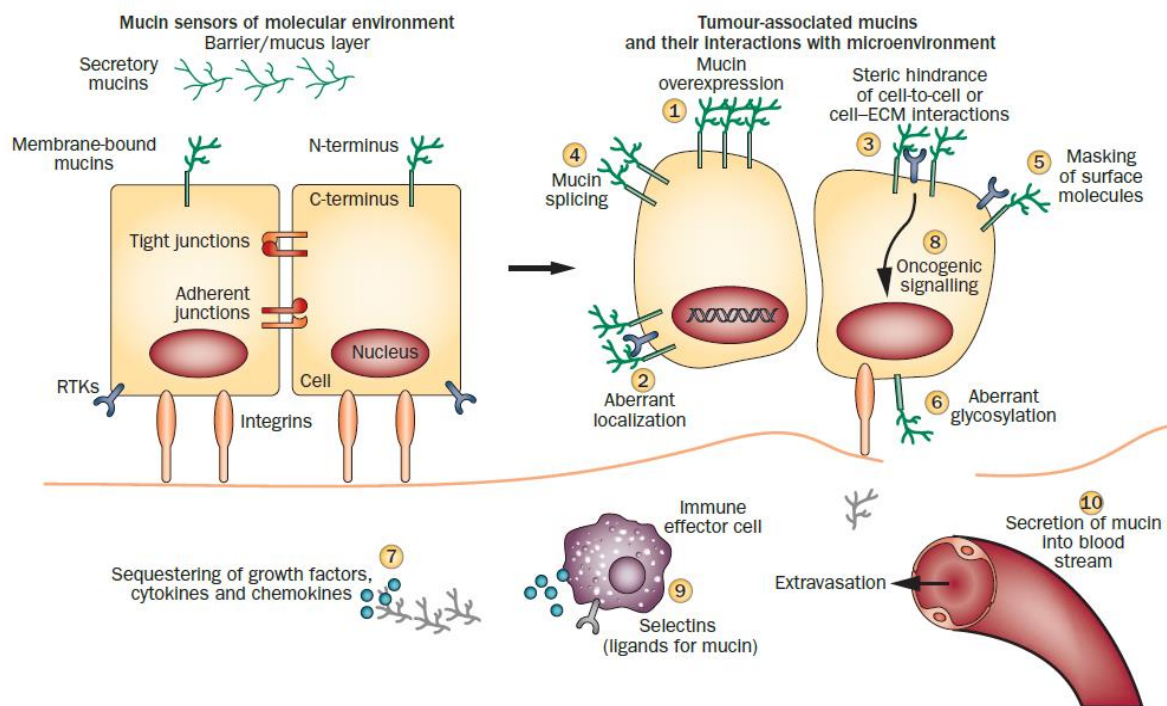


Figure 14 : Rôle des mucines dans la cancérogenèse (Kaur *et al.*, 2013)

Dans les cellules cancéreuses, les mucines, *via* notamment les anomalies du trafic intracellulaire, leur surexpression, leurs anomalies de glycosylation, et enfin leur perte de localisation apicale, interviennent dans la cancérogenèse en masquant par exemple des molécules de surface, en séquestrant des facteurs de croissance, en perturbant les interactions cellules-cellules ou cellules-matrice extracellulaire, en modulant la signalisation cellulaire et la réponse immunitaire.

a. Propriétés d'adhérence et d'anti-adhérence des mucines

Les mucines possèdent à la fois des propriétés pro- et anti- adhérence, à l'origine, dans un premier temps, du détachement des cellules cancéreuses de la tumeur primitive et permettant ensuite leur interaction avec un site métastatique distant.

Les charges négatives portées par les chaînes glycaniques des mucines créent une barrière répulsive autour de la cellule tumorale. De plus, l'encombrement stérique généré par les mucines perturbe les interactions cellules-cellules et cellules-matrice extracellulaire en masquant des molécules d'adhérence présentes à la surface des cellules tumorales, favorisant ainsi leur libération dans la circulation. Par ailleurs, d'autres mécanismes entrent en jeu. Par exemple, les propriétés anti-adhérence de MUC1 sont médiées par son interaction avec la β -caténine *via* un motif présent dans sa queue intra-cytoplasmique (figure 15). Dans la cellule, la β -caténine interagit avec la E-cadhérine au niveau des jonctions d'adhérence qui permettent aux cellules d'être jointives. La surexpression de MUC1 est associée à une augmentation de sa liaison à la β -caténine, ce qui perturbe les interactions β -caténine/E-cadhérine au niveau de ces jonctions. Par conséquent, ce phénomène interrompt les contacts entre cellules facilitant la dissémination des cellules tumorales (Kufe, 2009).

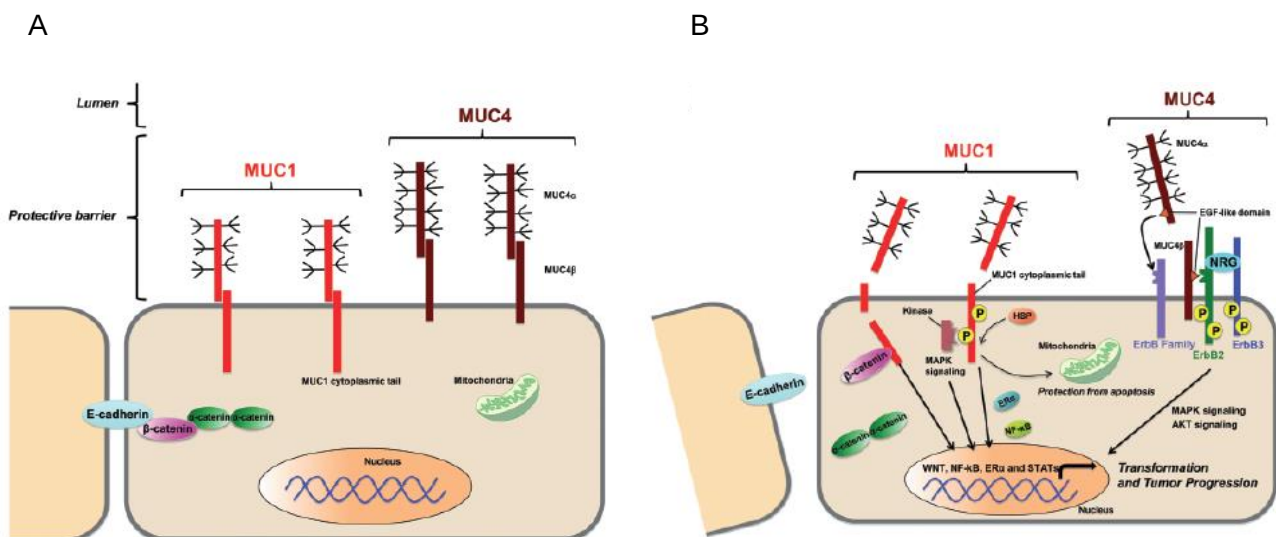


Figure 15 : MUC1 et jonctions d'adhérence (Yonezawa *et al.*, 2011).

(A) Situation normale : la β -caténine interagit avec la E-cadhérine et permet le maintien des jonctions d'adhérence (B) Mécanisme impliquant la surexpression de MUC1 dans la perte des jonctions d'adhérence via sa liaison à la β -caténine. Dans les cellules cancéreuses, l'interaction MUC1- β -caténine empêche l'interaction β -caténine-E-cadhérine et favorise ainsi la dissociation des cellules.

Les mucines possèdent également des propriétés adhésives. Ainsi, grâce à la structure modifiée de ses oligosaccharides, via le motif sLe^x, MUC1 peut se lier à des molécules d'adhérence telles qu'ICAM1 et la E-sélectine. Ce mécanisme facilite l'attachement des cellules cancéreuses à un site métastatique distant du site tumoral primitif (Andrianifahanana *et al.*, 2006; Kufe, 2009). De même, MUC16, par le biais de l'interaction de ses motifs N-glycaniques avec la mésothéline, participe au processus métastatique observé dans le cancer de l'ovaire (Das and Batra, 2015).

Une autre protéine, la galectine-3, contribue aux propriétés adhésives des mucines. En effet, étant présente dans le milieu extracellulaire à la surface des cellules endothéliales, elle permet, dans un premier temps, l'attachement réversible, des cellules cancéreuses exprimant MUC4 ou MUC1 par l'intermédiaire de sa fixation sur les motifs glycaniques présents sur ces mucines. Cette interaction entraîne une redistribution de MUC4 ou de MUC1 à la surface des cellules cancéreuses (phénomène de *capping*), permettant ainsi l'exposition de molécules d'adhérence normalement masquées par la grande taille des mucines (par exemple les intégrines). Ce mécanisme entraîne un attachement irréversible des cellules tumorales aux cellules endothéliales favorisant le phénomène métastatique (figure 16) (Senapati *et al.*, 2011; Zhao *et al.*, 2009). Il a également été démontré que MUC16 interagit avec la galectine-3 et la galectine-1 afin de favoriser la progression tumorale (Haridas *et al.*, 2014)

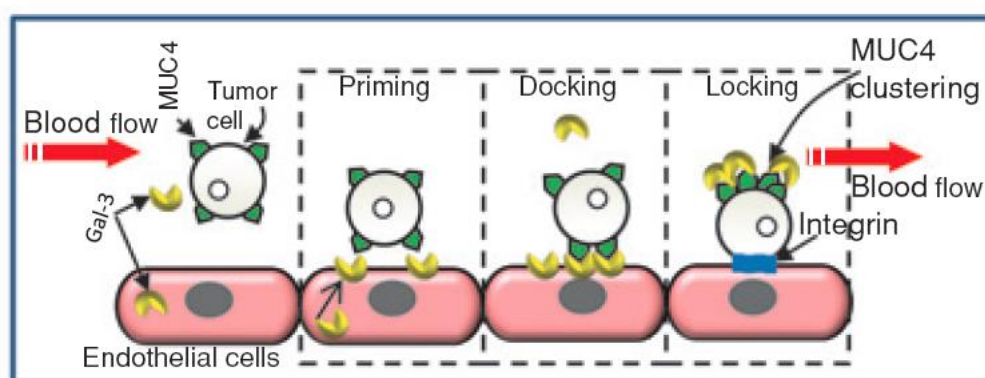


Figure 16 : Illustration de l'interaction entre MUC4 et la galectine-3 favorisant le phénomène métastatique (Senapati *et al.*, 2011).

Cette interaction favorise la redistribution de MUC4 à la surface cellulaire et l'exposition de molécules d'adhérence habituellement masquées par MUC4. Ceci permet l'attachement des cellules tumorales aux cellules endothéliales et contribue donc au processus métastatique.

b. Modulation de la réponse immunitaire

Les mucines perturbent la réponse immunitaire dirigée contre les cellules cancéreuses par différents mécanismes. Ainsi, la production importante de mucines solubles par les cellules tumorales et l'encombrement stérique généré par les mucines membranaires masquent les antigènes de surface cellulaire et empêchent l'accès des leucocytes, des cellules présentatrices d'antigènes et des cellules effectrices du système immunitaire aux cellules tumorales. De plus, la glycosylation différentielle des mucines modifie le profil des cytokines et des facteurs de croissance séquestrés dans la couche de mucus au profit de ceux favorisant la suppression de la réponse immunitaire (Hollingsworth and Swanson, 2004). Enfin, l'expression de MUC16 par les cellules tumorales perturbe la formation de la synapse immunologique entre la cellule cible et les cellules NK indispensable à la réponse cytotoxique (Haridas *et al.*, 2014).

c. Signalisation cellulaire

Les mucines membranaires participent au contrôle de la différenciation, de la prolifération et de la survie des cellules tumorales par le biais d'interactions récepteurs-ligands et de l'activation des voies de signalisation correspondantes. Les rôles des mucines membranaires MUC1 et MUC4 dans la signalisation cellulaire sont les plus documentés dans la littérature. Nous prendrons ici l'exemple de la mucine MUC1 pour illustrer ces fonctions. Cette mucine agit comme un oncogène par l'activation de voies de signalisation impliquées dans l'initiation et la progression tumorale. La queue intra-cytoplasmique de MUC1 compte sept tyrosines phosphorylables par des protéines kinases et des sites d'interaction pour des protéines à domaine de type SH2 (*Src Homology 2*). Ainsi, elle interagit avec les protéines *cellular-Src protein kinase* (c-Src), *Glycogen Synthetase Kinase 3 β* (GSK3 β), *Protein Kinase C δ* (PKC δ), *Tyrosine-protein kinase Lyn*, la β -caténine et les membres des récepteurs de la famille ErbB (figure 17). Ces interactions sont à l'origine de l'activation de différentes voies de signalisation comme les voies MAPkinases et PI3K (*Phosphatidylinositol 3-Kinase*) (Singh and Hollingsworth, 2006), ce qui explique le rôle de MUC1 dans la tumorigenèse. De plus, la queue intra-cytoplasmique de MUC1 peut subir un clivage, être transloquée dans le noyau pour former des complexes avec certains facteurs de transcription (*Estrogen receptor*

α , β -caténine...). Ceci a pour conséquence de moduler la transcription de leurs gènes cibles comme celui codant la cycline D1, influençant par exemple la prolifération cellulaire (Albrecht and Carraway, 2011). Par ailleurs, Merlin et al ont démontré que la galectine-3 module la voie de signalisation dépendante de MUC1 et de l'EGFR en influençant l'endocytose et la localisation subcellulaire de ces protéines dans les cellules cancéreuses pancréatiques (Merlin *et al.*, 2011).

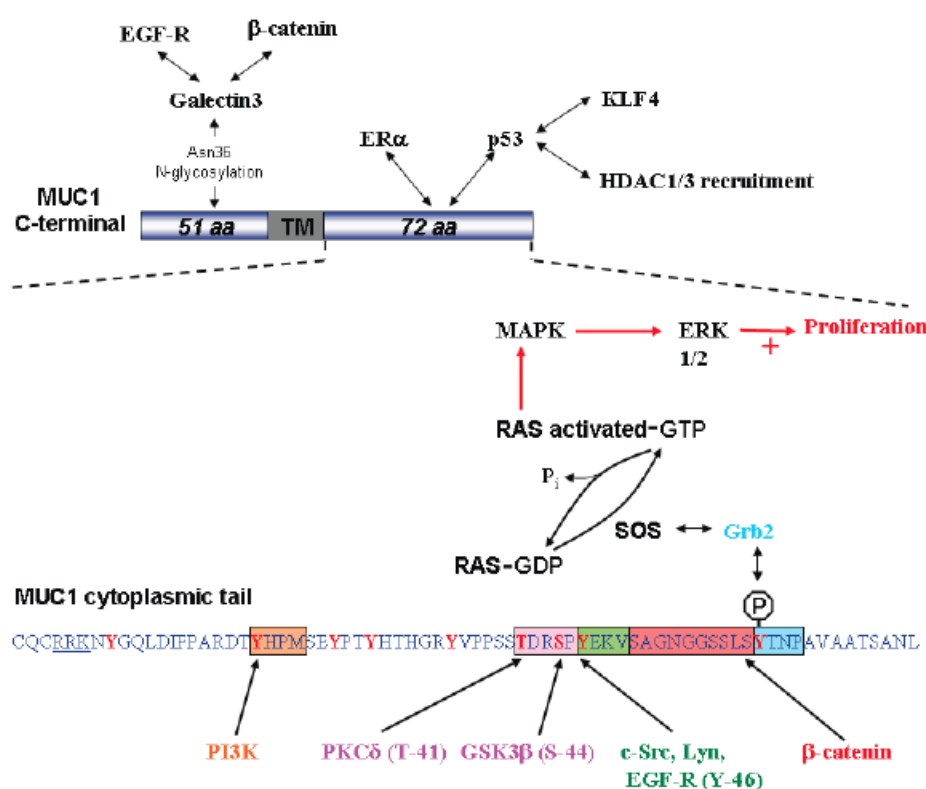


Figure 17 : Ligands de MUC1, fonctions de sa queue cytoplasmique et localisation des tyrosines phosphorylables (Y en rouge) au niveau de sa queue cytoplasmique (Jonckheere and Van Seuning, 2008).

d. Chimiorésistance

Enfin, les mucines contribuent au phénomène de chimiorésistance observé dans les cancers. Par exemple, MUC1 participe à ce processus par le biais de son action anti-apoptotique *via* la voie PI3K/Akt, Bcl-xl (*B-cell lymphoma-extra large*) et par le transfert de sa queue intra-cytoplasmique au niveau de la membrane mitochondriale externe ce qui atténue la libération de facteurs pro-apoptotiques (Albrecht and Carraway, 2011). MUC4 est

également impliquée dans ce processus. Ainsi, l'encombrement stérique créé par cette mucine est à l'origine de la résistance des cellules cancéreuses aux cellules immunitaires et aux agents thérapeutiques tels que les anticorps monoclonaux (comme le trastuzumab) qui n'ont plus accès à leurs molécules cibles (Price-Schiavi *et al.*, 2002). Elle participe aussi à la résistance des cellules tumorales à l'apoptose notamment en diminuant la libération du cytochrome c mitochondrial. Une étude a également démontré que l'expression du transporteur hCNT1 (*human concentrative nucleoside transporter 1*), permettant l'entrée de la gemcitabine dans les cellules cancéreuses, est diminuée dans les cellules exprimant MUC4 altérant ainsi la sensibilité de la tumeur à cet agent anti-cancéreux (Skrypek *et al.*, 2015). La figure 18 illustre l'implication de MUC4 dans la chimiorésistance à de multiples niveaux.

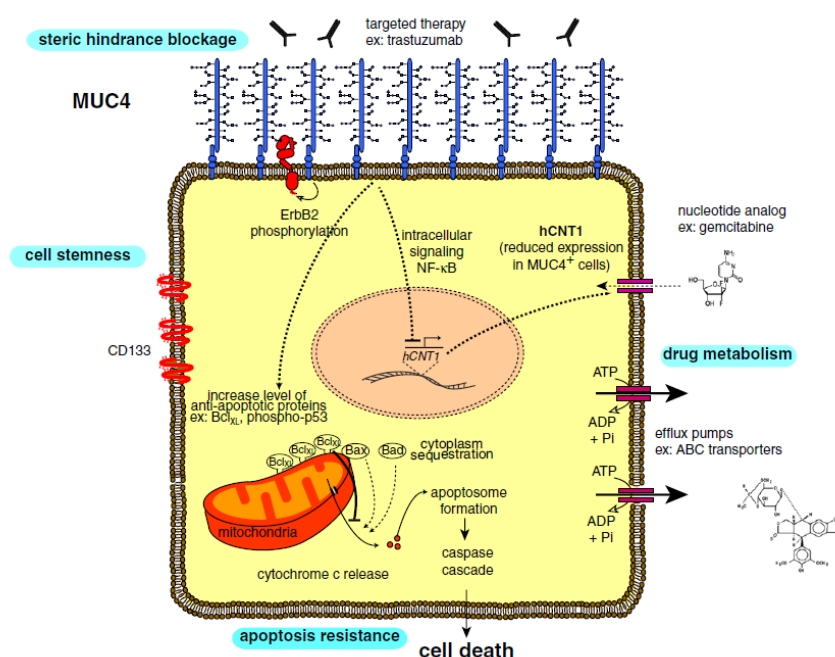


Figure 18 : Rôles de MUC4 dans la chimiorésistance des cancers épithéliaux.

Celle-ci intervient par exemple via son encombrement stérique, elle entraîne une résistance à l'apoptose, et modifie le métabolisme des agents cytotoxiques (Jonckheere *et al.*, 2014).

Tous ces éléments bibliographiques généraux sur les mucines nous seront très utiles pour discuter nos résultats, et dans la suite de cette introduction, nous allons nous intéresser plus en détail aux caractéristiques structurales et fonctionnelles des mucines membranaires MUC4, MUC1 et MUC16 qui ont été le point de départ de notre travail.

III. Les mucines MUC4, MUC1 et MUC16

III.1. MUC4

III.1.1. Structure

La mucine membranaire MUC4 est synthétisée sous forme d'un précurseur de 550 à 930 kiloDalton (kDa), en fonction de la taille du domaine central répété en tandem. Son domaine extra-cellulaire s'étend sur 2,12 μm au-delà de la surface cellulaire. La partie N-terminale de la protéine est formée d'un peptide signal de 27 aa suivi de trois motifs de répétitions imparfaites variant de 126 à 130 aa et d'une séquence unique de 554 aa. Le domaine central comprend un nombre variable de répétitions parfaites en tandem d'un motif de 16 aa (145 à 395 répétitions) générant un polymorphisme interindividuel de longueur de type VNTR. La partie C-terminale est constituée de 12 domaines distincts (CT1 à CT12) parmi lesquels se trouvent un domaine NIDO homologue du domaine angiogène de la protéine nidogène, un domaine AMOP, deux domaines riches en résidus cystéines, trois domaines EGF-like, un domaine transmembranaire et enfin un domaine intra-cytoplasmique. Un site de clivage protéolytique « GDPH » est situé entre les domaines CT4 et -5 au sein de la séquence glycine-acide aspartique-proline-histidine (figure 19) (Chaturvedi *et al.*, 2008). En effet, le précurseur de MUC4 est clivé en deux sous-unités associées de façon non covalente après sa sortie du réticulum endoplasmique: MUC4 α (sous-unité *mucin like* extra-cellulaire) et MUC4 β (sous-unité transmembranaire *growth factor like*) (Moniaux *et al.*, 1999).

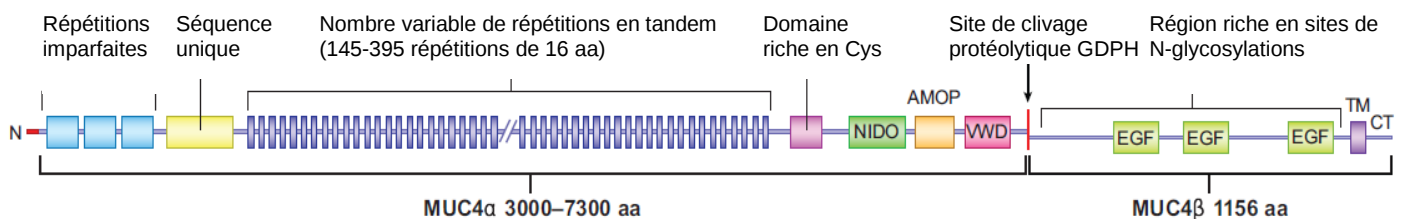


Figure 19 : Représentation schématique de la mucine MUC4 sous forme linéaire (Hattrup and Gendler, 2008).

MUC4 est l'homologue du complexe sialomucine membranaire du rat appelé *Sialomucin Complex* (SMC). Il s'agit d'un complexe protéique hétérodimérique composé d'une sous-unité ASGP1 (*Ascites Sialoglycoprotein*) et d'une sous-unité membranaire *growth factor like* ASGP2. Carraway *et al.* (1999) ont démontré l'interaction entre le domaine EGF1 de la sous-unité membranaire ASGP2 et le récepteur ErbB2. L'interaction entre MUC4 et ErbB2 a également été démontrée chez l'homme (Jonckheere *et al.*, 2012).

III.1.2. Gène et ARNm

Le gène *MUC4* a été cloné à partir d'une banque d'ADNc de muqueuse trachéobronchique humaine (Porchet *et al.*, 1991). Il est situé sur le chromosome 3, dans la région q29, s'étend sur plus de 60 kilobases (kb) et possède 26 exons (figure 20). Le premier exon code la région 5'UTR (*Untranslated Region*) ainsi que le peptide signal. Le deuxième exon qui est le plus grand du transcrit (environ 12 kb) contient une séquence unique suivie de la répétition d'un motif de 48 paires de bases (pb). Les domaines extracellulaire, transmembranaire et cytoplasmique sont codés par les 24 exons restants (Moniaux *et al.*, 1999).

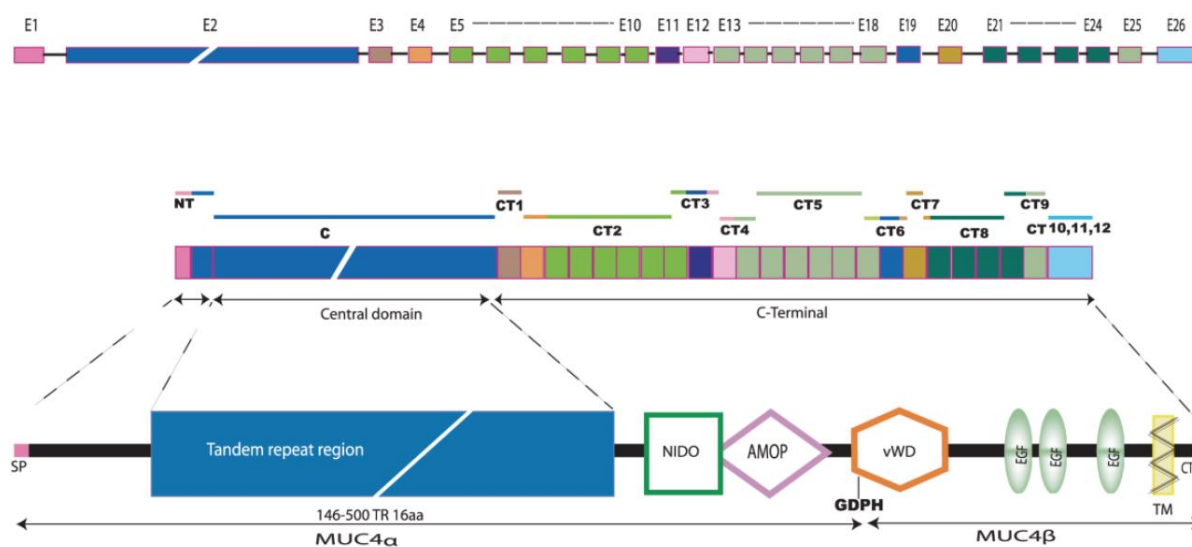


Figure 20 : MUC4 : correspondance entre la structure du gène, de l'ARNm et de la protéine (Chaturvedi *et al.*, 2008) .

La région *tandem repeat* de MUC4 est codée par l'exon 2.

L'épissage alternatif du pré-ARNm de *MUC4* génère 24 isoformes, dénommées sv0 à sv21-MUC4, MUC4/X et MUC4/Y (figure 21), dont les plus longs ont une taille comprise entre 16 et 24 kb (forme pleine longueur appelée sv0). Trois catégories de variants sont répertoriées selon la nature de l'isoforme protéique encodée, la majorité des évènements d'épissage se produisant en aval de la région TR :

- les formes sécrétées (sv1 à-7, sv9, sv11 à -20) ;
- les variants membranaires (sv0, sv8, sv10 et sv21) ;
- les formes dépourvues du domaine central de type mucine (MUC4/X et MUC4/Y).

Ces variants sont le plus souvent générés par l'utilisation d'exons alternatifs, de sites cryptiques d'épissage de type donneur ou de type accepteur (Escande *et al.*, 2002).

Le profil d'expression de ces différents variants d'épissage est dépendant des types cellulaires, des organes et de l'état physiopathologique (Escande *et al.*, 2002). En effet, Choudhury *et al.* (2000) ont montré une modification de l'expression des variants de *MUC4* dans les cellules cancéreuses pancréatiques. Leur étude indique que la forme membranaire de type sv10-MUC4 est observée dans 37 % des cas d'adénocarcinome pancréatique canalaire alors qu'elle est absente du tissu pancréatique sain. De même, Zhu *et al.* (2014) ont démontré que l'expression du variant MUC4/Y, dépourvu de la région *tandem repeat*, contribue à la progression du cancer pancréatique et que son expression est corrélée au pouvoir invasif de la tumeur ainsi qu'à la présence de métastases (Xie *et al.*, 2014; Zhu *et al.*, 2014).

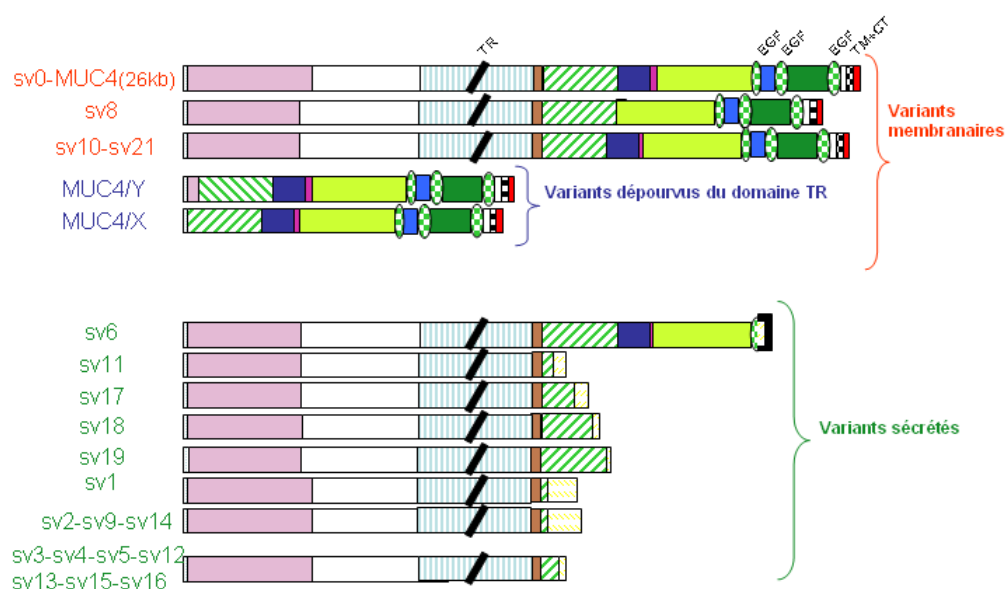


Figure 21 : Les variants d'épissage de *MUC4* (adapté de Escande *et al.*, 2002).

Il en existe trois types : sécrétés, membranaires (formes longues de type « mucine ») et dépourvus de la région *tandem repeat* (formes courtes dépourvues de la région caractéristique des mucines).

III.1.3. Caractérisation de la région 5'-flanquante de *MUC4* et régulation de son expression

La région 5' flanquante de *MUC4*, qui s'étend sur 4 kb en amont du premier codon ATG, contient deux régions actives appelées promoteurs proximal et distal (figure 22) (Perrais *et al.*, 2001).

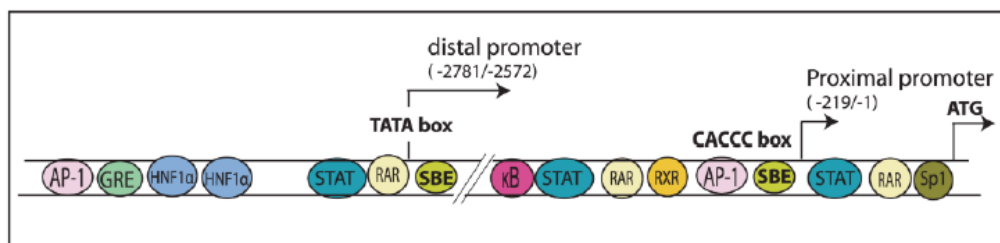


Figure 22 : Schématisation des éléments *cis* contenus dans le promoteur de *MUC4* (Chaturvedi *et al.*, 2008).

La région 5' flanquante de *MUC4* possède quatre sites d'initiation de la transcription : un dans le promoteur proximal et trois dans le promoteur distal. Le promoteur proximal de 2,7 kb environ ne contient pas de boîte TATA et se situe en aval de la première boîte TATA active (-2672/-2668). Ces deux régions promotrices sont riches en résidus GC et contiennent des sites de liaison pour de nombreux facteurs de transcription tels que AP-1 (*Activator Protein-1*), Sp1 (*Specific Protein 1*), PEA3 (*Polyomavirus Enhancer Activator 3*)... (Perrais *et al.*, 2001).

L'expression de *MUC4* est finement régulée et fait appel à différents mécanismes.

Au niveau épigénétique, l'hyperméthylation des îlots CpG situés dans la région promotrice de *MUC4* par les enzymes DNMT3A/3B (*DNA methyltransferase*) et la désacétylation des histones médiée par HDAC3 (*histone désacétylase*) répriment son expression dans le pancréas sain (Jonckheere *et al.*, 2015).

De plus, au niveau transcriptionnel différents facteurs peuvent agir afin de moduler son expression (Andrianifahanana *et al.*, 2006; Carraway *et al.*, 2009). En voici quelques exemples :

- Les facteurs de croissance comme l'EGF, le TGF α (*Transforming Growth Factor α*) induisent l'expression de *MUC4* par le biais de l'activation de tyrosines kinases. Le facteur de croissance TGF β intervient par l'intermédiaire de l'activation des facteurs Smad2 et Smad4 (*Mothers against decapentaplegic homolog*) ;

- L'interféron γ (IFN γ), cytokine pro-inflammatoire, régule l'expression de *MUC4* via STAT1 (*Signal transducer and activator of transcription 1*) ;
- Les acides biliaires, *via* la voie PI3K ou par l'intermédiaire de la liaison du facteur de transcription HNF-1 (*Hepatocyte Nuclear Factor 1*) sur le promoteur de *MUC4*, induisent également l'expression de *MUC4* ;
- Les interleukines (IL), telles que l'IL-9 et l'IL-4, régulent l'expression de *MUC4* respectivement *via* JAK3 (*Janus Kinase 3*) et l'activation du facteur de transcription STAT6 ;
- Le canal chlore CFTR (*Cystic fibrosis transmembrane conductance regulator*) est un régulateur négatif de l'expression de *MUC4*.

Les données concernant la régulation de l'expression de *MUC4* au niveau post-transcriptionnel sont peu nombreuses dans la littérature :

Fischer *et al.* (2003) ont mis en évidence une stabilisation des transcrits de *MUC4* par l'élastase produite par les polynucléaires neutrophiles lors du processus inflammatoire.

L'inhibition de l'expression de cette mucine peut aussi s'exercer *via* la fixation des miR-150 (Srivastava *et al.*, 2011), miR-219-1-3p (Lahdaoui *et al.*, 2015) et -211 (Xu *et al.*, 2016) dans la région 3'UTR de *MUC4* et du miR-200c au sein de l'exon 1 (Radhakrishnan *et al.*, 2013).

Enfin, des études ont également mis en évidence une régulation de l'expression de *MUC4* au niveau post-traductionnel. Ainsi, le CFTR, en plus de son action au niveau transcriptionnel, intervient *via* le contrôle de la dégradation de la protéine, la demi-vie de la protéine *MUC4* étant raccourcie lorsque CFTR est exprimé (figure 23) (Singh *et al.*, 2007).

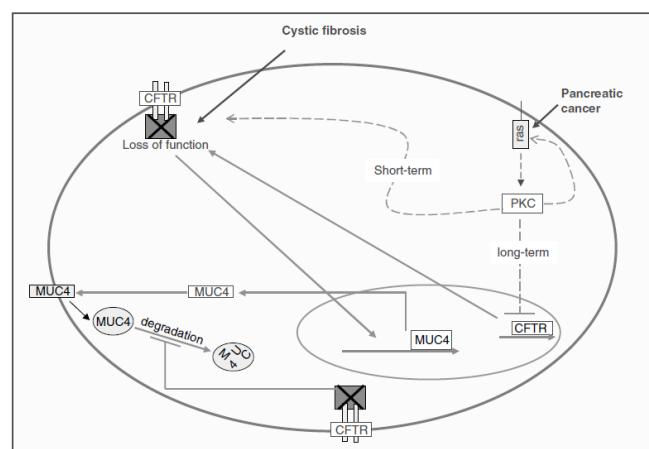


Figure 23 : Régulation de l'expression de *MUC4* par le CFTR au niveau transcriptionnel (régulation négative de l'expression de *MUC4*) et post-traductionnel (en favorisant la dégradation de *MUC4*) (Singh *et al.*, 2007).

III.2. MUC1

III.2.1. Structure

La mucine MUC1 est une *O*-glycoprotéine (250 à 500 kDa) qui était connue avant le clonage de son gène sous de nombreuses appellations parmi lesquelles : *Polymorphic Epithelial Mucin* (PEM), épisialine, Dupan-2, *Human Milk Fat Globule* (HMFG), *Epithelial Membran Antigen* (EMA), DF3 et CD227 (Apostolopoulos *et al.*, 2015).

L'apomucine MUC1 ne possède pas de domaines de polymérisation et est dépourvue de domaines de type EGF. Synthétisée sous forme d'un précurseur unique, elle est clivée dans le réticulum endoplasmique au niveau du motif glycine-sérine-valine-valine-valine (GSVVV) localisé entre son domaine SEA et son domaine transmembranaire. Les deux sous-unités formées sont associées de façon non covalente. La sous-unité extracellulaire de type mucine est constituée de la répétition en tandem d'un motif de 20 aa sur lequel viennent se lier de nombreux résidus glycaniques. La sous-unité transmembranaire est formée d'un domaine extracellulaire de 58 aa suivi d'une région transmembranaire de 28 aa et d'une queue intracytoplasmique de 72 aa. MUC1 possède des sites de N-glycosylation notamment au niveau du domaine SEA qui est situé après le domaine répété en tandem (figure 24) (Hatstrup and Gendler, 2008; Nath and Mukherjee, 2014)

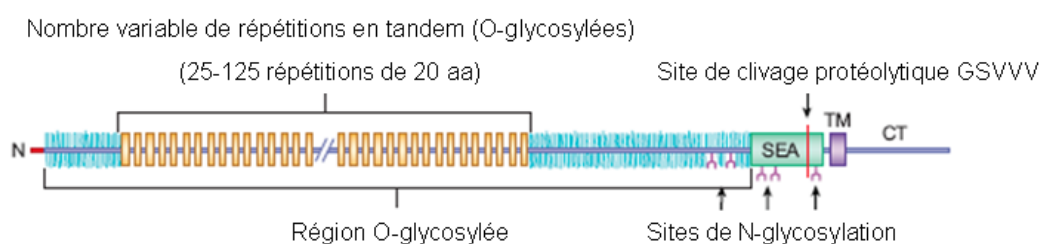


Figure 24 : Représentation de la structure de MUC1 sous forme linéaire (Hatstrup and Gendler, 2008).

Le motif *tandem repeat* de MUC1 est constitué d'un motif de 20 aa répété 25 à 125 fois.

III.2.2. Gène, ARNm et régulation de l'expression de MUC1

Le gène codant la mucine MUC1, situé sur le chromosome 1 en position q21-q24, compte sept exons et sa taille varie de 4 à 7 kb en fonction du nombre de répétitions en

tandem. L'exon 1 code le peptide signal, l'exon 2, le plus grand du transcrit (jusqu'à 5 kb) le domaine central répété en tandem et les exons 6 et 7 codent respectivement le domaine transmembranaire et la queue intra-cytoplasmique (Jonckheere and Van Seuning, 2008).

La régulation de l'expression de *MUC1* au niveau transcriptionnel a surtout été étudiée dans le cadre du cancer du sein. Bien que cette mucine soit surexprimée dans l'adénocarcinome pancréatique canalaire, aucun mécanisme de régulation transcriptionnelle spécifique du pancréas n'a été décrit jusqu'à présent. Au niveau épigénétique, son expression est modulée par l'intermédiaire de la méthylation ou l'acétylation des histones, de la méthylation de l'ADN. Concernant le promoteur de *MUC1*, celui-ci contient une boîte TATA située 25 nucléotides en amont du premier codon ATG et est riche en résidus GC. La région 5' flanquante, de 2,9 kb environ, possède des sites de liaison pour de nombreux facteurs de transcription comme Sp1, AP-1, AP-2, AP-3, ER (*Estrogen Receptor*), STAT3, GATA3 (*GATA Binding Protein 3*)... Ainsi, l'IFN γ et l'IL-6 (via les facteurs de transcription STAT1 et -3) et le TNF α (*Tumor necrosis factor α*), par le biais du facteur de transcription *Nuclear Factor Kappa B* (NF κ B), induisent l'expression de *MUC1* (figure 25) (Jonckheere and Van Seuning, 2010).

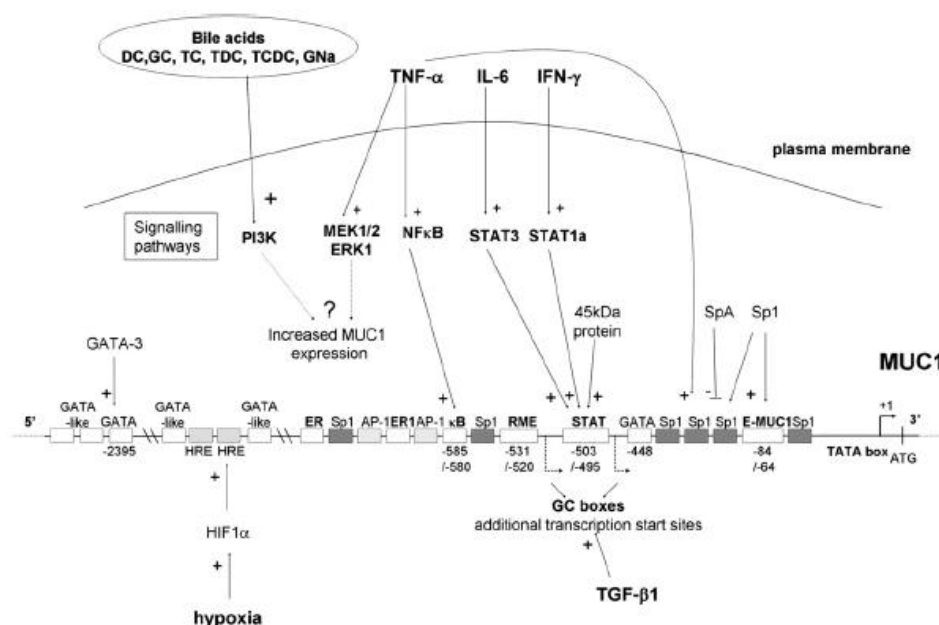


Figure 25 : Représentation schématique du promoteur de *MUC1* et facteurs intervenant dans sa régulation (Jonckheere and Van Seuning, 2010).

Les facteurs de transcription STAT1/3, NF κ B, Sp1... régulent l'expression de cette mucine via leur fixation sur leurs éléments *cis* cibles présents dans le promoteur de *MUC1*.

Peu de données dans la littérature concernent une régulation de *MUC1* au niveau post-transcriptionnel. Son expression est modulée par des micro-ARN tels que le miR-145 (Sachdeva and Mo, 2010), le miR-125b (Rajabi *et al.*, 2010), le miR-1226 (Jin *et al.*, 2010), le miR-29a et le miR-330-5p (Tréhoux *et al.*, 2015).

Environ 78 transcrits de *MUC1* issus d'un épissage alternatif ont été décrits (Nath and Mukherjee, 2014), les plus fréquents étant MUC1/A, /B, /C, /D et MUC1/Y (figure 26). Les variants de pleines longueurs MUC1/A, /B, /C et /D, qui diffèrent entre eux par la taille de l'exon 2, résultent d'un épissage alternatif au niveau de sites situés dans l'intron 1 et l'exon 2. Les formes MUC1/X, /Y et /ZD sont caractérisées par l'absence de l'exon 2. Le variant MUC1/SEC est, quant à lui, sécrété dans le milieu extracellulaire suite à un épissage alternatif conduisant à la perte du domaine transmembranaire et est capable de se lier à la forme MUC1/Y entraînant ainsi sa phosphorylation. Cependant, les fonctions de ces différents variants de MUC1 ne sont pas entièrement élucidées (Hanisch and Müller, 2000; Nath and Mukherjee, 2014).

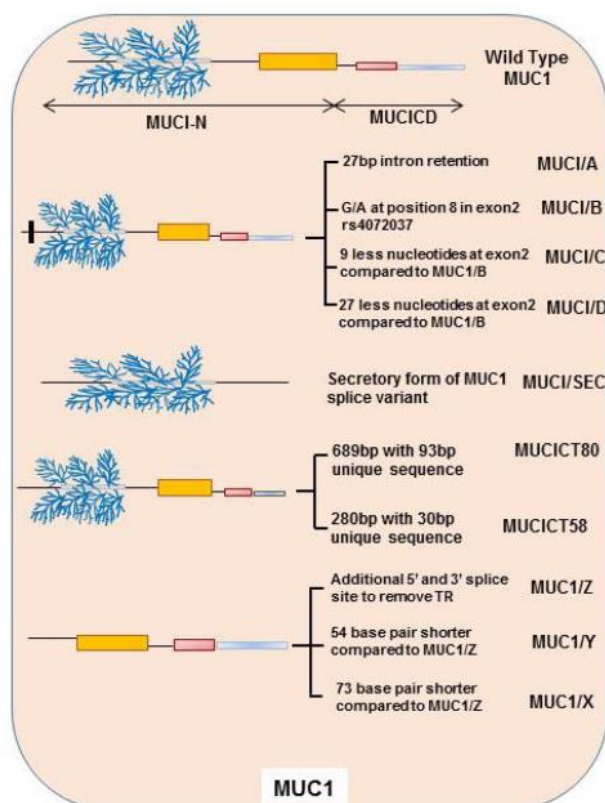


Figure 26 : Les différents variants d'épissage de MUC1 : sécrétés, membranaires, dépourvus de la région *tandem repeat* (Kumar *et al.*, 2016).

III.3. MUC16

III.3.1. Structure protéique

La mucine MUC16 a été initialement identifiée comme l'antigène carbohydre 125 (CA125), un antigène surexprimé dans le cancer de l'ovaire. Il s'agit de la mucine humaine la plus volumineuse, cependant ses fonctions et sa structure restent à ce jour peu étudiées.

Cette protéine de 22 152 aa a une masse moléculaire de 2,5 MDa et celle-ci atteint 20 MDa lorsqu'elle est glycosylée. Tout comme les autres mucines membranaires, elle est constituée d'une région N-terminale, d'un grand domaine TR interrompu par des domaines SEA (56 au total) et d'une région C-Terminale comprenant une partie transmembranaire et une queue intra-cytoplasmique de 32 AA possédant plusieurs sites de phosphorylation (figure 27). A la différence de MUC1 et MUC4, cette mucine ne possède pas de domaines EGF. La région TR de MUC16 est formée d'un motif de 156 AA répété jusqu'à 60 fois avec un polymorphisme de type VNTR. Chaque domaine SEA représente une grande partie du motif TR (de l'aa 1 à 128). La région N-terminale de MUC16 subit majoritairement une O-glycosylation et sa région TR un processus de O- et N-glycosylation. Par ailleurs, le schéma de glycosylation de cette mucine peut varier selon le tissu et l'organe où MUC16 est exprimée mais aussi selon les circonstances physio-pathologiques (Haridas *et al.*, 2014).

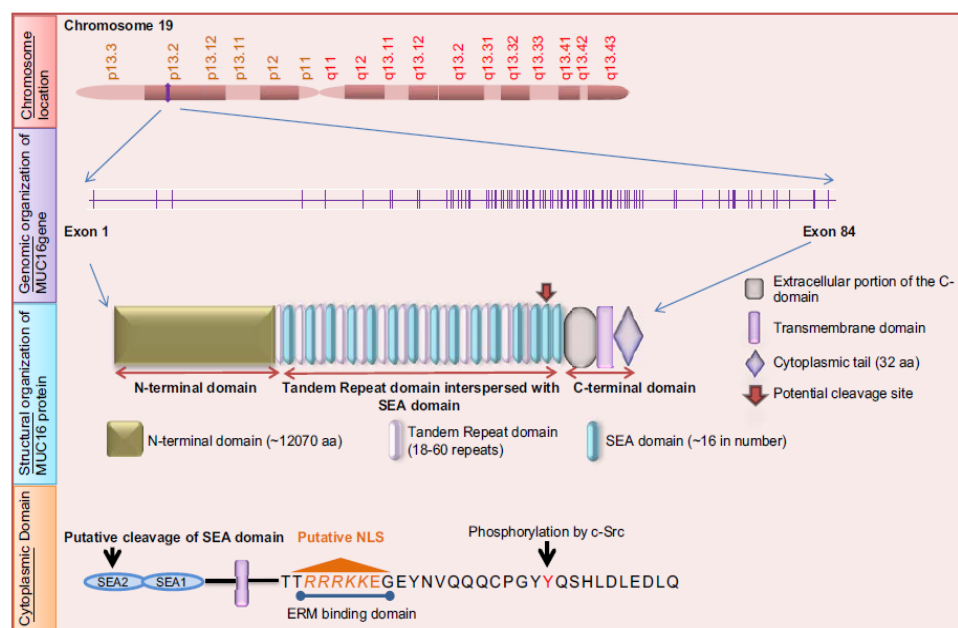


Figure 27 : Représentation schématique de la mucine MUC16 : du gène à la protéine (Haridas *et al.*, 2014).

Le gène codant cette mucine est situé sur le bras court du chromosome 19, en 13.3 et comporte 84 exons. L'exon 3 est le plus volumineux du transcrit (21,5 kb).

Il existe deux sites de clivage de MUC16 qui seraient localisés au niveau des avant-dernier et dernier domaines SEA. Suite à ce clivage, la partie extracellulaire de la protéine est libérée dans la circulation et la détection de cet antigène CA125 représente un marqueur intéressant de suivi de la réponse au traitement et de la récurrence des cancers de l’ovaire (figure 28) (Das and Batra, 2015).

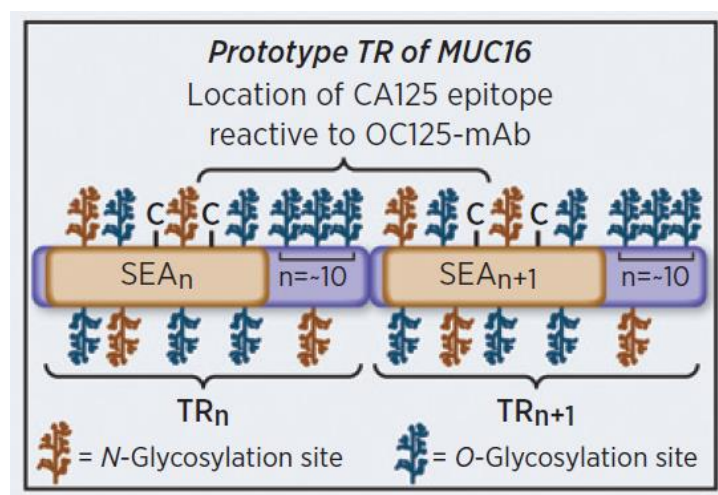


Figure 28 : Représentation schématique du TR de MUC16, le motif CA125 étant situé entre les deux résidus cystéines des motif TR adjacents (Das and Batra, 2015).

III.3.2. Régulation de l'expression de MUC16

A ce jour, peu d'études dans la littérature se sont intéressées à la régulation de l'expression de *MUC16*. Morgado *et al.* (2016) ont mis en évidence que l'expression de *MUC16* est stimulée par les cytokines pro-inflammatoires telles que le TNF α , l'IFN γ par l'intermédiaire de la translocation de NF κ B dans le noyau et l'activation du promoteur de *MUC16*. A l'inverse, PPAR α (*Peroxisome proliferator-activated receptor α*) régulerait négativement l'expression de *MUC16* via l'inhibition de NF κ B (Morgado and Carson, 2017).

Concernant la régulation de l'expression de *MUC16* au niveau post-transcriptionnel, là encore, peu de données sont disponibles dans la littérature. Son expression serait régulée par le miR-200c, tout comme *MUC4* (Radhakrishnan *et al.*, 2013).

III.3.3. Expression de MUC16 en situations physiologiques et pathologiques

MUC16 est habituellement exprimée au niveau des épithéliums bordant le système respiratoire, la cornée, les organes reproducteurs féminins, le péritoine, le péricarde, la cavité abdominale... Cette mucine a également été mise en évidence dans différentes sécrétions comme le lait et le liquide amniotique. Au niveau physiologique, MUC16 joue un rôle de protection au niveau de la surface oculaire en empêchant la fixation des éléments pathogènes (Haridas *et al.*, 2014).

Au niveau pathologique, l'altération de la O-glycosylation et de l'expression de MUC16 contribueraient au développement des maladies auto-immunes telles que le syndrome de Sjogren (Caffery *et al.*, 2008). L'expression de MUC16 est également modifiée dans différents cancers. Ainsi, la mucine MUC16 est surexprimée dans le cancer de l'ovaire et sa détection est utilisée dans son suivi. Tout comme MUC4, MUC16, qui est absente du pancréas sain, est néo-exprimée au cours du cancer pancréatique avec une augmentation progressive de son expression des stades PanIN au stade PDAC (Haridas *et al.*, 2014)

Parmi les partenaires des mucines, nous avons vu que la **galectine-3**, protéine également surexprimée dans le cancer pancréatique, est capable d'interagir avec MUC4, MUC1 et MUC16. La partie suivante de l'introduction sera consacrée à cette famille des galectines, protéines à propriétés de lectines endogènes.

IV. Galectines

IV.1. Généralités

IV.1.1. Classification et structure

Les galectines ont été progressivement caractérisées chez les vertébrés depuis 1976. Elles sont définies par le *Human Genome Nomenclature Committe* (HGCN) comme une famille de lectines endogènes solubles (type S) : « "Galactoside-binding lectins", "S-type lectins", "Lectins, galactoside-binding" » et sont très conservées dans le Règne Animal. Elles doivent leur nom aux 2 caractéristiques qu'elles partagent et qui sont (Barondes *et al.*, 1994):

- Leur capacité à se lier aux β -galactosides de type I ($\text{Gal}\beta 1 \rightarrow 4\text{GlcNAc}$) et de type II ($\text{Gal}\beta 1 \rightarrow 3\text{GlcNAc}$) avec une très forte affinité pour les chaînes poly N-acétyllactosaminiques ;
- La présence dans leur structure d'un « domaine de reconnaissance des carbohydrates » (CRD) de type lectine.

15 gènes codant des galectines (symbole *LGALS*) ont été identifiés chez les mammifères, certains d'entre eux partageant des caractéristiques structurales (nombre et taille des exons par exemple) (tableau VIII). Toutes les galectines de mammifères ont en commun la présence d'au moins un domaine CRD conservé de 130 aa environ, et la particularité de n'être pas elles-mêmes glycosylées (bien que certaines possèdent pourtant des sites potentiels de N-glycosylation (Dumic *et al.*, 2006; Hughes, 1997).

Cette famille est en rapide évolution, comme en témoignent des variations de répertoires entre espèces de mammifères. Par exemple, les galectines des sous-familles -5 et -6 existent chez les rongeurs mais pas chez l'homme, de même les galectines des sous-familles -11, -14, -15 existent chez les ovins mais pas chez l'homme. Par contre, le génome humain compte 3 membres de la sous-famille 9. Des variations de répertoires existent entre espèces très proches, rat et souris, et même entre lignées d'une même espèce, de souris, par exemple. De plus certaines galectines peuvent encore exister sous différentes isoformes d'épissage. (Cooper, 2002; Rabinovich, 1999)

Tableau VIII: Gènes humains de la famille des galectines (symbole *LGALS*) (extrait de *Human Genome Nomenclature Committee*).

Approved Symbol	Approved Name	Previous Symbols	Synonyms	Chromosome
LGALS1	galectin 1		GBP	22q13.1
LGALS2	galectin 2		HL14	22q13.1
LGALS3	galectin 3	LGALS2	MAC-2, GALIG	14q22.3
LGALS4	galectin 4		GAL4	19q13.2
LGALS7	galectin 7		GAL7, PIG1, TP53I1, LGALS7A	19q13.13
LGALS7B	galectin 7B		GAL7	19q13.2
LGALS8	galectin 8		PCTA-1	1q43
LGALS9	galectin 9		LGALS9A	17q11.2
LGALS9B	galectin 9B			17p11.2
LGALS9C	galectin 9C			17p11.2
CLC	Charcot-Leyden crystal galectin		Gal-10, LGALS10 , MGC149659	19q13.2
LGALS12	galectin 12		GRIP1	11q12.3
LGALS13	galectin 13		PP13, PLAC8	19q13.2
LGALS14	galectin 14		PPL13, CLC2	19q13.2
LGALS16	galectin 16			19q13.2

En fonction de l'organisation et du nombre de domaines CRD, les galectines des mammifères sont subdivisées en trois sous-types (figure 29) (Barondes *et al.*, 1994 ; Fukumori *et al.*, 2007) :

- Le groupe « prototype » comprend celles qui n'ont qu'un seul domaine CRD, soit les galectines-1, -2, -5, -7, -10, -11, -13, -14, -15 ;
- La galectine-3 (gal-3), seul membre du groupe « chimère », est formée d'un court domaine N-terminal suivi d'une séquence consensus répétée riche en résidus tyrosines, glycines et prolines rattachée au domaine CRD ;
- Le type double, ou en tandem, dans lequel se trouvent les galectines-4, -6, -8, -9, -12 est constitué d'une seule chaîne polypeptidique comportant deux domaines CRD distincts mais homologues et séparés par une séquence non conservée de 70 aa. Ces deux domaines CRD permettent la liaison à deux épitopes glycaniques.

L'affinité des galectines pour leurs ligands glycosylés est variable selon la composition et la longueur de leurs chaînes glycaniques (Advedissian *et al.*, 2015).

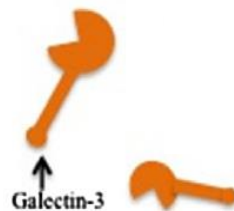
(a) Prototypic galectins (galectin-1, -2, -5, -7, -10, -11, -13, -14, -15)



(b) Tandem repeat type galectins (galectin-4, -6, -8, -9, -12)



(c) Chimera galectin (galectin-3)

**Figure 29 :** Classification des galectines (Illustration extraite de Hu *et al.*, 2017).

La première classification a été réalisée par Barondes *et al.*, (1994). Le groupe « prototype » regroupe les galectines formées d'un domaine CRD, le groupe « tandem repeat » les galectines possédant deux domaines CRD et le groupe « chimère » comprend la galectine-3 avec un domaine CRD rattaché à un domaine N-terminal.

IV.1.2. Expression tissulaire et cellulaire des galectines

Une fonction importante des galectines étant de réguler la réponse immunitaire et l'inflammation, elles sont exprimées dans de nombreux types cellulaires épithéliaux et dans les cellules du système immunitaire : lymphocytes T et B activés, cellules dendritiques, mastocytes, éosinophiles, monocytes/macrophages, et neutrophiles (Dumic *et al.*, 2006). Certaines galectines, comme la galectine-1 et la galectine-3, sont présentes dans une grande variété de tissus tandis que d'autres ont des territoires d'expression plus spécifiques. Par exemple, la galectine-2 est surtout exprimée dans les épithéliums digestifs, la galectine-7 dans

les épithéliums stratifiés, la galectine-5 dans les cellules sanguines et la galectine-4 est spécifique des cellules épithéliales du tractus gastro-intestinal (Liu *et al.*, 2002).

Les galectines peuvent être localisées dans tous les compartiments subcellulaires, tout particulièrement le cytosol, mais aussi le noyau, voire les mitochondries et les lysosomes, ainsi qu'à la surface cellulaire du pôle apical, avec une distribution variant selon le type cellulaire (Dumic *et al.*, 2006) mais également selon l'état de différenciation de la cellule (Thijssen *et al.*, 2015). Les galectines peuvent être également sécrétées dans le milieu extracellulaire, or elles ne présentent ni peptide signal, ni glycosylation, ce qui permet de les considérer comme des lectines animales très particulières (Advedissian *et al.*, 2015).

Les galectines sont biosynthétisées par les ribosomes libres et leur présence en dehors de la cellule est la conséquence d'une sécrétion par un mécanisme non conventionnel puisqu'elles sont dépourvues de la séquence signal permettant l'insertion dans le réticulum endoplasmique (Hughes, 1999). Différents mécanismes sont proposés pour expliquer la présence des galectines dans le milieu extracellulaire : translocation à travers la membrane, ectocytose par bourgeonnement de la membrane, sécrétion médiée par des vésicules spécifiques ou par les exosomes (Viguier *et al.*, 2014). Ce dernier mécanisme permettrait soit la libération extracellulaire de la galectine-3 suite à la lyse des vésicules, soit la fusion de la vésicule avec une autre cellule (figure 30) (Dumic *et al.*, 2006).

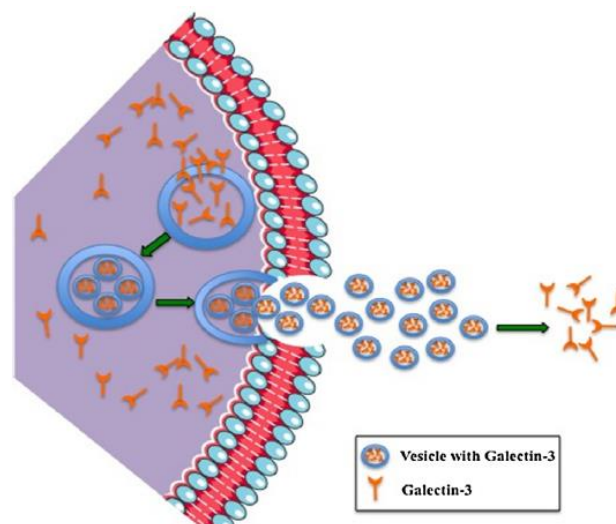


Figure 30 : Illustration de la sécrétion de la galectine-3 par l'intermédiaire de vésicules de sécrétion qui sont libérées dans le milieu extracellulaire (Hu *et al.*, 2017).

IV.1.3. Fonctions des galectines

Décrites à l'origine comme étant impliquées précocement au cours du développement, les galectines apparaissent à partir des années 1980 comme impliquées dans de très nombreuses fonctions biologiques telles que l'adhérence cellulaire, et la signalisation cellulaire. Aujourd'hui leur rôle dans la réponse immunitaire et inflammatoire est bien documenté dans trois domaines : reconnaissance des pathogènes, régulation de la réponse immunitaire adaptative et résolution de l'inflammation (Advedissian *et al.*, 2015; Viguiet *et al.*, 2014). L'étude des modèles murins porteurs de gènes de galectines inactivés permet de proposer pour ces protéines des rôles généralistes de molécules d'adaptation au stress cellulaire, c'est-à-dire impliquées dans le retour à l'homéostasie cellulaire, quel que soit le stress, soit causé par une blessure, un foyer infectieux, une exposition aux UVB ou à un toxique (Advedissian *et al.*, 2015).

Les fonctions variées des galectines sont en lien avec leur capacité à former des homodimères ou des oligomères, leur permettant ainsi d'interagir simultanément avec différents glycoconjugués. En plus de leur capacité à lier des motifs N-acétyllactosaminiques, les galectines peuvent aussi lier d'autres sucres : par exemple, la galectine-10 se lie de façon plus restreinte au mannose (Ruvolo, 2016).

Toutes ces propriétés favorisent de façon importante l'établissement d'interactions cellules-cellules mais aussi cellules-matrice extracellulaire (figure 31A) (Thijssen *et al.*, 2015). En plus de leurs interactions avec les glycanes, les galectines engagent aussi des interactions de type protéine-protéine (figure 31A), moins bien connues mais impliquées dans différents processus tels que dans la signalisation cellulaire, l'apoptose, la progression du cycle cellulaire, l'épissage des pré-ARNm (figure 32) (Thijssen *et al.*, 2015).

La multivalence des galectines leur permet de former des réseaux dynamiques à la surface cellulaire qui régulent la distribution des glycoprotéines avec lesquelles elles se lient (figure 31B) (Nabi *et al.*, 2015). Les complexes ainsi formés modulent les voies de signalisation intracellulaires et favorisent le « clustering » de récepteurs ainsi que l'inhibition de leur internalisation (Elola *et al.*, 2015).

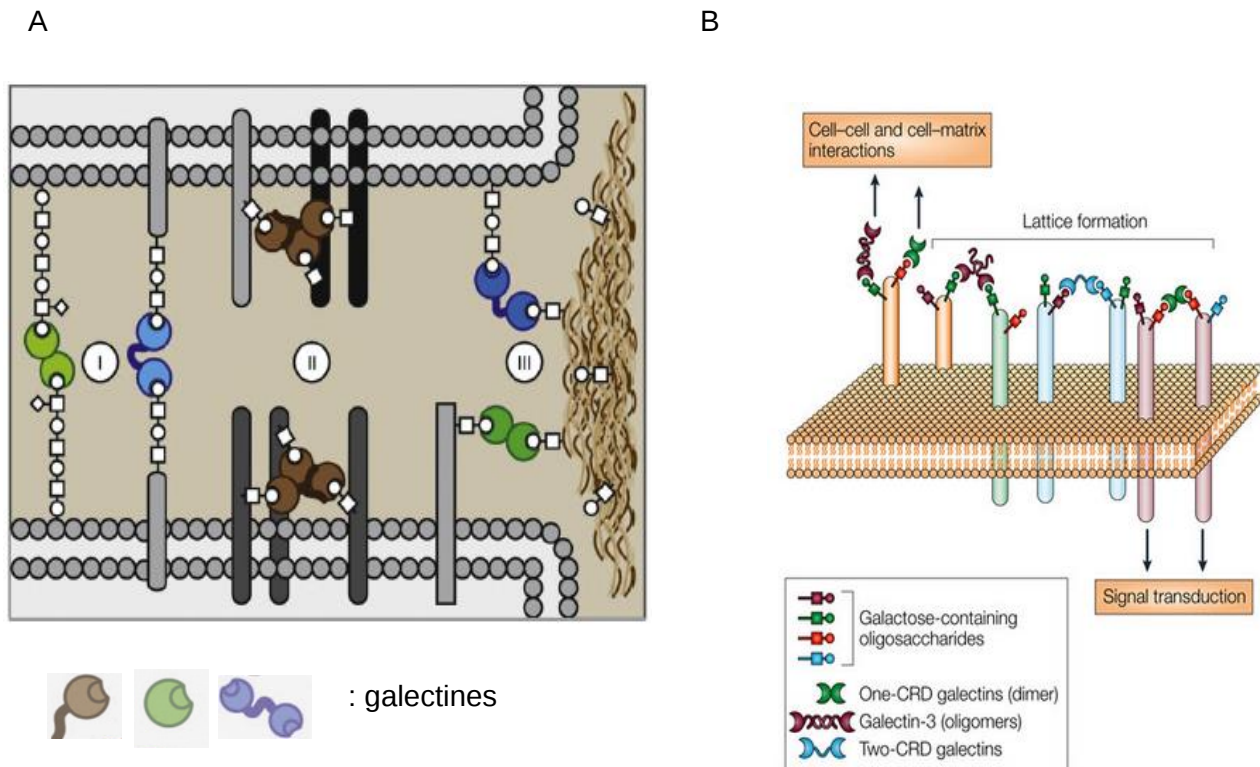


Figure 31 : Illustration des fonctions liées à la dimerisation et à l'oligomérisation des galectines (Thijssen *et al.*, 2015).

(A) I) interactions cellules-cellules, II) *clustering* des récepteurs, III) interactions cellules-matrice extracellulaire. (B) Les galectines interagissent avec des glycoprotéines transmembranaires, stimulant ainsi la signalisation intracellulaire. Le regroupement des récepteurs médiés par les galectines crée des réseaux au niveau des surfaces cellulaires (Liu and Rabinovich, 2010).

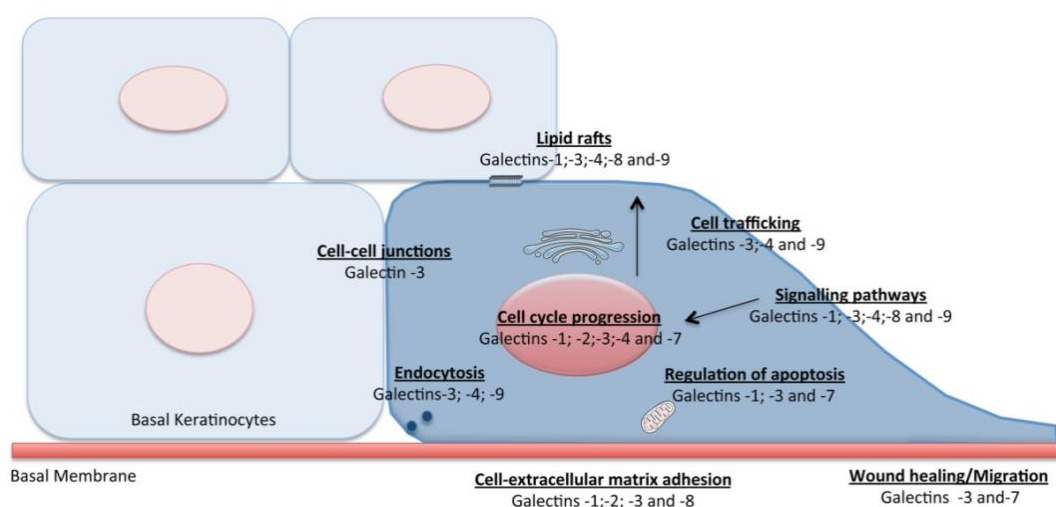


Figure 32 : Illustration des fonctions des galectines dans les cellules épithéliales : voies de signalisation, trafic intracellulaire, apoptose, cycle cellulaire... (Viguié *et al.*, 2014).

Capables de lier les glycanes, les protéines ou les pré-ARNm, Thijssen *et al.* (2015) les définissent les galectines comme les « interprètes cellulaire du glycome ». Le glycome, ensemble des glycanes exprimés par une cellule ou un organisme, représente avec le protéome et le génome l'un des répertoires majeurs de biomolécules agissant sur le fonctionnement d'une cellule ou d'un organisme. Les galectines feraient donc le lien entre ces répertoires en traduisant les informations portées par le glycome en fonctions cellulaires. Schnaar (2016) les résume comme une « famille fascinante » de protéines capables de lier toutes sortes de glycanes mais aussi d'exercer des fonctions importantes indépendantes des glycanes.

IV.1.4. Galectines et cancer

Les galectines interviennent dans le processus tumoral de différentes façons : elles modulent l'apoptose, l'angiogenèse et l'immunité anti-tumorale, et favorisent le processus métastatique (Advedissian *et al.*, 2015). Leur expression est fréquemment modifiée dans de nombreux types de cancers, digestifs, respiratoires, génitaux, neuroendocriniens ou encore hématologiques (Thijssen *et al.*, 2015). Par exemple, dans le cancer du foie et celui du poumon, l'augmentation de l'expression de la galectine-1 est plutôt de mauvais pronostic alors que celle de la galectine-9 est plutôt associée à un meilleur pronostic. Par ailleurs, les galectines étant sécrétées dans le milieu extracellulaire, une augmentation des taux circulants de galectine-1 et -3 a été observée au cours du cancer de la thyroïde sans pour autant permettre de différencier les formes bénignes et malignes (Thijssen *et al.*, 2015).

IV.2. Galectine-3

IV.2.1. Gène

La galectine-3 est codée par le gène *LGALS3*, dont la taille est de 17 kb environ. Localisé sur le chromosome 14 en position q21-22, il est constitué de 6 exons et 5 introns (figure 33) (Dumic *et al.*, 2006). L'exon 1 code une partie de la séquence 5' UTR de l'ARNm et l'exon 2 le reste de cette région ainsi que le site d'initiation de la traduction et les six premiers aa. L'exon 3 code la séquence du domaine N-terminal de la protéine. Les trois derniers exons codent le domaine CRD (Krześlak and Lipińska, 2004). Par épissage alternatif, plusieurs isoformes peuvent être générées.

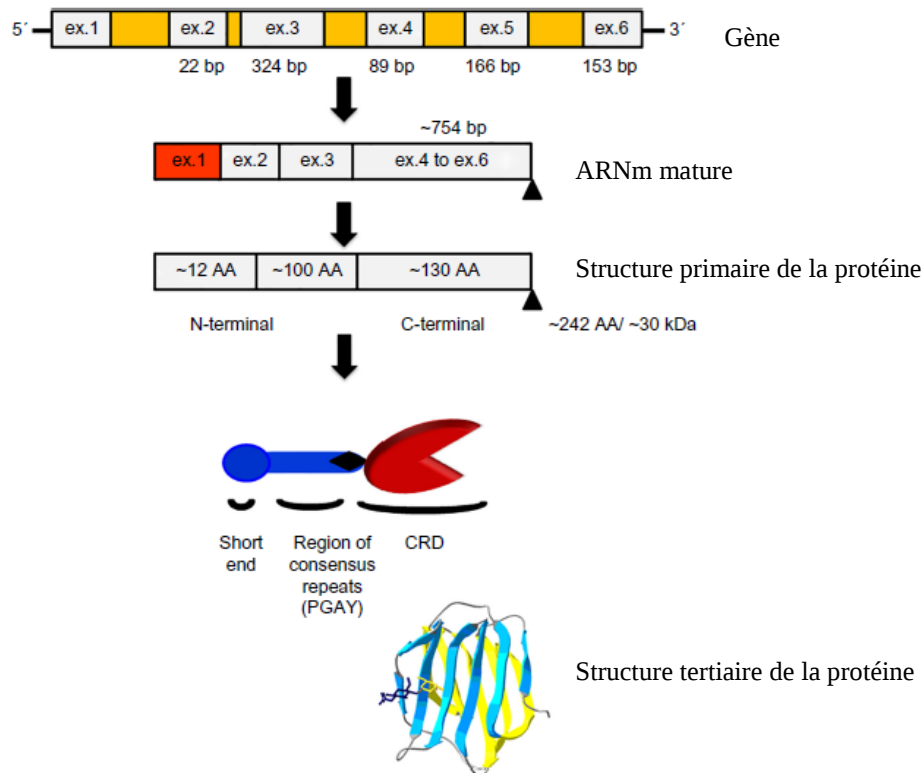


Figure 33 : La galectine-3 : correspondance entre les exons et les domaines protéiques (Krześlak and Lipińska, 2004).

LGALS3 est considéré comme un gène de réponse immédiate et son expression est régulée au niveau transcriptionnel par différents facteurs. Elle est par exemple induite par certaines cytokines pro-inflammatoires telles que le $\text{TNF}\alpha$, l'IL-2. Ainsi, la région promotrice du gène codant la galectine-3 comporte divers éléments régulateurs tels que cinq sites de liaison supposés pour le facteur de transcription Sp1, deux sites $\text{NF}\kappa\text{B}$ -like, une séquence consensus hélice-boucle-hélice, 4 sites AP-1, 5 sites CRE (*cAMP dependent response element*), 1 site AP-4 like et 1 site SIE (*sis-inducible element*) (Dumic *et al.*, 2006). Son promoteur possède également des sites de fixation pour les facteurs de transcription Runx1 (*Runt-related transcription factor 1*) et HIF1 α (*Hypoxia Inducible Factor 1 α*) (Wang and Guo, 2016). De plus, il existe une régulation au niveau épigénétique puisque la méthylation des îlots CpG présents dans son promoteur intervient dans la régulation de son expression (Dumic *et al.*, 2006). Par ailleurs, Ramasamy *et al.* (2007) ont mis en évidence que MUC1 régule l'expression de la galectine-3 par l'intermédiaire du miR-322, en supprimant l'expression de ce micro-ARN qui cible normalement les transcrits de la galectine-3. Ce mécanisme permet

donc une régulation positive de l'expression de la galectine-3 par le biais de MUC1 (Ramasamy *et al.*, 2007).

IV.2.2. Protéine

La galectine-3 est une protéine de 250 aa (30 kDa) environ, seul membre humain du groupe « chimère » des galectines de mammifères. Elle possède plusieurs domaines fonctionnels (figure 34) (Dumic *et al.*, 2006; Newlaczyl and Yu, 2011) :

- Une région N-terminale, fortement conservée entre les espèces (Cooper, 2002; Dumic *et al.*, 2006), formée d'un domaine N-terminal de 12 aa et d'un domaine *collagene like* constitué de 7 à 14 répétitions d'une séquence de 9 aa riche en résidus prolines, glycines et tyrosines (PGAYPGXXX) (Wang *et al.*, 2004). Cette région ne possède pas en elle-même d'activité de liaison aux glycanes mais elle est essentielle à cette fonction et à l'activité biologique générale de la galectine-3, en particulier aux propriétés anti-apoptotiques (Dumic *et al.*, 2006). En effet, elle est impliquée dans la sécrétion de cette lectine par un mécanisme non conventionnel et intervient dans sa polymérisation (formation de pentamères) en présence de ligands (figure 35) (Dumic *et al.*, 2006).

Cette région comporte des sites fonctionnels remarquables : un site d'acétylation pour l'alanine en 2 (signal potentiel de dégradation, la méthionine en 1 étant clivée) et deux sites de phosphorylation au niveau des sérines 6 et 12 (Argüeso *et al.*, 2015). Les 11 aa N-terminaux sont indispensables à la localisation nucléaire et aux fonctions de la galectine-3. En effet, la galectine-3 présente une localisation ubiquitaire au sein de la cellule, mais sa distribution subcellulaire est très dépendante du cycle cellulaire, majoritairement cytosolique dans la cellule quiescente et majoritairement nucléaire dans la cellule en réplication (Liu *et al.*, 2002). La phosphorylation/déphosphorylation de la sérine 6 est un événement-clé pour la régulation du trafic de la galectine-3 entre le noyau et le cytosol, elle module également ses propriétés de liaison aux carbohydrates (Wang *et al.*, 2004) et ses propriétés pro-apoptotiques (Liu *et al.*, 2002). La galectine-3 cytoplasmique est exclusivement phosphorylée, alors que la galectine-3 nucléaire peut être phosphorylée ou non (Liu *et al.*, 2002).

- Un domaine C-terminal de structure globulaire de 130 aa environ, contenant le CRD, site de liaison des carbohydrates (aa 113-148) et organisé selon deux feuillets bêta anti-parallèles formant une structure en sandwich (Dumic *et al.*, 2006). Il contient un motif asparagine-tryptophane-glycine-arginine (NWGR) fortement conservé, commun au domaine BH1 des

protéines de la famille Bcl-2 (*B-Cell lymphoma 2*) et impliqué dans l'activité anti-apoptotique de la galectine-3. Le CRD permet également la dimérisation de la galectine-3 en l'absence de ligands. Il contient encore un site de phosphorylation au niveau de la sérine 188. Le CRD joue également un rôle majeur pour la localisation subcellulaire de la galectine-3, car il contient à la fois des séquences non classiques d'import nucléaire (ou NLS pour *nuclear localization sequence*) et des séquences d'export nucléaire (aa 226-241) (ou NES pour *nuclear export sequences*) (Ruvolo, 2016; Funasaka *et al.*, 2014).

La galectine-3 peut être détectée sous une forme clivée en particulier dans les cellules cancéreuses. En effet, le domaine *collagen-like* de sa région N-terminale est la cible des métalloprotéases MMP-2 et -9 (*Matrix metalloproteinases*) entre les aa alanine 62 et tyrosine 63. Ce clivage génère un fragment C-terminal qui possède des capacités de fixation accrues aux carbohydrates et qui favorise la progression tumorale par le biais d'une augmentation de l'angiogenèse et d'une résistance à l'apoptose (Dumic *et al.*, 2006; Nangia-Makker *et al.*, 2008).

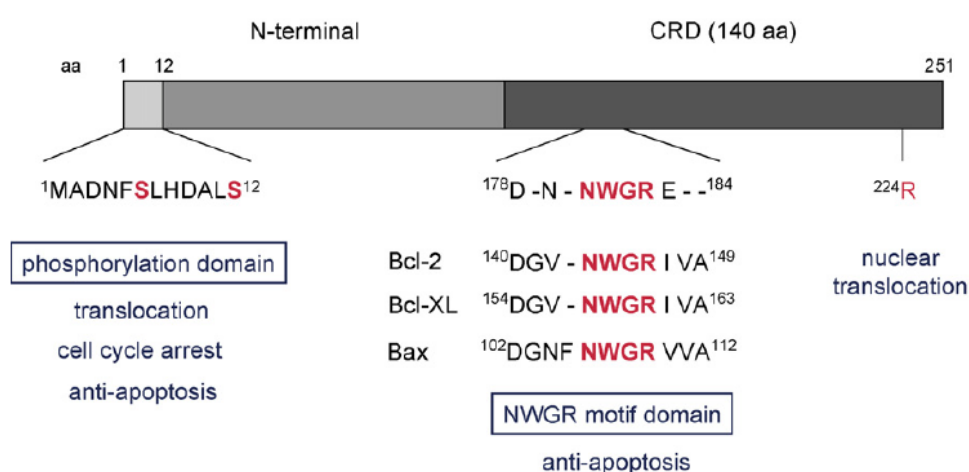


Figure 34 : Représentation des différents domaines de la galectine-3 avec les fonctions qui leur sont attribuées (Fukumori *et al.*, 2007).

La région N-terminale contient des sites de phosphorylation et est impliquée dans le cycle cellulaire par exemple, le motif NWGR localisé au niveau du CRD participe à ses fonctions anti-apoptotiques.

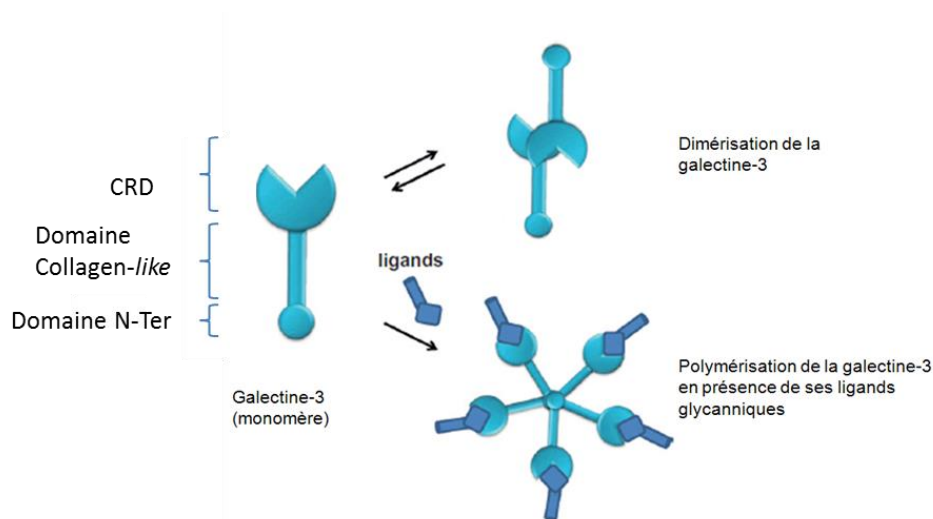


Figure 35 : Représentation schématique de la galectine-3 dans ses différents états (monomère, dimère et polymère) selon la fixation de ses ligands (d'après Newlaczyl and Yu, 2011).

IV.2.3. Sites tissulaires et cellulaires d'expression

La galectine-3 est exprimée de façon ubiquitaire dans les tissus humains mais à des niveaux variables selon le type tissulaire et le stade de développement. Durant le premier trimestre de l'embryogenèse, son expression est principalement détectée dans les épithéliums, le rein, le foie et les chondrocytes (Dumic *et al.*, 2006). Fortement exprimée par les cellules immunitaires et épithéliales chez l'adulte, la galectine-3 est par exemple présente dans la muqueuse gastrique, les épithéliums mammaires, digestifs, prostatiques, respiratoires, rénaux ainsi que dans les monocytes et les macrophages (Fukumori *et al.*, 2007). A l'inverse, la galectine-3 est faiblement, voire pas du tout, exprimée dans les lymphocytes au repos. Cependant, l'activation de ces cellules s'accompagne d'une augmentation de son expression (Dumic *et al.*, 2006). Au niveau extracellulaire, la galectine-3 a été détectée dans des exosomes dérivant de divers types cellulaires (cellules dendritiques, cellules tumorales épithéliales ou hématologiques) (Fei *et al.*, 2015).

IV.2.4. Ligands

Les galectines sont des lectines se liant de façon spécifique aux β -galactosides par l'intermédiaire de leur CRD, la présence de galactose étant indispensable pour leur liaison à leurs ligands. Cependant, l'affinité de liaison augmente si ce résidu est associé à d'autres

saccharides, par exemple à la *N*-acétyl-glucosamine (Dumic *et al.*, 2006). Il est intéressant de noter que beaucoup de ses ligands extracellulaires sont glycosylés à la différence de ses ligands intracellulaires (Advedissian *et al.*, 2015).

La galectine-3 possède de nombreux partenaires qui sont fonctionnellement et structurellement très divers, localisés aussi bien dans le milieu extracellulaire qu'intracellulaire (tableau IX). Ainsi, la galectine-3 se lie à la laminine, protéine de la matrice extracellulaire, à des protéines intracellulaires telles qu'Alix/AIP-1 (*ALG-2 interacting protein X*), Bcl-2 ou à des protéines membranaires comme LAMP1 et 2 (*Lysosome Associated Membrane Protein*) (Wang *et al.*, 2004). La galectine-3 liée à la synexine est capable de pénétrer la membrane mitochondriale (Newlaczyk and Yu, 2011). Au niveau nucléaire, la galectine-3 interagit avec des facteurs de transcription comme TTF-1 (*Thyroid Transcription Factor-1*) ou avec des constituants impliqués dans l'assemblage du spliceosome comme la protéine Gemin 4 ou la ribonucléoprotéine U1 (Haudek *et al.*, 2009, 2010a).

Tableau IX : Exemples de ligands intracellulaires de la galectine-3 (Haudek *et al.*, 2010a).
Abréviation : nk : not known, nd : not determined

Ligand	Subcellular Compartment	Gal3 Domain ¹	Carbohydrate Sensitivity ²	Direct or Indirect
Bcl-2	Cytoplasm	CRD	Yes	nk ³
β-catenin	Cytoplasm and nucleus	CRD	Yes	Direct
CBP70	Nucleus	CRD	Yes	nk
Chrp	Cytoplasm	CRD	no	Direct
Cytokeratin	Cytoplasm	CRD	Yes	Direct
Gemin4	Cytoplasm and nucleus	CRD	nd ⁴	Direct
hnRNP Q	Nucleus	nd	nd	nk
Nucling	Cytoplasm	nd	nd	nk
OCA-B	Cytoplasm	CRD	no	Direct
K-Ras	Cytoplasm and plasma membrane	nd	nd	nk
Sufu	Cytoplasm and nucleus	nd	nd	Direct
Synexin	Cytoplasm and mitochondria	nd	nd	Direct
TFII-I	Nucleus	CRD	Yes	nk
TTF-1	Nucleus	nd	nd	Direct

Par ailleurs, les interactions entre galectine-3 et mucines sont bien documentées dans la littérature (figure 36). Ainsi, Tanida *et al.* (2014). ont démontré que la galectine-3 se lie au motif glycanique Thomsen-Friedenreich (TF) présent sur la mucine MUC1, cette interaction entraînant le recrutement de la β-caténine par cette mucine. Plusieurs équipes ont également

mis en évidence que la liaison de la galectine-3 au niveau de ce motif présent sur MUC1 et MUC4 est à l'origine du regroupement de ces mucines à la surface cellulaire (Senapati *et al.*, 2011; Zhao *et al.*, 2009). La galectine-3 est également capable de se fixer sur la partie C-terminale extracellulaire de MUC1 au niveau du motif N-glycanique porté par l'Asparagine 36 pour favoriser l'interaction de cette mucine avec l'EGFR (Ramasamy *et al.*, 2007).

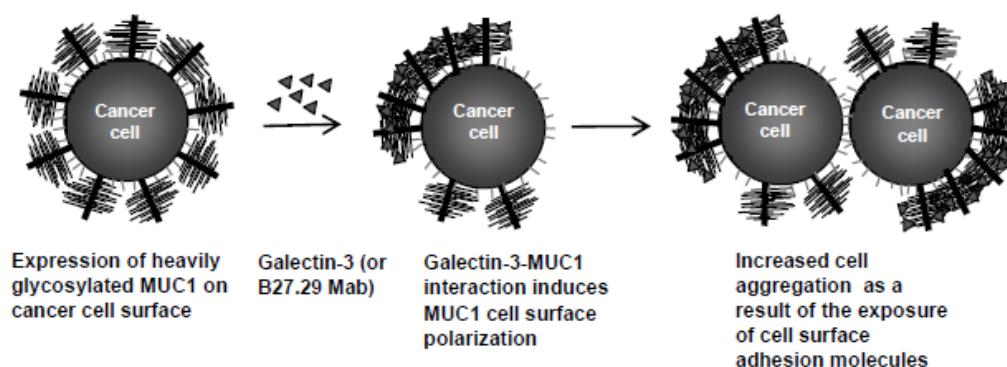


Figure 36 : Modèle de l'interaction entre MUC1 et la galectine-3 via les motifs glycaniques Tn porté par la mucine MUC1 (Zhao *et al.*, 2010).

L'interaction MUC1-galectine-3 entraîne le regroupement des mucines au niveau de la surface cellulaire et l'exposition de molécules d'adhérence favorisant ainsi la dissémination métastatique.

IV.2.5. Fonctions

La galectine-3 exerce de nombreuses fonctions biologiques qui dépendent de sa localisation intracellulaire ou extracellulaire.

Par exemple, au niveau **extracellulaire**, la galectine-3 :

- intervient dans la migration et l'adhérence cellulaire, les interactions cellules-cellules et cellules-matrice extracellulaire *via* ses propriétés de liaison aux glycanes présents sur de nombreuses glycoprotéines (laminine, fibronectine, intégrine...) (Dumic *et al.*, 2006).
- possède des propriétés de chemoattractant pour différents types cellulaires tels que les monocytes ou les macrophages et augmente la réponse oxydative des polynucléaires neutrophiles par le biais de sa fixation sur les récepteurs CD66a et CD66b (Dumic *et al.*, 2006).

- possède des fonctions pro-apoptotiques. Ainsi, sa liaison au complexe CD29/CD7 active des voies de signalisation intracellulaires aboutissant à l'activation des caspases (Dumic *et al.*, 2006).
- grâce à ses propriétés de liaison multivalentes, est capable de former des « clusters » de glycoconjugués formant ainsi un réseau galectine-3/glycoprotéines qui affecte le turnover et l'endocytose de ces récepteurs glycoprotéiques affectant ainsi la réponse à leur ligand et la signalisation intracellulaire (Fortuna-Costa *et al.*, 2014; Liu and Rabinovich, 2005).
- interagit, au niveau de la membrane apicale, avec les mucines MUC1 et MUC16 afin de maintenir la barrière épithéliale de la surface oculaire (Argüeso *et al.*, 2009).

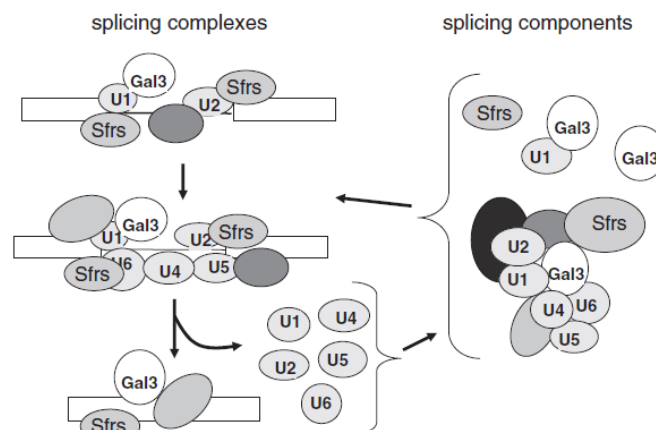
Dans la cellule, la galectine-3 est présente majoritairement au niveau nucléaire et cytoplasmique, sa localisation variant selon le type cellulaire, le statut de prolifération des cellules, les conditions de culture et la transformation cellulaire (Wang *et al.*, 2004).

Au niveau **nucléaire**, elle intervient dans la régulation de certaines voies de signalisation telles que la voie Wnt (*Wingless*) (Liu *et al.*, 2002) et dans l'épissage des pré-ARNm via son interaction avec la ribonucléoprotéine U1 (figure 37A) (Dagher *et al.*, 1995; Haudek *et al.*, 2009). Cette dernière est un constituant du spliceosome, complexe dynamique de particules ribonucléoprotéiques.

En plus de ce rôle, Lin *et al.* (2002) ont démontré le rôle de la galectine-3 nucléaire dans la régulation de la transcription de gènes (figure 37B) (Lin *et al.*, 2002). En effet, elle favorise le pouvoir transactivateur des facteurs de transcription CREB (*cAMP-Response Element Binding protein*) et Sp1, induisant ainsi l'activité du promoteur du gène de la cycline D1. Song *et al.* (2005) ont mis en évidence que la galectine-3 participe à la régulation de l'expression de *MUC2* dans les cellules cancéreuses coliques au niveau transcriptionnel par le biais de la stabilisation de la liaison de protéines nucléaires au niveau du site AP-1 du promoteur de *MUC2*. Par ailleurs, elle interagit également avec le facteur de transcription TTF-1 stimulant ainsi l'activité transcriptionnelle de ce facteur (Paron *et al.*, 2003) (figure 37B).

Lorsqu'elle est retrouvée au niveau nucléaire, la galectine-3 possède des fonctions pro-apoptotiques. Celles-ci seraient liées à son rôle dans la réponse cellulaire face aux dommages de l'ADN. En effet Carvalho *et al.* (2014) ont mis en évidence une augmentation de la résistance cellulaire face aux altérations de l'ADN en l'absence de galectine-3.

A



B

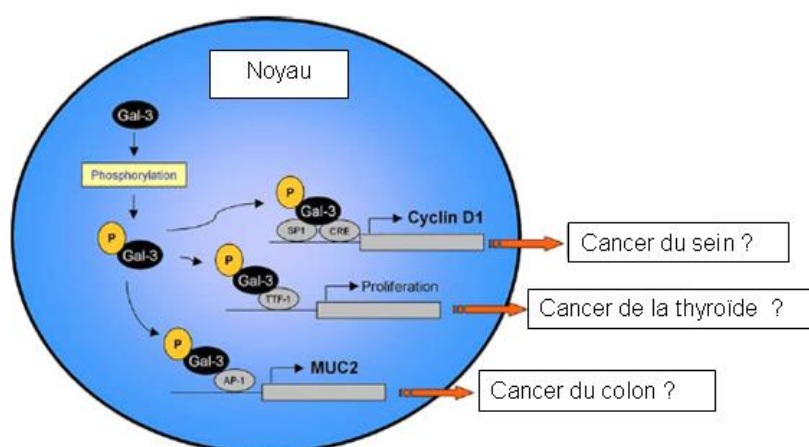


Figure 37 : Illustration de quelques fonctions nucléaires de la galectine-3.

Exemple du rôle de la galectine-3 dans (A) l'épissage des pré ARNm via son interaction avec la ribonucléoprotéine U1 (Haudek *et al.*, 2010b) et dans (B) la régulation de l'expression de gènes comme *MUC2* ou la *cycline D1* en favorisant la fixation des facteurs de transcription sur les promoteurs de ces gènes (d'après Nakahara and Raz, 2007).

Au niveau **cytoplasmique**, la galectine-3 intervient dans différentes voies de transduction de signaux comme par exemple celle faisant intervenir K-Ras (Elad-Sfadia *et al.*, 2004). Elle participe aussi au trafic intracellulaire des glycoprotéines et à l'adressage apical indépendant des rafts lipidiques de certaines glycoprotéines par le biais de la formation de clusters de haut poids moléculaire. Ainsi, l'adressage apical de la *lactase phlorizin hydrolase* (LPH) est dépendant d'une interaction directe avec la galectine-3 (Delacour *et al.*, 2009). Par ailleurs, Merlin *et al.* (2011) ont montré que cette lectine régule l'endocytose et la localisation subcellulaire de MUC1 et de l'EGFR (*EGF receptor*) dans les cellules cancéreuses

pancréatiques. A l'inverse de la galectine-3 nucléaire, la forme cytoplasmique exerce une action anti-apoptotique par le biais de l'inhibition de la libération du cytochrome c suite à sa translocation dans la membrane mitochondriale ou par activation de la voie PI3K/Akt (Dumic *et al.*, 2006) (figure 38).

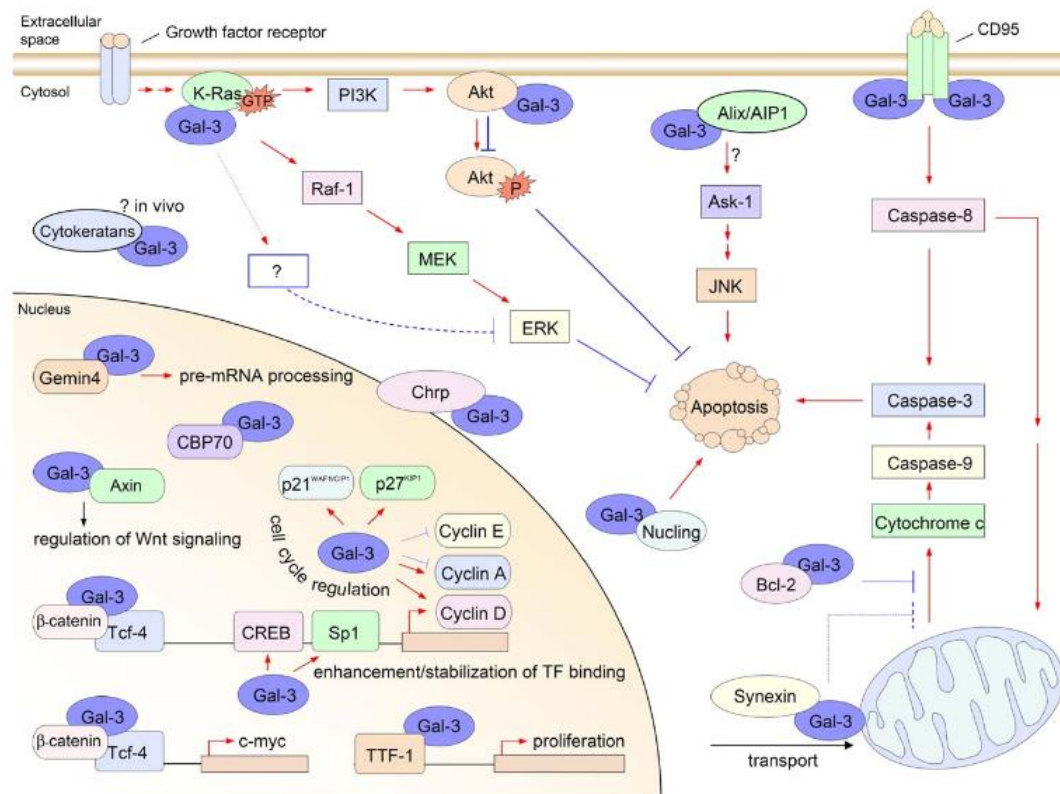


Figure 38 : Illustration des fonctions intracellulaires de la galectine-3 : au niveau cytoplasmique : apoptose, voies de signalisation par exemple et au niveau nucléaire régulation de l'expression des gènes et épissage des pré-ARNm... (Dumic *et al.*, 2006).

Abréviations : Akt : the serine/threonine kinase Akt, Ask-1 :apoptosis signal-regulating kinase 1, CBP70:carbohydrate binding protein 70, Chrp:cysteine- and histidine-rich protein, CREB:cAMP-response element-binding protein, ERK:extracellular signal-regulated kinase, Gal-3:galectin-3, GTP:guanosine triphosphate, JNK:c-Jun NH2-terminal kinase, MEK:mitogen-activated protein/ERK kinase, P:phosphate, PI3K:phosphatidylinositol 3-kinase, Raf-1:the serine/threonine kinase Raf-1, Tcf-4:T cell factor 4, TF:transcription factor,TTF-1:thyroid-specific transcription factor.

Il est intéressant de noter que la galectine-3 possède des caractéristiques communes avec les protéines de la famille SR (*serine/arginine-rich*, protéines riches en résidus sérines et arginines). En effet, ces deux groupes de protéines se lient aux motifs glycaniques, participent à l'épissage des pré-ARNm, présentent une distribution à la fois nucléaire et cytoplasmique et peuvent être retrouvées aussi à la surface de certains types cellulaires (Haudek *et al.*, 2010b).

IV.2.6. Galectine-3 et cancer

La galectine-3 est exprimée dans une grande variété de tumeurs (tableau X) avec une intensité dépendante de la progression tumorale, du pouvoir invasif et métastatique (Danguy *et al.*, 2002). Ainsi, la galectine-3 est surexprimée dans le cancer du foie, de la thyroïde, de l'estomac et à l'inverse, son expression est diminuée dans le cancer du sein, de l'ovaire, de l'utérus, de la prostate ou de l'endomètre (Wang and Guo, 2016). De plus, sa valeur pronostique est dépendante du type tumoral (Thijssen *et al.*, 2015).

Berberat *et al.* (2001) ont montré par *Western blot* et immunohistochimie que la galectine-3 est surexprimée au cours de l'adénocarcinome pancréatique. Les taux sériques de galectine-3 sont également augmentés chez les patients atteints de PDAC par rapport aux sujets sains (Xie *et al.*, 2012) sans pour autant permettre de distinguer les pathologies bénignes et malignes (Coppin *et al.*, 2016). Le dosage de la galectine-3 circulante serait utile pour suivre l'efficacité des traitements. En effet, son taux chute chez les patients atteints de cancer pancréatique après chirurgie (Thijssen *et al.*, 2015). De plus, sa distribution intracellulaire est modifiée dans certaines tumeurs, ce qui fait d'elle un marqueur potentiel pour le diagnostic et le pronostic (Liu and Rabinovich, 2005). Ainsi, dans le cancer de la prostate, une délocalisation de la galectine-3 est observée avec un passage du noyau vers le cytoplasme (Haudek *et al.*, 2010a).

Tableau X : Modifications de l'expression de la galectine-3 au cours de différents cancers (Wang and Guo, 2016).

	Up-regulation	Down-regulation	Conflicted reports
Cancer Tissue	Pancreatic cancer; Hepatic carcinoma Tongue cancer; Thyroid carcinoma; Gastric carcinoma; Intrahepatic cholangio- carcinoma (early stage)	Breast carcinoma; Prostate carcinoma; Endometrial carcinoma; Intrahepatic cholangioc-arcinoma (later stage);	Colon carcinoma Up-regulation in later stage Down-regulation in during neoplastic progression
Patients' Serum	Breast carcinoma; Gastrointestinal carcinoma; Lung carcinoma; Head and neck carcinoma;		

La galectine-3 participe au processus tumoral en intervenant dans (figure 39A et B) :

- 1) **la transformation cellulaire** via interaction avec la protéine K-Ras, ce qui stabilise cette dernière sous forme activée (associée au GTP) au niveau de la membrane

plasmique. Ce phénomène favorise la transformation cellulaire induite par cet oncogène par activation des voies de signalisation en aval (Liu and Rabinovich, 2005). De plus, la galectine-3 module l'expression du mir let-7, régulateur négatif de la transcription de K-Ras (Levy *et al.*, 2011).

- 2) **la régulation du cycle cellulaire** à différents points (arrêt en G1 ou G2/M). La galectine-3 stimule l'expression des protéines p21 et p27 (Dumic *et al.*, 2006) et diminue celle des cyclines A et E. A l'inverse, la galectine-3 active le promoteur de la cycline D1 (Lin *et al.*, 2002) et stabilise le complexe cycline D1/CDK4 (*cyclin-dependent kinase 4*) favorisant ainsi la phosphorylation de la protéine du rétinoblastome (Kim *et al.*, 2014).
- 3) **la résistance à l'apoptose** par sa translocation au niveau de la mitochondrie, ce qui stabilise le potentiel de membrane mitochondrial et limite la libération du cytochrome c mitochondrial et l'activation des caspases. De par son interaction avec Bcl-2, elle favorise également la localisation mitochondriale de cette protéine. De plus, la galectine-3 active les voies anti-apoptotiques MAPK et Akt/PI3K (Liu and Rabinovich, 2005). Elle entraîne aussi une résistance à l'apoptose induite par TRAIL (*tumor-necrosis-factor related apoptosis inducing ligand*) en altérant le trafic des récepteurs DR4 et DR5 (*Death Receptors*) ce qui favorise leur *clustering* en réseau empêchant ainsi la formation du complexe DISC (*Death-inducing signaling complex*) (Mazurek *et al.*, 2012).
- 4) **le processus métastatique** par la modification de l'adhérence cellulaire en favorisant dans un premier temps le détachement des cellules tumorales via la perturbation des interactions existantes ou par des interactions avec les glycoprotéines de la matrice extra-cellulaire. Ainsi, la galectine-3, par le biais de son interaction avec la N-cadhérine déstabilise les jonctions cellulaires (Fortuna-Costa *et al.*, 2014). Dans un second temps, le processus métastatique implique l'attachement des cellules tumorales à un site distant. La galectine-3 participe à ce processus via ses propriétés pro-adhésives (Liu and Rabinovich, 2005). Ainsi, la galectine-3 reconnaît l'antigène Thomsen-Friedenreich, habituellement masqué par les acides sialiques dans les cellules normales mais démasqué à la surface de la plupart des cellules cancéreuses en raison d'un processus de O-glycosylation incomplet (Guha *et al.*, 2013). Ce

mécanisme intervient dans l'interaction entre galectine-3 et mucines, entraînant une redistribution et un *clustering* de ces protéines à la surface cellulaire. Ceci entraîne l'exposition de molécules d'adhérence, comme les intégrines, habituellement masquées par la grande taille de MUC1 et MUC4 (Senapati *et al.*, 2011; Zhao *et al.*, 2009) permettant ainsi l'interaction entre les cellules cancéreuses et l'endothélium. La galectine-3 intervient également dans la migration et la mobilité cellulaire par le biais du remodelage du cytosquelette. Elle participe à l'invasion des cellules tumorales en recrutant notamment les intégrines et d'autres protéines de surface ou par le biais de l'activation de métalloprotéinases telles que MMP-1 via l'activation de son facteur de transcription AP-1 (Braeuer *et al.*, 2012). Jiang *et al.* (2008) ont démontré dans leur étude que la galectine-3 sécrétée par les cellules cancéreuses pancréatiques stimule les cellules stellaires pancréatiques responsables de la production de stroma, favorisant ainsi le processus métastatique. Enfin, la galectine-3 stimule la production de cytokines pro-inflammatoires telles que l'IL-6, le G-CSF (*granulocyte colony-stimulating factor*), le GM-CSF (*Granulocyte-macrophage colony-stimulating factor*) qui induisent l'expression de molécules d'adhérence au niveau des cellules endothéliales favorisant l'adhérence des cellules tumorales (Fortuna-Costa *et al.*, 2014).

- 5) **l'angiogenèse** en favorisant la migration des cellules endothéliales et en influençant l'activité angiogénique du VEGF (*Vascular Endothelial Growth Factor*) et bFGF (*Basic fibroblast growth factor*). Ainsi, la galectine-3 favorise le *clustering* des récepteurs VEGFR 2 (*Vascular endothelial growth factor receptor 2*) ce qui accentue la signalisation cellulaire et réduit l'internalisation de ces récepteurs (Fortuna-Costa *et al.*, 2014).
- 6) **l'échappement de la tumeur** face à la surveillance du système immunitaire par la modulation de l'activité de ses cellules effectrices (induction de l'apoptose des lymphocytes T, blocage de la différenciation des lymphocytes B en plasmocytes) et des taux de cytokines (diminution de la production d'IL-5) (Liu and Rabinovich, 2005; Newlaczyl and Yu, 2011).

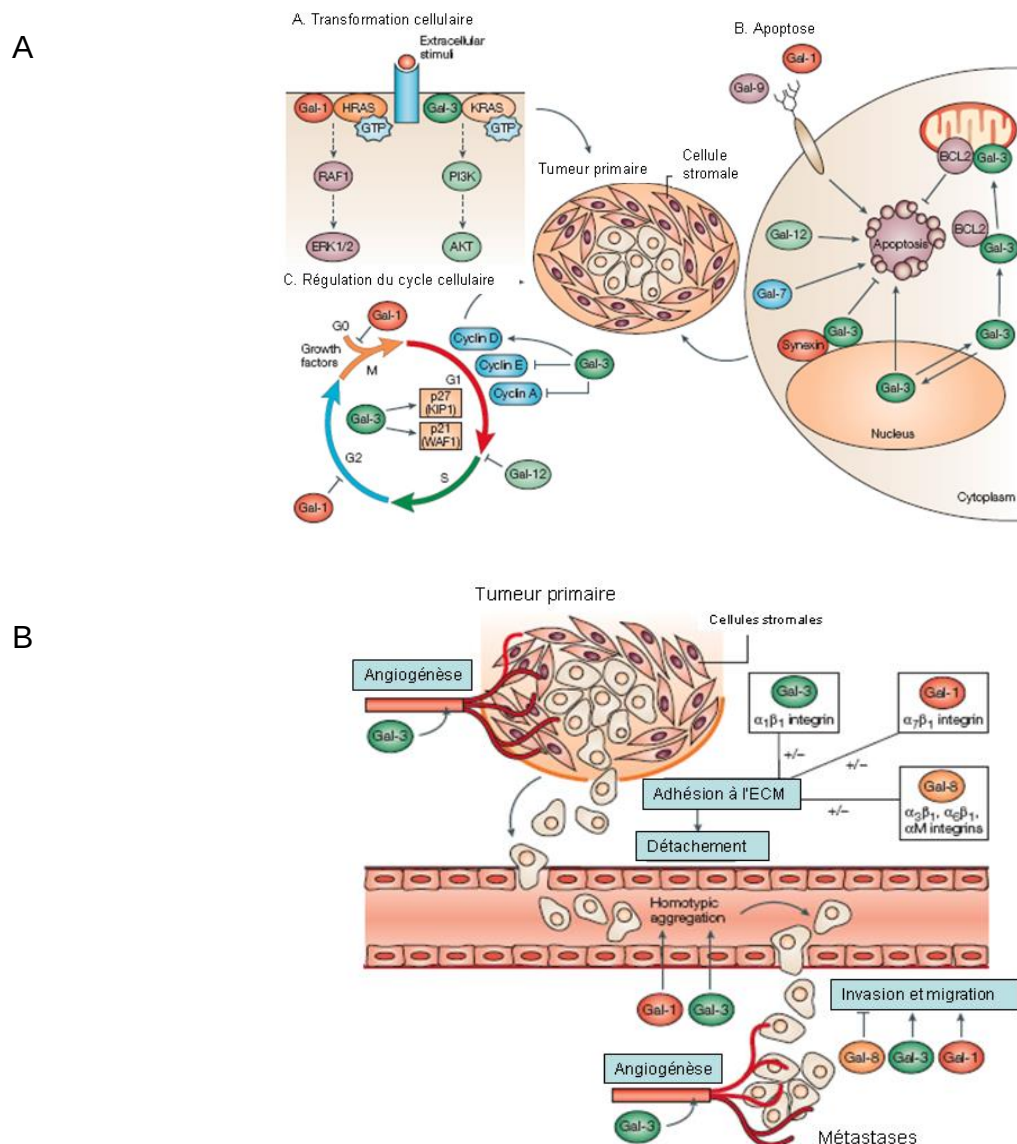


Figure 39 : Illustration de l'intervention de la galectine-3 dans les différentes étapes de la cancérogenèse (Liu and Rabinovich, 2005).

(A) Dans l'initiation du processus tumoral : la galectine-3 affecte la transformation cellulaire, l'apoptose et modifie le cycle cellulaire. (B) Dans le processus métastatique : la galectine-3 module l'angiogenèse, la migration et l'invasion cellulaire.

IV.2.7. Les inhibiteurs de galectine-3

Ces dernières années, de nombreux inhibiteurs de galectine-3 ont été développés et testés dans le cadre de différents types de cancers tels que le mélanome, le cancer colorectal. Il s'agit de molécules correspondantes soit à des inhibiteurs de type glucidique ciblant le CRD comme la pectine citrique et ses dérivés (GCS-100 testé dans les rechutes de leucémie lymphoïde chronique, GM-CT-01 testé dans le cancer colorectal) (Blanchard *et al.*, 2014) soit à des inhibiteurs non glucidiques (Thijssen *et al.*, 2015). Ainsi, le G3-C12, un inhibiteur

peptidique spécifique de la galectine-3, inhiberait le processus métastatique dans le cancer du sein mais des études complémentaires sur ce composé sont encore nécessaires (Blanchard *et al.*, 2014).

Récemment Zhang *et al.* (2016) ont montré que le RN1, un inhibiteur polysaccharidique de la galectine-3 riche en β -galactose, inhibe la croissance des cellules cancéreuses pancréatiques par blocage de l'interaction de la galectine-3 avec l'EGFR, les récepteurs *bone morphogenetic protein* (BMPR) et les intégrines entraînant l'inactivation des voies de signalisation sous-jacentes : EGFR/ERK/Runx1, BMPR/Smad/Id-3 et intégrine/FAK/JNK (figure 40A).

De son côté, en mimant l'antigène TF, le dérivé glucidique Lactulose-L-Leucine inhibe l'interaction entre la galectine-3 et l'antigène TF, perturbant ainsi l'adhérence des cellules tumorales entre elles ou avec les cellules endothéliales. De plus, cette molécule interfère les fonctions anti-apoptotiques de la galectine-3 (figure 40B) (Glinskii *et al.*, 2012).

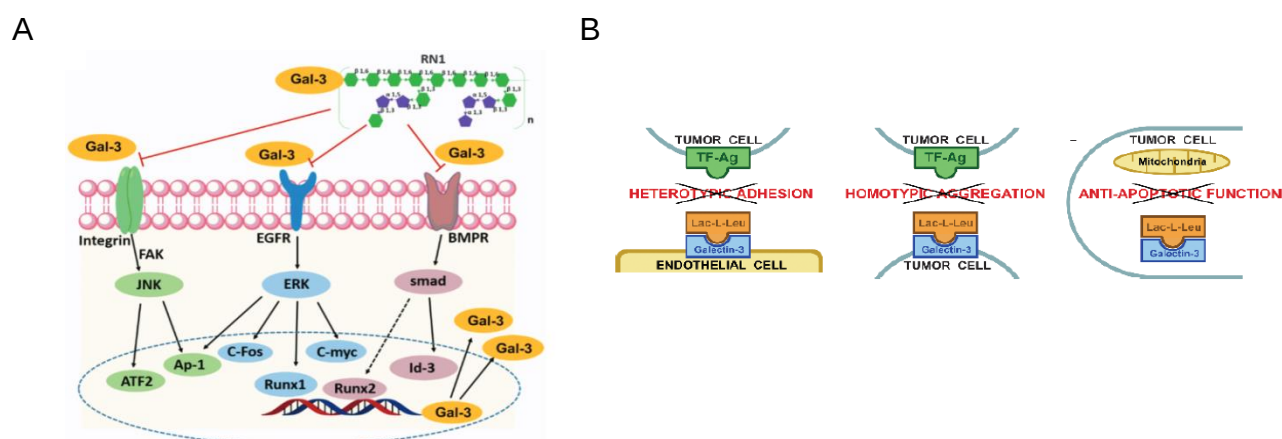


Figure 40 : Mécanismes d'action des deux inhibiteurs de la galectine-3 : (A) le RN1 (Zhang *et al.*, 2016) bloque l'interaction de la galectine-3 avec l'EGFR, BMPR et les intégrines et (B) le Lactulose-L-Leucine inhibe l'interaction entre la galectine-3 et l'antigène TF (Glinskii *et al.*, 2012).

IV.2.8. Galectine-3 et autres pathologies

Les galectines sont également impliquées dans d'autres pathologies comme l'asthme, les maladies inflammatoires, les fibroses et les maladies métaboliques. Par ailleurs, de par son rôle dans la fibrose cardiaque, la galectine-3 est devenu un marqueur pronostic d'insuffisance cardiaque approuvé par la FDA (*Food and Drug Administration*) (Li *et al.*, 2014).

La galectine-3 exerce un rôle dans l'inflammation par le biais de ses propriétés à la fois pro-et anti-inflammatoires (figure 41). Ainsi, lorsqu'elle est présente dans le milieu

extracellulaire au cours de la phase aigüe de l'inflammation, elle recrute, active et induit le *burst* respiratoire des polynucléaires neutrophiles et favorise leur adhérence aux cellules endothéliales. Elle stimule également la libération de médiateurs de l'inflammation par les mastocytes et exerce des propriétés similaires aux chemoattractants pour les monocytes et les macrophages (Dumic *et al.*, 2006). Ainsi, des taux élevés de galectine-3 sont retrouvés chez des patients atteints de pathologies inflammatoires comme l'arthrite. Dans ce dernier exemple, la galectine-3, produite intensément par la membrane synoviale enflammée participe à l'initiation et à la progression de la pathologie par ses propriétés pro-inflammatoires et pro-cataboliques (induction de l'expression des métalloprotéase dans les chondrocytes...) (Hu *et al.*, 2017).

Cependant, la galectine-3 participe à la résolution de l'inflammation au cours d'un processus chronique en vue de limiter les dommages tissulaires (Menini *et al.*, 2016). Ainsi, elle favorise l'acquisition d'un phénotype de type M2 anti-inflammatoire pour les macrophages et induit l'apoptose des lymphocytes T en se liant au complexe CD29/CD7 qui entraîne l'activation de voies de signalisation pro-apoptotiques conduisant à la libération du cytochrome c mitochondrial et à l'activation de la caspase 3 (Fortuna-Costa *et al.*, 2014). De plus, la galectine-3 participe au processus d'efferocytose des polynucléaires neutrophiles apoptotiques, c'est-à-dire à leur élimination. En effet, la galectine-3 est externalisée des polynucléaires neutrophiles et, en se fixant à leur surface, agit comme une opsonine en constituant un signal de clairance pour les cellules phagocytaires (Wright *et al.*, 2016).

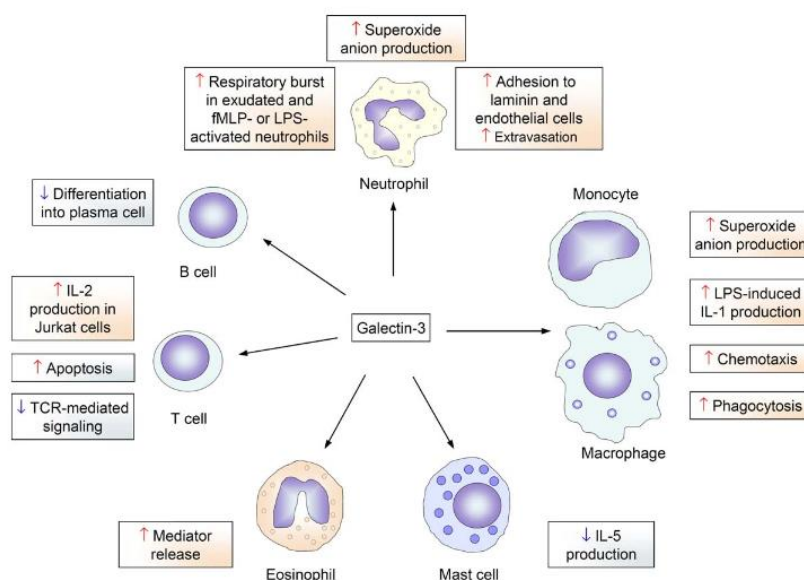


Figure 41 : Cibles de la galectine-3 dans l'inflammation (*via* son action sur les monocytes, macrophages, polynucléaires neutrophiles, éosinophiles, mastocytes et lymphocytes) (Dumic *et al.*, 2006). Abréviations : fMLP :formyl-méthionyl-leucyl-phenylalanine, LPS :lipopolysaccharide, TCR :T cell receptor.

Il existe un continuum entre inflammation et fibrose. En effet, l'inflammation chronique conduit à l'apparition d'un tissu cicatriciel, les macrophages stimulant la production de protéines de la matrice extracellulaire, comme la fibronectine, par les fibroblastes (Henderson and Sethi, 2009). La galectine-3 est donc impliquée dans le processus fibrotique et dans le remodelage tissulaire, notamment par la stimulation de la prolifération de fibroblastes, la synthèse des constituants de la matrice extracellulaire et par la libération de facteurs pro-inflammatoires. Son expression est ainsi augmentée au cours de la pancréatite chronique (Wang *et al.*, 2000). L'implication de la galectine-3 dans la fibrogenèse hépatique, rénale, pulmonaire et myocardique a été démontrée par différentes études (Li *et al.*, 2014). Elle participe ainsi à la fibrose hépatique par stimulation des cellules stellaires hépatiques produisant du collagène (figure 42) (Henderson *et al.*, 2006).

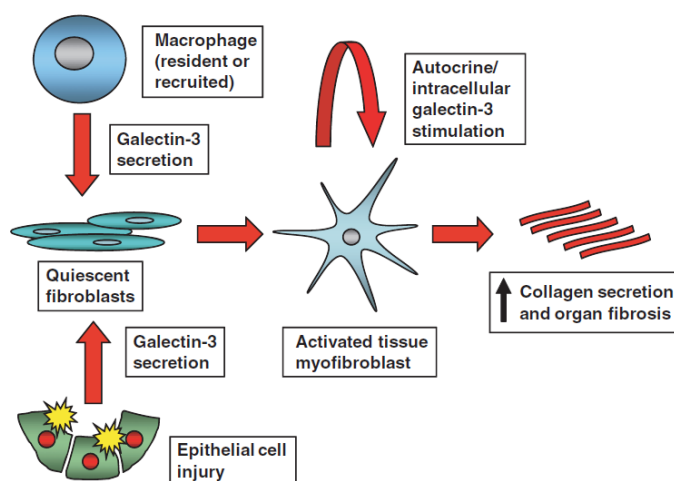


Figure 42 : Intervention de la galectine-3 dans le processus fibrotique *via* son action sur les fibroblastes (Henderson and Sethi, 2009).

L'activation des fibroblastes par la galectine-3 favorise la production de collagène et la fibrose.

La galectine-3 est également impliquée dans le développement des maladies métaboliques et cardiovasculaires comme le diabète de type 2 et l'obésité par le biais de sa participation à l'homéostasie glucidique, son rôle régulateur dans la méta-inflammation du tissu adipeux et des îlots pancréatiques (Menini *et al.*, 2016). En effet, la galectine-3 produite par les macrophages favorise d'une part le recrutement des monocytes et d'autre part, se lie sur le récepteur à l'insuline interférant ainsi avec l'activation des voies de signalisation sous-jacentes à l'origine du développement d'une insulino-résistance (figure 43) (Siwicki *et al.*, 2016). Par sa liaison aux produits de glycation avancée, elle exerce un rôle protecteur contre

les dommages rénaux et tissulaires créés par ces entités en favorisant leur élimination. Plusieurs études ont par ailleurs démontré que les taux circulants de galectine-3 sont augmentés chez les sujets obèses, diabétiques, souffrant d'hypertension ou d'hypercholestérolémie, ces taux étant inversement corrélés à ceux de l'hémoglobine A1c (Menini *et al.*, 2016).

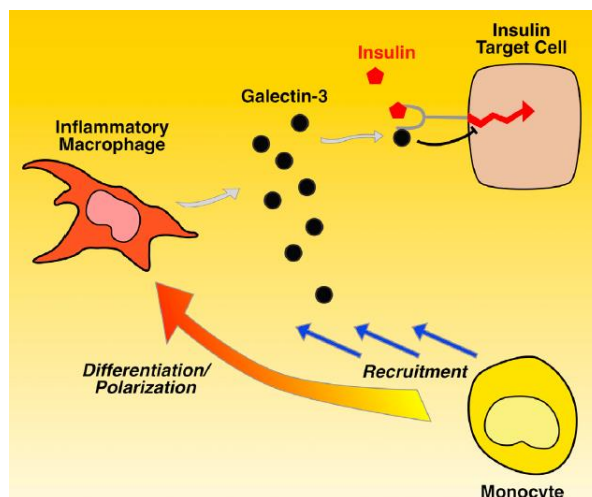


Figure 43 : Rôle de la galectine-3 dans l'insulino-résistance, dans le contexte de l'obésité (Siwicki *et al.*, 2016).

La galectine-3 produite par les macrophages agit en favorisant le recrutement des monocytes et de plus, *via* sa liaison sur le récepteur à l'insuline, interfère avec l'activation des voies de signalisation sous-jacentes à l'origine du développement d'une insulino-résistance.

IV.2.9. Modèles animaux

Afin d'étudier le rôle de cette galectine, des modèles murins invalidés pour le gène *LGALS3* ont été mis en place (KO, fonds génétique C57Bl/6 ou 129/Sv selon les études). Ces modèles sont viables et vivent aussi longtemps que les souris sauvages (WT). De plus ces souris ne développent pas spontanément de tumeur (Eude-Le Parco *et al.*, 2009).

a. Souris KO et anomalies du développement

Les souris KO présentent des anomalies de l'ossification endochondrale (c'est à dire la transformation du tissu cartilagineux en os) (Colnot *et al.*, 2001). Au niveau rénal, elles développent une hypertrophie rénale et leur nombre de glomérules rénaux est plus faible par rapport aux souris WT (Bichara *et al.*, 2006).

b. Souris KO et anomalies des cellules épithéliales

Ces modèles animaux ont permis de souligner l'importance du rôle de la galectine-3 dans la polarité cellulaire, le trafic intracellulaire des glycoprotéines ainsi que dans la morphogenèse épithéliale. En effet, Delacour *et al.* (2008) ont constaté chez les souris KO des altérations de la membrane basolatérale des entérocytes avec apparition de microvillosités à ce niveau (figure 44). De plus, la β -actine et la villine, deux marqueurs de la bordure en brosse, se retrouvent localisées au niveau de ces membranes basolatérales anormales. De la même façon, Koch *et al.* (2010) ont mis en évidence des anomalies de la ciliogénèse rénale chez ces souris KO avec apparition de cils plus longs et dysmorphiques. De plus, leurs travaux démontrent l'association de la galectine-3 avec le centrosome durant le processus de polarisation cellulaire. Ces deux études soulignent donc l'importance de la galectine-3 dans le maintien de la polarité des cellules épithéliales.

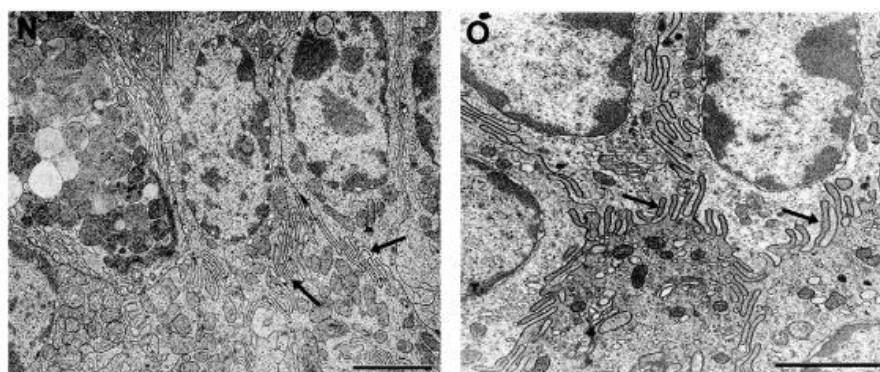


Figure 44 : Altération de la membrane basolatérale des entérocytes chez les souris invalidées pour le gène *LGALS3* avec apparition de microvillosités (Delacour *et al.*, 2008).

c. Souris KO et cancer

Afin d'évaluer le rôle de la galectine-3 dans la cancérogenèse, Eude-Le Parco *et al.* (2009) ont croisé les souris KO avec des souris génétiquement programmées pour développer un cancer (gène *Apc* non fonctionnel (*Adenomatosis Polyposis Coli*) ou gène *PyMT* (*polyomavirus middle T-antigen*) exprimé sous le contrôle du promoteur MMTV). Leurs travaux démontrent que l'absence de galectine-3 n'affecte pas l'évolution tumorale. Les auteurs expliquent ces observations par le fait que la galectine-3 ne jouerait peut être pas un rôle prédominant dans les cancers étudiés (sein, colon) ou qu'une galectine ayant des

fonctions redondantes à celle de la galectine-3 compenserait son absence dans les modèles invalidés.

d. Souris KO et fibrose

Les souris KO ont également servi à étudier le rôle de la galectine-3 dans la fibrose. Ainsi, les souris KO traitées par un agent induisant des dommages hépatiques développent moins de fibrose que les souris contrôles. Ceci serait en lien avec le rôle de la galectine-3 dans l'activation des myoblastes et dans les dépôts de collagène (Henderson *et al.*, 2006). De la même façon, les souris KO soumises à une alimentation athérogène sont protégées de l'apparition de la stéatose hépatique non alcoolique (NASH) (stéatose hépatique, inflammation, fibrose atténuées par rapport aux souris WT) (Iacobini *et al.*, 2011).

e. Souris KO et métabolisme

Les modèles murins sont contradictoires quant au rôle de la galectine-3 dans le métabolisme glucidique et le développement de l'obésité. Ainsi, Pang *et al.* (2013) ainsi que Pejnovic *et al.* (2013) indiquent que les souris KO soumises à une alimentation riche en graisses (*high fat*) présentent une altération du métabolisme glucidique associée à une augmentation de la masse grasse et une inflammation du tissu adipeux viscéral. L'accélération de l'obésité chez ces souris KO serait due à une inflammation systémique amplifiée et liée à l'effet anti-inflammatoire de la galectine-3 en situation chronique. A l'inverse, Baek *et al.* (2015) ont montré que les souris KO ont un poids plus faible et un tissu adipeux moins développé que les souris contrôles. De plus, dans leur étude, le taux de glucose n'est pas modifié entre les deux groupes. Les effets observés seraient liés à la diminution de l'expression PPAR γ en absence de galectine-3 et donc à la baisse de son activité transcriptionnelle.

f. Souris KO et inflammation

Enfin, le rôle de la galectine-3 dans l'inflammation est souligné par le fait que l'induction d'un remodelage des voies respiratoires par des allergènes est atténuée chez les souris KO par rapport aux souris WT avec une baisse de la production de mucus, de la fibrose épithéliale et de l'infiltration des polynucléaires éosinophiles (Ge *et al.*, 2010). De plus, le

recrutement des polynucléaires neutrophiles faisant suite à une infection par le streptocoque est diminué en absence de cette lectine (Nieminen *et al.*, 2008). Wright *et al.* (2016) ont également démontré que l'élimination des polynucléaires apoptotiques par les macrophages est réduite en absence de cette lectine.

Ces différents modèles nous montrent donc le rôle de la galectine-3 dans de nombreux processus et dans de nombreux tissus. C'est surtout dans la réponse à un stimulus que les différences entre les souris KO et WT sont observées.

Après avoir présenté les différents acteurs de notre étude, mucines membranaires et galectine-3, ainsi que les contextes cellulaires et physiopathologiques de leurs rôles connus respectifs, inflammation et cancer, nous allons nous intéresser aux mécanismes que nous avons explorés dans notre étude, à savoir la régulation post-transcriptionnelle en détaillant, dans la dernière partie de l'introduction, **le cycle de vie des ARNm**.

V. Le métabolisme et le devenir des ARN pré-messagers

La régulation de l'expression des gènes est un processus finement contrôlé et celle-ci intervient à différents niveaux : depuis la synthèse des ARNm à la traduction en protéines et au catabolisme de celles-ci en passant par la dégradation des ARNm.

Tout d'abord, la régulation de l'expression des gènes peut s'effectuer au niveau transcriptionnel par l'intermédiaire de la structure de l'ADN. En effet, la double hélice peut se présenter sous une forme lâche, appelée l'euchromatine, ou sous une forme compactée, l'hétérochromatine. Sous cette dernière forme, les gènes ne sont pas accessibles aux ARN polymérases. L'ADN est enroulé autour d'un cœur protéique constitué d'histones pour former un nucléosome. Les histones peuvent être principalement méthylées, acétylées et/ou phosphorylées. Ces modifications entraînent un changement dans la structure de la chromatine qui devient plus ou moins lâche, modulant ainsi l'accès des promoteurs aux facteurs de transcription (Pollard *et al.*, 2004).

L'ADN peut aussi être méthylé au niveau des répétitions de cytosines et de guanines appelées îlots CpG. Cette méthylation induit la condensation de l'ADN et entraîne l'inactivation transcriptionnelle des gènes (Kaplan and Delpech, 1989).

La régulation transcriptionnelle passe également par le recrutement et l'activation de facteurs de transcription. Ces protéines, capables d'interagir avec l'ADN, agissent en contrôlant l'assemblage de la machinerie transcriptionnelle. Cette dernière est un complexe comprenant l'ARN polymérase, enzyme synthétisant l'ARN pré-messager, et des facteurs de transcription dits « généraux » puisque nécessaires à la transcription de l'ensemble des gènes. Pour moduler de façon fine le niveau de transcription de certains gènes, les gènes eucaryotes possèdent, au niveau de leurs régions promotrices, des éléments *cis*-régulateurs reconnus par des facteurs de transcription dits « spécifiques » (activateurs ou répresseurs) de certains types cellulaires et qui augmentent ou diminuent le niveau de transcription de ces gènes (Watson *et al.*, 2009).

La quantité d'ARNm présents dans la cellule à un moment donné est la résultante du pool des ARNm nouvellement synthétisés (taux de transcription) et du pool des ARNm dégradés. La régulation de l'expression des gènes s'exerçant aux niveaux des ARN messagers met en jeu des mécanismes co- et post-transcriptionnels et intervient à différentes étapes de la

vie des ARNm : maturation, transport, localisation subcellulaire, traduction et stabilité/dégradation (figure 45). Cette régulation du métabolisme des ARNm permet une réponse cellulaire adaptative rapide face à des stimuli variés (Adeli, 2011).

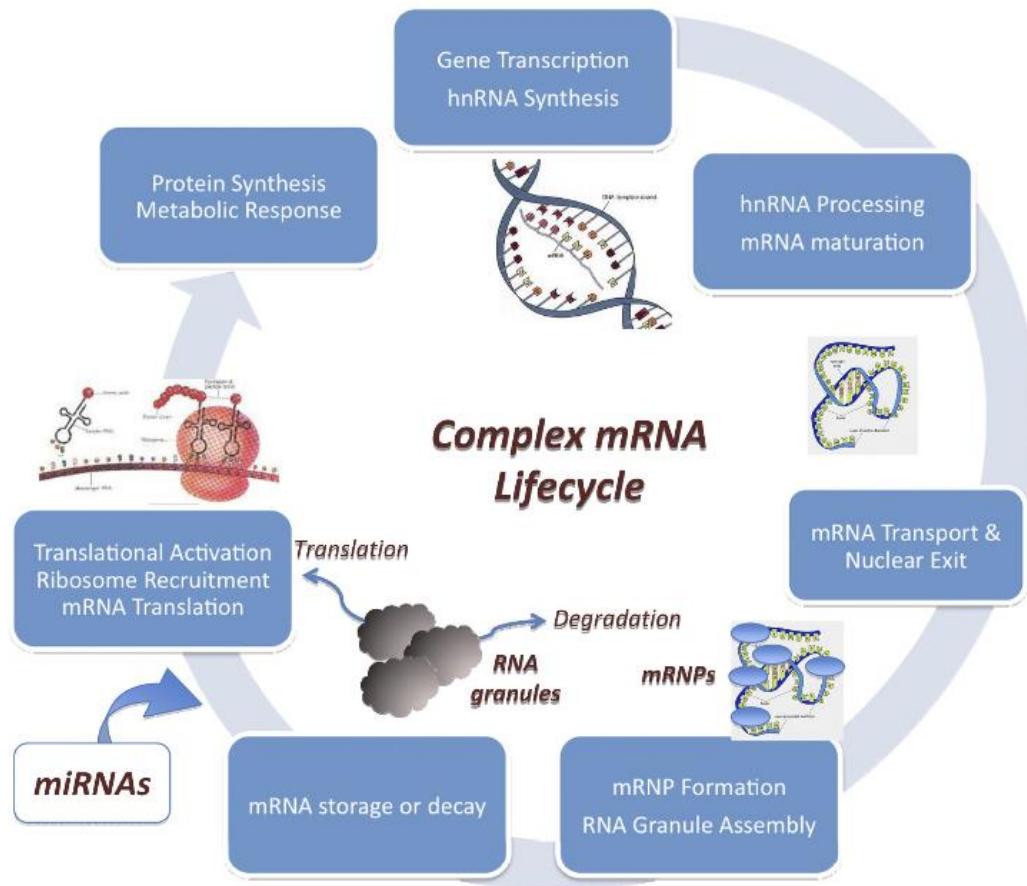


Figure 45 : Cycle de vie ARNm : de la transcription des gènes à la synthèse protéique (Adeli, 2011).

Ce schéma résume les différents moments-clés de la vie des ARNm intervenant dans la régulation de l'expression des gènes.

Lorsque l'ARNm mature atteint le cytoplasme, sa traduction en protéine est, elle aussi, régulée. L'initiation de la traduction peut être mise en suspens par des protéines régulatrices qui se lient au niveau de la région 5'UTR de l'ARNm et qui interfèrent par conséquent avec les fonctions des ribosomes. Un exemple classique est celui, en cas de déficit en fer, de l'*iron regulatory protein* (IRP) : celle-ci se lie à l'ARNm de la ferritine et empêche sa traduction en bloquant la progression des ribosomes le long de l'ARNm (Staton *et al.*, 2000).

En dehors de certains aspects de l'initiation de la traduction, la régulation de l'expression des gènes au niveau traductionnel et post-traductionnel ne sera pas abordée dans ce chapitre, nous nous intéresserons en effet plus particulièrement au cycle de vie des ARNm.

V.1. Les acteurs du métabolisme des ARN : les RNA-binding proteins (RBP)

V.1.1. Les RBP

Les ARN, quels qu'ils soient, sont accompagnés tout au long de leur cycle de vie par des protéines. Ils possèdent dans leur séquence nucléotidique des éléments *cis* reconnus par des protéines appelées *RNA-binding proteins* (RBP, facteurs *trans*) qui peuvent se fixer au niveau de leurs régions codantes, ou des 3'UTR et 5'UTR (*Untranslated Transcribed Region*). Des éléments structuraux particuliers tels que par exemple la coiffe (*cap*) ou la queue poly(A) des ARNm peuvent également être reconnus par ce type de protéines. L'association des ARN à ces protéines forme les particules ribonucléoprotéiques appelées RNP, et mRNP lorsqu'il s'agit d'ARN messagers (Singh *et al.*, 2015).

Le génome humain compte au moins 1542 gènes codant des RBPs, soit plus de 7,5 % des gènes humains codant des protéines et leur inventaire n'est pas encore terminé. On peut même dire que la connaissance des RBP humaines n'en est qu'à ses débuts. Un travail important a été réalisé par des curateurs des données du génome humain, relatives aux RBP, et des éléments possibles de classification ont été mis en lumière. Les protéines retenues à l'inventaire l'ont été sur la base de données bioinformatiques (obtenues à l'aide de logiciels d'analyse et de comparaison des séquences des gènes et des protéines), et d'éléments bibliographiques d'analyses fonctionnelles (isolement par immunoprécipitation de complexes ribonucléoprotéiques (RNP) et analyse protéomique par spectrométrie de masse) (Calabretta and Richard, 2015; Gerstberger *et al.*, 2014a, 2014b).

Au total, il est estimé aujourd'hui qu'environ 50% des RBP identifiées dans le génome humain peuvent avoir un effet direct ou indirect sur les voies de régulation post-transcriptionnelle des gènes et donc sur le destin intracellulaire métabolique des ARN messagers (Gerstberger *et al.*, 2014a). Une autre approche pour la connaissance du rôle et des fonctions des RBP est celle de la génétique. En effet, plus de 150 RBP apparaissent dans la

base OMIM (*Online Mendelian Inheritance in Man*) des maladies génétiques (Gerstberger *et al.*, 2014b). Il s'agit majoritairement de maladies génétiques neurologiques, de maladies du développement, de retard mental, de susceptibilités aux cancers, de maladies métaboliques.

On peut individualiser des classes fonctionnelles de RBP en fonction **du type d'ARN** auquel elles se fixent en distinguant celles liant les ARNm, les ARNt (ARN de transfert), les pré-rRNA (preribosomal RNA), les snRNA (small nuclear RNA), les snoRNA (small nucleolar RNA), les ncRNA (non-coding RNA), mais il reste encore celles de fonctions inconnues, en particulier celles liant plusieurs types d'ARN ou encore celles liant des régions d'ARN particulières (double brin par exemple, présentes dans les pré-microARN et les ARNm) (figure 46) (Gerstberger *et al.*, 2014a).

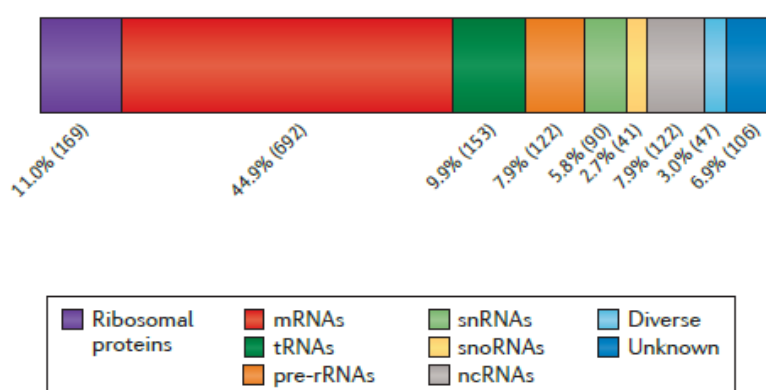


Figure 46 : Type d'ARN associés aux RBP (Gerstberger *et al.*, 2014a).

Les RBP peuvent aussi être classées sur **la base structurale de leurs RNA-binding domains** : taille du motif ARN reconnu (4-6 nucléotides, plus pour certains, moins pour d'autres, le domaine PUF (*Pumilio and FBF homology protein*) reconnaît un seul nucléotide), taille et séquence des motifs peptidiques, mode d'interaction avec l'ARN (simple brin ou double brin) (Gerstberger *et al.*, 2014a). Voici quelques domaines à titre d'exemple (figure 47) :

-le domaine RRM (*RNA recognition domain*) reconnaissant une séquence courte de 4 à 6 nucléotides, comme dans la famille ELAVL (*Embryonic Lethal, Abnormal Vision, Drosophila-Like*), impliquées dans l'épissage des ARNm) ;

- le domaine KH (*heterogeneous nuclear RNP K-homology domain*, présent par exemple dans la protéine IGF2BP1 (*insulin-like growth factor 2 mRNA-binding protein*)) ;
- le motif DSRM (*double-stranded RNA-binding motif*, présente par exemple dans la famille Staufen impliquée dans la dégradation des ARNm) ;
- le domaine en doigt de zinc (exemple : dans la famille MBNL, *Muscleblind-like proteins*, impliquées dans l'épissage alternatif des ARNm).

L'association de plusieurs domaines/motifs renforce la spécificité de l'interaction ARN-protéine. Collectivement les *RNA-binding domains* représentent un vaste répertoire d'environ 600 classes de domaines de liaison distincts. Et parmi ces 600 classes, seulement 20 classes comportent plus de 10 membres, ce sont celles connues pour lier les ARN messagers (mRBP) (Gerstberger *et al.*, 2014a, 2014b).

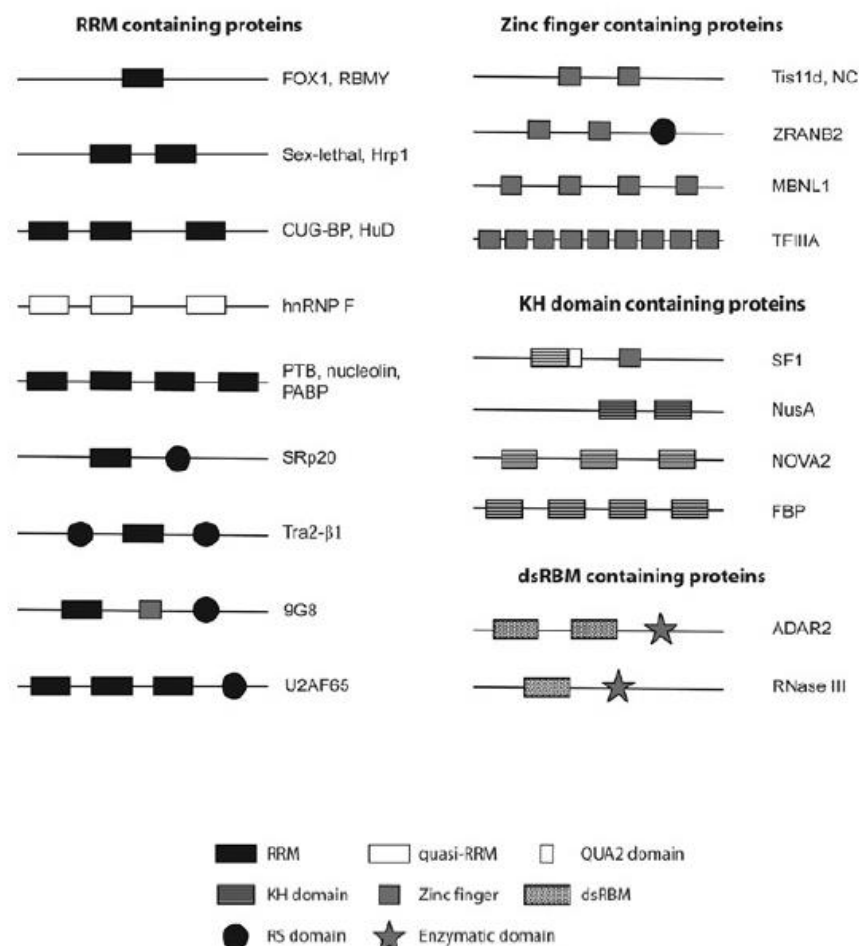


Figure 47 : Classification des RBP selon leur type de domaine de liaison à l'ARN (Cléry and Allain, 2013).

D'autres éléments sont à prendre en considération pour la classification des RBP, comme **l'existence de motifs supplémentaires, ayant un rôle catalytique**, comme par exemple de type ribonucléase, endonucléase, télomérase, hélicase ou encore ayant un rôle encore inconnu (*vault RNA* : fonction *silencing* d'ARN ? ; *Y RNA* : fonction de chaperonne ?), ces motifs pouvant d'ailleurs amener certaines RBP à agir également sur l'ADN. A titre d'exemple, nous pouvons citer la large classe des DExD/H hélicases ATP-dépendantes, qui peuvent contenir 2 types de domaines de liaison à l'ARN : les motifs DEAD (Asp-Glu-Ala-Asp) ou DEAH (Asp-Glu-Ala/X) qui jouent un rôle de remodelage des RNP (Gerstberger *et al.*, 2014a, 2014b).

Un autre critère possible de classification est **d'ordre cellulaire** et concerne la localisation subcellulaire de la RBP, soit nucléaire, avec un rôle supposé dans l'épissage, ou soit, au contraire, cytoplasmique, avec un rôle supposé post-transcriptionnel. Le problème est que bon nombre de RBP se comportent comme des navettes et sont capables d'un trafic entre noyau et cytoplasme et jouent vraisemblablement des rôles multiples (Gerstberger *et al.*, 2014a, 2014b).

Les complexes entre ARN et RBP sont des structures dynamiques qui peuvent jouer un rôle à un moment ou un autre du cycle de vie des ARN, au niveau de leur maturation, de leur transport, de la régulation de leur métabolisme, du remodelage des complexes, ou encore de chaperonne, lors du franchissement des membranes. Certaines protéines qui n'ont pas de site direct de liaison à l'ARN participent au métabolisme des ARN via des liaisons de type protéine-protéine qui leur permettent de jouer un rôle de « plateforme » ou « d'échafaudage » dans la constitution d'édifices macromoléculaires actifs. Tous ces éléments permettent de comprendre la complexité de la catégorisation des RBP (Gerstberger *et al.*, 2014a).

V.1.2. Les mRBP

Sur les 692 RBP se liant aux ARNm (mRBP) estimées dans le génome humain, 405 contiennent un ou plusieurs types de domaines/motifs de liaison à l'ARNm qui peuvent être répétés chacun en nombre variable ou en combinaison variable (Gerstberger *et al.*, 2014a). Elles sont classées selon le type de *RNA-binding domain* qu'elles possèdent et leur liaison directe à l'ARN est attendue. A côté de ces mRBP « classiques », il existe des mRBP dites « non conventionnelles », qui ne possèdent pas ces domaines, mais au contraire, des régions

désordonnées (sans structure tertiaire) telles que des régions de faible complexité contenant des motifs répétés riches en résidus glycines, lysines, arginines et même tyrosines (Calabretta and Richard, 2015). Cette absence de structure tertiaire confère aux protéines une grande flexibilité pour leur interaction avec l'ARN, l'ADN ou avec d'autres protéines. D'autre part, les propriétés de multimérisation de certaines mRBP favorisent la reconnaissance simultanée de plusieurs ARNm possédant ce motif ou l'interaction avec plusieurs motifs séparés mais présents sur le même ARNm.

D'autres niveaux de complexité pour leur étude sont encore à souligner :

- un type d'ARN messager peut être associé à de nombreuses mRBP et chaque type de mRBP peut avoir des centaines d'ARN-cibles différents (Müller-McNicoll and Neugebauer, 2013) ;
- pour un même transcrit, la composition des protéines au sein des mRNP le concernant peut varier car influencée par différents paramètres physiopathologiques ;
- l'utilisation de promoteurs alternatifs ou de sites alternatifs de polyadénylation, peut encore constituer vis-à-vis de l'ARN une source de gain ou de perte de séquences de fixation pour certaines RBP (Singh *et al.*, 2015).

Les mRBP et les autres protéines qui leur sont associées, par des interactions de type protéine-protéine, sont des éléments importants pour la régulation de l'expression des gènes. Parmi les mRBP, les mieux connues, nous pouvons citer les protéines hnRNP (*Heterogeneous nuclear ribonucleoproteins*, Geuens *et al.*, 2016), les protéines SR (possédant un domaine *serine/arginine-rich*, Twyffels *et al.*, 2011), les ARE-BP (*AU-rich element binding proteins* (Barreau *et al.*, 2005) dont nous reparlerons dans l'exposé des travaux.

Le rôle des mRBP est important à toutes les étapes de la vie de l'ARNm : depuis la transcription et la maturation des transcrits primaires dans le noyau jusqu'à leur traduction et leur dégradation dans le cytoplasme. La composition des RBP associées aux ARNm est dynamique et change constamment au cours des différentes étapes du cycle de vie des ARNm. Elle détermine donc le devenir de l'ARNm de façon spatiotemporelle : immédiatement traduit ou transporté dans des zones de stockage ou de traduction localisées (Anderson *et al.*, 2015). De plus, une même mRBP est souvent impliquée dans de nombreuses étapes de la vie des ARNm. La protéine HuR (*human antigen R*) illustre bien ce phénomène (figure 48) : il s'agit d'une protéine RBM (*RNA binding motif*) (gène *ELAVL1*), qui participe

notamment à l'épissage, à la stabilisation de nombreux types d'ARNm-cibles et à leur traduction (Srikantan and Gorospe, 2011). De la même façon, les protéines hnRNP interviennent à différentes étapes du cycle de vie de l'ARNm : transcription, épissage, export nucléaire, stabilisation des transcrits et traduction (Geuens *et al.*, 2016). Certaines protéines hnRNP ont une localisation strictement nucléaire et sont donc détachées de l'ARNm avant son export alors que d'autres restent associées à l'ARNm lors de sa sortie du noyau et sont donc également retrouvées dans le cytoplasme.

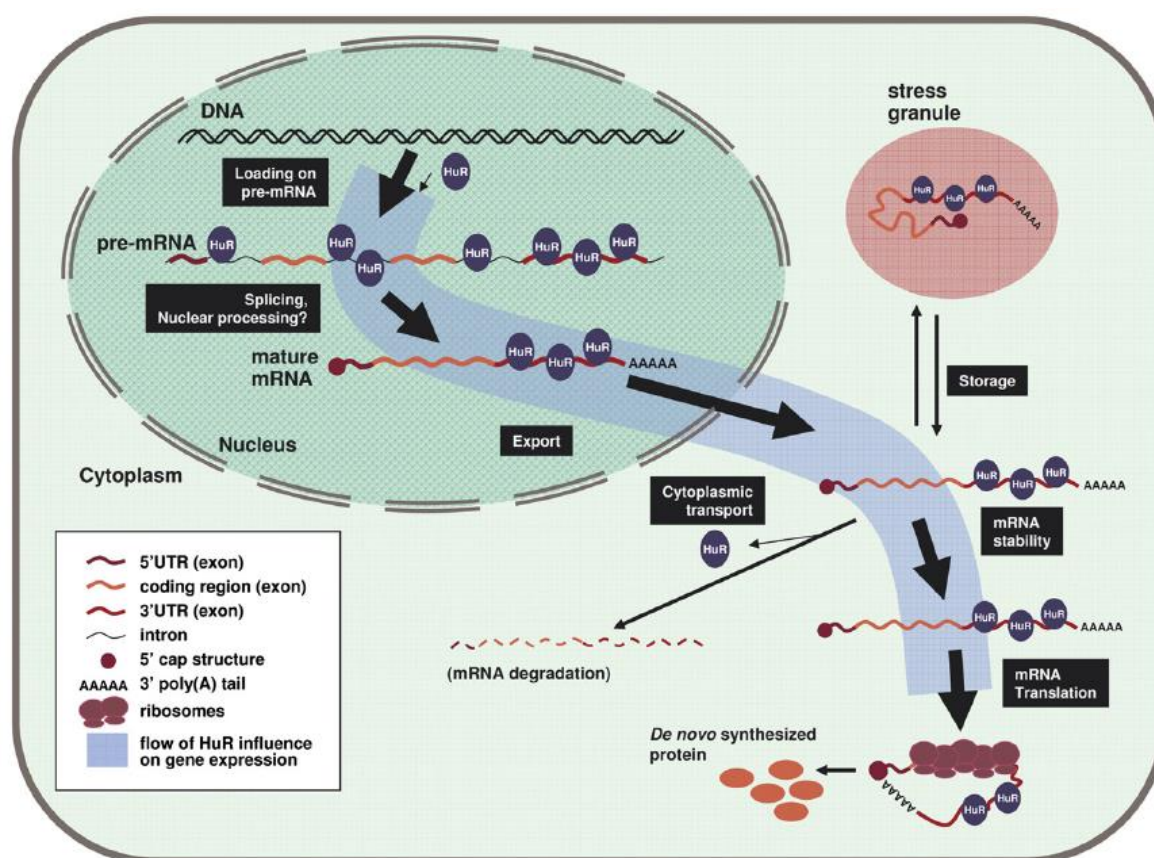


Figure 48 : Illustration de l'intervention de la protéine HuR à différents niveaux du métabolisme des ARNm : épissage, export, stabilité, traduction... (Srikantan and Gorospe, 2011).

Des mRBP classiques et des protéines non conventionnelles, candidates à cette fonction, peuvent coopérer et agir collectivement sur le destin intracellulaire des ARN messagers. Elles peuvent être identifiées ensemble au sein d'interactomes dont les différents partenaires protéiques possèdent le même registre fonctionnel, définissant ainsi des réseaux. Les enzymes métaboliques à fonction de RBP, comme l'aconitase, décrits par Castello *et al.* (2012) sur le modèle cellulaire HeLa en sont un exemple (tableau XI). Les réseaux REM

(ARN, enzyme, métabolite) identifiés par ces auteurs et impliqués dans les voies majeures du métabolisme intermédiaire (métabolisme des sucres, acides aminés, lipides, nucléotides) permettent d'établir des connexions entre le métabolisme intermédiaire, la biologie des ARN et la régulation post-transcriptionnelle.

Tableau XI : Illustration de quelques enzymes métaboliques à fonction de RBP (Castello *et al.*, 2012).

Protein	Class	Pathway	Cofactor
ALDH18A1	mRNA interactome	amino acids and derivatives	dinucleotide
PKM2	mRNA interactome	carbohydrates	nucleotide and anionic substrate
ENO1	mRNA interactome	carbohydrates	
LTA4H	mRNA interactome	lipids and lipoproteins; prostanoid metabolism	
ALDH6A1	mRNA interactome	amino acids and derivatives	dinucleotide
DUT	mRNA interactome	nucleotides; pyrimidine metabolism	
ASS1	mRNA interactome	amino acids and derivatives	nucleotide and anionic substrate
TXN	mRNA interactome	nucleotides	
HADHB	mRNA interactome	fatty acids; triacylglycerol and ketone body	dinucleotide
MDH2	mRNA interactome	pyruvate and TCA	dinucleotide
ADK	mRNA interactome	nucleotides; purine	nucleotide and anionic substrate
FDPS	mRNA interactome	lipids and lipoproteins	
SUCLG1	mRNA interactome	pyruvate and TCA cycle	nucleotide and anionic substrate
FASN	mRNA interactome	fatty acids, ketone, vitamins, and cofactors	dinucleotide
NQO1	mRNA interactome	amino acids and derivatives	dinucleotide
P4HB	mRNA interactome	lipids and lipoproteins	
NME1	mRNA interactome	nucleotides	nucleotide and anionic substrate
SHMT2	candidate RBP	amino acid and folate	
AK2	candidate RBP	nucleotides	nucleotide and anionic substrate
CPS1	candidate RBP	amino acids and derivatives	nucleotide and anionic substrate
SDHA	candidate RBP	pyruvate and TCA cycle	dinucleotide
CAD	candidate RBP	nucleotides; pyrimidine	
AKR1B1	candidate RBP	lipids and lipoproteins; steroid hormones, and vitamins A and D	dinucleotide
BLVRB	candidate RBP	porphyrins	dinucleotide
DLD	candidate RBP	amino acids and pyruvate; TCA cycle	dinucleotide
MTHFD1	candidate RBP	vitamins and cofactors; folate and pterines	dinucleotide
GMPT2	candidate RBP	nucleotides; purine	dinucleotide
PGK1	candidate RBP	carbohydrates: glucose	nucleotide and anionic substrate
DECR1	candidate RBP	fatty acid, triacylglycerol, and ketone body	dinucleotide
ENO3	candidate RBP	carbohydrates; glucose	

V.2. Le cycle de vie de l'ARNm

V.2.1. Maturation

Les ARNm sont d'abord synthétisés dans le noyau sous une forme immature. Ces pré-ARNm vont subir une maturation comportant l'ajout d'une coiffe en 5', d'une queue poly(A) en 3', un épissage et parfois un processus d'*editing* (modification de certains nucléotides). Chacune de ces étapes est sujette à régulation (Parent and Bisailon, 2006). Ces processus surviennent de façon co- ou post-transcriptionnelle et sont interconnectés grâce à différents facteurs protéiques qui interagissent entre eux. Cela fait notamment intervenir le domaine carboxy-terminal (CTD) de l'ARN polymérase de type II qui leur sert de « plateforme » d'interaction (Parent and Bisailon, 2006).

a. Epissage

L'épissage des pré-ARNm est un processus participant au contrôle de l'expression des gènes et qui se produit de façon co-transcriptionnelle. Il correspond à l'élimination des introns présents dans les pré-ARNm associée à une ligation des exons adjacents. Ce processus est assuré par le *spliceosome*, complexe dynamique en constant remodelage au cours du processus d'épissage (Sibley, 2014). Celui-ci est formé de 5 snRNP (*small nuclear ribonucleoproteins*) différentes (U1snRNP à U6snRNP) couplées à de nombreux cofacteurs protéiques. Chaque snRNP est elle-même formée d'un snRNA associé à des protéines (Matera and Wang, 2014).

Le processus d'épissage est basé sur la reconnaissance par le *spliceosome* de courtes séquences conservées au niveau des jonctions intron-exon. Le site 5' d'épissage est localisé au début de l'intron (AG/GURAG) et le site 3' d'épissage (enchaînement polypyrimidique suivi de AG) à la fin de cet intron. Le point de branchement (YURAY) est situé 15 à 50 nt en amont du site 3' d'épissage. Le choix des sites d'épissage est régulé par de nombreux éléments *cis* présents dans le pré-ARNm (Matera and Wang, 2014).

Le processus d'épissage fait intervenir plusieurs étapes (figure 49) :

- U1snRNP reconnaît l'extrémité 5' d'épissage grâce à un appariement entre U1snRNA et le pré-ARNm, ce qui forme le complexe E. Ceci est facilité par CTD qui interagit avec U1snRNP. L'interaction entre U1snRNP et l'ARNm est stabilisée par les protéines SR et le *cap binding complex* (CBC) se liant à la coiffe. Les facteurs

d'épissage SF1 (*splicing factor 1*) et U2AF (*U2 auxiliary factor*) reconnaissent quant à eux le point de branchement et le site 3' d'épissage.

- Ce 2 derniers facteurs permettent ensuite le recrutement de U2snRNP au point de branchement, puis U2snRNP interagit avec U1snRNP ce qui rapproche les extrémités 5' et 3' d'épissage et forme le complexe A.
- Puis le complexe U4snRNP-U6snRNP-U5snRNP est recruté pour former le complexe B. Le spliceosome ainsi formé subit des réarrangements puisqu'il y a libération de U4snRNP et de U1snRNP. Le complexe formé des snRNP restantes devient actif et réalise une première réaction de transestérification : le groupement 2'hydroxyl de l'adénosine de branchement attaque la liaison phosphodiester du site 5' d'épissage. Ceci permet la coupure de l'extrémité 5' de l'intron et la formation d'un lasso. Le complexe C est formé.
- Une seconde réaction de transestérification permet le clivage de l'extrémité 3' de l'intron. Pour cela, le groupement 3'hydroxyl de l'exon en 5' attaque la liaison phosphodiester du site 3' d'épissage permettant la ligation des exons et la libération de l'intron. L'intron est donc excisé, les 2 exons associés et U2snRNP-U5snRNP-U6snRNP sont ensuite libérés du complexe.

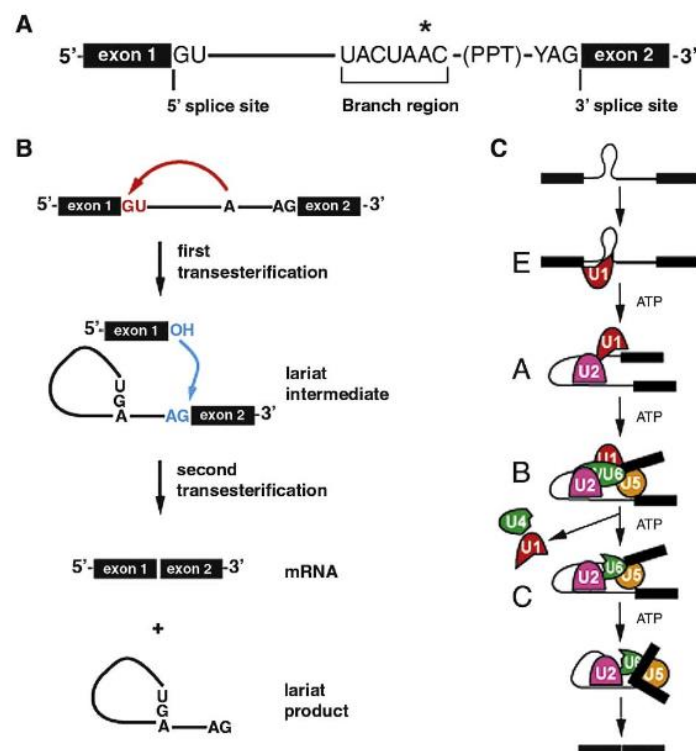


Figure 49 : Illustration du processus d'épissage des pré-ARNm avec formation du lasso (à gauche) et protéines intervenant dans ce processus (à droite) (Ritchie *et al.*, 2009).

Au cours du processus d'épissage, le complexe *Exon junction complex* (EJC) dont les protéines UPF2 et UPF3 (*up-frameshift suppressor homolog*) font parties, est déposé au niveau des jonctions exon/exon 24 nt en amont de la jonction épissée (figure 50). Il sert de «trace» de ce processus, participe à l'export nucléaire des ARNm et reste assemblé sur l'ARNm jusqu'au premier cycle de traduction (Le Hir *et al.*, 2016; Sibley, 2014).

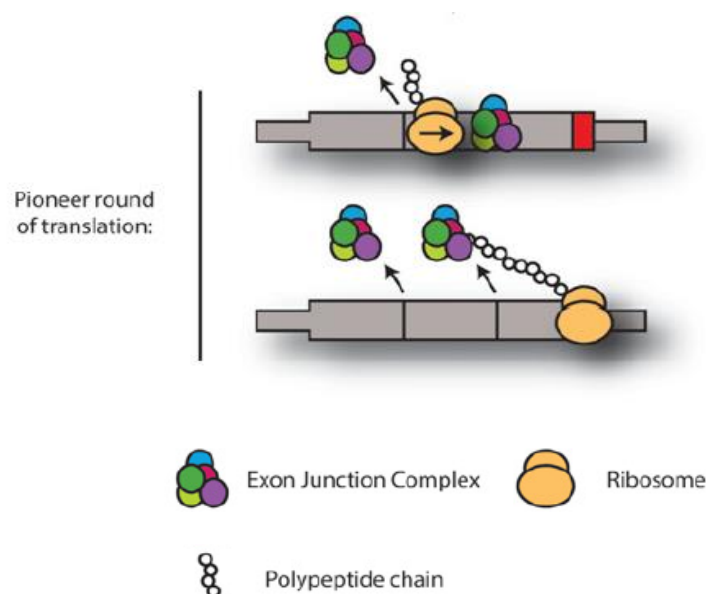


Figure 50 : Dépôt de l'EJC au niveau des jonctions exons/exons et élimination au premier cycle de traduction (Sibley, 2014).

L'épissage est appelé « constitutif » quand un exon est systématiquement inclus dans l'ARNm mature. A l'inverse, l'épissage dit « alternatif » est un processus permettant de générer, à partir d'un seul pré-ARNm, des transcrits qui diffèrent entre eux par leurs régions codantes ou leurs UTR en fonction de la sélection des régions lors du processus l'épissage. Ils résultent par exemple de l'utilisation de sites alternatifs d'épissage en 5' et 3' ou sont caractérisés par une rétention d'intron, un saut d'exon... Les sites d'épissage peuvent être « forts » ou « faibles » en fonction de leur rapprochement à la séquence consensus et ces variations influencent l'affinité pour les facteurs d'épissage. C'est la proximité d'un site « fort » et d'un site « faible » entrant en compétition qui est à l'origine de l'épissage alternatif (Kornbliht *et al.*, 2013). Un site « fort » conduit à un épissage constitutif alors que le recours à un site « faible » va dépendre d'éléments *cis* présents dans la séquence du pré-ARNm qui sont soit des *silencers* ou soit des *enhancers*. En fonction de leur localisation exonique ou intronique, ces séquences sont appelées ESE (*exonic splicing enhancer*) ou ESS (*exonic splicing silencer*), ISE (*intonic splicing enhancer*) ou ISS (*intonic splicing silencer*). Ces

motifs sont reconnus par des RBP telles que les protéines SR et les protéines hnRNP qui activent ou inhibent le processus d'épissage en affectant l'assemblage du spliceosome (figure 51) (Kornblihtt *et al.*, 2013). Les protéines hnRNP liées à une séquence *silencer* ont un effet répresseur sur l'épissage en empêchant l'accès aux constituants du spliceosome. Les protéines SR associées à une séquence *enhancer* ont quant à elles une fonction activatrice d'épissage en favorisant le recrutement des snRNP (Fu and Ares, 2014). D'autre part, la présence de structures secondaires dans le pré-ARNm ainsi que les RBP interviennent dans l'épissage alternatif en affectant l'accessibilité des sites d'épissage.

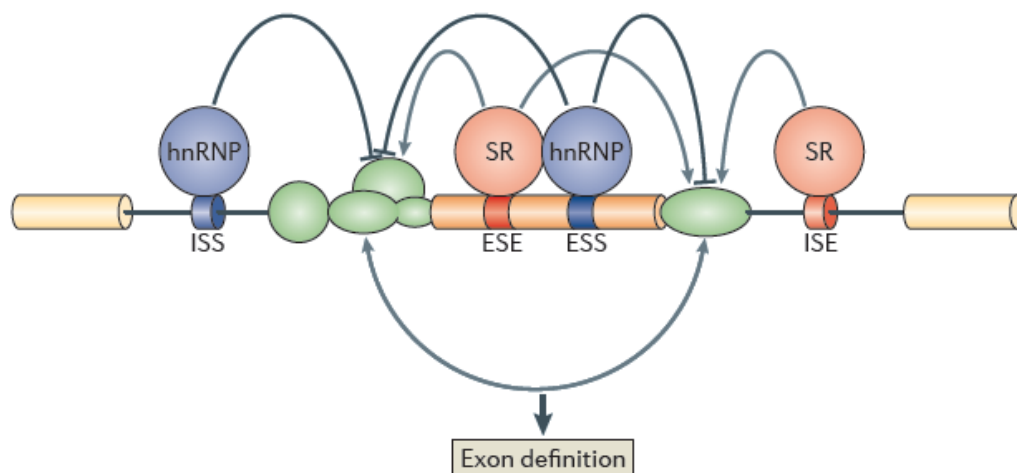


Figure 51 : Influence des séquences ESS, ESE, ISS, ISE (*exon/intron splicing enhancer/silencer*) reconnues par les protéines SR et hnRNP sur la sélection des sites d'épissage (Kornblihtt *et al.*, 2013).

La vitesse d'élongation de l'ARN polymérase II influence également l'épissage alternatif : plus sa vitesse est ralentie, plus l'épissage alternatif est efficace et inversement. Ceci s'explique par le fait qu'un ralentissement de la vitesse d'élongation au niveau d'un site d'épissage « faible » laisse plus de temps aux facteurs d'épissage de le reconnaître avant qu'un site compétiteur fort ne soit transcrit en aval favorisant ainsi la rétention de l'exon (Moehle *et al.*, 2014). A l'inverse, une vitesse rapide favorisera le saut de ce même exon (figure 52).

Différents facteurs peuvent moduler la vitesse d'élongation de l'ARN polymérase :

- la structure de la chromatine, via les modifications des histones, impacte la vitesse d'élongation de l'ARN polymérase ;

- la présence des nucléosomes, qui imposent des pauses à l'ARN polymérase II, participent à l'épissage alternatif ;
- la méthylation de l'ADN qui induit une structure chromatinienne dense, limite la vitesse d'élongation.

D'autres éléments influencent l'épissage alternatif. Ainsi, les histones peuvent recruter des protéines qui interagissent avec des facteurs régulateurs d'épissage tels que les protéines SR. La signalisation cellulaire intervient par le biais de modifications post-traductionnelles de protéines intervenant dans ce processus. Ainsi la voie MAPK régule l'épissage alternatif de *CD44* par phosphorylation du facteur d'épissage SAM68 (*Src-Associated substrate in Mitosis of 68 kDa*). De plus, la localisation cellulaire de régulateurs d'épissage peut influencer ce processus (Kornblihtt *et al.*, 2013).

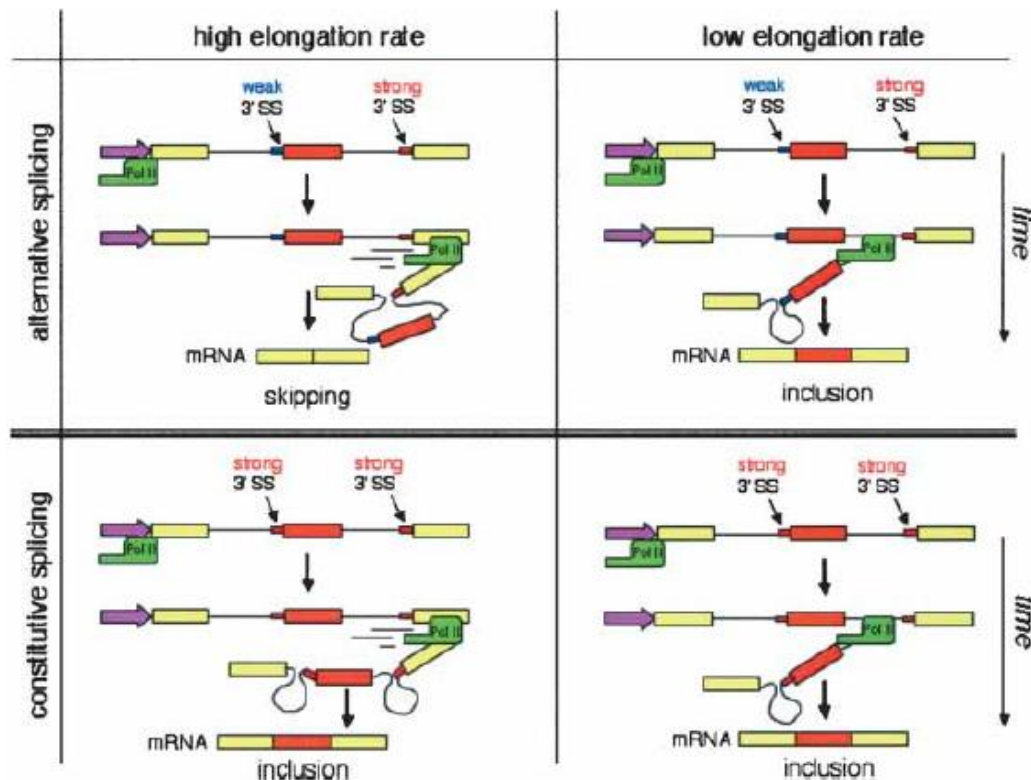


Figure 52 : Influence de la vitesse d'élongation de l'ARN polymérase lors du processus d'épissage (Kornblihtt *et al.*, 2004).

Une vitesse rapide favorise le saut d'exon lors de l'épissage alternatif car le site d'épissage faible n'a pas le temps d'être reconnu par les facteurs participant au processus d'épissage. A l'inverse, une vitesse lente favorise la reconnaissance de ce motif par ces facteurs lors de l'épissage alternatif et entraîne donc une rétention de cet exon. Au cours de l'épissage constitutif (sites d'épissage forts), l'exon est inclus quelque soit la vitesse d'élongation de l'ARN polymérase.

b. Ajout de la coiffe et de la queue poly(A)

Lorsque le transcrit nouvellement synthétisé est formé de 20 à 30 nt, une coiffe est ajoutée à son extrémité 5' afin de la protéger de la dégradation par les 5'-3' exonucléases. Ce processus (figure 53) consiste en l'addition d'un résidu 7-méthylguanosine sur le premier nucléotide de l'ARNm. L'ajout de la coiffe fait appel à 3 enzymes : une ARN triphosphatase, une ARN guanylyltransférase et une ARN (guanine-7)-méthyltransférase (Parent and Bisailon, 2006). La coiffe intervient dans le processus d'épissage puisque le complexe CBC (*cap binding complex*), formé des protéines CBP20 et 80 (*cap-binding protein*) et qui se fixe à la coiffe, recrute U1snRNP et U6snRNP. De plus, la coiffe participe à l'export nucléaire des ARNm et est impliquée dans la traduction par son interaction avec eIF4E (*Eukaryotic Initiation Factor 4E*) qui permet le recrutement des ribosomes (Parent and Bisailon, 2006).

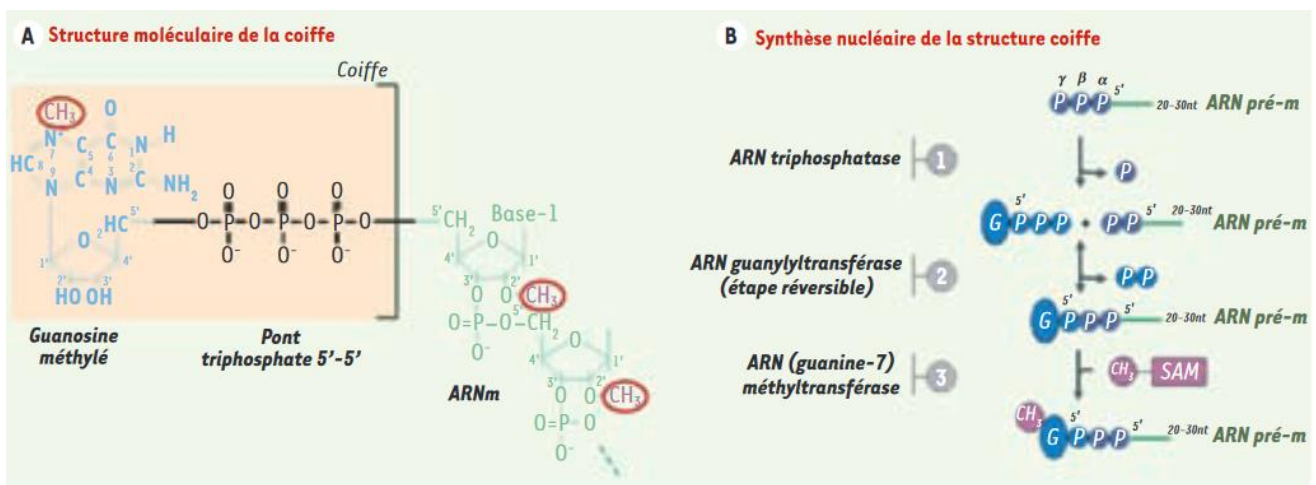


Figure 53 : La coiffe 7-méthylguanosine en 5' des transcrits (Parent and Bisailon, 2006).

(A) Structure de la coiffe et (B) processus de synthèse avec les enzymes impliquées.

Les ARNm possèdent à leur extrémité 3' une queue poly(A) et l'ajout de celle-ci se fait en 2 étapes : les transcrits possèdent un site de polyadénylation (motif AAUAAA) situé à 10 à 30 nt de l'endroit à partir duquel l'ARNm est clivé en aval, puis une poly(A) polymérase vient ajouter la queue poly(A) (en moyenne 200 à 250 résidus). Celle-ci est reconnue par des *poly(A) binding protein* (PABP), ce qui protège l'ARN des 3'-5' exonucléases et des déadénylases. La queue poly(A) participe également à l'export nucléaire des ARNm, à leur stabilité et leur traduction (Millevoi and Vagner, 2010)

V.2.2. Export nucléaire

La maturation des ARNm dans le noyau est une étape importante car seuls ceux qui sont complètement matures pourront être exportés, les autres seront dégradés. De plus, au cours de ces différentes étapes, sont également recrutées des protéines participant à leur export nucléaire (Carmody and Wente, 2009). L'export nucléaire assure la transition entre la transcription des ARNm dans le noyau et leur traduction dans le cytoplasme. Ce mécanisme peut parfois être sélectif et ainsi contribuer à la régulation de l'expression des gènes (Wickramasinghe and Laskey, 2015).

L'export nucléaire des mRNP se fait au travers de pores nucléaires enchâssés dans l'enveloppe nucléaire et nécessite la présence des protéines NXF1 (*Nuclear Export Factor 1*) et de son co-facteur NXT1 (*Nuclear Transport Factor 2 Like Export Factor 1*) ou du récepteur d'export CRM1 (*Chromosomal Maintenance 1*, appelé également *exportin 1*) (figure 54) (Wickramasinghe and Laskey, 2015). Les mRNP sont en effet exportés selon différentes voies en fonction de la taille des transcrits : les longs ARNm sont exportés par le complexe binaire NXF1-NXT1 et les courts ARNm (snRNA <200 nt) par CRM1. Cependant, lorsque les complexes mRNP sont trop gros pour passer par les pores nucléaires, ils sont exportés par bourgeonnement de l'enveloppe nucléaire (Müller-McNicoll and Neugebauer, 2013).

L'enveloppe nucléaire est une structure très poreuse qui comporte près de 4000 pores nucléaires et permet les échanges entre le nucléoplasme et le cytoplasme. Le pore nucléaire, structure très volumineuse (125MDa), a une organisation radiale et symétrique faisant appel à 3 éléments fonctionnels complexes : on parle donc de complexe de pore nucléaire (NPC, *Nuclear pore complex*). Le NPC comporte (Wickramasinghe and Laskey, 2015) :

- un anneau cytoplasmique et un anneau nucléaire constitués chacun de 8 granules et de filaments épais permettant l'ancrage et la mise en attente des molécules à transporter ;
- un anneau central qui peut s'ouvrir par le glissement sur eux-mêmes de 8 rayons épais mobiles, il interagit avec les transporteurs et garantit l'étanchéité du pore ;
- le panier ou « nuclear basket » qui est une structure dépendante du calcium et qui guide les protéines et les ARN au cours de leur transport.

Le transport est actif ou passif et bidirectionnel, il s'effectue à haute vitesse, toute molécule dépassant une masse de 60-70 kDa doit recourir à un transporteur pour son passage au travers du pore. Sa perméabilité est contrôlée par des protéines appelées FG-nucléoporines possédant

dans leur structure des régions riches en phénylalanine et glycine (F et G) et qui sont reconnues par les facteurs NXF1 et CRM1 (Wickramasinghe and Laskey, 2015).

Les complexes TREX (*transcription-export complex*), formés par l'association de plusieurs protéines (dont ALY, UAP56 (*U2AF65-associated protein 56 KDa*)) et recrutés au niveau de la coiffe via le complexe CBC, interviennent dans le processus d'export nucléaire puisque le complexe NXF1-NXT1 est recruté sur les mRNP grâce au facteur ALY et escorte les ARNm et leurs protéines associées jusqu'au NPC. De plus, NXF1 interagit également directement avec les protéines SR associées aux ARNm. Le facteur CRM1 interagit quant à lui avec des protéines adaptatrices possédant un motif NES et se lie au FG-nucleoporine pour permettre le passage des ARNm au travers le NPC (Wickramasinghe and Laskey, 2015).

Lorsque les mRNP atteignent la face cytoplasmique du complexe NPC, les protéines GLE1 (nucléopore liée à IP6 (inositol hexaphosphate)) et DBP5 (*DEAD-box protein*) s'associent au mRNP. DBP5 possède une fonction ATPase qui transforme l'ATP associé à DBP5 en ADP. Ceci induit un changement de conformation entraînant une modification de la composition des protéines associées au mRNP. Ainsi, le complexe CBC est remplacé par eIF4E (Carmody and Wentz, 2009).

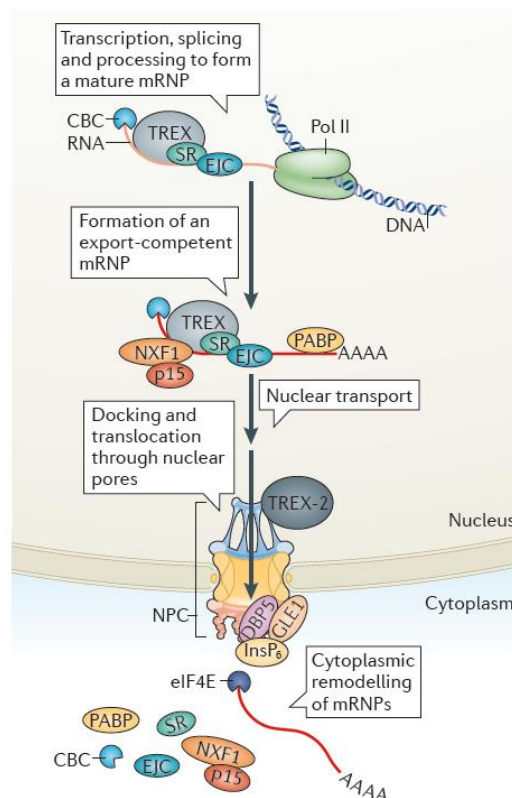


Figure 54 : Mécanisme et protéines intervenant dans l'export nucléaire des ARNm (Wickramasinghe and Laskey, 2015).

Les complexes TREX recrutés au niveau de la coiffe interagissent avec le complexe NXF1-NXT1, ce qui permet le transport des ARNm jusqu'au NPC.

V.2.3. Mécanisme de dégradation intra-cytoplasmique des ARNm

La dégradation intra-cytoplasmique des ARNm chez les eucaryotes peut emprunter deux voies qui débutent toutes les deux par une étape de déadénylation (figure 55). La déadénylation se déroule en deux phases : dans une première étape, la queue poly(A) est raccourcie de 250 à 110 nt environ par le complexe PAN2-PAN3 (*poly(A) nuclease 2 et 3*) puis le complexe CAF1-CCR4-NOT poursuit la déadénylation de la queue oligo(A) résiduelle (Brook and Gray, 2012). Pour cette dernière étape, il existe chez les mammifères (et non chez la levure) un autre complexe de déadénylation appelé PARN (*Poly(A)-Specific Ribonuclease*). Une fois déadénylés, les ARNm peuvent soit (Eulalio *et al.*, 2007) :

- être dégradés dans le sens 3' vers 5' par l'exosome, (un complexe multiprotéique qui recrute des RNases et des co-facteurs), le complexe SKI régulant son activité ;
- soit subir une élimination de la coiffe (*decapping*) par l'action de DCP2 (*Decapping Protein 2*) (et son coactivateur DCP1) suivie d'une dégradation exonucléolytique dans le sens 5' vers 3' effectuée par XRN1 (*5'-3' Exoribonuclease 1*).

Lors du processus de *decapping*, différentes protéines interviennent en tant que co-activateurs comme EDC3 (*Enhancer Of MRNA Decapping 3*), Ge-1 (*Enhancer Of MRNA Decapping 4*), le complexe heptamérique LSM1 à 7 (*U6 snRNA-associated Sm-like protein*), Dhh1 (*DEAD/H-box ATPase RNA helicase*) et Pat1 (*DNA topoisomerase 2-associated protein*). Ainsi, lorsque la queue poly(A) atteint une taille inférieure à 12 nucléotides, ces protéines sont recrutées et elles-mêmes favorisent le recrutement de DCP1 et 2, stimulant ainsi le *decapping* (Norbury, 2013).

La plupart des transcrits sont dégradés par cette voie dépendante de la déadénylation. Cependant, des voies alternatives peuvent être empruntées, comme par exemple celle faisant appel à un clivage interne endonucléolytique médié par l'*inositol-requiring enzyme 1* (IRE1) (Garneau *et al.*, 2007).

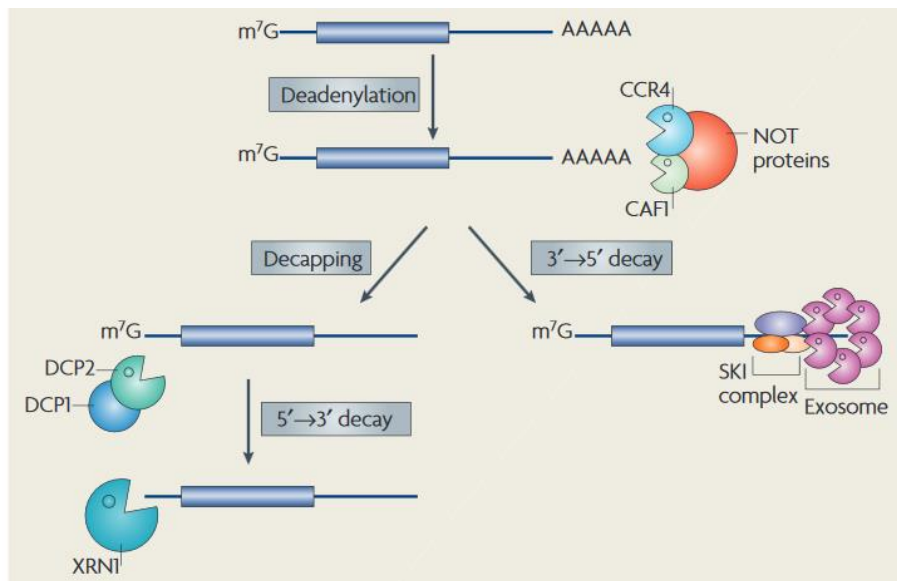


Figure 55 : Mécanismes de dégradation cytoplasmique des ARNm débutant par la déadenylation (Eulalio *et al.*, 2007).

Deux voies peuvent être empruntées après la déadenylation : le *decapping* suivi de la dégradation dans le sens 5'→3' ou la dégradation dans le sens 3'→5'.

Il existe des systèmes de contrôle de qualité des ARNm permettant uniquement aux ARNm, complètement matures et ne comportant pas d'anomalie, d'être accessibles à la machinerie de traduction. En effet, la détection d'aberrations lors de la traduction conduit à l'élimination de ces ARNm par des voies de dégradation spécifiques.

Le *nonsense mediated decay* (NMD) fait partie de ces systèmes de surveillance et permet la dégradation d'ARNm possédant un codon stop prématuré (PTC, *Premature Termination Codon*) (figure 56). Après l'épissage, le complexe EJC est positionné 20 à 24 nucléotides en amont de la jonction d'épissage. Ce complexe est éliminé lors du premier passage des ribosomes sur les ARNm « normaux ». Cependant, si un codon stop prématuré est situé à plus de 50-55 nucléotides en amont du complexe EJC, ce complexe n'est pas éliminé par les ribosomes et le NMD se met en place. Si la distance est inférieure à 50-55nt, l'EJC étant trop proche du PTC, il sera éliminé par les ribosomes. Dans le cas contraire, les ribosomes recrutent le complexe SURF (constitué des facteurs SMG1-UPF1-eRF1-eRF3) et UPF1 rejoint le EJC en interagissant avec les composants UPF2 et UPF3 de l'EJC pour former le complexe DECID (*decay-inducing*). Ceci active la phosphorylation de UPF1 qui conduit au recrutement de protéines SMG (*Smaug*), SMG6 puis le complexe SMG5-SMG7,

responsables respectivement de l'initiation du clivage endonucléolytique ou du recrutement de facteurs de dégradation classiques des ARNm (Le Hir *et al.*, 2016). Dans le cas du NMD, les ARNm peuvent aussi subir le *decapping* de façon indépendante de la déadénylation et ensuite être dégradés selon les modalités exposées précédemment (Eulalio *et al.*, 2007).

Il existe un mécanisme de NMD non dépendant des EJC qui repose sur la détection d'une distance anormale entre le codon stop prématuré et la queue poly(A). En effet, ce pseudo-3'UTR de longueur augmentée entraîne un arrêt des ribosomes trop loin de la queue poly(A) et stimule le NMD (Garneau *et al.*, 2007; Le Hir *et al.*, 2016).

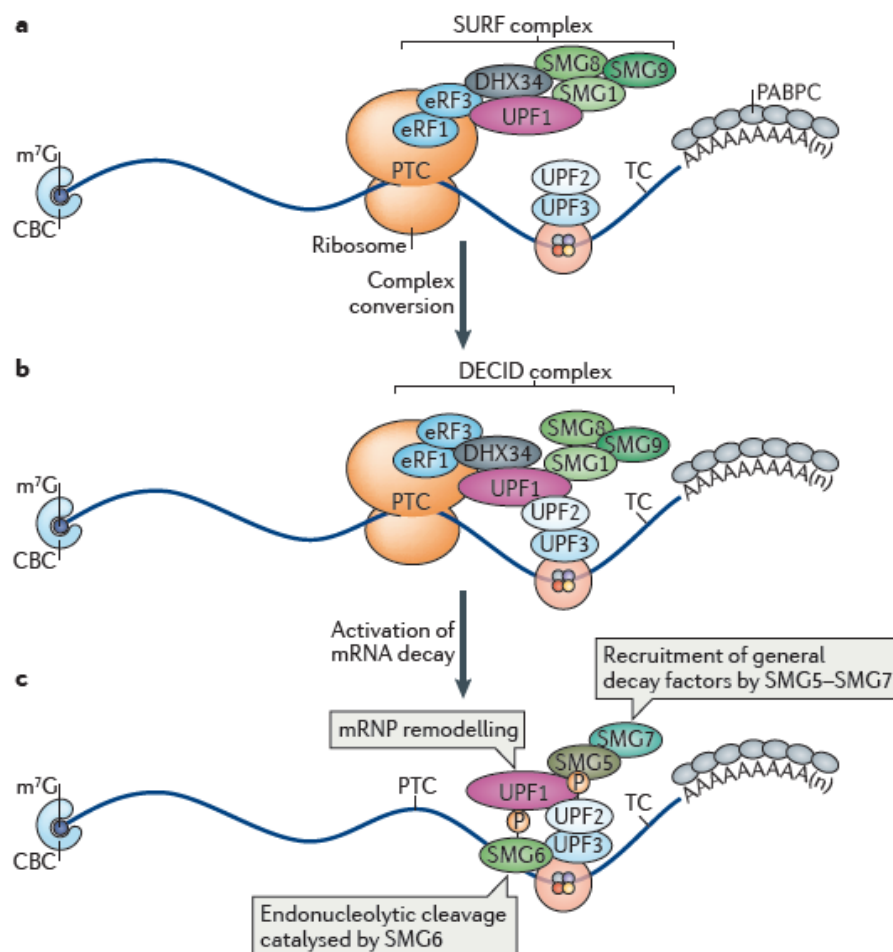


Figure 56 : Mise en place du NMD suite à la présence d'un codon stop prématuré (Le Hir *et al.*, 2016).

L'arrêt des ribosomes sur le codon stop prématuré permet l'assemblage du complexe SURF (a) puis le facteur UPF1 du complexe SURF interagit avec UPF2-UPF3 de l'EJC pour former le complexe DECID (b). S'en suit un remodelage permettant le recrutement des facteurs de dégradation des ARNm (c).

D'autres systèmes de surveillance permettent un contrôle de la qualité des ARNm (Graille and Séraphin, 2012) (figure 57):

- le *No-go decay*, phénomène qui peut par exemple être lié à des anomalies de structure des transcrits, conduit à un clivage endonucléolytique des ARNm sur lesquels les ribosomes se sont arrêtés pendant la traduction. Ce clivage est suivi d'une dégradation par l'exosome ou Xrn1.
- Le *Non-stop decay* consiste en une dégradation dans le sens 3'5' des ARNm qui ne contiennent pas de codon stop dans leur séquence.

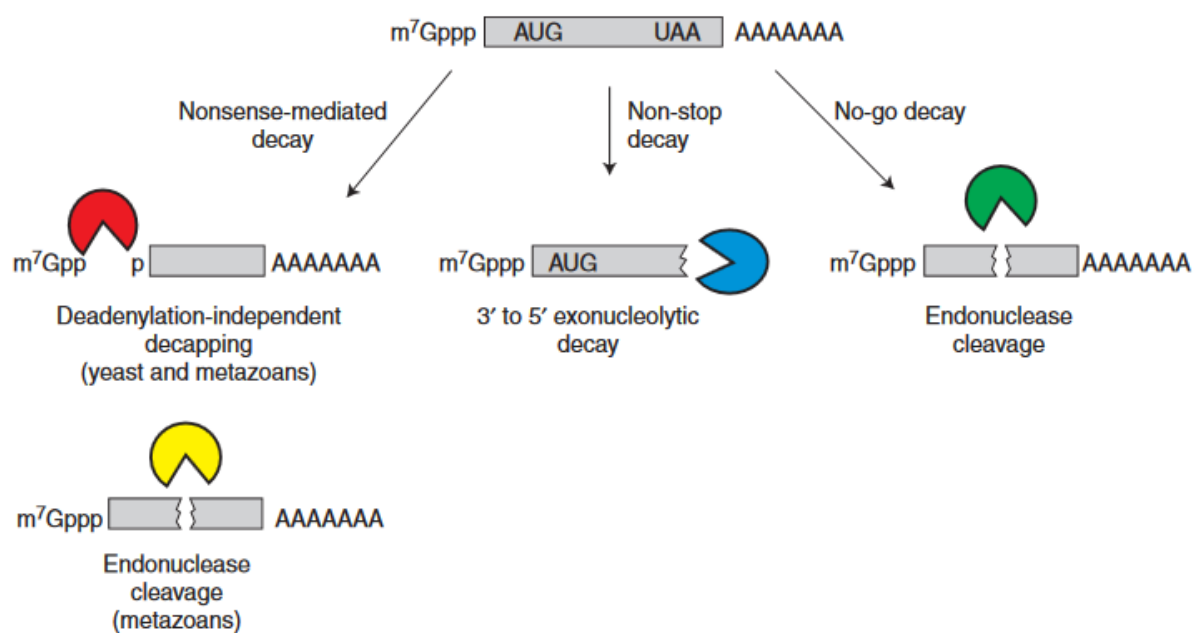


Figure 57 : Schéma résumant les différentes voies de dégradation des ARN liées à une anomalie constatée lors de la traduction : *Non stop*, *No-go* et *Nonsense mediated decay* (Decker and Parker, 2012).

Le *Non-stop decay* concerne les ARN ne possédant pas de codon stop et le *No-go decay* les ARN sur lesquels les ribosomes sont à l'arrêt suite par exemple à une anomalie structurale de l'ARN.

V.2.4. Facteurs influençant la stabilité des ARNm

Les éléments *cis* intervenant dans la stabilité des ARNm sont très nombreux et le plus souvent localisés dans les régions UTR, et surtout en 3'UTR. Nous nous intéresserons ici à la régulation exercée par la queue poly(A) et les *poly(A)-binding proteins* (PABP), les sites de fixation des micro-ARN (miR) et les *AU rich elements* (ARE).

a. Queue poly(A) et PABP

1. Rappel sur l'initiation de la traduction

L'initiation de la traduction consiste en une série d'interactions entre ARNm, facteurs d'initiation de la traduction (eIF) et sous-unités ribosomiques conduisant à l'assemblage du ribosome 80S au sein duquel le codon d'initiation de la traduction présent sur l'ARNm est assemblé avec l'anticodon de l'ARN de transfert initiateur. L'initiation de la traduction se passe de la façon suivante (figure 58) : le complexe de pré-initiation 43S (formé de la sous-unité ribosomique 40S, du complexe eIF2-GTP-Méthionine ARNt et des facteurs eIF3, eIF1, eIF1A, eIF5) s'attache à la coiffe, scanne le 5'UTR et reconnaît le codon d'initiation. Il y a alors formation du complexe 48S. Puis, les facteurs eIF5 et eIF5B hydrolysent le GTP, ce qui entraîne la libération des différents facteurs eIF et le recrutement de la sous-unité 60S pour former le ribosome 80S. L'élongation de la traduction peut ensuite commencer (Jackson *et al.*, 2010).

Le recrutement du complexe de pré-initiation 43S au niveau de la coiffe est facilité par le complexe eIF4F. Celui-ci est constitué de 3 éléments :

- eIF4E se liant à la coiffe ;
- eIF4A ayant une activité *RNA-dependant ATPase* et hélicase ;
- eIF4G qui est une protéine se liant à eIF4E et eIF4A et qui recrute la sous-unité 40S des ribosomes en se liant à eIF3.

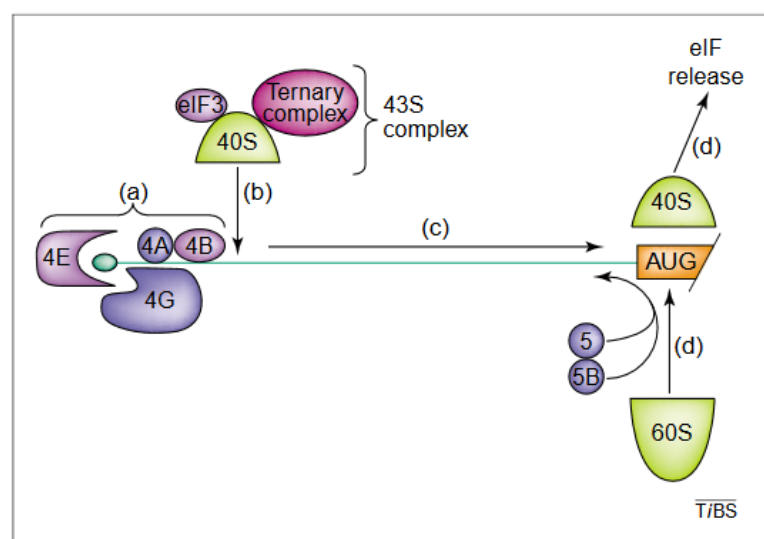


Figure 58 : Processus d'initiation de la traduction dépendant de la coiffe avec recrutement du complexe 43S, libération des facteurs eIF et arrivée de la sous-unité 60S (Wilkie *et al.*, 2003).

A côté de ce schéma classique d'initiation de la traduction, il existe un mécanisme beaucoup plus rare, surtout destiné aux ARNm viraux et faisant appel aux séquences « *internal ribosome entry site* » (IRES) qui entraînent un recrutement des ribosomes de façon indépendante de la coiffe, dans la région d'initiation de la traduction (Sonenberg and Hinnebusch, 2009).

La régulation de l'expression des gènes au niveau post-transcriptionnel peut s'effectuer en modulant l'initiation de la traduction, ce qui en permet un contrôle rapide, réversible et local.

2. Influence de la queue poly(A) sur la stabilité et la traduction des ARNm

La queue poly(A) est reconnue par une famille de protéines appelées « *poly(A) binding proteins* » (PABP). Il existe 4 PABP, l'une étant nucléaire (PABPN1) à haute affinité pour les chaînes poly(A) naissantes, et les 3 autres étant cytosoliques (PABPC1, PABPC3, PABPC4). La PABPC1 cytosolique est la mieux caractérisée.

Au niveau nucléaire, PABPN1, en se liant aux premiers résidus ajoutés de la queue poly(A), stimule l'activité de la poly(A) polymérase permettant ainsi l'élongation de la queue poly(A). Elle contrôle également le nombre de résidus A ajoutés et participe à l'export des ARNm (Mangus *et al.*, 2003).

Les PABP cytosoliques (PABPC), influencent la stabilité des ARNm et leur traduction. Tout d'abord, celles-ci stimulent l'initiation de la **traduction** par le biais de la formation d'une boucle regroupant les extrémités 5' et 3' des ARNm (figure 59), ce qui stimule l'assemblage du complexe d'initiation de la traduction au niveau de la coiffe en 5'. Cette boucle est réalisée grâce à l'interaction entre les PABP et eIF4G. De plus ces interactions augmentent l'affinité de eIF4E pour la coiffe et de PABP pour la queue poly(A) et donc la traduction. Par ailleurs, les PABP stimulent aussi l'arrivée de la sous-unité 60S (Sonenberg and Hinnebusch, 2009). A l'inverse, la déadénylation diminue la traduction des ARNm en limitant la fixation des PABP, ce qui a pour conséquence de limiter le recrutement des ribosomes au niveau de l'extrémité 5' de l'ARNm (Tourrière *et al.*, 2002).

Ensuite, les PABP peuvent interagir avec eRF3 (*eukaryotic Release Factor 3*, figure 59), facteur de terminaison de la traduction. Cette interaction augmente l'efficacité de cette étape et permet le recyclage des ribosomes vers l'extrémité 5' de l'ARNm pour engager plusieurs cycles de traduction (Goss and Kleiman, 2013 ; Gorgoni and Gray, 2004).

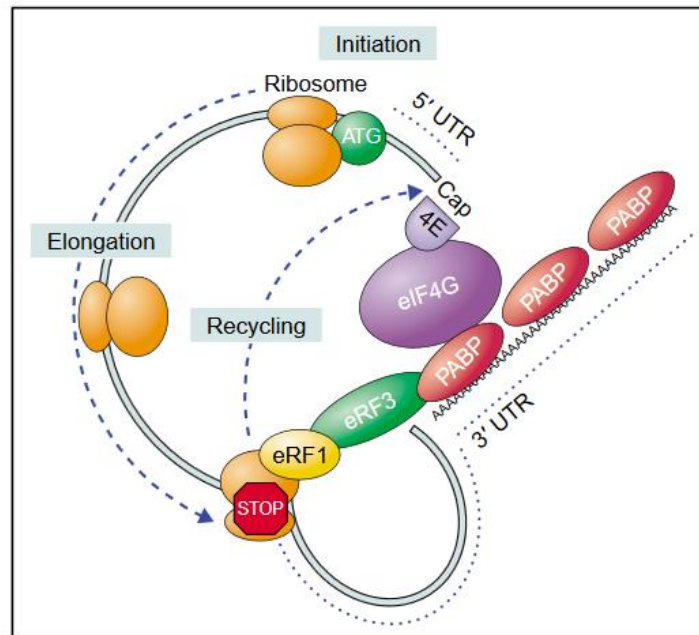


Figure 59 : Circulation de l'ARNm induit par l'interaction entre PABP et eIF4G (Sonenberg and Dever, 2003).

Ce processus permet de favoriser la traduction en facilitant notamment l'assemblage du complexe d'initiation de la traduction.

En ce qui concerne la **stabilité** des ARNm, les PABP préservent la queue poly(A) de la déadénylation en empêchant l'accès des déadénylases. Elles protègent également l'extrémité 5' de l'ARNm du *decapping* par l'intermédiaire de la formation du complexe PABP-eIF4G-eIF4E (Goss and Kleiman, 2013 ; Smith *et al.*, 2014). Les PABP régulent également le recrutement des endonucléases. Ainsi, la région 3'UTR du transcrit de l' α -globine contient un site CRE (*C-rich element*) sur lequel se fixe le complexe α CP1- α CP2 (α -globin poly(C)-binding protein) qui interagit avec PABP1. Cette interaction augmente la liaison de PABP et du complexe α CP1- α CP2 sur leurs éléments *cis*, ce qui protège l'ARNm de l'action d'endonucléases et de la déadénylation (Brook and Gray, 2012; Mangus *et al.*, 2003). Enfin, les PABP peuvent interagir avec les ARE-BP. Ainsi, l'interaction entre PABP1

et HuR augmente la protection de la queue poly(A) de l'ARNm de la β -caséine (Brook and Gray, 2012).

b. Micro-ARN

Les micro-ARN (miRNA) sont de petits ARN non codants d'environ 21 nucléotides. Leur synthèse débute par la transcription d'un pri-miR par l'ARN polymérase de type II (figure 60). Ils possèdent une coiffe en 5' et une queue poly(A) ainsi qu'une structure en épingle à cheveux. Ils sont ensuite transformés par une endonucléase appelée Drosha (et par son cofacteur DGCR8 (*DiGeorge syndrome critical region 8*)), dans le noyau, en pré-miR de 70 nt à structure en épingle à cheveux. Les pré-miR sont ensuite exportés dans le cytoplasme pour y achever leur maturation et sont transformés par l'endonucléase Dicer en structure de microARN double brin constitué 21 à 25 nt qui contient le microARN mature associé à son brin complémentaire. Le brin mature est ensuite incorporé dans un complexe ribonucléoprotéique appelé miRISC (*miRNA-induced silencing complex*) et l'autre brin est dégradé. Au sein de ce complexe, les composants protéiques majeurs sont les protéines GW182 (*glycine-tryptophan protein of 182 kDa*) et celles de la famille Argonaute (Ago). Parmi les membres de la famille Ago, seule Ago2 a une activité *RNAseH-like PIWI domain* permettant le clivage de l'ARNm cible (He and Hannon, 2004).

La spécificité de reconnaissance de la cible ARNm par les microARN est assurée par leur séquence *seed* (située au niveau des nucléotides 2 à 8 en 5' du microARN). Si la complémentarité avec la cible ARNm est parfaite, celle-ci est dégradée. Cependant, la plupart des microARN s'apparient de façon imparfaite avec leur cible. Ce type d'interaction ne conduit généralement pas à la dégradation de l'ARNm par clivage endonucléolytique mais néanmoins inhibe la synthèse des protéines soit par répression de la traduction, soit par une augmentation de la déadénylation et de la dégradation des ARNm qui lui est associée (Krol *et al.*, 2010). De plus, les microARN sont aussi capables d'inhiber l'expression des gènes au niveau transcriptionnel via une liaison directe à l'ADN. Enfin, les microARN peuvent se lier à certaines RBP possédant des sites de liaison à l'ARN, ce qui empêche la fixation de ces RBP sur leurs ARN-cibles (Garzon *et al.*, 2010).

Les microARN interagissent le plus souvent avec les régions 3'UTR des ARNm, mais des sites de fixation sont également présents dans les régions 5'UTR et même dans les régions

codantes. Par ailleurs, la liaison dans la région 5'UTR est plutôt associée à une activation de la traduction plutôt qu'à une répression (Fabian *et al.*, 2010).

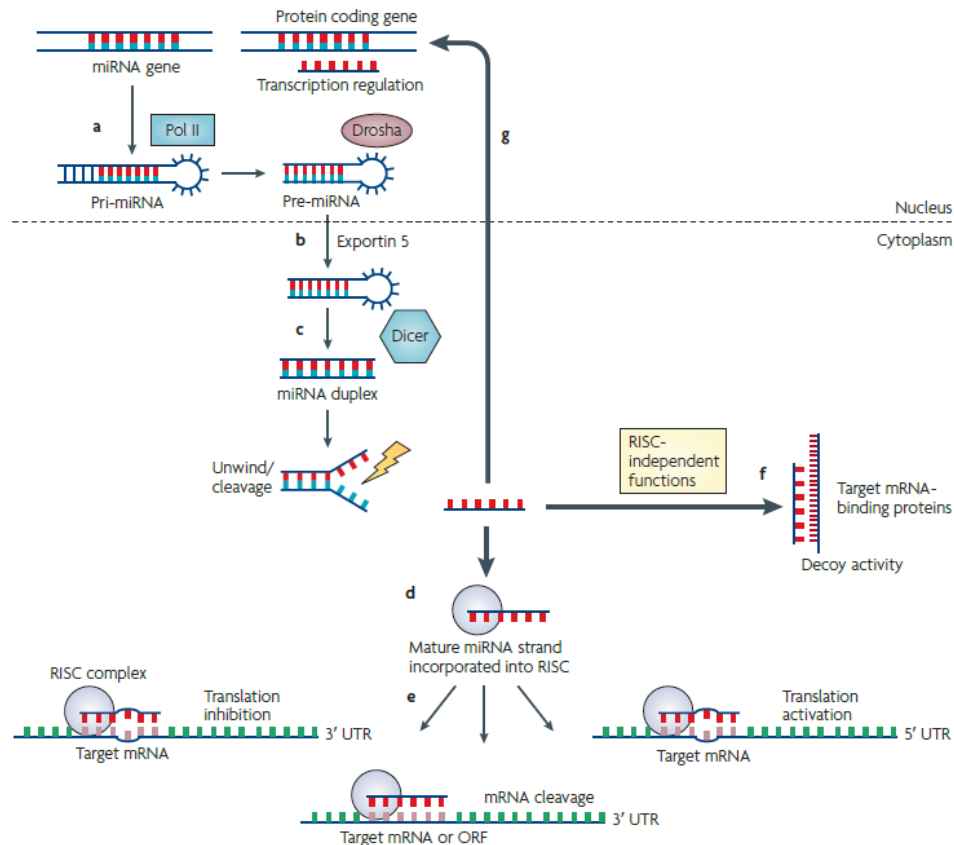


Figure 60: Synthèse et mode d'action des micro-ARN (Garzon *et al.*, 2010).

Les microARN sont synthétisés sous la forme de précurseurs (pri-miRNA) eux même transformés en pré-miRNA et enfin en miRNA matures. Ils agissent de différentes façons : en favorisant la dégradation de leurs ARNm cibles, en inhibant ou activant leur traduction et certains d'entre eux régulent la transcription.

La dégradation des ARNm exercée via les microARN requiert des constituants localisés dans des granules cytoplasmiques appelés *P-Bodies*. En effet, les constituants du complexe miRISC (Ago et GW182) ainsi que ceux intervenant dans la dégradation des ARNm sont enrichis dans ces structures. Afin de dégrader l'ARNm-cible (figure 61), GW182 recrute le complexe de déadénylation CCR4-NOT1 et favorise le *decapping* par l'intermédiaire du complexe DCP1-DCP2. De plus, GW182 interagit avec les PABP fixées à la queue poly(A) facilitant ainsi leur déplacement, ce qui stimule le processus de déadénylation. La dégradation des ARNm avec lesquels les microARN s'apparient est possible également via un clivage réalisé par la protéine Ago (Fabian *et al.*, 2010).

Les microARN peuvent également inhiber la traduction. Pour se faire, le complexe mi-RISC en s'associant avec le facteur eIF6 (inhibiteur de l'initiation de la traduction), empêcherait l'association de la sous-unité ribosomique 60S à la sous-unité ribosomique 40S. De plus, les protéines Ago (Argonaute), composants catalytiques du complexe RISC, entrent en compétition avec eIF4E pour la liaison à la coiffe, ce qui empêche la formation du complexe eIF4F (Shyu *et al.*, 2008). Par ailleurs, GW182, en interagissant avec PABP empêche la formation de la boucle entre les extrémités 3' et 5' de l'ARNm issue de la liaison entre eIF4G et PABP, ce qui inhibe l'initiation de la traduction. Les microARN peuvent également agir après l'initiation de la traduction en interférant avec l'élongation des ribosomes et provoquer leur dissociation prématurée de l'ARNm (Fabian *et al.*, 2010).

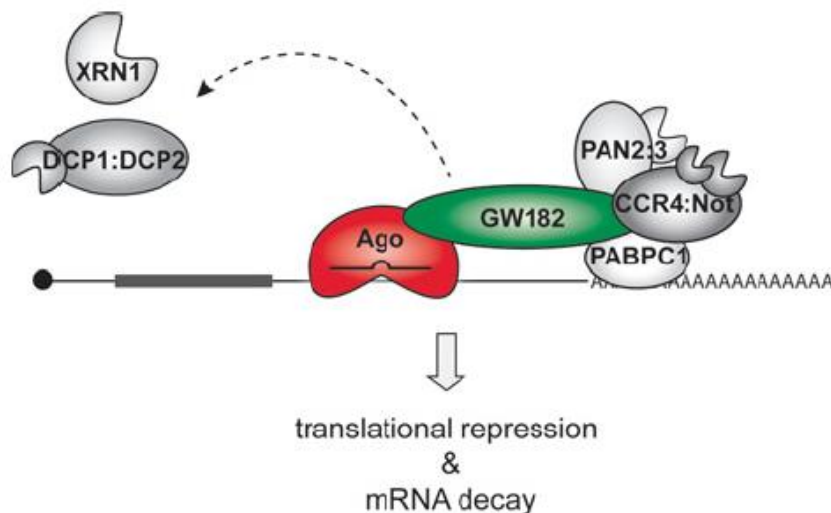


Figure 61 : Illustration de l'interaction entre ARNm-miRNA-Ago-GW182 et leur rôle dans la dégradation des ARNm (Pfaff and Meister, 2013). GW182 recrute le complexe de déadénylation CCR4-NOT1 et favorise le deccaping par l'intermédiaire du complexe DCP1-DCP2.

c. ARE

Parmi les éléments *cis* les mieux caractérisés pour influencer la stabilité des ARNm figurent les *AU rich elements* (ARE). La régulation médiée par les ARE peut être très spécifique, ciblant seulement les ARNm possédant ces séquences, ou toucher plusieurs transcrits avec des fonctions liées et dont l'expression est co-réglée.

Les ARE sont présents dans la région 3'UTR des ARNm généralement à demi-vie courte, codant pour des protéines dont l'expression est transitoire (cyclines, facteurs de

croissance...). Cependant, les effets des ARE sur les ARNm dépendent du type de ARE-BP qui s'y fixe (tableau XII). Par exemple, la protéine TTP (*Tristetrapolin*) favorise une dégradation rapide des ARNm du GM-CSF alors que la protéine HuR stabilise les transcrits du *VEGF* (Eulalio *et al.*, 2007). Lorsque les ARE sont déstabilisateurs, ces séquences dirigent les ARNm vers une dégradation rapide par un mécanisme dépendant de la déadénylation et appelée la voie AMD (*ARE mediated RNA decay*). La déstabilisation des ARNm reconnus par les ARE-BP est basée sur le recrutement de la machinerie classique de dégradation des transcrits. Ainsi, l'ARE-BP TTP interagit avec la déadénylase CCR4-NOT et l'exosome (Barreau *et al.*, 2005).

Lorsque les ARE-BP agissent en inhibant la traduction, ceci est lié à la dégradation de la queue poly(A) qui entraîne la libération des PABP, déstabilisant ainsi l'interaction entre eIF4E et la coiffe. Il existe un autre modèle dans lequel les ARE-BP se lient à eIF4E ce qui empêche sa fixation à eIF4G. Enfin, certaines ARE-BP dirigent les ARN vers des granules cytoplasmiques appelés granules de stress dans lesquels les ARNm sont bloqués à l'étape initiale de la traduction (Barreau *et al.*, 2005).

La stabilisation médiée par les ARE-BP peut être exercée par leur entrée en compétition avec la liaison d'un facteur déstabilisateur sur le transcrit. Par exemple, HuR entre en compétition avec les protéines déstabilisatrices AUF1 (*AU-rich binding factor-1*), KSRP (*KH splicing regulatory protein*) et TTP pour sa liaison sur les ARE. Si la fixation est en faveur de HuR, l'ARN cible est stabilisé (Garneau *et al.*, 2007).

D'autre part, les modifications post-traductionnelles affectant les ARE-BP modulent leur affinité pour les ARE (Wilusz and Wilusz, 2004) ou leur localisation cellulaire. Ainsi, la phosphorylation de l'ARE-BP HuR par la protéine kinase p38 MAPK entraîne son accumulation dans le cytoplasme et augmente sa liaison aux transcrits de p21 à l'origine de leur stabilisation (Lafarga *et al.*, 2009).

Tableau XII : Exemples d'ARE-BP, de leurs fonctions et modes d'action (Garneau *et al.*, 2007):

RNA-binding protein	Function	RNA-binding domain	Mode of action	Modifications	Other functions
AUF1 (hnRNP D) and its four splice isoforms (p37, p40, p42, p45)	Usually destabilizing	RRM	Recruit the exosome; remodel mRNA to allow other proteins to bind	Phosphorylation allows isomerization by PIN1 leading to dissociation from RNA; interacts with 14-3-3 proteins	DNA binding
CUG-BP	Destabilizing	RRM	Recruits PARN; modulates ARE function	Phosphorylated by myotonic dystrophy protein kinase	Splicing; translation
ELAV proteins, for example, HuR and HuD	Stabilizing	RRM	Compete with destabilizing proteins for ARE-binding; might relocalize mRNAs away from decay machinery	CARM1-mediated methylation reduces stabilizing function	Translation; RNA localization
KSRP	Destabilizing	KH domain	Recruits decay enzymes: PARN and the exosome	Phosphorylation by p38-MAPK pathway leads to reduced RNA-binding affinity	Splicing
RHAU	Destabilizing	RNA helicase	Recruits decay enzymes: PARN and the exosome	Not known	Not known
TIA-1, TIAR	Translational silencing	RRM	Induce aggregation into stress granules	Phosphorylated by FAST	Alternative splicing
Tristetraprolin (TTP; TIS11, ZFP36), BRF1 (TIS11B, ZFP36L1), BRF2 (TIS11D, ZFP36L2)	Destabilizing	CCCH-type zinc finger	Recruit decay enzymes: CCR4, DCP1, PM-Sci75, RRP4	Phosphorylation by p38-MAPK pathway leads to association with 14-3-3 proteins	Transcription

Abbreviations :AUF1, AU-rich binding factor-1; CUG-BP, CUG-binding protein; ELAV, embryonic lethal abnormal vision; FAST, Fas-activated serine/threonine kinase; hnRNP, heterogeneous nuclear ribonucleoprotein; KSRP, KH splicing regulatory protein; MAPK, mitogen-activated protein kinase; PARN, poly(A)-specific ribonuclease; RHAU, RNA helicase associated with AU-rich element; RRM, RNA-recognition motif; TIA-1, T-cell-restricted intracellular antigen-1; TIAR, TIA-1-related; TIS, TPA-induced sequence; ZFP, zinc finger protein

L'effet des ARE-BP sur les ARNm peut être modulé par les microARN qui se lient sur le même transcrit et ce de façon coopérative ou compétitive (figure 62). Ainsi, la *compétition* entre HuR et un microARN s'accompagne soit d'une augmentation de l'expression du gène si la fixation HuR-mRNA prévaut face à l'interaction du transcrit avec le mir et soit d'une diminution si l'interaction est favorable au mir. HuR et les microARN peuvent entrer directement en compétition pour les mêmes sites de fixation. Cependant ceux-ci étant souvent distants, HuR agit soit en créant un encombrement stérique soit en induisant un changement de conformation de l'ARNm, ce qui masque ou démasque les sites de fixation pour les complexes miRISC. La *coopération* entre HuR et le microARN se traduit par une diminution de l'expression de l'ARNm. En effet, lors d'un effet coopératif, la fixation de HuR peut induire un changement de conformation de l'ARNm qui devient ainsi accessible au microARN qui exerce alors son action (Srikantan *et al.*, 2012).

Les interactions entre HuR et les microARN sont régulées par l'abondance de la production de HuR (régulation transcriptionnelle), par sa localisation cytoplasmique et ses modifications post-traductionnelles (phosphorylation, méthylation) (Srikantan *et al.*, 2012). Par exemple, l'ARNm de *CAT1* (*cationic amino acid transporter 1*) peut être reconnu par le mir-122. Lors d'un stress, HuR est dirigé du noyau vers le cytoplasme et se lie sur l'ARE présent dans la région 3'UTR de l'ARNm de *CAT1*, empêchant le microARN d'exercer son action et induisant une relocalisation des ARNm des *P-Bodies* vers le pool des transcrits traduits (Garneau *et al.*, 2007). A l'inverse, HuR facilite l'interaction du complexe let-7-RISC avec le 3'UTR de *MYC* (*v-myc myelocytomatosis viral oncogene homolog*) conduisant à la déstabilisation des transcrits (Srikantan *et al.*, 2012)

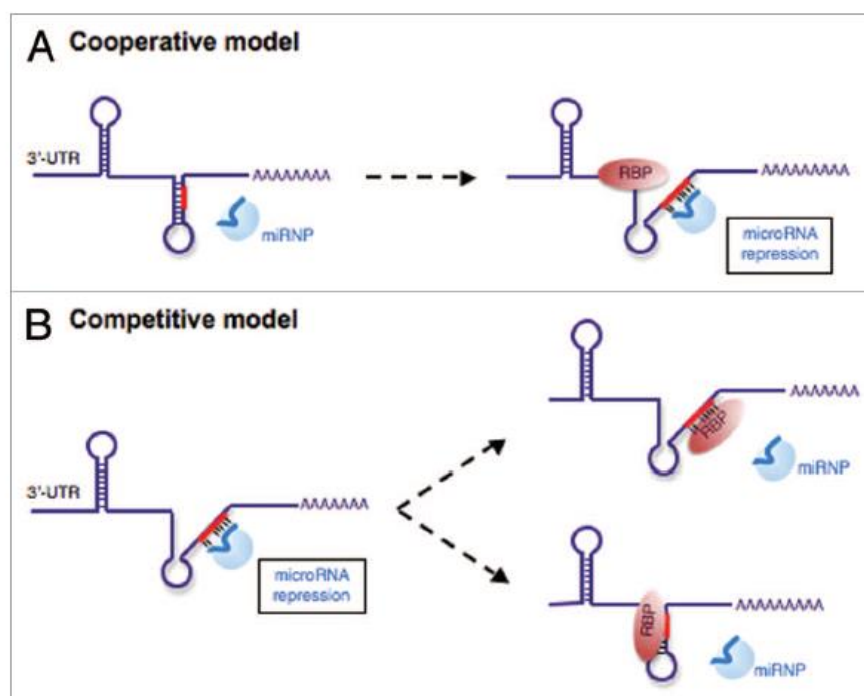


Figure 62 : microARN et ARE-BP (Ciafrè and Galardi, 2013).

Ceux-ci peuvent agir de façon coopérative lorsque la fixation de la RBP favorise le recrutement du microARN (A), ou de façon compétitive si la RBP et le microARN se fixent sur le même site ou si la liaison de la RBP induit un changement de conformation de l'ARN et perturbe la fixation du microARN sur le transcrit (B).

V.2.5. Stockages dans les granules

Le contrôle de la traduction et de la dégradation des ARNm cytoplasmiques peut s'effectuer via le stockage des transcrits au sein de granules à ARN (Kedersha and Anderson,

2007). Ces granules interagissent avec le cytosquelette, ce dernier jouant un rôle important dans leurs propriétés dynamiques (Rajgor and Shanahan, 2014).

Lors d'un stress induisant la libération des ARNm des polysomes, les transcrits peuvent être stockés et concentrés dans des structures non délimitées par une membrane, interconnectées, formant un pool d'ARNm non traduits, et appelées, selon leur composition, P-Bodies (PB) et Granules de stress (SG). Ces granules ont la propriété de moduler la traduction des ARNm (*via* leur stockage, dans les SG) ou leur dégradation (*via* les PB) (Kedersha and Anderson, 2007). Ces deux structures contiennent des facteurs communs, interagissent ensemble et les ARNm peuvent passer d'un type à l'autre de granules (figure 63). Lorsque le stress est résolu, les ARNm peuvent soit retourner au sein du pool actif de la traduction soit être dégradés.

Les RBP interagissant avec les ARNm au sein des granules doivent avoir 2 propriétés essentielles (Kato *et al.*, 2012) :

- être capables de se lier spécifiquement à un ARNm-cible ;
- contenir des éléments de régulation, afin d'être adressées et incorporées dans ces granules, le contrôle de l'assemblage des granules étant réalisé par les modifications post-traductionnelles de ces protéines (Decker and Parker, 2012).

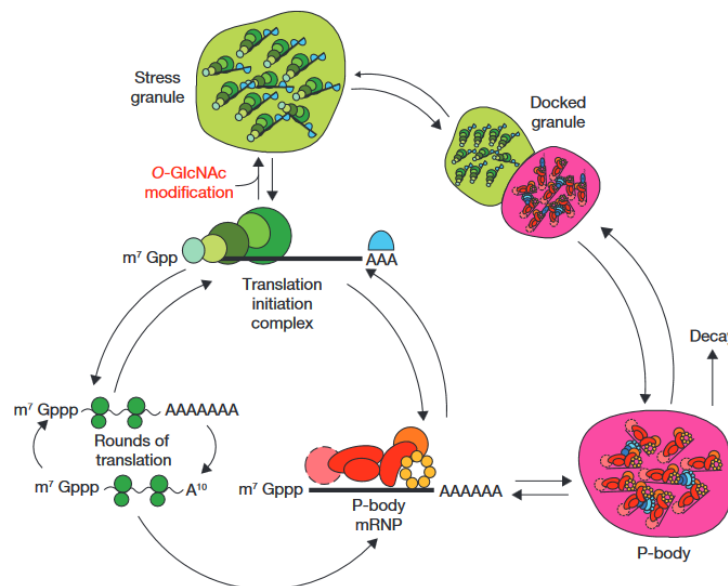


Figure 63 : Dynamique des ARNm entre les différents types de granules (P-Bodies, granules de stress) et le pool d'ARNm en traduction (Hilliker and Parker, 2008). Au cours d'un stress, les ARNm peuvent quitter le pool actif de la traduction et se localiser dans les P-Bodies ou les granules de stress dont l'assemblage est favorisé par des changements de modifications post-traductionnelles des RBP entrant dans leur composition (par exemple modifications de la glycosylation).

a. P-bodies

Les études de microscopies électroniques ont caractérisées les PB comme étant des structures dynamiques de 100 à 300 nm, non délimitées par une membrane et présentes dans les cellules à l'état basal (figure 64). La réponse à un stress passe par la régulation de leur nombre et de leur taille (Aizer *et al.*, 2014). Ces deux paramètres sont également influencés par le cycle cellulaire : ils sont absents durant la mitose par exemple, et abondants pendant la phase G2 (Eulalio *et al.*, 2007).

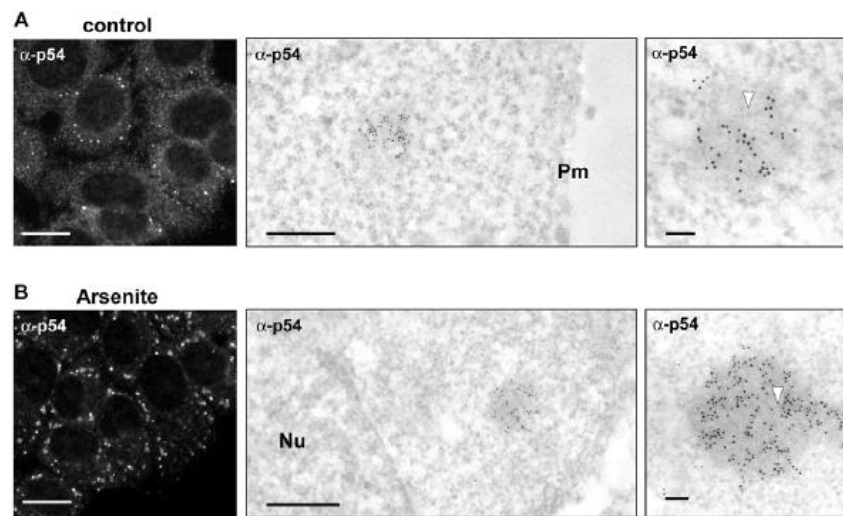


Figure 64 : Influence du stress (Arsenite) sur la formation des P-Bodies (Souquere *et al.*, 2009).

(A) : contrôle, (B) stress par arsenite. A droite : étude par microscopie électronique, à gauche : par immunofluorescence avec un anticorps anti-p54 (spots blancs).

L'assemblage des PB débute par la formation de mRNP, dont la traduction est réprimée, et qui s'agrègent par des interactions protéines-protéines pour former des structures plus grandes (Balagopal and Parker, 2009). Leur formation dépend de la présence des ARN et augmente quand la proportion d'ARN non traduits ou subissant un *decapping* s'accroît. Les ARNm doivent donc sortir du cycle de traduction pour entrer dans les PB (Eulalio *et al.*, 2007).

Les PB contiennent des ARNm inactifs sur le plan traductionnel ainsi que des facteurs impliqués dans la répression traductionnelle et la dégradation des transcrits (le *decapping*, le

NMD, la répression médiée par les microARN) (Anderson *et al.*, 2015; Hilliker and Parker, 2008). On y trouve donc :

- les enzymes de *decapping* Dcp1/Dcp2 ;
- les activateurs du *decapping* RCK/p54, Pat1, Lsm1-7, Scd6/RAP55, EDC3 ;
- l'exonucléase Xrn1 ;
- le complexe de déadénylation CCR4/CAF1/NOT ;
- des protéines impliquées dans le NMD (UPF1, UPF2, UPF3, SMG5, SMG6, SMG7) ;
- la machinerie associée au miRNA (GW182, Argonaute) ;
- des facteurs de décroissance associés aux ARE comme la protéine TTP ;
- des facteurs associés à la régulation de la traduction (eIF4E, eIF4E-T, CPEB, PCBP2, FASTK).

Par contre, les facteurs impliqués dans l'initiation de la traduction (en dehors de eIF4E et eIF4E-T (*eIF4E binding protein 4E-T*)) et les protéines ribosomiques en sont absentes.

Les PB ne doivent pas être considérés comme des sites de stockage des effecteurs de la dégradation des ARNm mais comme un site où cette dégradation peut se produire (Eulalio *et al.*, 2007). Les PB permettraient la dégradation des ARNm codant pour des protéines « inutiles » à la cellule lors d'une réponse au stress, afin de privilégier la traduction de protéines permettant, elles, urgemment, de faire face à cette situation critique pour la survie de la cellule (Rajgor and Shanahan, 2014).

b. Granules de stress

Les granules de stress sont des structures dynamiques associées aux microtubules qui apparaissent rapidement dans les cellules exposées à un stress puis disparaissent une fois que celui-ci est résolu (Anderson *et al.*, 2015; Panas *et al.*, 2016). A la différence des PB, ils ne sont pas présents à l'état basal dans les cellules (figure 65). Ils contiennent des ARNm bloqués à l'étape de l'initiation de la traduction et sont composés de différentes protéines telles que des RBP, des facteurs d'initiation de la traduction... Leur contenu n'est pas complètement inventorié. On y trouve par exemple eIF4E, eIF4A, G3BP (*RasGAP SH3 domain Binding Protein*), TIA-1 (*T-cell-restricted intracellular antigen-1*), TIAR (*TIA-related*) et HuR. La caractéristique des SG par rapport au PB est l'absence de certains

composants de la machinerie de dégradation des ARNm et la présence de la petite sous-unité ribosomique 40S (Balagopal and Parker, 2009)

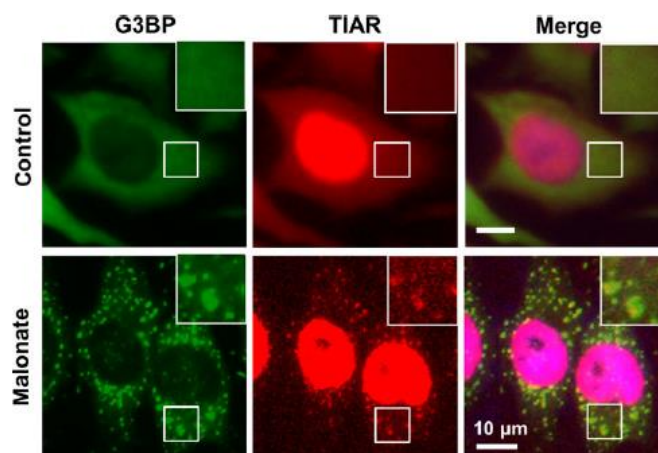


Figure 65 : Induction de la formation des GS suite à une exposition à un stress (malonate) (Fu *et al.*, 2016).
Les marqueurs utilisés sont G3BP et TIAR.

Suite à la survenue d'un stress, les SG se forment par condensation des mRNP (association d'ARNm et de protéines) via des interactions protéines-protéines médiées par les RBP dont certaines impliquent des séquences de faible complexité (figure 66). Ainsi, la méthylation, phosphorylation et glycosylation de ces protéines peut affecter la formation des SG en modifiant les interactions entre protéines (Protter and Parker, 2016). Lorsque les cellules sont exposées à un stress, la traduction protéique est provisoirement suspendue, les polysomes se désassemblent. Les ARNm libérés des ribosomes sont reconnus par certaines RBP telles que TIA-1, TIAR (*TIA-1-related*) et ces complexes mRNP non traduits sont les « nucléateurs » des granules de stress (Panas *et al.*, 2016). Lorsque l'état de stress disparaît, les ARNm quittent les SG pour retourner soit vers la traduction, c'est-à-dire les polysomes, soit vers les PB, afin d'être dégradés (Decker and Parker, 2012; Fu *et al.*, 2016)

La dissociation des polysomes et donc la formation des SG est initiée par la phosphorylation de eIF2 α réalisée par la kinase GCN2 (*general control nonderepressible 2*). Ce phénomène induit une diminution de la disponibilité du complexe eIF2-GTP-tRNA^{met} et par conséquent une baisse de la formation du complexe 48S de pré-initiation qui conduit à l'inhibition de la traduction (Kedersha and Anderson, 2007). De plus, lors d'un stress métabolique, la protéine de liaison eIF4E-BP (*eIF4E binding protein*) n'est plus phosphorylée

par la sérine-thréonine kinase mTOR (*mammalian target of rapamycin*). Ceci favorise donc sa liaison à eIF4E et le blocage de l'initiation de la traduction (Panas *et al.*, 2016).

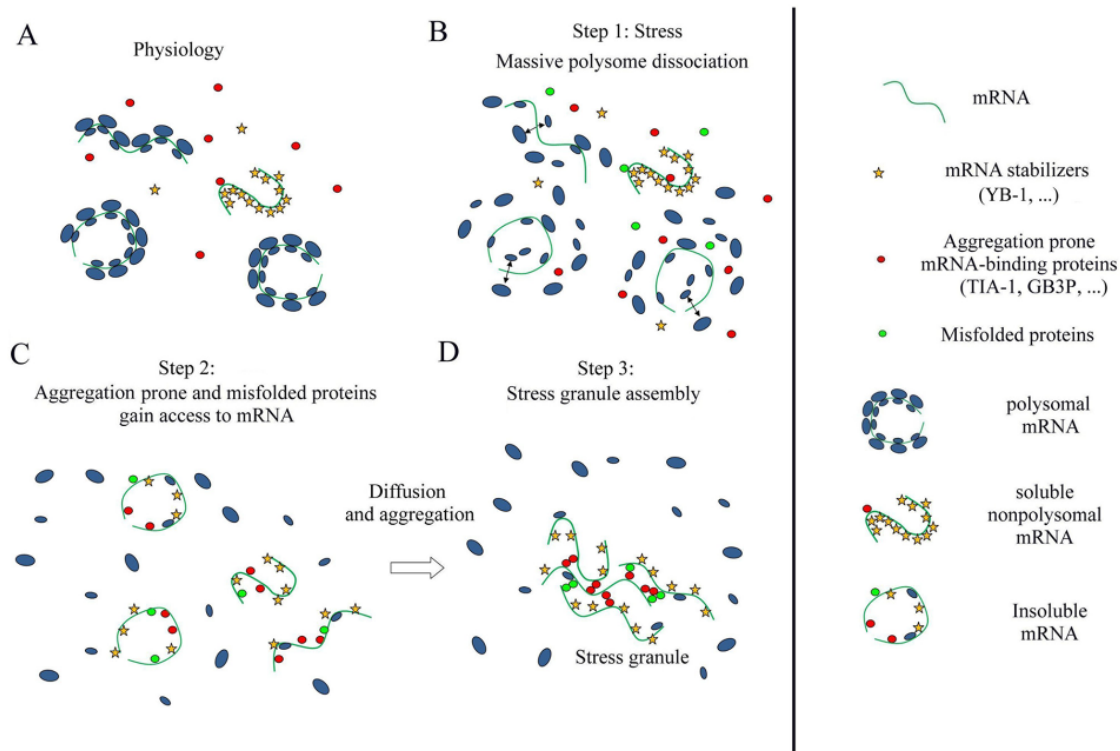


Figure 66 : Processus de formation des granules de stress après exposition à un stress (Bounedjah *et al.*, 2014).

Après exposition à un stress, les polysomes se dissocient, les ARNm sont reconnus par des RBP pour former des complexes mRNP, puis l'agrégation de ces complexes conduit à la formation des granules de stress.

Il existe une classe particulière d'ARNm possédant juste après la coiffe en 5' un enchaînement de 4 à 15 pyrimidines (séquence TOP (*Terminal Oligopyrimidine Tract*)) suivi immédiatement d'un domaine riche en CG. Ces motifs sont présents spécifiquement dans les ARNm codant pour les protéines ribosomiques et les facteurs de traduction. En effet, dans des conditions de stress, la traduction de ces ARNm est inhibée en vue de limiter la synthèse des composants de la machinerie traductionnelle. Les protéines TIA-1 et TIAR ont la particularité de se fixer sur ces motifs, ce qui a pour conséquence le blocage des ARNm à l'étape d'initiation de la traduction, leur libération des polysomes et leur accumulation dans les SG (Ivanov *et al.*, 2011).

c. Autres granules

A côté de ces deux catégories de granules, il en existe également d'autres tels que les granules des cellules germinales qui stockent les ARNm maternels lors des étapes précoces du développement. Le nombre d'une douzaine est avancé par DeGracia *et al.* (2008) dans un modèle cellulaire de neurone.

Une étude a également décrit des granules à IMP1 (*Insulin-like growth factor 2 mRNA-binding protein 1*) qui contiendraient des ARNm quiescents au niveau de la traduction et codant des protéines participant au contrôle de qualité du réticulum endoplasmique, au métabolisme dépendant de l'ubiquitine, et impliquées dans les voies sécrétoires. Ces granules seraient donc des opérons post-traductionnels permettant une production coordonnée et locale des différentes protéines codées par les ARNm présents dans ces granules (Jønson *et al.*, 2007).

D'autre part, il existe des granules de transport permettant de diriger les ARNm vers un site subcellulaire précis afin de permettre une traduction localisée. Ces granules se forment par oligomérisation des RNP via des interactions protéine-protéine. Ces granules sont également associés à des protéines motrices permettant le transport de ces granules le long des éléments du cytosquelette (Martin and Ephrussi, 2009), la traduction étant réprimée durant le transport. Par exemple, au niveau des neurones, le transport des ARNm dans des granules neuronaux permet une production localisée au niveau de la synapse (Elvira *et al.*, 2006).

Par ailleurs, hors du cadre des granules inactifs au plan traductionnel, Lui *et al.* (2014) ont récemment démontré que certains types de granules sont des lieux actifs de la traduction, en dehors de toute situation de stress.

Problématique et objectifs du travail

Les ARNm des mucines possèdent des propriétés inhabituelles par rapport aux autres transcrits humains : ils sont en effet caractérisés par une très grande taille, des séquences codantes répétées mais aussi une longue demi-vie (Porchet and Aubert, 2004). Peu de travaux dans la littérature se sont consacrés à l'investigation des mécanismes à l'origine de l'hyperstabilité des ARNm de mucines. Le profil d'expression d'un grand nombre de transcrits sur 947 lignées cellulaires cancéreuses (Affymetrix Genome-Wide Human SNP Array) a récemment été publié (Barretina *et al.*, 2012). Nous avons constaté dans ces résultats qu'un ensemble de lignées cancéreuses issues de la sphère pancréato-biliaire surexpriment concomitamment *MUC1*, *MUC4*, *MUC16* ainsi que les transcrits de la galectine-3 (*LGALS3*). Nous avons donc décidé d'utiliser ce type de modèle pancréato-biliaire pour aborder l'étude du rôle potentiel de la galectine-3 dans la régulation de l'expression des mucines membranaires.

Afin de mener à bien cette étude, nous avons utilisé la lignée cellulaire cancéreuse pancréatique humaine CAPAN-1 dans laquelle l'expression de la galectine-3 a été supprimée par *sh*-RNA (cellules Sh1) ainsi que la lignée contrôle correspondante (cellules Sc) (Merlin *et al.*, 2011). Cette lignée représente pour nous un modèle pertinent de cancérogenèse pancréatique adapté à l'étude de *MUC4* au regard du profil d'expression qui la caractérise, c'est-à-dire une expression importante de *MUC4* mais aussi de *MUC1* et de *MUC16*. Des travaux préalables à mon arrivée avaient mis en évidence une diminution de la quantité des transcrits de *MUC1* et de *MUC4* dans la lignée Sh1 (Merlin *et al.*, 2011).

Comme nous l'avons décrit au cours de l'introduction du manuscrit, l'expression des gènes eucaryotes est un mécanisme finement régulé pouvant s'exercer à différentes étapes, depuis la séquence d'ADN jusqu'aux produits finaux des gènes. Ainsi sont concernés la structure de la chromatine, la transcription, la maturation des ARN pré-messagers, la traduction et la stabilité des ARNm ainsi que celle des protéines. Nous avons donc eu comme objectif de déterminer le(s) niveau(x) où la galectine-3 intervient dans la régulation de l'expression des mucines membranaires et plus particulièrement de *MUC4* ainsi que les mécanismes mis en jeu.

La démarche de notre étude s'est basée sur les connaissances récentes relatives aux mécanismes impliqués dans la régulation de l'expression des gènes chez les eucaryotes, ainsi que sur les données concernant les mucines et la galectine-3, disponibles au sein de l'équipe ou issues de la littérature.

Nous avons donc envisagé deux hypothèses :

- l'hypothèse d'une régulation transcriptionnelle : la galectine-3 modifierait l'activité du promoteur des gènes de mucines;
- l'hypothèse d'une régulation post-transcriptionnelle : la galectine-3 modifierait la stabilité des transcrits des mucines.

Ensuite, au vu de nos résultats qui ont fait l'objet d'un premier article dans le journal *Scientific Reports*, il nous a semblé très pertinent d'élargir nos investigations en étudiant de façon plus large le rôle de la galectine-3 dans le métabolisme des ARNm autres que ceux codant les mucines. Nous avons obtenu, dans ce domaine des premiers résultats encore non publiés qui seront présentés dans la seconde partie des résultats de notre étude.

Résultats

1^{ère} partie : Galectine-3 et ARNm de mucines

I. Etude de l'expression des gènes et protéines d'intérêt dans les lignées cellulaires Sh1 et Sc

Nous avons réalisé notre étude sur la lignée cellulaire Sh1, précédemment mise en place au laboratoire à partir de la lignée cancéreuse pancréatique humaine CAPAN-1 dans laquelle l'expression de la galectine-3 a été supprimée par méthode de *sh*-RNA (Merlin *et al.*, 2011). Nous avons utilisé également la lignée cellulaire-contrôle correspondante (Sc).

Nous avons quantifié l'expression des ARNm des trois gènes codant les mucines membranaires *MUC1*, *MUC4* et *MUC16* par PCR quantitative en temps réel (RT-qPCR). Les amorces utilisées pour la quantification ont été choisies de façon à détecter la majorité des variants d'épissage (amorces localisées dans la région 3'UTR). Les résultats indiquent une baisse importante de l'expression de *MUC4* et *MUC1*, associée à celle de la galectine-3 : l'expression relative de *MUC4* n'est plus que de 16,7% dans les cellules Sh1 par rapport aux cellules Sc (100 %) ; cette diminution est statistiquement significative au seuil de 1% ($p < 0,01$) (figure 67). En ce qui concerne *MUC1*, l'expression résiduelle est à un niveau de 21,8 % dans les cellules Sh1, quantité significativement plus faible par rapport à la quantité présente dans les cellules contrôles ($p < 0,01$). L'expression de *MUC16* reste quant à elle inchangée entre les deux types de cellules (93,9 % d'expression relative dans les cellules Sh1).

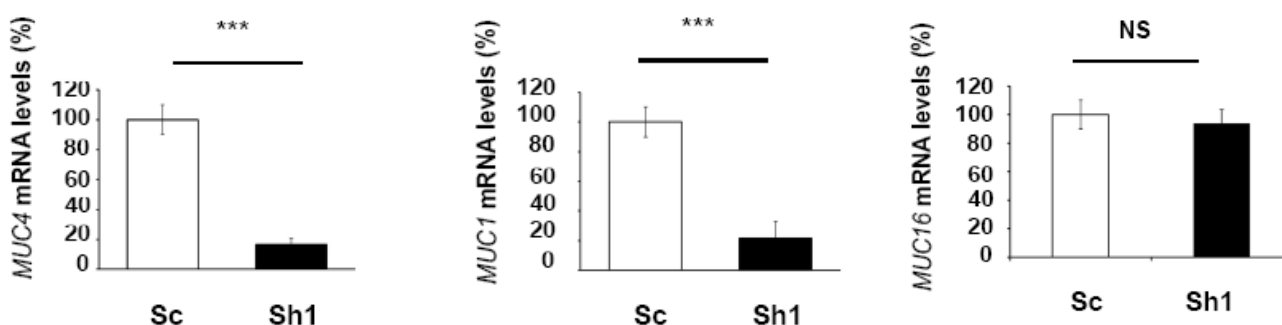


Figure 67: Expression relative de *MUC4*, *MUC1* et *MUC16* dans les cellules Sc et Sh1 évaluée par RT-qPCR.

Le graphique représente la moyenne $\pm$ l'écart-type de trois expériences indépendantes où les échantillons ont été analysés en triple (*** : $p < 0,001$, NS : non significatif).

La galectine-3 exerce donc un contrôle positif sur l'expression des ARNm de *MUC4* et de *MUC1* mais aucun contrôle sur ceux de *MUC16*.

Pour conforter ce premier résultat, nous avons également réalisé une expérience de restauration de phénotype. Il s'agissait de déterminer si l'ajout de galectine-3 recombinante dans le milieu de culture permet aux cellules Sh1 de retrouver des niveaux d'expression des ARNm de *MUC1* et *MUC4* comparables à ceux observés dans les cellules-contrôles Sc. Pour cela, les cellules Sh1 ont été traitées par de la galectine-3 recombinante (rGal-3) pendant 36 h puis les ARN ont été étudiés par RT-PCR semi-quantitative (figure 68).

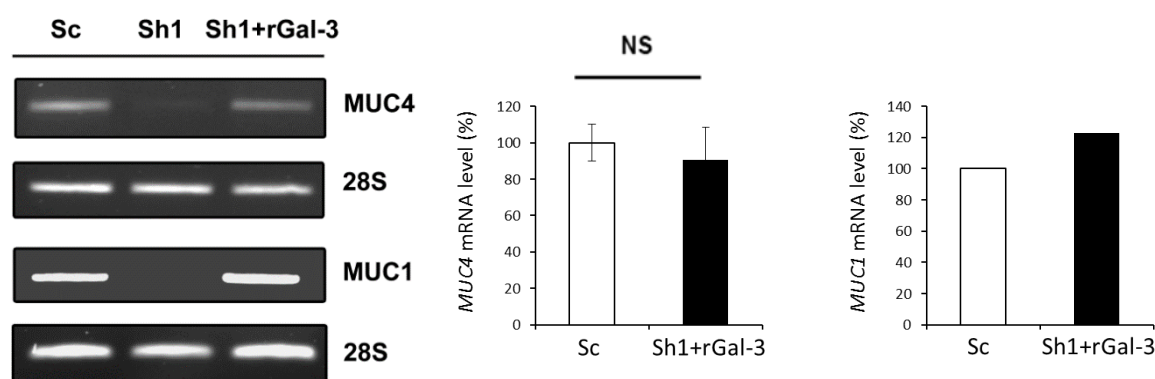


Figure 68 : Etude de l'expression de *MUC4* (n=3) et de *MUC1* (n=1) dans les cellules Sc, Sh1 et Sh1 traitées par 4 μ M de galectine-3 recombinante (36 h) par RT-PCR semi-quantitative.

L'ARN 28S est utilisé comme contrôle interne. Les graphiques de droite représentent l'intégration densitométrique des bandes de *MUC4* (n=3) et de *MUC1* (n=1) visualisées par PCR et normalisées par rapport aux résultats d'intégration des bandes correspondantes aux transcrits 28S.

Les résultats indiquent que, dans les cellules Sh1 traitées par la galectine-3, les niveaux d'expression des ARNm de *MUC4* et de *MUC1* redeviennent similaires à ceux présents dans les cellules Sc. L'expression des ARNm de ces mucines est donc stimulable expérimentalement suite à un apport de galectine-3 exogène dans les cellules Sh1 qui en sont dépourvues. **Ceci confirme la relation entre galectine-3 et expression des mucines.**

II. Etude du rôle de la galectine-3 dans la régulation de l'expression des mucines membranaires MUC1 et MUC4

II.1. Etude de l'activité du promoteur de MUC1 et MUC4 sous l'influence de la galectine-3

Afin d'expliquer le rôle de la galectine-3 sur l'expression des ARN messagers de *MUC1* et *MUC4* nous avons tout d'abord exploré l'hypothèse d'une régulation transcriptionnelle exercée par la galectine-3 via la modification de l'activité des promoteurs de *MUC1* et de *MUC4*.

Nous avons tout d'abord validé par Western-blot (figure 69) l'induction de l'expression de la galectine-3 après transfection transitoire des cellules Sh1 par des quantités croissantes du vecteur d'expression codant la galectine-3 (pCMV6-XL4, Origene) par rapport au vecteur « vide » servant de contrôle.

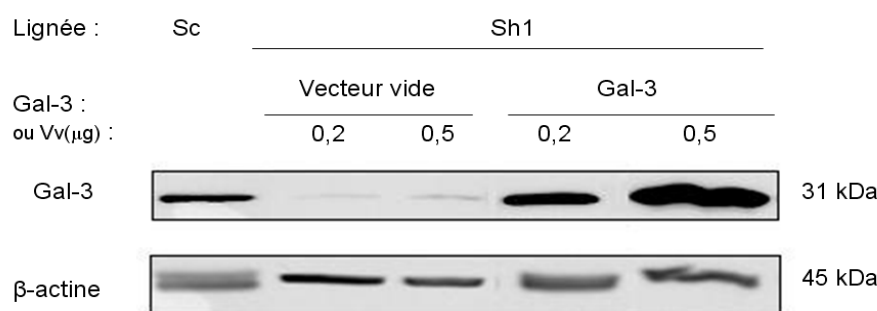
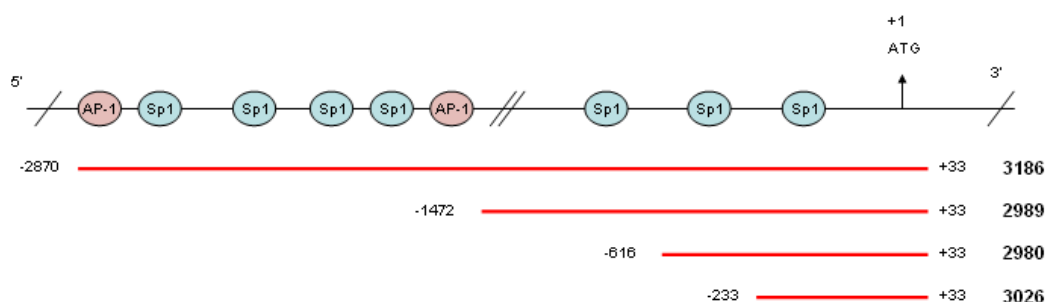


Figure 69 : Etude de l'expression de la galectine-3 dans les cellules Sh1 transfectées transitoirement par le vecteur d'expression de la galectine-3. Celle-ci a été réalisée par *Western Blot*, à partir d'extraits cellulaires totaux de cellules Sc et Sh1 transfectées par le vecteur d'expression de la galectine-3 ou par le vecteur vide (Vv). La β -actine sert de contrôle de charge.

Puis, nous avons sélectionné des constructions contenant le promoteur de *MUC1* ou de *MUC4*, cloné en amont du gène codant la luciférase, parmi celles disponibles au laboratoire (Perrais *et al.*, 2001). Notre choix s'est porté sur celles qui contenaient des éléments *cis*-régulateurs pour les facteurs de transcription AP-1 et Sp1. En effet, les données de la littérature indiquent que la galectine-3 régule l'expression de certains gènes-cibles par le biais de la stabilisation des facteurs de transcription AP-1 et Sp1 sur leurs promoteurs (Lin *et al.*,

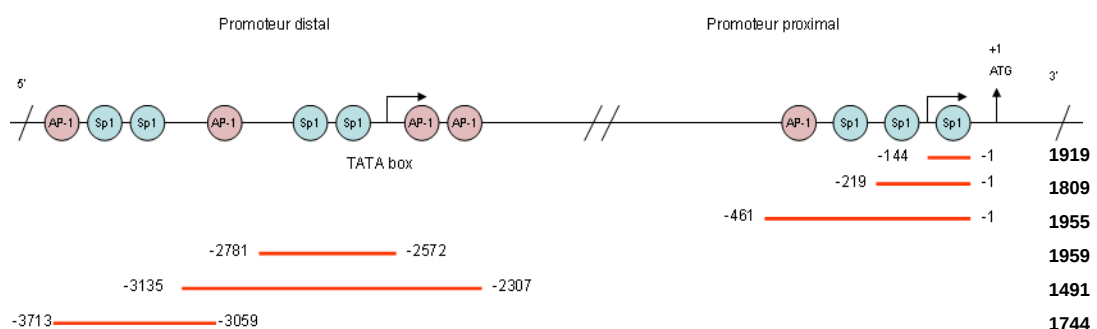
2002; Song *et al.*, 2005). Nous avons donc sélectionné pour *MUC1*, 4 constructions et pour *MUC4*, 6 constructions recouvrant le promoteur proximal et distal. La figure 70 schématise les promoteurs de ces deux mucines, les éléments *cis* régulateurs pour les facteurs de transcription AP-1 et Sp1 ainsi que la position des constructions sélectionnées.

A MUC1



N° des constructions du promoteur de *MUC1*

B MUC4



N° des constructions du promoteur de *MUC4*

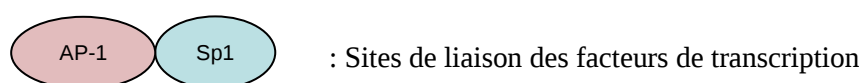


Figure 70 : Positionnement des éléments *cis*-régulateurs pour les facteurs de transcription Sp1 et AP-1 au sein des promoteurs de *MUC1* (A) et de *MUC4* (B). Les constructions utilisées dans la suite des expériences sont représentées par un trait rouge. La numérotation se réfère au A du premier ATG (Perrais *et al.*, 2001).

Les expériences de co-transfections cellulaires transitoires de ces différentes constructions en présence du vecteur d'expression codant la galectine-3 (ou du vecteur-

contrôle vide) ont été réalisées sur cellules Sh1. Les résultats obtenus révèlent que les activités luciférase relatives mesurées tant pour le promoteur de *MUC1* (figure 71A) que pour celui de *MUC4* (figure 71B) ne sont pas significativement différentes de celles obtenues pour les contrôles.

Par conséquent, la galectine-3 ne semble pas moduler l'activité des promoteurs de *MUC1* et de *MUC4*.

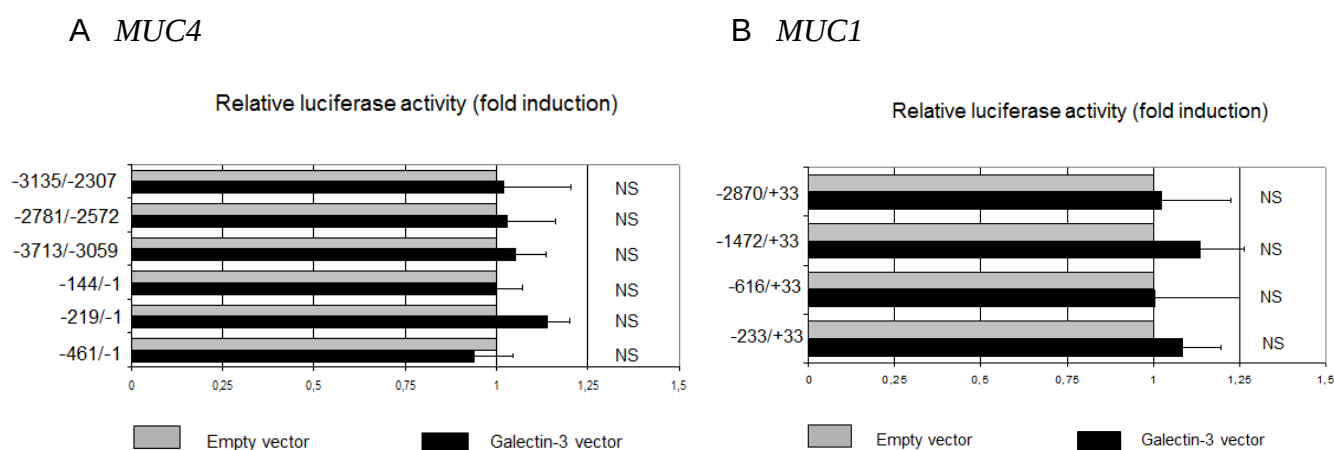


Figure 71 : Effets de la galectine-3 sur l'activité des différentes régions des promoteurs de *MUC1* et de *MUC4*.

Les activités luciférase relatives ont été obtenues lors des co-transfections transitoires des cellules Sh1 par le vecteur d'expression de la galectine-3 et les constructions du promoteur de *MUC4* (A) ou de *MUC1* (B). Ces résultats sont exprimés en niveau d'induction d'activité luciférase relative par rapport au vecteur d'expression vide (dont la valeur est posée à 1). Les résultats représentent la moyenne $\pm$ l'écart-type de deux expériences indépendantes.

II.2. Etude de l'influence de la galectine-3 sur la stabilité des transcrits de *MUC1* et de *MUC4*

II.2.1. Etude de la demi-vie des transcrits de *MUC1* et de *MUC4*

Après avoir écarté l'hypothèse d'une régulation transcriptionnelle, nous avons recherché une influence de la galectine-3 sur la régulation de l'expression des mucines au niveau post-transcriptionnel. Par conséquent, nous nous sommes demandés en première intention si la

galectine-3 pouvait influencer la stabilité des ARNm de *MUC1* et de *MUC4*. Pour cela, nous avons déterminé la demi-vie de ces transcrits en traitant les cellules Sc et Sh1 par l'actinomycine D (8 µg/ml) (Debailleul *et al.*, 1998). Cet antibiotique exerce un effet intercalant sur l'ADN au niveau des séquences riches en GC en inhibant l'élongation de la transcription (Bensaude, 2011). Il est donc possible de suivre la cinétique de décroissance des ARNm déjà transcrits au moment du traitement. Afin de mesurer la demi-vie des transcrits, les ARN totaux ont été récupérés après différents temps d'incubation en présence de la drogue (0h, 3h, 6h, 12h, 24h et 30h) et quantifiés par RT-qPCR afin de tracer les courbes de décroissance des transcrits de *MUC1* et de *MUC4* (quantité restante en % par rapport au 100% fixé au temps 0). De ces courbes nous avons déduit la demi-vie des transcrits, qui correspond au temps nécessaire pour que la quantité des ARNm présents au temps 0 soit divisée par 2 (figure 72).

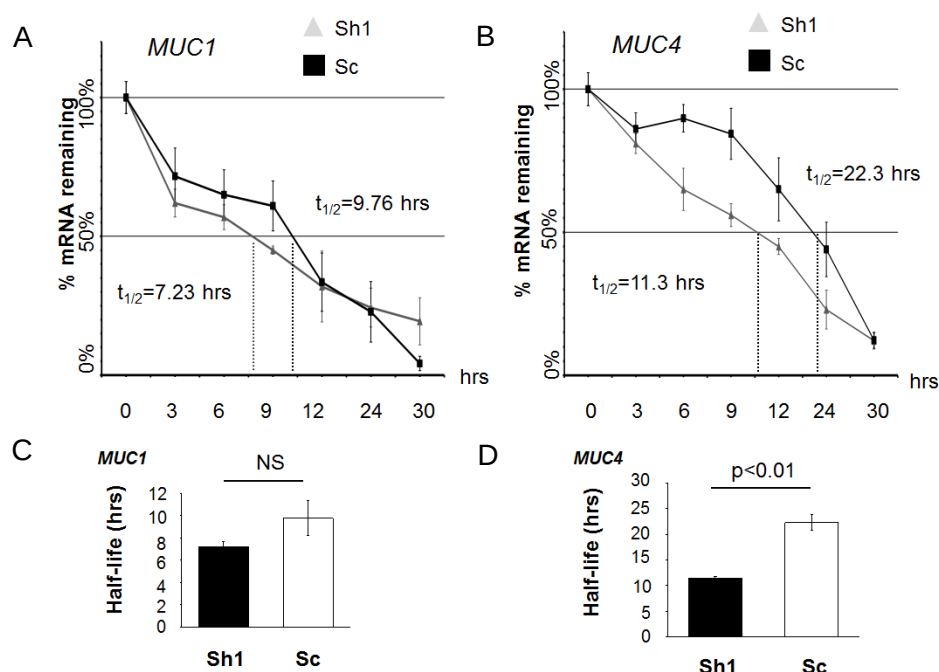


Figure 72 : Etude de la demi-vie des transcrits de *MUC1* et de *MUC4*.

(A et B) Courbes traduisant la diminution des ARNm de *MUC1* (A) et de *MUC4* (B), déterminées par qPCR dans les cellules Sc et Sh1 (quantité normalisée grâce à la quantité d'ARN 18S de chaque échantillon). Chaque point des graphiques représente la moyenne $\pm$ l'écart-type de trois expériences indépendantes.

(C et D) Demi-vie des transcrits de *MUC1* (C) et de *MUC4* (D) dans les cellules Sc et Sh1. Les graphiques représentent la moyenne $\pm$ l'écart-type de la demi-vie des ARNm de *MUC1* et de *MUC4* dans les cellules Sc et Sh1 (n.s.: non significatif).

La demi-vie des transcrits de *MUC4* est de $22,3 \text{ h} \pm 1,6 \text{ h}$ (moyenne $\pm$ SEM, $n=3$) dans les cellules Sc et de $11,3 \text{ h} \pm 0,5 \text{ h}$ dans les cellules Sh1, différence statistiquement significative ($p<0,01$) (figure 72). En ce qui concerne *MUC1*, la demi-vie des transcrits n'est pas modifiée entre les cellules Sc et Sh1 ($9,8 \text{ h} \pm 2,7 \text{ h}$ dans les cellules Sc et $7,2 \text{ h} \pm 0,8 \text{ h}$ dans les cellules Sh1 ; différence statistiquement non significative). Le premier temps pour lequel une différence significative de la quantité des transcrits de *MUC4* entre les cellules Sc et Sh1 est observée se situe à 6 h ($p<0,05$).

Ces résultats indiquent donc que la galectine-3 stabilise spécifiquement les ARNm de *MUC4* mais pas ceux de *MUC1*. La régulation de l'expression de *MUC4* par la galectine-3 s'effectue donc au niveau post-transcriptionnel.

II.2.2. Confirmation de la régulation de l'expression de MUC4 par la galectine-3 au niveau post-translationnel

Afin de confirmer le rôle régulateur exercé par la galectine-3 sur l'expression de *MUC4* au niveau post-transcriptionnel, nous avons combiné un traitement par Act. D ($8 \mu\text{g/ml}$) à un ajout de galectine-3 recombinante dans le milieu de culture des cellules Sh1 pendant 6 h. Cette durée de traitement de 6 h par Act. D a été choisie puisque c'est le tout premier temps pour lequel une différence significative de la quantité des ARNm de *MUC4* résiduels est observée entre les cellules Sc et Sh1. L'expression des mucines a été quantifiée par RT-qPCR. Les résultats, présentés sur la figure 73, indiquent que l'ajout de galectine-3 recombinante dans le milieu de culture s'accompagne bien d'un ralentissement significatif de la décroissance des transcrits de *MUC4* alors que ce phénomène n'est pas observé pour les ARNm de *MUC1*.

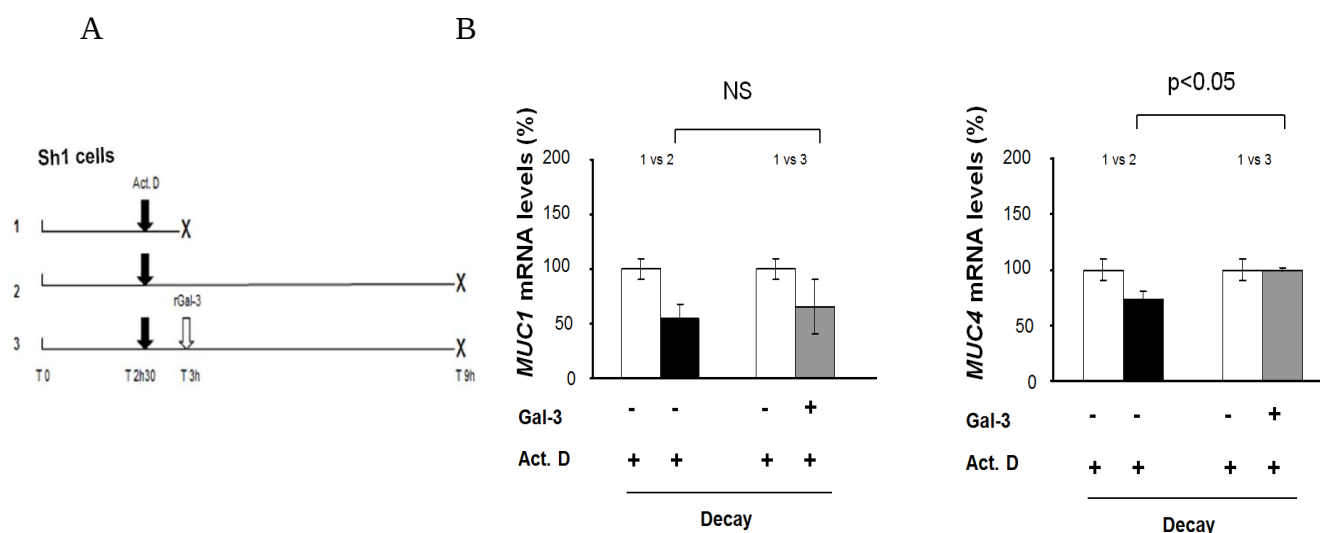


Figure 73 : Influence de la galectine-3 recombinante sur la décroissance des transcrits de mucines.

(A) Schéma de l'expérience, la croix représente la récupération des ARNm.

(B) Les cellules ont été traitées pendant 6 h par l'Act. D couplé à un traitement par la galectine-3 recombinante (rGal-3) puis les ARNm ont été quantifiés par RT-qPCR. Les résultats représentent la moyenne $\pm$ écart-type de trois expériences indépendantes et sont exprimés par rapport à la quantité d'ARNm de *MUC4* présente dans la condition 1 (arbitrairement fixée à 100 %) (NS : non significatif).

De façon complémentaire, grâce à une collaboration établie avec le Docteur Cyrille Grandjean (Faculté des Sciences et Techniques, Nantes), nous avons obtenu une molécule inhibitrice de la galectine-3 correspondant à un dérivé saccharidique et ciblant le CRD ainsi qu'un analogue structural dépourvu de propriété inhibitrice envers cette galectine. Nous avons traité les cellules Sc par l'Act. D pendant 6 h couplé à un traitement soit par l'inhibiteur chimique de la galectine-3 (Gal-3 Pos), soit par l'analogue structural dépourvu de propriétés inhibitrices de la galectine-3 et servant de contrôle (Gal-3 Neg) soit par le système de reprise de ces molécules (véhicule : eau) (figure 74). La quantification des transcrits de *MUC4* par RT-qPCR indique que le véhicule et l'analogue structural ne modifient pas la quantité des transcrits de *MUC4* par rapport aux cellules non traitées. A l'inverse, le traitement par l'inhibiteur spécifique de la galectine-3 s'accompagne d'une diminution statistiquement significative de ces ARNm ($p < 0,05$) après 6 h de traitement par Act. D.

L'ensemble de ces résultats confirment donc le rôle de la galectine-3 dans la régulation de *MUC4* au niveau post-transcriptionnel.

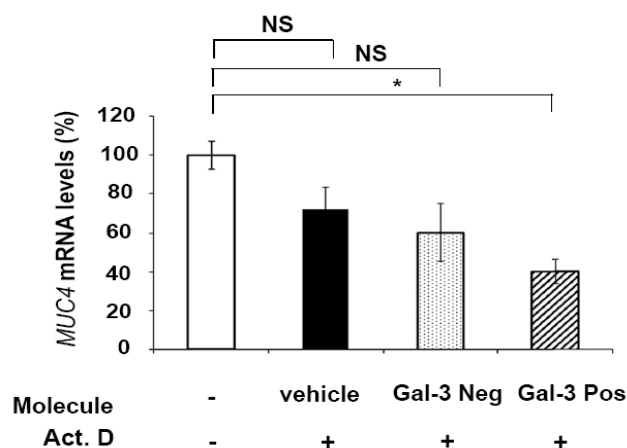


Figure 74 : Influence du traitement par un inhibiteur de la galectine-3 sur la quantité des transcrits de *MUC4*.

Les cellules Sc ont été traitées pendant 6 h par l'Act. D (8 µg/ml), en couplage à un traitement par : soit un inhibiteur chimique de la galectine-3 (Gal-3 Pos), soit un analogue structural dépourvu de propriétés inhibitrices de la galectine-3 et servant de contrôle (Gal-3 Neg) (200 µM) ou soit par le système de reprise de ces molécules (véhicule : eau). La quantification des transcrits de *MUC4* a été réalisée par RT-qPCR (moyenne ± écart-type) et est exprimée en fonction de la quantité présente dans la condition sans Act. D où le niveau est arbitrairement fixé à 100 % (NS : non significatif, * : $p < 0,05$).

II.3. Relevance des résultats dans des tissus d'adénocarcinomes pancréatiques canaux humains

Afin de déterminer si nos résultats obtenus sur une lignée cellulaire pancréatique cancéreuse avaient bien une relevance à la pathologie humaine, nous avons étudié 10 couples de tissus humains sélectionnés en tant que « adénocarcinomes pancréatiques canaux (PDAC)-tissus sains associés » (collaboration avec la Tumorothèque du CHRU de Lille et les pathologistes de l'équipe). Nous y avons quantifié l'expression des transcrits de *MUC4*, *MUC1*, *MUC16*, *LGALS3* et *18S* par RT-qPCR et nous avons constaté qu'il existe une corrélation positive entre l'expression de *MUC4* et celle de *LGALS3* ($r^2 = 0,7331$, $p < 0,001$) (figure 75). En effet, ces 10 couples de tissus peuvent être divisés en deux sous-groupes : un premier (groupe B) où une forte expression des transcrits de la galectine-3 dans le tissu tumoral est associée à une forte expression de *MUC4* par rapport au tissu sain correspondant et un second où l'expression de *MUC4* et *LGALS3* est faible dans le tissu tumoral (groupe A). En ce qui concerne les transcrits de *MUC1* et *MUC16*, nous n'avons pas mis en évidence de

corrélation entre leur expression et celle de *LGALS3* (respectivement $r^2=0,0022$ et $r^2=0,0701$) confirmant l'ensemble des résultats obtenus jusqu'à présent.

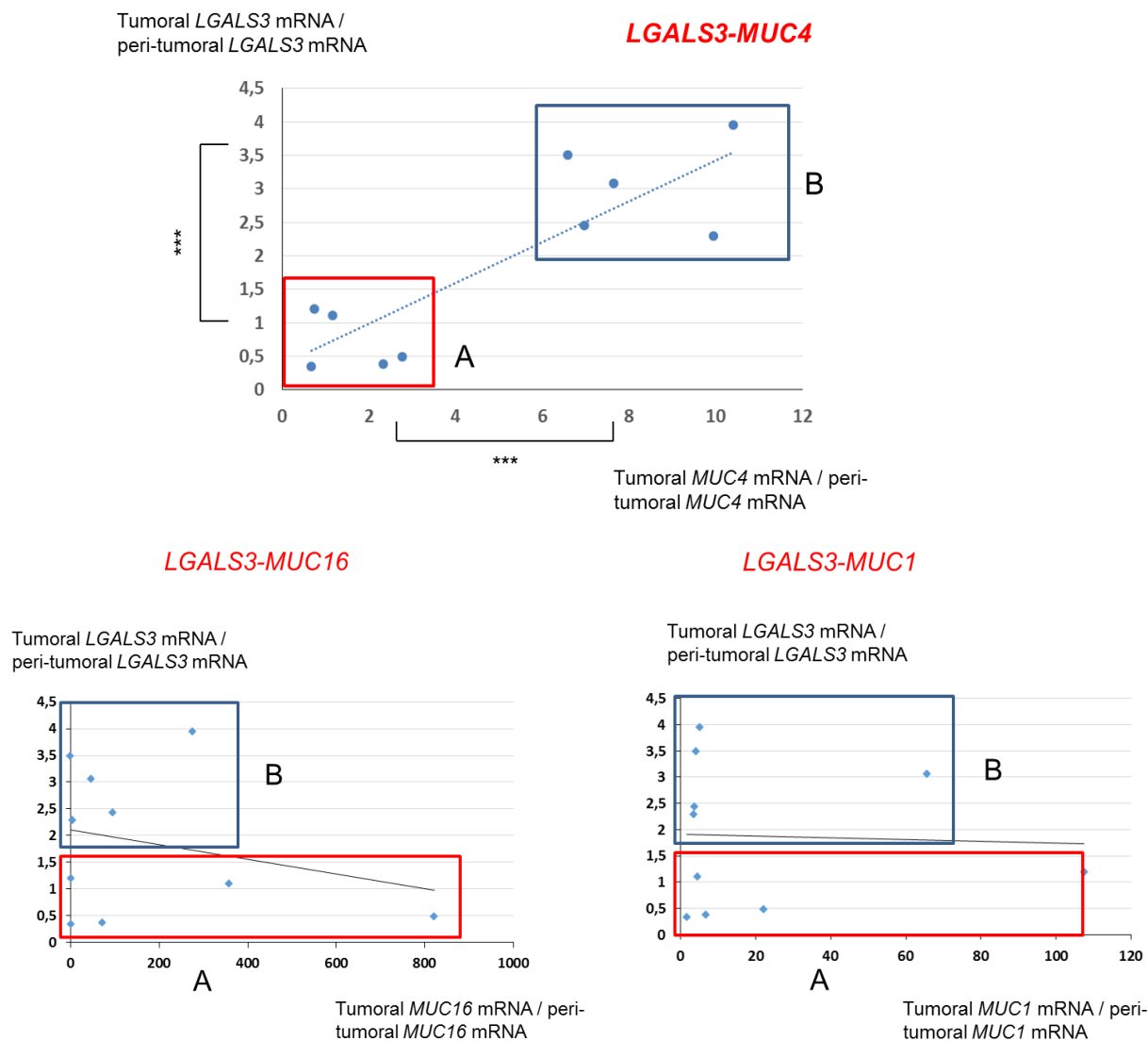


Figure 75 : Etude de l'expression des transcrits de *MUC4*, *MUC1*, *MUC16* et *LGALS3* dans 10 couples de PDAC-tissus sains correspondants. Les transcrits ont été quantifiés par RT-qPCR. Les résultats sont exprimés en facteur d'induction d'expression dans le tissu tumoral par rapport au tissu sain (niveau arbitrairement fixé à 1) pour chacun des 10 tissus (** $p < 0,001$).

III. Mécanisme de stabilisation des transcrits de MUC4 par la galectine-3

III.1. Rôle du 3'UTR de MUC4 dans l'effet stabilisateur de la galectine-3

Nous avons ensuite cherché à déterminer le mécanisme par lequel la galectine-3 agit sur la stabilisation des transcrits de *MUC4*. La stabilité des ARNm étant souvent dépendante de motifs particuliers, notamment de sites de fixation pour des *RNA-binding proteins* présents dans leur région 3'UTR, nous avons donc analysé la région 3'UTR de *MUC4* à l'aide de l'outil RBPDB (*RNA-Binding Protein DataBase*) (Cook *et al.*, 2011). Le 3' UTR de *MUC4* contient notamment 3 motifs ARE (*AU-rich element*) potentiellement reconnus par la protéine HuR, une région *CA repeat* (CARE) ainsi que des sites potentiels de fixation pour des micro-ARN (figure 76).

Homo sapiens MUC4 gene, 3' flanking region (AJ010901.1)

```

TAC TTC AGA TGC GAT GGC TAC AAG GGC TAC GAC CTG GTC TAC AGC CCC CAG
Y F R C D G Y K G Y D L V Y S P Q
AGC GGC TTC ACC TGC GTG TCC CCG TGC AGT AGG GGC TAC TGT GAC CAT GGA GGC CAG TGC
S G F T C V S P C S R G Y C D H G G Q C
CAG CAC CTG CCC AGT GGG CCC CGC TGC AG-INTRON 24-C TGT GTG TCC TTC TCC ATC TAC
Q H L P S G P R C S C V S F S I Y
ACG GCC TGG GGC GAG CAC TGT GAG CAC CTG AGC ATG AAA CTC GAC GCG TTC TTC GGC ATC
T A W G E H C E H L S M K L D A F F G I
TTC TTT GGG GCC CTG GGC GGC CTC TTG CTG CTG GGG GTC GGG ACG TTC GTG GTC CTG CGC
F F G A L G G L L L L L G V G T F V V L R

TTC TGG GGT TGC TCC GGG GCC AGG TTC TCC TAT TTC CTG AAC TCA GCT GAG GCC TTG CCT
F W G C S G A R F S Y F L N S A E A L P

TGAAAGGGGCAGCTGTGGCCTAGGCTACCTCAAGACTCACCTCATCCTTACCGCACATTAAAGCGCCATTGCTTTTGG
X ARE1 MIR 150

GAGACTGGAAAAGGGAAGGTGACTGAGGGCTGTGCTCAGGATTCTT|CAAGGAGAATGAATACTGGGAATCAAGACAAGACT
CARE PROBE START
ATACCTTATCCATAGGCGCAGGTGCACAGGGGGAGGCCATAAAGATCAAACATGCATGGATGGGTCTCTACGCAGACAC
MIR 210 CA-REPEAT
ACCCACAGAAGGACACTAGCCTGTGCACGGCGCGCTGACACACACACACACACGAGTTCATAATGTGGTGTATGG
AS2 AS1

CCCTAAGTTAAGCAAAATGCTTCTGCACACAAAACCTCTCTGGTTTACTTCAAATTAACTCTATTTTAAATAAAGTCTCTC
"ARE-LIKE" ARE2

3866
TGACTTTTGTGCTCTCAAAAAA
ARE3 CARE PROBE STOP

```

Figure 76 : Analyse du 3'UTR de MUC4 à l'aide de l'outil RBPDB (Cook *et al.*, 2011).

Le codon stop est en caractère noir gras, le CARE est en rose, les premiers et derniers nucléotides de la sonde *CA repeat* sont en caractères gras rouge, les régions ciblées par les oligonucléotides antisens sont soulignées en rouge et vert, les éléments ARE sont en caractères bleus, les sites de fixation des micro-ARN sont en violet et le site de polyadénylation est en caractères rouges.

L'alignement du 3'UTR des orthologues de *MUC4* de différentes espèces de mammifères indique qu'un ARE (l'ARE2 sur la figure 76) et que la région CARE sont conservés dans le 3'UTR de *MUC4* de rat et de souris notamment (figure 77).

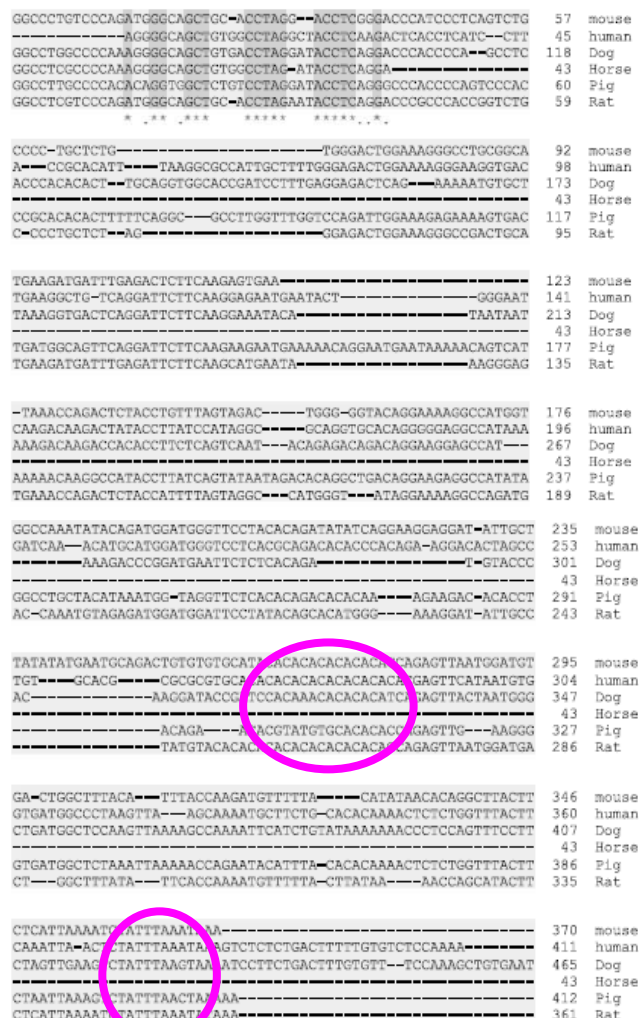


Figure 77 : Alignement du 3'UTR des gènes orthologues de *MUC4* de différentes espèces de mammifères (uniprot.org/align/).

Les éléments d'intérêt conservés (le CARE et un ARE) sont entourés en rose.

Afin d'explorer le possible rôle du 3'UTR de *MUC4* dans l'effet stabilisateur exercé par la galectine-3, nous avons sélectionné une construction disponible au laboratoire contenant cette région clonée dans le vecteur pGL3 promoter codant la luciférase (Lahdaoui *et al.*, 2015) (figure 78). Le vecteur pGL3 promoter dépourvu du 3'UTR de *MUC4* nous a servi de contrôle lors des transfections.

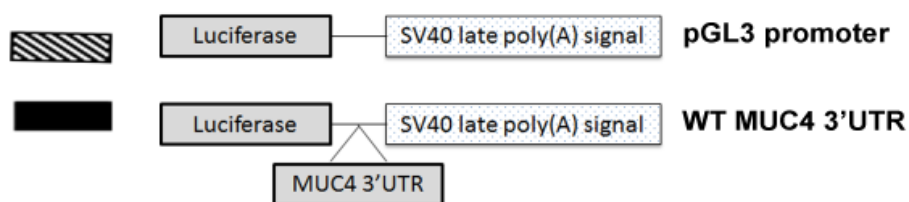


Figure 78 : Schéma représentant le vecteur pGL3 et la construction contenant le 3'UTR de *MUC4* sous forme *wild type* (WT MUC4 3'UTR).

Nous avons ensuite réalisé des co-transfections cellulaires transitoires de l'une ou l'autre de ces deux constructions ainsi que du vecteur d'expression codant la galectine-3 dans les cellules Sh1 suivies soit d'une quantification de l'activité luciférase au niveau protéique soit d'une quantification des transcrits *luciférase*. Nos résultats indiquent que les transfections réalisées en présence d'un vecteur d'expression de la galectine-3 augmentent significativement l'activité luciférase en présence du 3'UTR de *MUC4* ($p < 0,001$; barres noires) alors qu'aucune induction n'est observée en son absence (pGL3 promoteur, barres hachurées) (figure 79).

Nous avons effectué des co-transfections de ces deux constructions avec un vecteur d'expression de la protéine HuR (pcDNA3, fourni par J.A. Steitz, Yale University) qui se fixe sur les éléments ARE. Nous avons constaté une augmentation de l'activité luciférase en présence du vecteur pGL3 promoteur d'un facteur 2 environ, liée à la présence d'éléments ARE dans le 3'UTR du vecteur natif. Cependant, une induction plus forte (d'environ un facteur 4) est observée avec le vecteur possédant le 3'UTR de *MUC4* qui contient 3 ARE additionnels par rapport au vecteur-contrôle (figure 79).

Cependant, la réalisation de co-transfections en présence des vecteurs d'expression de la galectine-3 et de HuR n'augmente pas l'effet inducteur de HuR sur le 3'UTR de *MUC4*.

Ces résultats laissent donc penser que la galectine-3 n'intervient pas *via* les éléments ARE pour exercer son activité stabilisatrice sur les transcrits de *MUC4*.

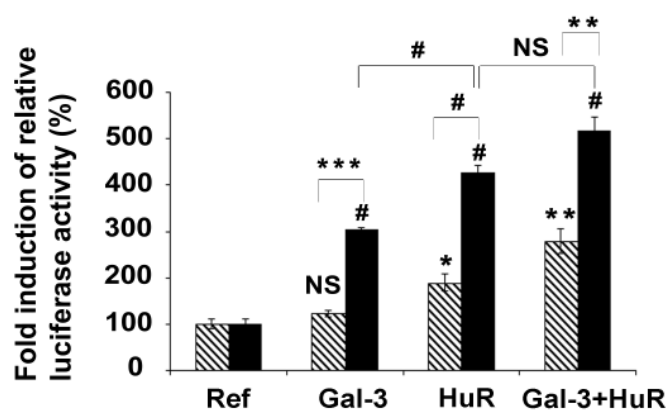


Figure 79 : Etude du 3'UTR de *MUC4* dans l'effet stabilisateur médié par la galectine-3.

Des co-transfections ont été réalisées en présence du vecteur pGL3 promoter (barres hachurées) ou du vecteur pGL3 promoter contenant le 3'UTR de *MUC4* en présence du vecteur d'expression de HuR (pCDNA3), de la galectine-3 (pCMV6-XL4) ou des deux. L'activité luciférase a été mesurée 48 h après les transfections. Un ratio de 100 % a été fixé pour les transfections réalisées avec chaque référence. Les résultats (moyenne $\pm$ écart-type) ont été obtenus lors de trois expériences indépendantes (* : $p < 0,05$; ** : $p < 0,02$; # : $p < 0,01$; *** : $p < 0,001$; NS : non significatif).

Nous nous sommes ensuite intéressés au *CA repeat element* (CARE) présent dans le 3'UTR de *MUC4*, élément qui est également conservé chez le rat et la souris. Ce motif est décrit dans la littérature comme étant reconnu par la protéine hnRNP-L (Shih and Claffey, 1999). Afin d'étudier le rôle potentiel de ce motif, nous avons réalisé une construction du vecteur pGL3 contenant le 3'UTR de *MUC4* dans laquelle la région CARE a été délétée par mutagenèse dirigée (CARE-deleted *MUC4* 3'UTR).

Nous avons ensuite réalisé des expériences de co-transfection du vecteur d'expression de la galectine-3 en présence de la construction WT *MUC4* 3'UTR ou en présence de la construction dépourvue du CARE (figure 80). Comme attendu, l'ajout du vecteur d'expression codant la galectine-3 s'accompagne d'une augmentation des transcrits *luciférase* en présence de la construction WT *MUC4* 3'UTR ($p < 0,01$). Cependant, en absence du CARE, la galectine-3 perd son effet inducteur sur la production des transcrits *luciférase*.

La galectine-3 module donc le niveau des transcrits *luciférase* de manière CARE dépendante.

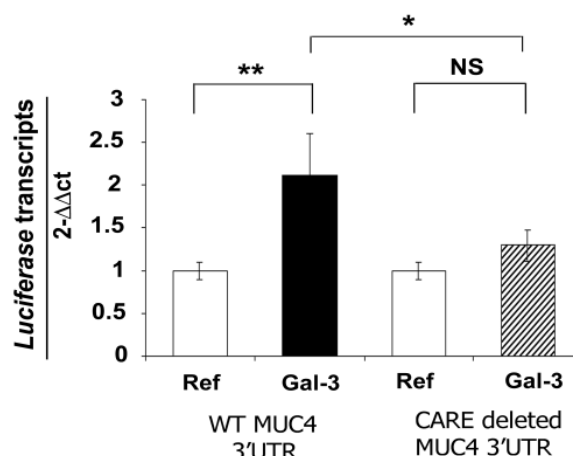


Figure 80 : Etude de l'effet du CARE dans l'effet stabilisateur de la galectine-3.

Les co-transfections ont été réalisées en présence de la construction pGL3 promoter contenant le 3'UTR de *MUC4* délété ou non du CARE et du vecteur d'expression de la galectine-3. Les transcrits *luciférase* et *GAPDH* ont été quantifiés par RT-qPCR 48 h après les transfections. Les résultats représentent la moyenne $\pm$ écart-type de trois manipulations indépendantes (NS : non significatif, * : $p < 0,05$; ** : $p < 0,02$).

Pour compléter et rechercher un rôle éventuel de la protéine hnRNP-L qui est connue pour interagir avec le motif CARE, nous avons utilisé une stratégie faisant appel à des oligonucléotides antisens ciblant :

- soit le site de l'interaction potentielle entre la protéine hnRNP-L et le motif CARE présent dans le 3'UTR de *MUC4* (oligo AS1) ;
- soit une région-contrôle négatif du 3'UTR de *MUC4* dépourvue d'éléments régulateurs, et choisie par analyse bio-informatique (AS2) (figure 76).

La quantification des transcrits de *MUC4* par RT-qPCR indique que l'utilisation du nucléotide *no target* (AS2) ne modifie pas la quantité de transcrits de *MUC4* par rapport à la condition sans oligonucléotide (Ctl) (figure 81). Cependant, le blocage de l'accès de hnRNP-L au CARE par l'oligonucléotide AS1 diminue l'expression de *MUC4* endogène et ceci de façon plus marquée dans les cellules Sh1 ($p < 0,01$).

Ces résultats confirment donc que la galectine-3 régule les transcrits de *MUC4* de manière CARE dépendante.

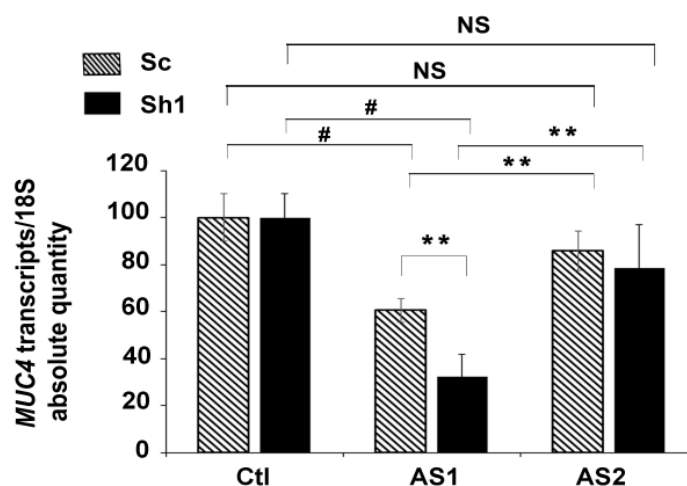


Figure 81 : Confirmation du rôle du CARE du 3'UTR de *MUC4* dans l'effet de la galectine-3. Les transfections ont été réalisées 1/en présence d'oligonucléotides antisens ciblant le CARE (AS1) ou 2/ en présence d'oligonucléotides ciblant une région dépourvue d'éléments régulateurs (AS2) ou 3/ sans oligonucléotides (Ctl). Les transcrits de *MUC4* et *18S* ont été quantifiés par RT-qPCR. Le niveau d'expression des transcrits de *MUC4* dans la condition Ctl a été fixé arbitrairement à 100 %. Les résultats représentent la moyenne $\pm$ écart-type de trois manipulations indépendantes (** : $p < 0,02$; # : $p < 0,01$; NS : non significatif).

Nous nous sommes inspirés des données de la littérature ayant montré que la *RNA-binding protein* hnRNP-L se fixe sur la région CARE présente dans les ARNm du *VEGF* (Shih and Claffey, 1999) pour caractériser une éventuelle interaction entre hnRNP-L et le CARE de *MUC4*. Pour ce faire, nous avons réalisé dans les cellules Sc et Sh1 une expérience de *RNA pull down*. Cette technique permet de détecter une interaction entre une sonde ARN biotinylée correspondant à la région d'intérêt (CARE dans le cas présent) et des extraits cellulaires totaux contenant les protéines supposées s'y fixer. Après élution des protéines fixées, celles-ci sont analysées par Western-blot (figure 82). Nos résultats indiquent que hnRNP-L se fixe sur le CARE de *MUC4* et ceci de façon plus marquée en présence de galectine-3 (augmentation de la fixation d'un facteur 1,6, $p < 0,05$) alors que le niveau d'expression de hnRNP-L reste inchangé entre les cellules Sc et Sh1. De plus, nous avons réalisé l'expérience avec des tampons de faible force ionique (NaCl 60 mM), condition favorisant l'interaction de la galectine-3 avec les ARN (Wang *et al.*, 2006). Le Western-blot révèle que la galectine-3 ne se fixe pas directement sur la sonde.

Nous pouvons donc conclure que :

- hnRNP-L se lie sur la CARE présent dans le 3'UTR de *MUC4* ;

- la galectine-3 ne se fixe pas directement sur cette région ;
- la galectine-3 potentialise donc la fixation de hnRNP-L sur le CARE.

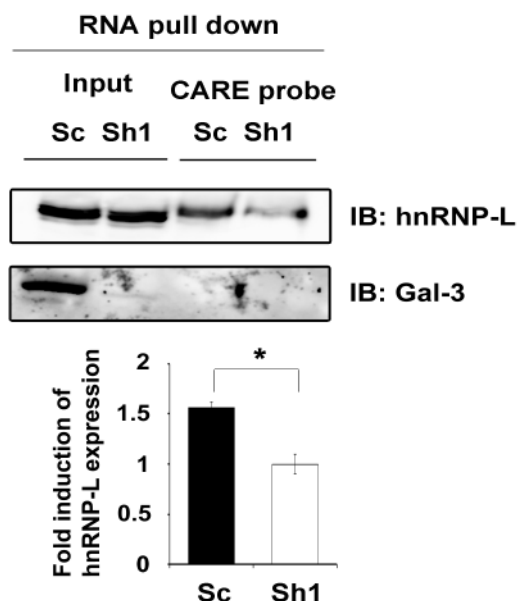


Figure 82 : Etude de l'interaction entre le CARE et hnRNP-L ou la galectine-3 par *RNA pull-down*.

Une sonde ARN biotinylée a été mise en présence d'extraits cellulaires totaux de cellules Sc et Sh1. Les protéines fixées ont été analysées par *Western-blot*. Les bandes ont été quantifiées par analyse densitométrique à l'aide du logiciel ImageQuantTL, les résultats représentent la moyenne $\pm$ écart-type des résultats d'intégration de 2 expériences indépendantes (* : $p < 0,05$).

Afin d'étudier la possibilité d'une synergie entre les protéines hnRNP-L et galectine-3 sur l'effet stabilisateur des transcrits de *MUC4* médié par le CARE, nous avons réalisé des co-transfections cellulaires transitoires en présence du vecteur pGL3 promoter contenant le 3'UTR de *MUC4* WT et :

- du vecteur d'expression de la galectine-3 ou
- du vecteur d'expression de hnRNP-L (pCMV6-XL5) ou
- des deux vecteurs d'expression.

Les activités luciférase mesurées 48 h après ces différentes co-transfections indiquent que les vecteurs d'expression de hnRNP-L et de la galectine-3 pris individuellement stimulent chacun l'activité luciférase ($p < 0,001$) mais l'ajout du vecteur d'expression de la galectine-3 amplifie l'effet stimulateur de hnRNP-L sur le 3'UTR de *MUC4* ($p < 0,02$) (figure 83).

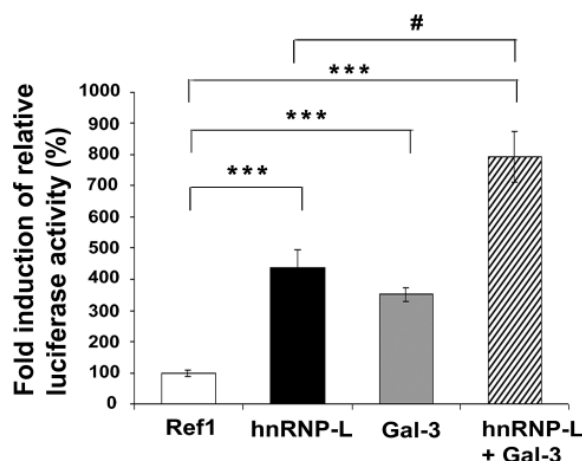


Figure 83 : Etude de l'effet de hnRNP-L et de la galectine-3 sur le 3'UTR de *MUC4*.

Des co-transfections ont été réalisées en présence du vecteur pGL3 promoter contenant le 3'UTR de *MUC4* et des vecteurs d'expression codant soit hnRNP-L (pCMV6-XL5), soit la galectine-3 (pCMV6-XL4), ou des deux vecteurs simultanément. L'activité luciférase a été mesurée 48 heures après les transfections. Un ratio de 100 % a été fixé pour les transfections réalisées avec chaque référence. Les résultats (moyenne $\pm$ écart-type) ont été obtenus lors de trois expériences indépendantes (# : $p < 0,02$; *** : $p < 0,001$).

Par ailleurs, l'inhibition modérée de l'expression de hnRNP-L d'un facteur de 40 % par l'action de *si*-ARN est suffisante pour induire une réduction significative de la quantité des transcrits de *MUC4* endogènes dans les cellules Sc et Sh1 (figure 84). Cette diminution est d'ailleurs plus marquée en absence de galectine-3 confirmant ainsi l'implication partagée de hnRNP-L et de la galectine-3 dans la régulation de l'expression de l'ARNm de *MUC4*.

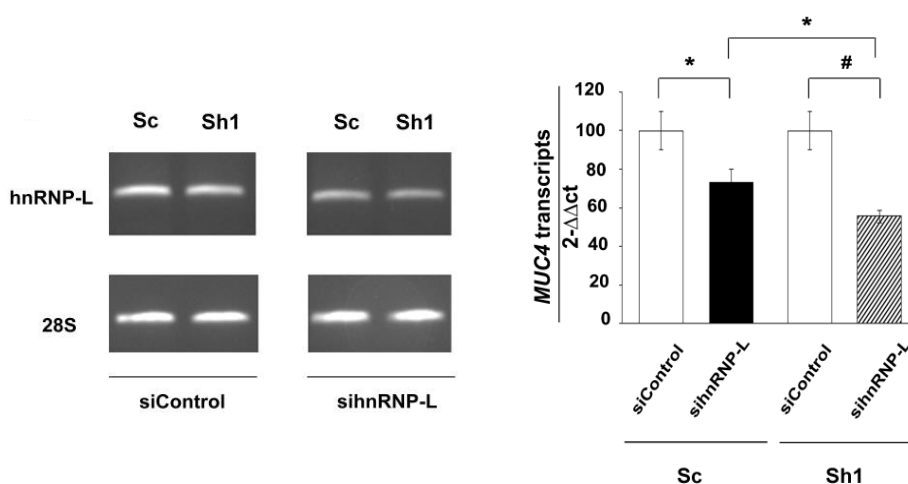


Figure 84 : Inhibition de l'expression de hnRNP-L : effet sur la quantité des ARNm de *MUC4*.

L'effet inhibiteur des *si*-ARN hnRNP-L a été quantifié par RT-PCR semi-quantitative des transcrits de *MUC4* et de la *GAPDH* après traitement des cellules Sc et Sh1 par des *si*-ARN-contrôles ou ciblant hnRNP-L. Les résultats sont exprimés par rapport à la quantité d'ARNm de *MUC4* obtenue avec les *si*-ARN-contrôles (fixée arbitrairement à 100 %). Les résultats (moyenne $\pm$ écart-type) ont été obtenus lors de trois expériences indépendantes (# : $p < 0,01$; * : $p < 0,05$).

Nous venons de voir que la protéine hnRNP-L se fixe directement sur le CARE de *MUC4* mais pas la galectine-3. Nous avons donc cherché à mettre en évidence une interaction entre galectine-3 et hnRNP-L. Sachant que les interactions qui caractérisent la liaison de la galectine-3 avec ses ligands sont des liaisons faibles (Wang *et al.*, 2006), nous avons choisi des conditions expérimentales qui puissent les stabiliser. La co-immunoprécipitation a été réalisée après *cross-linking* des interactions protéine-protéine par du formaldéhyde. Nos résultats (figure 85) montrent que l'anticorps anti-galectine-3 est capable d'immunoprécipiter la protéine hnRNP-L dans les cellules Sc, ce qui confirme l'existence d'une interaction entre ces deux protéines.

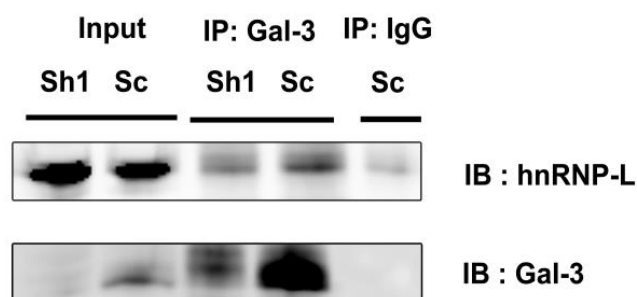


Figure 85 : Etude de l'interaction entre hnRNP-L et la galectine-3 par co-immunoprécipitation.

Les extraits cellulaires totaux des cellules Sc et Sh1 ont été mis en présence de l'anticorps anti-galectine-3 ou de l'isotype correspondant. Une étape préalable de *crosslinking* par le formaldéhyde a été réalisée avant la préparation des extraits cellulaires totaux. Les protéines immunoprécipitées ont été analysées par *Western-blot*.

Pour confirmer ces données de co-immunoprécipitation, nous avons réalisé des expériences de *Proximity ligation assay* (PLA) et de co-immunofluorescence (co-IF) afin de visualiser par imagerie cellulaire, l'interaction entre la galectine-3 et hnRNP-L. En effet, dans l'expérience de PLA, les spots apparaissent uniquement lorsque deux protéines sont distantes au maximum de 40 nm, la colocalisation étant évaluée dans la co-IF par superposition des images. Ainsi, la colocalisation des deux protéines est matérialisée pour le PLA par des spots rouges et pour la co-IF par des spots blancs.

Afin de pouvoir quantifier l'interaction entre la galectine-3 et hnRNP-L, nous avons préalablement réalisé des expériences permettant d'évaluer la sensibilité et la spécificité de

cette technique en utilisant des cibles pour lesquelles les interactions sont validées dans la littérature (MUC4-ErbB2) (Jonckheere *et al.*, 2012) et des cibles n'interagissant pas entre elles en raison d'une différence de localisation cellulaire (MUC4-ZEB1). Les résultats obtenus indiquent que l'interaction MUC4-ErbB2 se traduit par environ 9,6 spots ($\pm 2,1$ spots) par cellule et le témoin négatif (MUC4-ZEB1) ne présente aucun spot (figure 86).

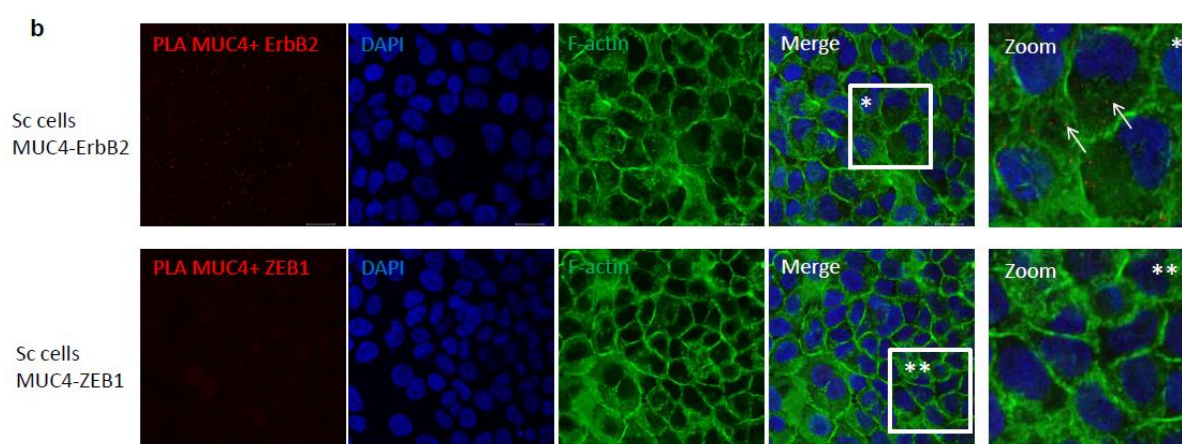


Figure 86 : Contrôle de sensibilité et spécificité de l'expérience de PLA.

Un contrôle positif (panel supérieur) a été effectué en analysant une interaction protéine-protéine validée (MUC4 et ErbB2) (Jonckheere *et al.*, 2012). Le panel inférieur est un contrôle négatif réalisé en ciblant deux protéines qui n'interagissent pas (MUC4 et ZEB-1). Un zoom est réalisé dans la région délimitée par un cadre blanc. Les spots rouges indiquent l'interaction entre les deux protéines, les noyaux sont marqués au DAPI (bleu) et les filaments d'actine sont marqués par la phalloïdine (vert) (échelle : 20 μ m). Les flèches indiquent des spots de PLA.

Afin d'étudier l'interaction entre la galectine-3 et hnRNP-L, les expériences de PLA ont ensuite été réalisées à la fois en situation basale et sur des cellules bloquées en phase G2/M. En effet, les protéines de la famille hnRNP sont des protéines capables de lier de nombreux ligands en lien avec des rôles variables à la fois dans le noyau et dans le cytoplasme, et ceci en fonction du cycle cellulaire. Leurs fonctions et liaisons à des partenaires dans le cadre de la régulation post-transcriptionnelle sont attendues se dérouler en phase G2 plutôt qu'en phase S (Lane *et al.*, 2013). Pour étudier ceci, nous avons donc effectué un traitement préalable des cellules par le nocodazole (Dawkins *et al.*, 2016), un agent antimitotique (Methyl [5-(2-thienylcarbonyl)-1H-benzimidazol-2-yl]carbamate) qui cible le cytosquelette en stimulant la dépolymérisation des microtubules. Il a pour effet de

synchroniser les cellules et de les bloquer en phase G2/M comme indiquent les résultats de cytométrie en flux (figure 87).

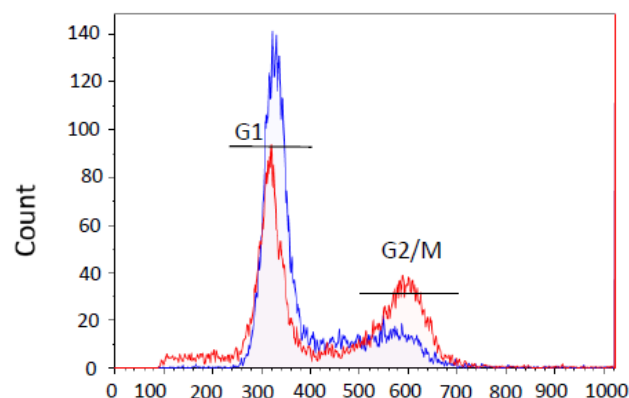


Figure 87 : Evaluation de l'efficacité du nocodazole (400 ng/ml) pour bloquer les cellules Sc en phase G2/M.
Les cellules Sc ont été analysées par cytométrie en flux après marquage par iodure de propidium.

Dans les expériences de PLA, nous observons à l'état basal (*non treated*) que la plupart des situations de co-localisation entre la galectine-3 et hnRNP-L est observée dans le cytoplasme, en région périnucléaire (figure 88A). Cette interaction est significative puisque le nombre de spots est d'environ 11 par cellule ($\pm$ écart-type 3,3), ce qui est équivalent au nombre de spots obtenus pour l'interaction validée entre MUC4 et ErbB2 (figure 86). Il est intéressant de noter aussi que ces interactions sont plus abondantes dans les quelques cellules en mitose. Cette dernière observation est confirmée lorsque les cellules sont traitées par le nocodazole, ce traitement se traduisant effectivement par une augmentation significative du nombre de spots de PLA (29,8 spots par cellules $\pm$ écart-type 9,9 ; $p < 0,001$) et de la colocalisation évaluée par co-immunofluorescence (figure 88 A et B).

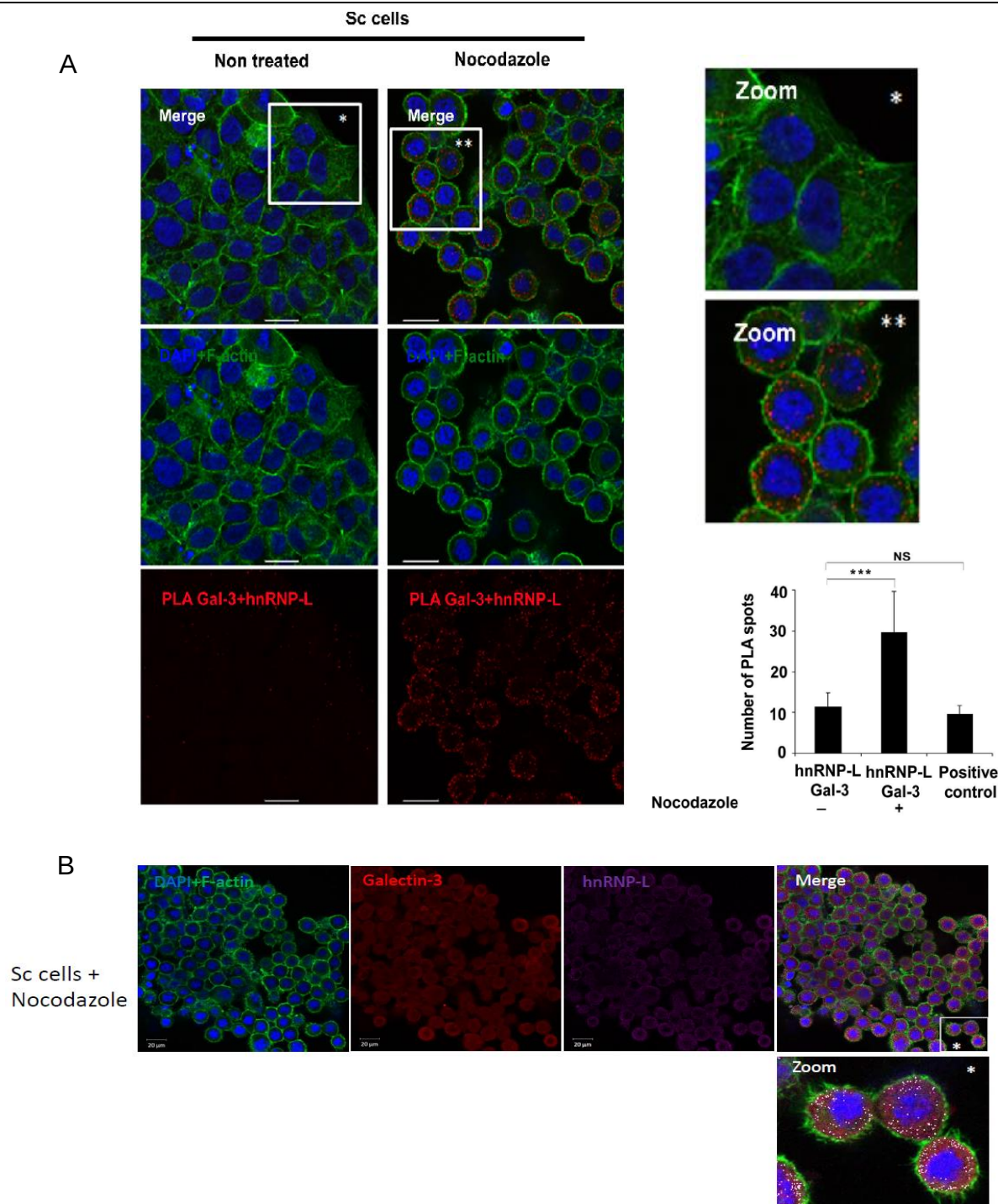


Figure 88 : Evaluation de l'interaction entre hnRNP-L et la galectine-3 par imagerie sur les cellules Sc traitées ou non par du nocodazole.

(A) Dans l'expérience de PLA, les spots rouges indiquent l'interaction entre les deux protéines, les noyaux sont marqués au DAPI (bleu) et les filaments d'actine sont marqués par la phalloïdine (vert) (échelle : 20 μ m). Les spots ont été comptés avec le logiciel ImageJ sur au moins 13 cellules (*** : $p < 0,001$; NS : non significatif)

(B) Dans l'expérience de co-immunofluorescence, la galectine-3 est marquée en rouge et hnRNP-L en violet. Les noyaux sont marqués au DAPI (bleu) et les filaments d'actine sont marqués par la phalloïdine (vert) (échelle : 20 μ m). La colocalisation a été évaluée en superposant les images (merge) et elle est indiquée par des points blancs.

Nous avons également confirmé cet effet par co-immunoprécipitation (figure 89), l'interaction entre la galectine-3 et hnRNP-L étant plus importante lorsque toutes les cellules sont synchronisées et bloquées en phase G2/M dans les cellules Sc. C'est à ce moment du cycle que l'on peut proposer pour hnRNP-L et la galectine-3 des fonctions partagées régulatrices, dans un cadre post-transcriptionnel, sur des ARNm-cibles comme ceux de *MUC4*.

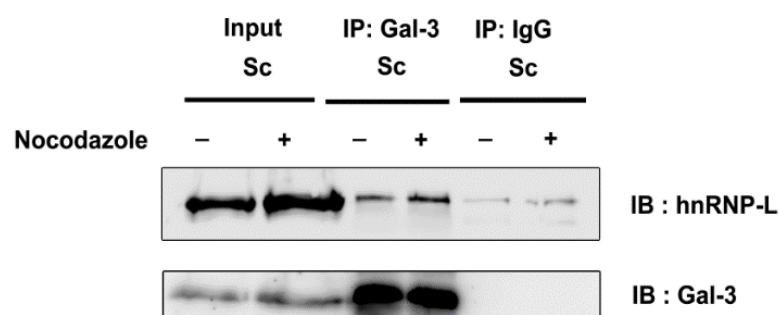


Figure 89 : Evaluation de l'interaction entre galectine-3 et hnRNP-L lors d'un traitement par nocodazole.

La co-immunoprécipitation a été réalisée à l'aide d'un anticorps anti-galectine-3 sur des extraits cellulaires totaux de cellules Sc traitées par nocodazole (400 ng/ml). Une étape préalable de cross-linking par le formaldéhyde a été réalisée avant la préparation des extraits cellulaires totaux. Les protéines immunoprécipitées ont été analysées par *Western-blot*.

A ce stade de l'étude, nous avons donc démontré que la galectine-3 interagit avec hnRNP-L dans le cytoplasme des cellules. De plus, l'effet stabilisateur de la galectine-3 sur les transcrits de *MUC4* s'exerce *via* la potentialisation de la fixation de hnRNP-L sur le CARE présent dans le 3'UTR de *MUC4*.

Nous avons ensuite souhaité déterminer si la stabilisation des transcrits de *MUC4* par la galectine-3 s'accompagne d'une augmentation de la traduction de l'ARNm de cette mucine. Les résultats des Western-blot (figure 90) mettent en évidence une très faible diminution de *MUC4* dans les cellules Sh1 par rapport aux cellules contrôles.

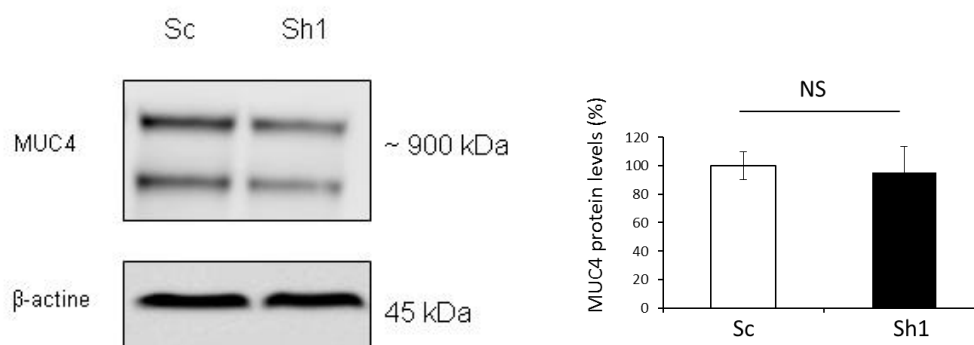


Figure 90 : Etude de l'expression de MUC4 par *Western Blot* dans les cellules Sc et Sh1. Les western-blot ont été réalisés sur des extraits cellulaires totaux, la β -actine sert de contrôle de charge. Le graphique représente l'intégration densitométrique de deux expériences avec la moyenne et l'écart-type du niveau d'expression de MUC4 (%) (NS : non significatif).

La stabilisation des transcrits de *MUC4* par la galectine-3 ne s'accompagne donc pas d'une augmentation de la synthèse protéique. Le rôle de la galectine-3 sur la régulation de l'expression de MUC4 est donc bien post-transcriptionnel et non traductionnel. La galectine-3 favoriserait le stockage des transcrits de MUC4 dans un compartiment où il n'est pas traduit.

III.2. Relevance des résultats *in vivo* étudiée sur des tissus murins

Afin de déterminer la relevance de nos résultats au niveau physiologique, nous avons quantifié par RT-qPCR les transcrits de *Muc4* dans différents tissus provenant de souris sauvages (WT) et invalidées (KO) pour le gène codant la galectine-3 (fond génétique C57BL/6), obtenus en collaboration avec le Dr D. Delacour. Les résultats indiquent que la quantité des transcrits de *Muc4* est plus faible dans le jéjunum des souris KO par rapport aux souris WT alors qu'aucune variation de l'expression de *Muc4* n'est observée au niveau du tissu pulmonaire (figure 91A). Nous avons ensuite évalué par immunohistochimie l'expression de hnRNP-L dans ces mêmes tissus. Celle-ci est plus faible dans le jéjunum des souris KO par rapport aux souris WT. Là encore, aucune variation d'expression d'hnRNP-L au niveau protéique n'est constatée entre les souris WT et KO au niveau pulmonaire (figure 91B). La faible expression de hnRNP-L contribue donc au taux faible des transcrits de *Muc4* observés dans les souris KO pour la galectine-3.

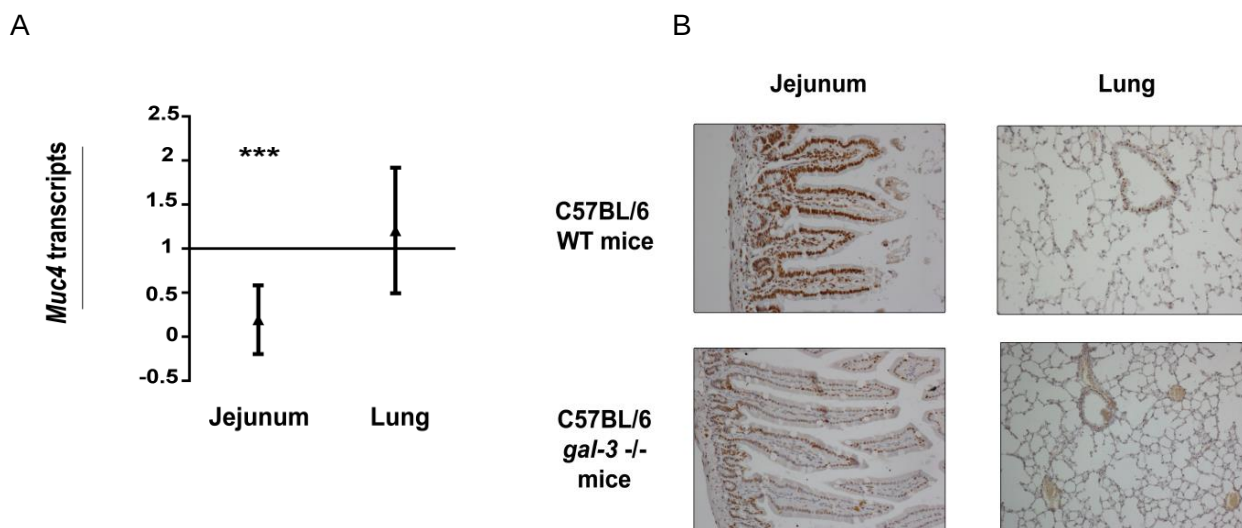


Figure 91 : Etude des tissus murins WT et KO pour *LGALS3* (fond génétique C57BL/6).
 (A) Quantité des transcrits de *Muc4* évalués par RT-qPCR présents dans les souris KO pour la galectine-3 exprimée par rapport au souris WT (pour lesquelles le niveau d'expression est fixé arbitrairement à 1) au niveau du jéjunum (WT n=9, KO n=12) et du poumon (WT n=14, KO n=12). Les transcrits *Gapdh* ont été utilisés comme référence interne (***: $p < 0,001$; NS : non significatif).
 (B) Analyses immunohistochimiques de l'expression de hnRNP-L sur des tissus de jéjunum et de poumon de souris WT et KO pour *LGALS3* (n=4 par condition).

L'ensemble de nos résultats confirment donc une action synergique entre hnRNP-L et la galectine-3 dans le contrôle de la stabilité des transcrits de *MUC4* dans les tissus épithéliaux digestifs normaux (murins) et cancéreux (humains).

IV. Etude de la liaison de la galectine-3 aux ARNm matures dans le cytoplasme.

Nos résultats de *RNA pull down* ont montré que la galectine-3 ne se lie pas directement sur le CARE présent dans la région 3'UTR de *MUC4*. Or la galectine-3 étant connue pour interagir avec les pré-ARNm (Haudek *et al.*, 2009), nous avons donc cherché à caractériser une liaison avec les transcrits de mucines.

Pour cela, nous avons fait appel à une expérience d'immunoprécipitation des complexes ARNm-protéines (RNA-IP), à l'aide d'un anticorps anti-galectine-3 et après *crosslinking* de ces complexes par le formaldéhyde afin d'éviter les possibles réarrangements artefactuels des complexes après la lyse des cellules (Jensen and Darnell, 2008). Les cellules

Sc ont été utilisées pour modèle d'étude. Nous avons analysé par transcriptomique de type RNA-seq les transcrits immunoprécipités par cet anticorps. Nous avons également analysés les transcrits immunoprécipités par un anticorps anti-HuR pour réaliser une expérience « témoin positif », dans la mesure où des répertoires de transcrits immunoprécipités par HuR sont disponibles dans la littérature (Uren *et al.*, 2011) et dans la mesure où le 3'UTR de *MUC4* possède plusieurs sites potentiels de fixation pour cette *RNA-binding protein*. Notre transcriptome « HuR » s'est révélé satisfaisant car partageant une communauté d'ARNm statistiquement proche de celui obtenu par Uren *et al.*, pourtant réalisé sur un autre modèle cellulaire (Uren *et al.*, 2011). Ceci valide donc nos protocoles et résultats de RNA-IP.

Nos résultats indiquent que l'anticorps anti-galectine-3 immunoprécipite 906 transcrits parmi lesquels se trouvent *MUC4*, *MUC16*, *MUC2* mais pas ceux de *MUC1* (tableau XIII).

Tableau XIII: Analyse par RNA-seq des transcrits de mucines immunoprécipités par les anticorps anti-HuR ou anti-galectine-3.

Après préparation des bibliothèques d'ADNc, celles-ci ont été analysées sur le séquenceur PGM. Les *reads* obtenus ont été alignés sur les transcrits de mucines à l'aide de l'outil Bowtie2 (Langmead and Salzberg, 2012) et normalisés en divisant le nombre de *reads* obtenus pour le transcrits par le nombre total de *reads*.

GenBank Acc.	Gene	Description	mRNA size (kbp)	Number of reads		Normalized reads (‰)	
				with Gal-3	with HuR	with Gal-3	with HuR
NM_024690	<i>MUC16</i>	Mucin 16, cell surface associated	≈ 43	29	50	2.25102849	1.00431857
NM_002457	<i>MUC2</i>	Mucin 2, oligomeric mucus/gel-forming	14-16	2	10	0.15524334	0.20086371
NM_152673	<i>MUC20</i>	Mucin 20, cell surface associated	≈ 2	1	3	0.07762167	0.06025911
NM_001098516	<i>MUC20</i>	Mucin 20, cell surface associated		1	1	0.07762167	0.02008637
NM_018406	<i>MUC4</i>	Mucin 4, cell surface associated	14-26	15	27	1.16432508	0.54233203
NM_001018017	<i>MUC1</i>	Mucin 1, cell surface associated	4-7	0	3	0	0.06025911
NM_002456	<i>MUC1</i>	Mucin 1, cell surface associated		0	1	0	0.02008637
NM_001044391	<i>MUC1</i>	Mucin 1, cell surface associated		0	1	0	0.02008637
NM_005961	<i>MUC6</i>	Mucin 6, oligomeric mucus/gel-forming	16,5-18	0	1	0	0.02008637

Nous nous sommes ensuite intéressés aux sites d'interaction de la galectine-3 avec les transcrits de mucines en analysant les lectures (*reads*) obtenues par RNA-seq. Concernant les transcrits de *MUC4*, les *reads* sont alignés au niveau de la jonction exon2-exon3 suggérant d'une part une interaction avec un ARNm épissé et d'autre part une liaison au niveau du plus gros exon du transcrit. Pour *MUC16* et *MUC2* les *reads* sont également localisés au niveau du plus gros exon des transcrits (exon 30 : 7 kb pour *MUC2*, exon 2 : 12 kb pour *MUC4*, exon 3 : 43 kb pour *MUC16*) (figure 92).

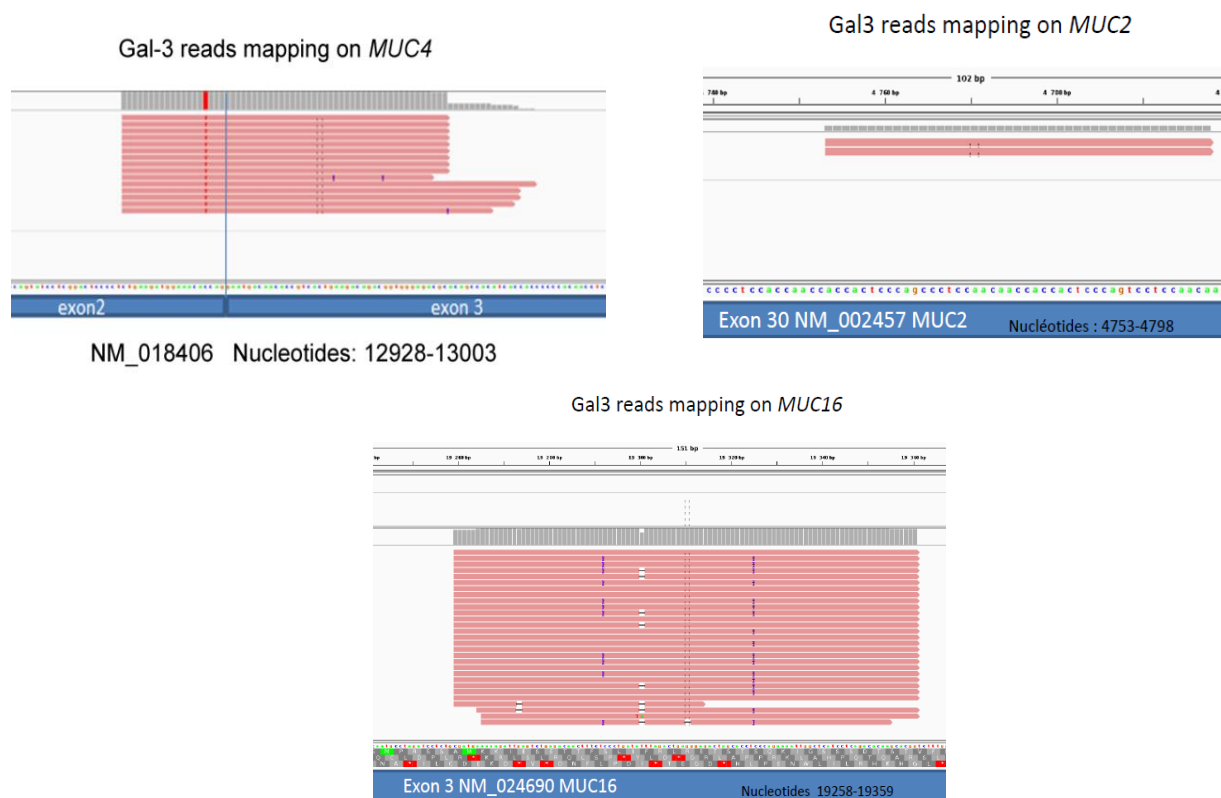


Figure 92 : Alignement des *reads* obtenus avec l'anticorps anti-galectine-3 sur les transcrits de *MUC4*, *MUC2* et *MUC16*.
Chaque *read* est représenté par un trait rouge.

Nous avons ensuite comparé les séquences nucléotidiques des 46 *reads* correspondants aux ARNm des mucines *MUC4*, *MUC16*, *MUC2* qui ont été immunoprécipités par l'anticorps anti-galectine-3, dans le but d'y rechercher un site commun putatif de liaison à la galectine-3 (figure 93). Ceci nous a permis d'identifier un motif hexamérique commun riche en nucléotides C et A et contenant un dinucléotide CA central. Ce motif est proche de celui décrit récemment par Bolisetty *et al.* comme étant impliqué dans la rétention des grands exons lors du processus d'épissage (Bolisetty and Beemon, 2012).

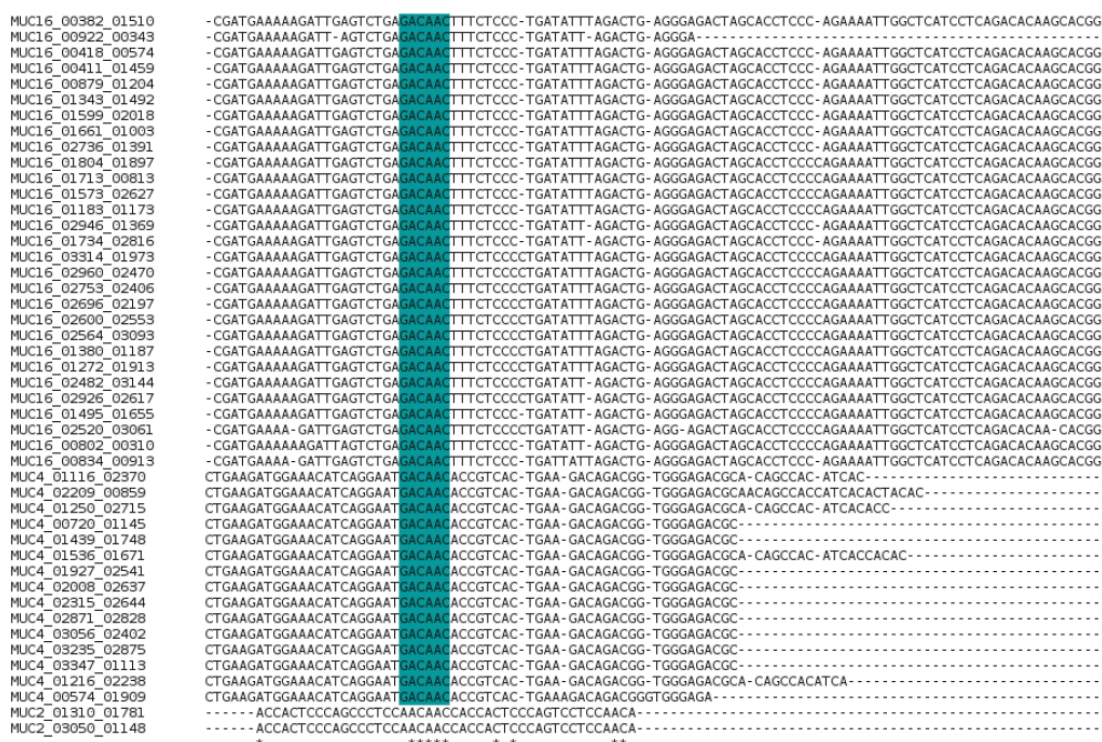


Figure 93 : Motif commun aux séquences des transcrits des mucines. *MUC4*, *MUC16*, *MUC2*.

L'ensemble des *reads* alignés sur les transcrits de mucines obtenus lors de l'expérience de RNA-Seq ont été alignés à l'aide de l'outil ClustalW. Les* représentent les nucléotides conservés dans les 46 *reads*. Le motif hexamérique partagé par 44 *reads* est surligné en vert et est indiqué sous l'alignement.

Afin d'évaluer l'influence de ce motif putatif de liaison à la galectine-3 sur l'effet stabilisateur exercé par la galectine-3 sur les transcrits de *MUC4*, nous avons cloné ce motif en amont du gène codant la luciférase dans la construction contenant le 3'UTR de *MUC4* en vecteur pGL3 promoter. Par transfection cellulaire, nous avons démontré que l'ajout du motif hexamérique ne modifie pas l'activité du 3'UTR de *MUC4* (figure 94).

Ces résultats suggèrent donc deux rôles distincts de la galectine-3 :

- une interaction avec les ARNm de mucines au niveau de ce motif ;
- une stabilisation des transcrits de *MUC4* médiée par la région *CA repeat* et hnRNP-L.

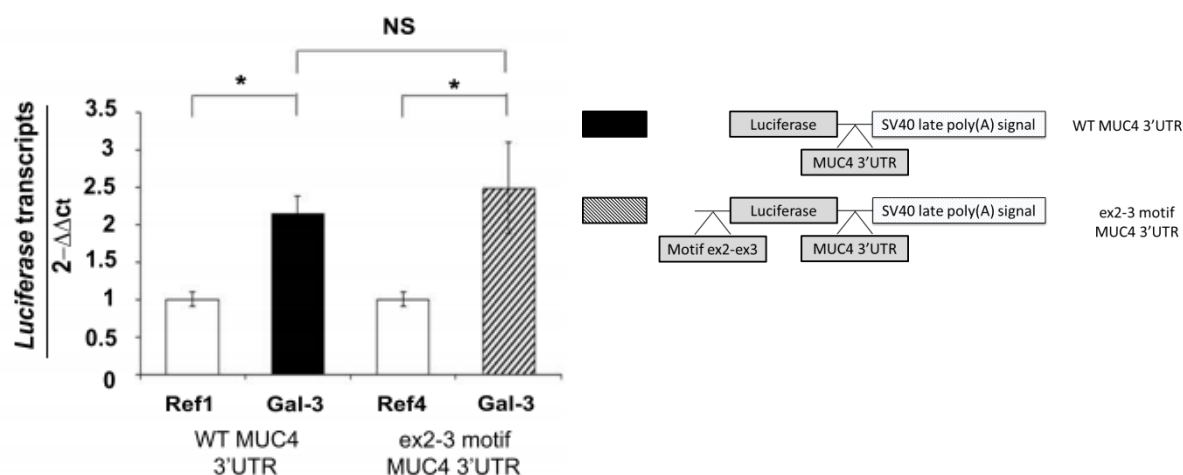


Figure 94 : Influence du motif hexamérique sur l'effet stabilisateur du 3'UTR de *MUC4*.

Les cellules Sh1 ont été co-transfectées par le vecteur d'expression codant la galectine-3 et le vecteur pGL3 promoter contenant le 3'UTR de *MUC4* seul ou comportant le motif exon 2-exon 3 cloné en amont du gène codant la luciférase. Les transcrits *luciférase* et de la *GAPDH* ont été quantifiés par RT-qPCR. Les résultats représentent la moyenne $\pm$ l'écart-type obtenus lors de deux expériences indépendantes (* : $p < 0,05$; NS : non significatif).

Après cette étude fonctionnelle, nous avons souhaité visualiser, dans les cellules Sc, le compartiment subcellulaire où se produit l'interaction entre la galectine-3 et les transcrits de *MUC4* par technique d'hybridation *in situ* couplée à l'immunofluorescence (FISH-IF). Une sonde oligod(T) a servi de contrôle positif pour détecter l'ensemble des ARN possédant une queue poly(A). La technique de FISH-IF nous a permis de visualiser par des couleurs différentes les transcrits de *MUC4* (rouge), la galectine-3 (vert) et leur colocalisation (orange). Les résultats présentés dans la figure 95 indiquent une colocalisation de la galectine-3 et des transcrits de *MUC4* au niveau périnucléaire des cellules Sc.

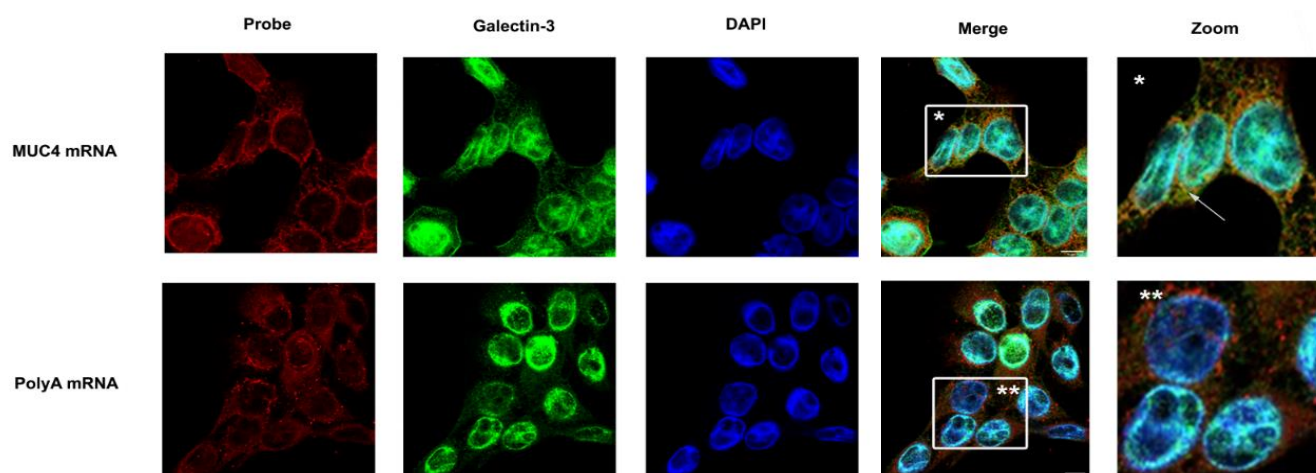


Figure 95 : Interaction entre la galectine-3 et les transcrits de *MUC4* dans le cytoplasme. Les ARNm de *MUC4* et des ARNm poly(A) ont été détectés par FISH à l'aide de sondes marquées par le cy3.5 (rouge) dans les cellules Sc. La galectine-3 a été marquée par immunofluorescence (vert) et les noyaux au DAPI (bleu). La colocalisation a été évaluée en superposant les images et est matérialisée par les spots orange. La flèche illustre un site de colocalisation

L'ensemble des résultats de RNA-IP couplé au RNA-Seq et d'imagerie indiquent donc que la galectine-3 se lie à des ARNm matures de *MUC4* dans le cytosol.

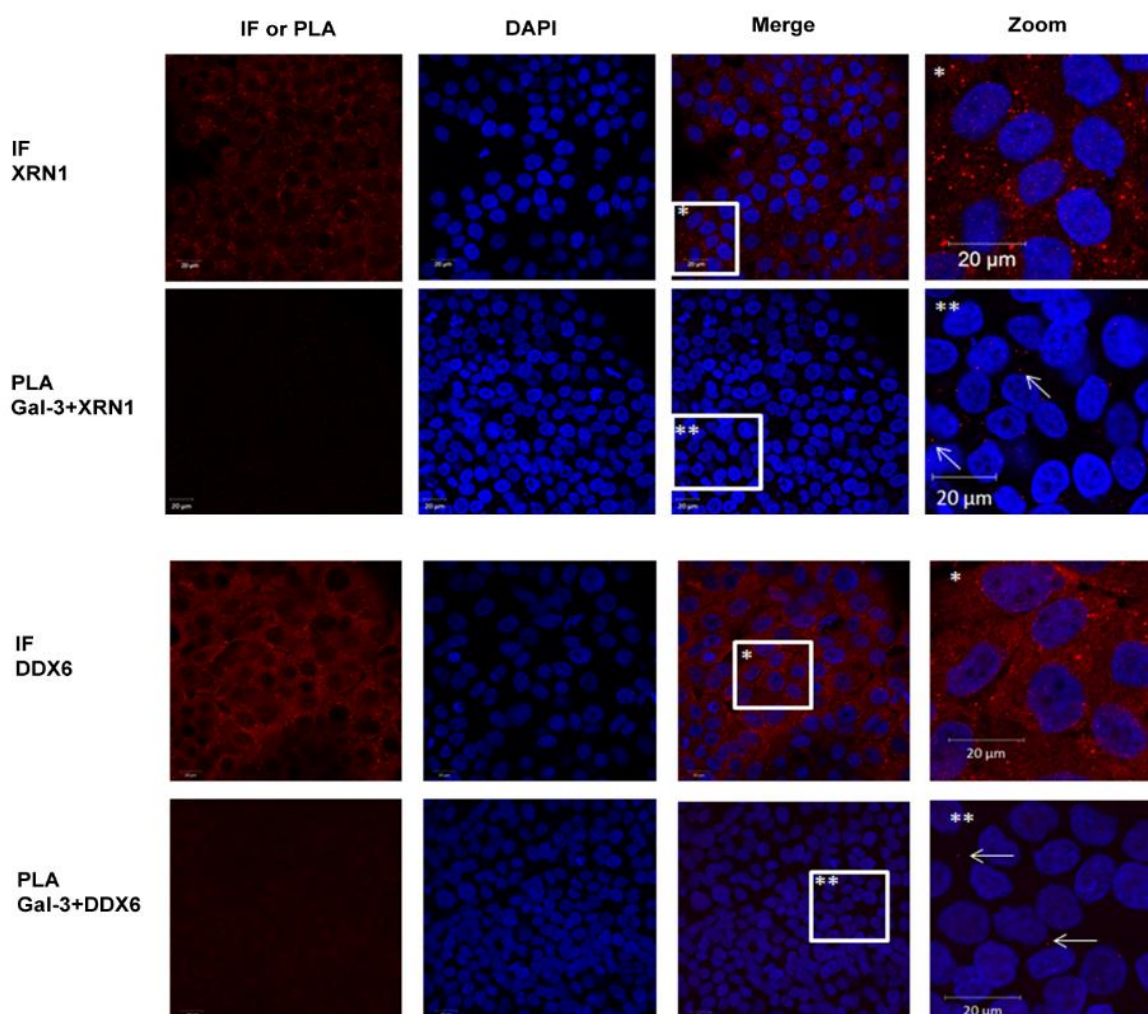
Comme nous l'avons décrit dans l'introduction, les ARNm matures sont en situation d'échanges dynamiques entre différents types de granules, au sein du cytosol.

Afin de déterminer si la galectine-3 entre dans la composition des P-Bodies, nous avons évalué dans les cellules Sc son interaction par PLA avec différents marqueurs de ces structures. Une analyse de la littérature et notamment la revue établie par Kedersha *et al.* (tableau XIV), a permis de choisir deux marqueurs : XRN1, qui est une 5'3' exonucléase, et DDX6 (*DEAD (Asp-Glu-Ala-Asp) box helicase 6*), une hélicase. Ces deux protéines sont classées comme étant des majoritairement présentes dans les P-Bodies (Kedersha and Anderson, 2007).

Tableau XIV : Liste des marqueurs de P-Bodies (Kedersha and Anderson, 2007)

Detects	Protein	Catalog number/ Commercial source	Marker for	Host species dilution	Comments
PB	DCP1a	none known	PBs only		
	4E-T	sc-13453, sc-13454 Santa Cruz Biotech	PBs only	Goat 1/200	Weak signal in U2OS cells
	GE-1/hedls	sc- 8418, Santa Cruz Biotech	PBs only	Mouse 1/1000	Robust PB marker but crossreacts with nuclear S6 kinase; use at 1/1000.
	P54/RCK	BL2142 A300-461A Bethyl Labs	PB>>SG	Rabbit 1/1000	Robust. Sees C- terminus
	Xrn1	Abcam A300-443A-1	PB>>SG	Rabbit 1/500	Use at 1/500

Nous nous sommes assurés, préalablement à la réalisation de l'expérience de PLA, de la présence des P-Bodies par immunofluorescence dans notre modèle cellulaire, les cellules Sc (figure 96).

**Figure 96** : Galectine-3 et P-Bodies.

Les PB ont été marqués dans les cellules Sc par immunofluorescence à l'aide d'un anticorps anti-XRN1 ou anti-DDX6 (rouge) puis l'interaction entre la galectine-3 et ces marqueurs a été évaluée par PLA (spot rouge). Les noyaux sont marqués au DAPI (bleu). Les flèches indiquent des exemples de spots rouges de PLA. Echelle : 20 μm.

Les r  sultats d'imagerie indiquent que de nombreuses structures granulaires marqu  es par des composants des P-Bodies sont pr  sentes dans les cellules Sc. A l'inverse, les spots de PLA sont relativement peu nombreux dans le cytoplasme (environ 1 par cellule) sugg  rant une interaction rare entre la galectine-3 et les marqueurs de P-Bodies.

De la m  me fa  on nous avons d  termin   si la galectine-3   tait pr  sente au sein des granules de stress dans notre mod  le cellulaire, en r  ponse    un stress exp  rimental. Pour cela nous avons visualis   ces structures    l'  tat basal et apr  s un choc thermique r  alis   pendant 30 min    42  C, par immunofluorescence    l'aide de diff  rents anticorps dirig  s contre des constituants des granules de stress comme TIA-1, eIF2   et G3BP1, choisis en fonction des donn  es de la litt  rature (Kedersha and Anderson, 2007, tableau XV).

Tableau XV : Analyse pr  paratoire au choix des marqueurs caract  ristiques des granules de stress : liste des marqueurs majoritairement pr  sents dans les granules de stress (Kedersha and Anderson, 2007).

Detects	Protein	Catalog number/ Commercial source	Marker for	Host species dilution	Comments
SG	p-eIF2��	KAP-P130 StressGen Biotech	SGs only	Rabbit 1/1000	Phospho-protein, causes SG assembly
	eIF3, p116	N20 sc-16377 Santa Cruz Biotech	SGs only	Goat 1/200	Strong signal, robust SG marker
	eIF4AI	sc-14211 Santa Cruz Biotech	SGs	Goat 1/200	Helicase, impairment causes SG assembly
	eIF4B	3592-Cell Signaling Technologies	SGs	Rabbit 1/100	Weak signal
	eIF4G	sc-11373 Santa Cruz Biotech		Rabbit 1/200	May not work well in mouse cells
	FXR1	sc-10544 Santa Cruz Biotech		Goat 1/100	Signal weak in U2OS; stronger when in SGs
	G3BP-1	BD-TL 611126 BD Transduction Labs		Mouse 1/200	Robust SG marker
	PABP-1	sc-32318 Santa Cruz Biotech	SGs only	Mouse 1/200	Strong signal human cells, does not detect mouse PABP.
SG>>PB	HuR	sc-5261 Santa Cruz Biotech		Mouse 1/200	Rarely detectable at PBs in cells lacking SGs
	TIA-1	goat sc-1751 Santa Cruz Biotech		Goat 1/200	Rarely detectable at PBs in cells lacking SGs
	TIAR	goat sc-1749 Santa Cruz Biotech		Goat 1/200	Rarely detectable at PBs in cells lacking SGs

La figure 97A confirme bien que des granules de stress apparaissent dans les cellules Sc cons  cutivement    un stress (choc thermique) car visualis  s par des anticorps eIF2  , G3BP1. L'anticorps TIA-1 n'a quant    lui pas donn   de marquage satisfaisant.

Nous nous sommes donc ensuite int  ress  s    l'interaction entre galectine-3 et eIF2   ou galectine-3 et G3BP1 par PLA (figure 97B). L   encore, le nombre de spots reste faible par

cellule (1 à 3, interaction rare) indiquant que la galectine-3 est peu localisée dans ces structures après un stress thermique.

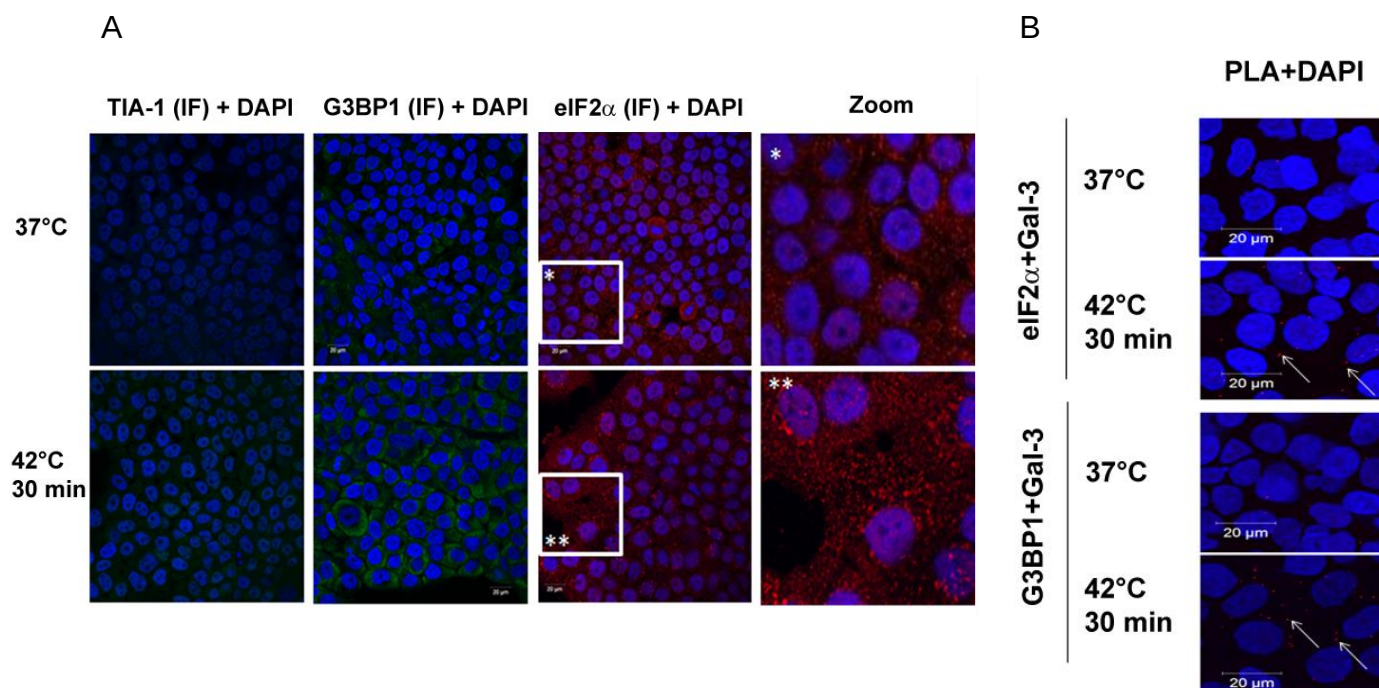


Figure 97 : Galectine-3 et granules de stress.

(A) Ces structures ont été marquées à l'état basal et après un stress thermique (42°C, 30 min) à l'aide d'anticorps anti TIA-1 (vert), eIF2α (rouge), G3BP1 (vert). Les noyaux sont marqués au DAPI (bleu).

(B) L'interaction entre la galectine-3 et eIF2α, G3BP1 a été évaluée par PLA (spot rouge) dans les cellules Sc à l'état basal et après un stress. Les flèches soulignent des exemples de spots rouges de PLA. Echelle : 20 µM.

Nous avons souhaité déterminer si les transcrits de *MUC4* sont stockés dans les granules de stress après un stress thermique et évaluer l'impact de ce stress sur l'interaction entre la galectine-3 et les transcrits de *MUC4*. Pour ce faire, nous avons réalisé une RNA-IP à l'aide d'un anticorps anti-galectine-3 et anti-TIA1 (marqueur des granules de stress) suivie d'une quantification des transcrits de *MUC4* RT-qPCR. En dehors de tout stress, nous observons un enrichissement des ARNm de *MUC4* associés à la galectine-3 par rapport à la condition correspondant à l'isotype (figure 98). Cependant, l'exposition au stress induit une diminution de la quantité des ARNm de *MUC4* associés à la galectine-3. De plus, le choc thermique ne modifie pas la quantité des ARNm de *MUC4* associés à l'anticorps anti-TIA-1. Ceci suggère donc que le stress par choc thermique ne s'accompagne pas d'un enrichissement des transcrits de *MUC4* dans les granules de stress.

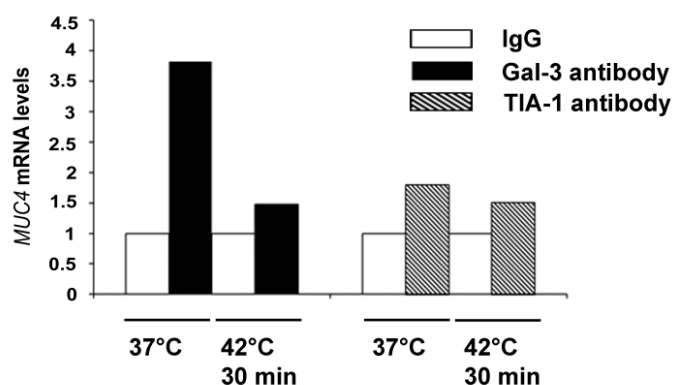


Figure 98 : Influence du choc thermique sur les transcrits de *MUC4*.

L'interaction des ARNm de *MUC4* avec la galectine-3 ou TIA-1 a été étudiée par RNA-IP couplé à de la RT-qPCR à l'état basal ou après un choc thermique de 30 min à 42 °C dans les cellules Sc. Les résultats sont exprimés par rapport à l'isotype correspondant (niveau de transcrits de *MUC4* fixé arbitrairement à 1).

Pour compléter ces résultats, nous avons visualisé l'interaction entre les transcrits de *MUC4* et la galectine-3 par hybridation *in situ* couplée à l'immunofluorescence dans les cellules Sc à l'état basal ou après un stress (figure 99). Nous retrouvons bien une interaction périnucléaire à l'état basal. Le stress s'accompagne d'une diminution de l'interaction entre les transcrits de *MUC4* et la galectine-3 (visualisée par les points de colocalisation blancs) et concorde avec la diminution des transcrits de *MUC4* observée dans l'expérience de RNA-IP (figure 98). Par ailleurs, dans ces conditions de stress thermique, l'interaction entre galectine-3 et transcrits de *MUC4* a lieu principalement dans le noyau des cellules Sc.

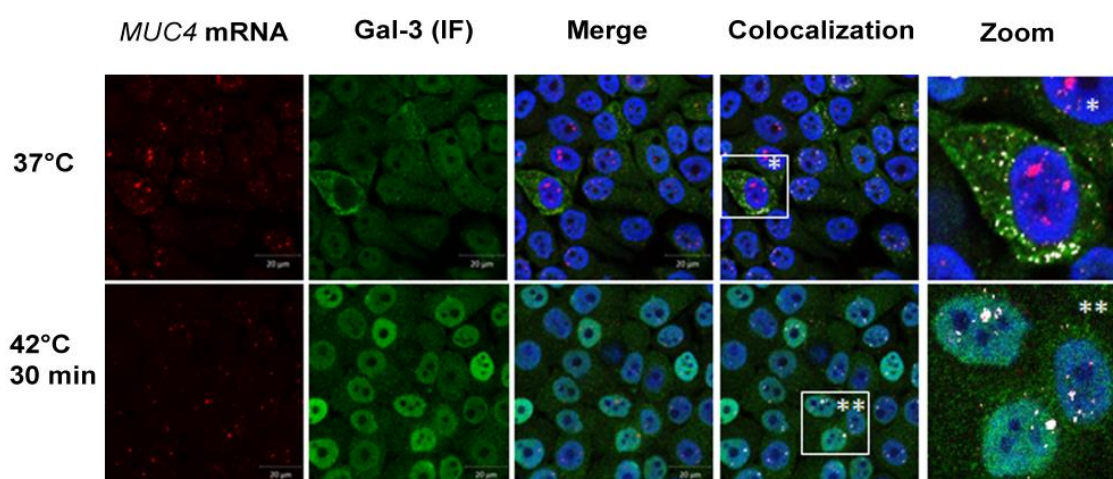


Figure 99 : Etude de la localisation des ARNm de *MUC4* et de la galectine-3 lors d'un stress.

L'interaction entre la galectine-3 (vert) et les transcrits de *MUC4* (rouge) a été étudiée par une expérience de FISH couplée à de l'immunofluorescence dans les cellules Sc à l'état basal ou après un stress (42°C pendant 30 min). Les noyaux sont marqués au DAPI (bleu). La colocalisation a été évaluée en superposant les images (points blancs) Echelle : 20 μm.

L'ensemble de nos résultats suggèrent donc que les transcrits de *MUC4* sont associés à la galectine-3 dans le cytoplasme en situation basale mais dans des structures différentes des P-Bodies et que lors d'un stress, les transcrits de *MUC4* ne sont pas stockés dans les P-Bodies ou les granules de stress et se dissocient de la galectine-3.

En conclusion, cette première partie de résultats a permis de démontrer plusieurs rôles de la galectine-3 dans le métabolisme des ARNm :

- (i) Stabilisation des ARNm de *MUC4* mais pas ceux de *MUC1*, autre mucine membranaire surexprimée dans le PDAC, par le biais de la potentialisation de la fixation de hnRNP-L sur le *CA repeat element* présent dans la région 3'UTR de *MUC4* et absent des régions 3'UTR de rat et souris ;
- (ii) Stabilisation des transcrits de *MUC4* par la galectine-3 ne s'accompagnant pas d'une augmentation de la traduction ;
- (iii) Interaction de la galectine-3 avec les ARNm matures de *MUC4* dans le cytoplasme ;
- (iv) Interaction de la galectine-3 avec les transcrits de *MUC4* dans d'autres structures cytoplasmiques que les P-Bodies et granules de stress ;
- (v) Une modification du site de l'interaction de galectine-3 et des ARNm de *MUC4* en réponse à un stress thermique : passage du cytoplasme vers le noyau et moindre interaction.

Résultats

2^{ème} partie : Galectine-3 et autres ARNm

Cette partie rassemble les résultats non encore publiés que nous avons obtenus en étudiant la capacité de la galectine-3 à stabiliser, à déstabiliser et/ou interagir avec d'autres ARNm que ceux codant les mucines et en recherchant une signification physiopathologique à ce rôle nouveau non décrit dans la littérature. Pour aborder cette problématique, nos questions et nos réponses ont été les suivantes :

I. Y a-t-il d'autres ARNm régulés de manière post-transcriptionnelle par la galectine-3 ?

Afin de déterminer si la galectine-3 est capable de stabiliser d'autres ARNm que ceux de *MUC4*, nous avons réalisé une expérience de *whole transcriptome stability* en collaboration avec la plateforme de génomique fonctionnelle (Emile Ait Yahya et les bioinformaticiens du CHRU de Lille dirigés par Martin Figeac, Caroline Thuillier et les ingénieurs du CHRU de Lille). Pour cela, nous avons analysé par RNA-seq l'ensemble des ARN des cellules Sc et Sh1 au cours d'une cinétique de traitement par actinomycine D sur une durée de 24 heures (T0h, T3h, T6h, T9h, T12h et T24h). Après extraction des transcrits, les bibliothèques d'ADNc correspondantes ont été préparées et analysées par séquençage à haut débit sur le séquenceur HiSeq 2500 Illumina de l'équipe du Pr P Froguel (CNRS UMR 8199, Génomique et épigénomique des maladies métaboliques, Lille). Le niveau d'expression d'un transcrit donné est exprimé en FPKM (*fragment per kb per Mb*) afin de s'affranchir d'un éventuel biais lié à la taille du transcrit donné. Trois transcrits dont l'expression est restée stable au cours des 24 h de traitement (RPLP0 (NM_001002 ; *Ribosomal Protein Lateral Stalk Subunit P0*), PGK1 (NM_000291 ; *Phosphoglycerate Kinase 1*) and TRAP1 (NM_016292 ; *TNF Receptor Associated Protein 1*)) ont été utilisés pour normaliser la quantité de transcrits présents à chaque temps et ce pour chacun des transcrits analysés (figure 100). Nous savions que la méthode de séquençage utilisée (Illumina, 2 x 75 pb) ne nous permettrait pas de caractériser les séquences répétitives de type VNTR des ARNm des mucines, en particulier ceux de *MUC4*. D'autres équipes avaient rencontré les mêmes problèmes en utilisant ce type de méthode de séquençage (Eisenstein, 2015; Guo *et al.*, 2014).

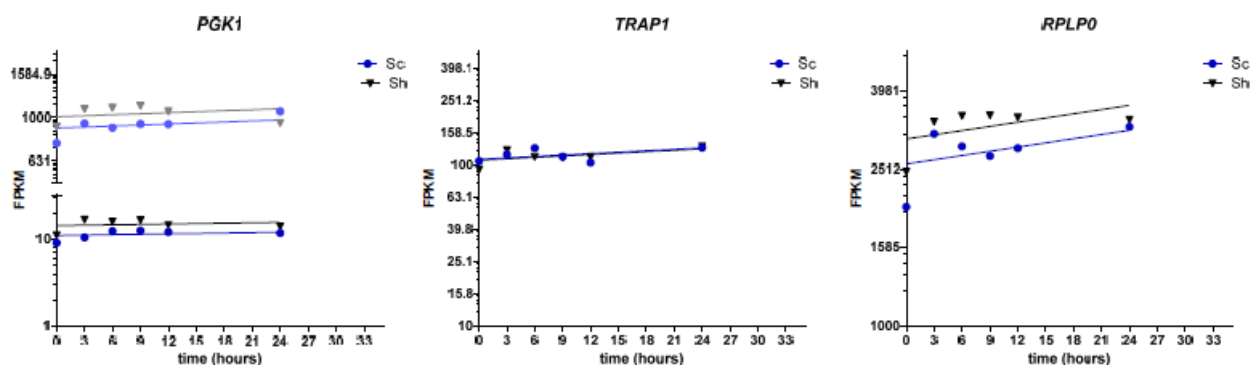


Figure 100 : Evolution de la quantité des transcrits choisis comme référence sur une durée de 24 h de traitement par Actinomycine D et évaluée par RNA-Seq.

Les données obtenues indiquent que trois familles de transcrits peuvent être distinguées dans la lignée Sh1 : celle dont la demi-vie n'est pas significativement modifiée, celle dont la demi-vie est allongée et celle dont la demi-vie est raccourcie en l'absence de la galectine-3.

L'analyse par RNA-Seq a identifié un total de **9549 ARNm exprimés** dans nos lignées cellulaires (tableau XVI). En excluant les ARNm stables et ceux à décroissance rapide (demi-vie < 2,5 h) dont la demi-vie ne pouvait pas être déterminée de façon précise dans la configuration de l'expérience, **385 ARNm** ont une demi-vie jugée comme très significativement raccourcie en absence de galectine-3 (c'est-à-dire dont la demi-vie est inférieure d'au moins un facteur 2 par rapport aux cellules Sc), suggérant **un rôle stabilisateur** exercé par cette lectine.

Le rôle stabilisateur de la galectine-3 n'est donc pas dirigé spécifiquement vers les ARNm des mucines, il s'étend donc à d'autres familles d'ARNm.

Néanmoins ce rôle ne concerne qu'environ **4 %** de l'ensemble des transcrits exprimés dans notre modèle (c'est-à-dire **385 sur 9549**).

Concernant maintenant les transcrits stabilisés en l'absence de galectine-3, c'est-à-dire ceux sur lesquels la galectine-3 pourrait exercer **un rôle déstabilisateur**, leur nombre est faible (**55 sur 9549**) soit **0,6 %** environ de l'ensemble des transcrits exprimés.

Au final, **440 transcrits sur 9549 ont une demi-vie impactée** par la présence ou l'absence de la galectine-3, et sont soit stabilisés soit déstabilisés suggérant, pour cette lectine,

un rôle non généraliste envers le métabolisme des ARNm, mais plutôt un rôle ciblé sur des fonctions cellulaires particulières qui seraient partagées par 440 cibles.

Tableau XVI : Nombre des transcrits stabilisés ou déstabilisés par la galectine-3 dans les cellules Sc (expérience de *whole transcriptome stability*) et classés selon leur demi-vie (supérieure ou inférieure à 22 h).

	Half-Life ≤ 22h	Half-Life > 22h	Whole	%
Gal-3 Stabilized RNAs	212	173 (44.9%)	385	4%
Gal-3 Destabilized RNAs	51	4 (7,3%)	55	0,58%

II. Y-a-t-il, parmi les ARNm régulés de manière post-transcriptionnelle par la galectine-3 certains ARNm-cibles qui soient aussi liés par la galectine-3 ?

Au total, **906 ARNm** sont immunoprécipités par l'anticorps anti-galectine-3 dans l'expérience de RNA-IP. Le recoupement des résultats de *whole transcriptome stability* et de RNA-IP nous a permis d'identifier **15 ARNm** à la fois stabilisés et liés par la galectine-3 (figure 101 et tableau XVII). Le faible nombre des ARNm-cibles stabilisés qui soient aussi liés suggère à nouveau une dissociation des fonctions de « liaison » et de « stabilisation », ou un rôle particulier de ces 15 cibles dans les fonctions recherchées, ou encore l'intervention dans ces fonctions d'autres protéines-partenaires de la galectine-3. Par contre, parmi les ARNm déstabilisés par la galectine-3, aucun n'est immunoprécipité, suggérant donc une fonction déstabilisatrice de la galectine-3 qui ne s'exerce pas par liaison directe avec l'ARNm-cible.

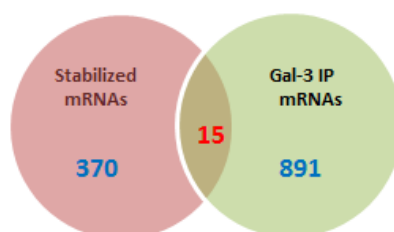


Figure 101 : Recouplement du nombre des ARNm à la fois liés (RNA-IP) et stabilisés (*whole transcriptome stability*) par la galectine-3.

Tableau XVII : Liste des 15 ARNm à la fois liés et stabilisés par la galectine-3.

Gene	Description	Half life Sc (h)	Half life Sh(h)	Functions (Uniprot-Pubmed-Genecards)
<i>CCDC64</i>	Bicaudal D related protein	8,35	3,59	Regulation of Intracellular vesicular trafficking, Rab GTPase-BP, dynactin-BP
<i>MAPK11P1L</i>	MAPK interacting and spindle stabilizing protein-like	14,42	5,08	Spindle orientation. Cell division. Maintenance of tissue architecture (enhancer for gal-3)
<i>KRT19</i>	Keratin, type I cytoskeletal 19	5,45	2,42	Structural integrity of epithelial cells, Cell signaling
<i>EZR</i>	Ezrin	48	5,3	Cell shape. Cytoskeletal anchoring at plasma membrane, endosomes
<i>ACTR3</i>	Actin-related protein 3	39	9,28	Cytoskeleton response to extracellular signals, Cell shape, mobility, adhesion, regulation of epithelial cell migration
<i>FAM83H</i>	Protein FAM83H	9,43	2,50	Cytoskeletal organization and disassembly. Biomineral tissue development
<i>PSMD3</i>	Proteasome 26S subunit	32,7	13,1	MAPK cascade, protein degradation
<i>UBE2S</i>	Ubiquitin-conjugating enzyme E2S	12,22	1,97	Degradation of nuclear substrates, promotes mitotic exit
<i>STC2</i>	Stanniocalcin 2	3,60	1,16	Ion-BP, calcium homeostasis, ER, response to oxidative stress, UPR
<i>SEMA3A</i>	Semaphorin-3A (secreted protein)	24,28	6,22	Cell adhesion, migration and survival, ERK signaling, apoptosis
<i>TONSL</i>	IKB and NFkB related protein	5,71	2,43	Genome caretaker. Replication forks, transcription corepressor
<i>KPNB1</i>	Importin subunit beta 1, Ran GTPase	48	7,45	Nuclear import of cargo, apoptotic cleavage of proteins
<i>XAB2</i>	Pre-mRNA splicing factor SYF1	12,3	5,5	Pre-mRNA splicing. Transcription regulation, TC-NER
<i>PITX1</i>	Pituitary homeobox 1	2,95	1,13	Transcription regulation. Helps to orchestrate cytoskeletal arrangement
<i>GFPT1</i>	Glutamine fructose 6 phosphate aminotransferase isomerizing 1	48	11,4	Glucose metabolism. Important role for glycolipids and glycoproteins synthesis, N- and O-glycans synthesis, UPR, transport to Golgi, exosome

Nous avons cherché à savoir si l'effet stabilisateur de la galectine-3 avait un impact sur la quantité des transcrits présents dans la cellule. Pour cela, nous avons quantifié par RT-qPCR 9 cibles parmi les 15 précédentes et avons recherché une diminution de l'expression dans les cellules Sh1 par rapport aux cellules Sc (n=2, figure 102). Les données confirment une baisse significative de l'expression pour 8 transcrits. Le résultat non significatif retrouvé pour l'ARNm *STC2* pourrait être lié à une incertitude sur la détermination de sa demi-vie puisque celle-ci est courte. Aussi, des expériences complémentaires sont nécessaires pour valider l'ensemble de ces 15 cibles.

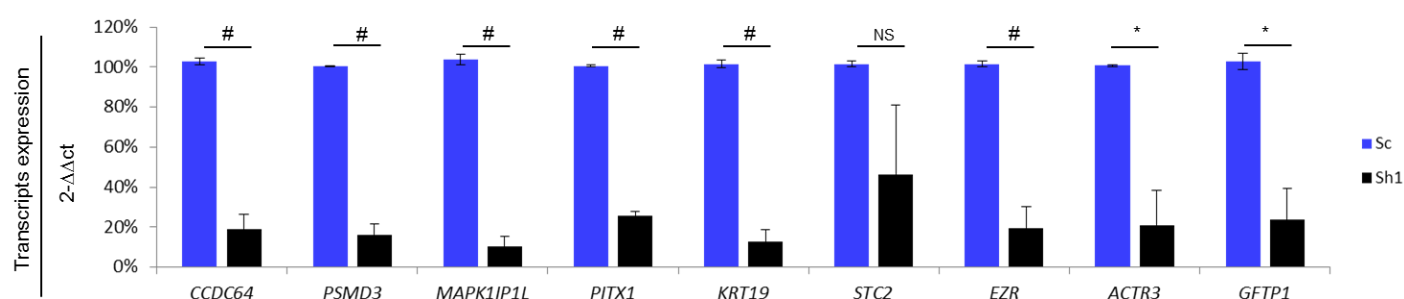


Figure 102 : Analyse de l'expression de 9 cibles-ARNm liées et stabilisées par la galectine-3. Les transcrits ont été quantifiés par RT-qPCR dans les cellules Sc et Sh1. Les résultats représentent la moyenne et l'écart type de deux expériences indépendantes (* : $p < 0,05$; # : $p < 0,01$; NS : non significatif).

III. Analyses des cibles stabilisées et déstabilisées par la galectine-3

Nous avons ensuite cherché à rassembler les fonctions connues des 440 transcrits dont la demi-vie est modifiée (ARNm stabilisés ou déstabilisés) par la galectine-3 grâce à l'outil bioinformatique *Gene Set Enrichment Analysis (GSEA, Broad Institute)* qui permet de classer en groupes distincts les transcrits partageant une même fonction biologique puis de déterminer les catégories sur-représentées dans l'échantillon à partir de 50 répertoires de gènes. Les résultats obtenus sont présentés sur la figure 103 et indiquent que 19 catégories de fonctions sont sur-représentées dans notre analyse avec un enrichissement significatif ($p < 0,05$) pour 16 d'entre elles (couleur bleue).

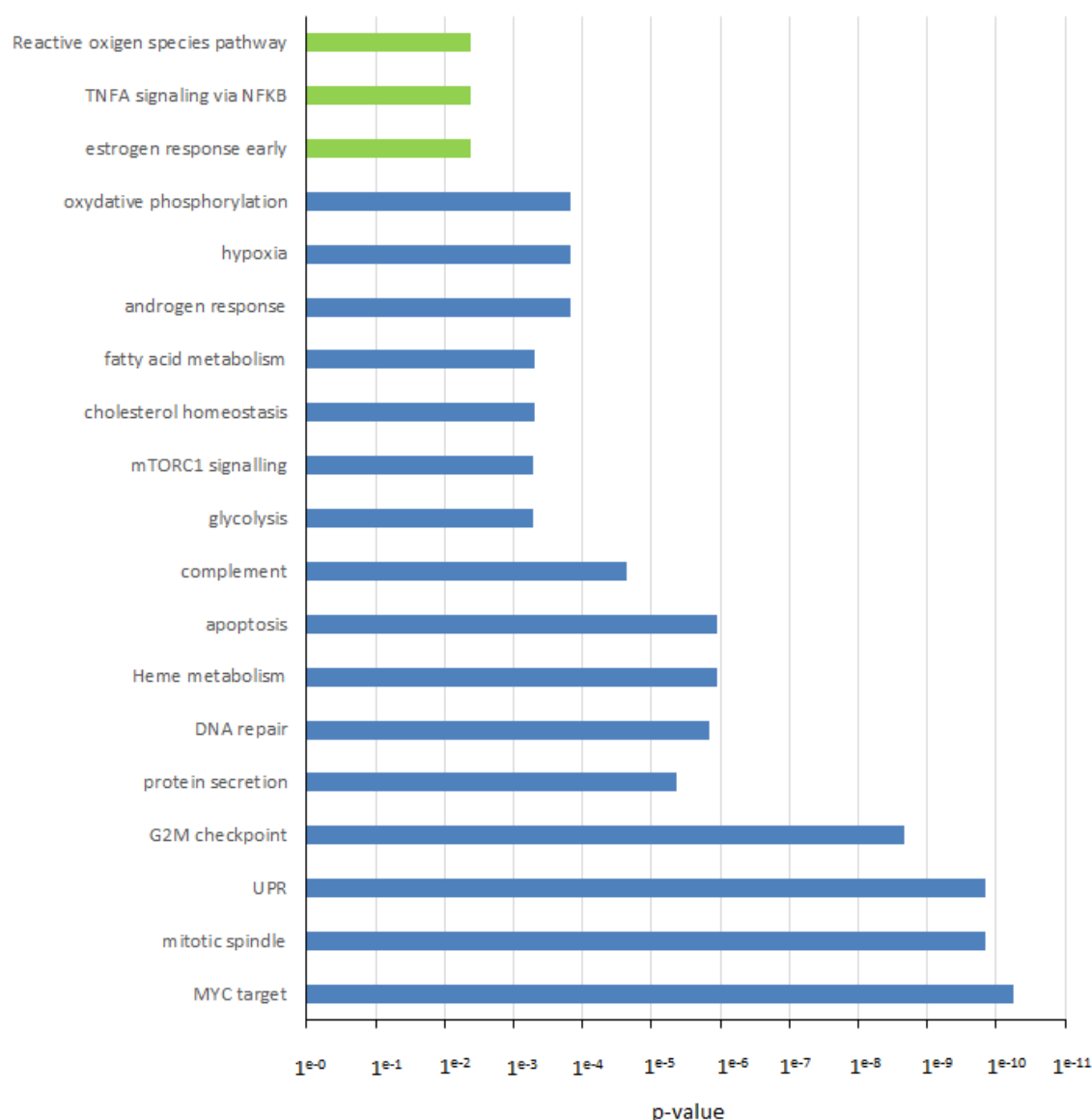


Figure 103 :Résultats de l'analyse des transcrits stabilisés et déstabilisés par la galectine-3 avec l'outil GSEA.

Les *p-values* (abscisse) sont indiquées pour chaque catégorie de fonction biologique (GO annotation, ordonnée). Les barres bleues représentent les catégories avec un enrichissement significatif ($p < 0,05$) et les vertes celles dont la *p-value* est supérieure à 0,05.

Après analyse de l'ensemble de ces fonctions biologiques, nous avons regroupé certaines d'entre elles en 7 ensembles cohérents :

- les cibles de MYC, qui regroupent 17 ARNm impliqués dans de nombreuses fonctions : le métabolisme intermédiaire, la régulation du cycle cellulaire, la transcription, la traduction, le transport des molécules entre les compartiments subcellulaires, la régulation de la voie des MAPK, l'apoptose ;

- le noyau et le cycle cellulaire (fuseau mitotique, réparation de l'ADN, cibles de MYC, contrôle G2/M, réponse aux androgènes) ;
- le stress cellulaire et UPR (voie UPR, voie des ROS) ;
- le métabolisme cellulaire intermédiaire (homéostasie du cholestérol, métabolisme des acides gras, de l'hème, phosphorylation oxydative, hypoxie, glycolyse) ;
- le métabolisme des protéines (sécrétion des protéines, synthèse protéique voie mTORC1) ;
- l'inflammation (TNF α et NF κ B, le complément);
- l'apoptose.

Nous avons ensuite complété cette étude par une analyse exhaustive et individuelle de chacune des cibles stabilisées ou déstabilisées par la galectine-3, d'abord par interrogation de bases de données (Pubmed, Genecards, Uniprot, principalement, mais également AmiGO2, PathCards, Weizmann Institute of Science database, accessibles à partir de GeneCards) puis par enfin par une étude bibliographique. Ceci a été effectué pour ne pas biaiser l'analyse par un choix de cibles *a priori* et également parce que les informations disponibles dans les bases de données sont de valeur inégale selon les gènes : annotations très anciennes parfois, annotations corrigées ou non par des spécialistes curateurs.

Nous nous sommes intéressés tout d'abord aux 385 transcrits stabilisés par la galectine-3. Nous avons pu les regrouper en **12 grandes classes de fonctions**, un transcrit pouvant être classé dans plusieurs catégories, en nous appuyant sur les annotations GO (*Gene Ontology*) concernant les fonctions moléculaires et les processus biologiques (Genecards, Uniprot) (tableau XVIII) :

- Fonctions, maintenance et biogenèse des organelles, transport vésiculaire et trafic intracellulaire ;
- Transcription, régulation de la traduction ;
- Remodelage du cytosquelette, régulation du cytosquelette par les Rho GTPases ;
- Membrane apicale, jonctions cellulaires ;
- Métabolisme ;
- ARNm : *processing*, transport ;
- Cycle cellulaire, réparation de l'ADN, remodelage de la chromatine ;
- UPR (*Unfolded Protein Response*), réponse au stress (thermique, oxydatif, calcium, radiations ionisantes, virus osmotique, hypoxie, drogues) ;

- Conformation des protéines et catabolisme, protéasome/ubiquitine, autophagie ;
- Apoptose ;
- Fonctions intracellulaires du calcium et magnésium (signalisation, Réticulum endoplasmique, cytosquelette) ;
- Voies de signalisation de NFκB, du TNFα, inflammation.

Tableau XVIII : Fonctions des transcrits stabilisés par la galectine-3.

Fonctions	% de transcrits
Métabolisme	8,32
Remodelage du cytosquelette, régulation du cytosquelette d'actine par les Rho GTPases	9,46
Fonctions, maintenance et biogenèse des organelles, transport vésiculaire et trafic intracellulaire	20,75
Conformation des protéines, catabolisme protéasome/ubiquitine, autophagie	6,27
ARNm : <i>processing</i> , transport	7,30
UPR, réponse au stress (thermique, oxydatif, calcium, radiations ionisantes, virus osmotique, hypoxie, drogues...)	6,61
Apoptose	6,16
Cycle cellulaire, réparation de l'ADN, remodelage de la chromatine	7,30
Transcription, régulation de la traduction	13,91
Voies de signalisation de NFκB, du TNFα, inflammation	2,51
Fonctions intracellulaires du calcium et magnésium (signalisation, RE, cytosquelette)	2,74
Membrane apicale, jonctions cellulaires	8,67

Ensuite, nous avons réalisé l'analyse de la liste des 55 transcrits déstabilisés par la galectine-3, sur la base des 12 classes préalablement définies, ceci nous a permis de retrouver des fonctions relatives aux mêmes domaines physiopathologiques que ceux précédemment définis pour les transcrits stabilisés par cette lectine, certaines cibles pouvant s'inscrire dans plusieurs fonctions (tableau XIX):

- UPR, réponse au stress (thermique, oxydatif, calcium, radiations ionisantes, virus osmotique, hypoxie, drogues) : **17 cibles**
- Transcription, régulation de la traduction : **11 cibles**
- Fonctions, maintenance et biogenèse des organelles, transport vésiculaire et trafic intracellulaire : **10 cibles**
- Apoptose : **8 cibles**
- Cycle cellulaire, réparation de l'ADN, remodelage de la chromatine : **6 cibles**
- Métabolisme : **5 cibles**

- Remodelage du cytosquelette, régulation du cytosquelette d'actine par les Rho GTPases : **4 cibles**
- Conformation des protéines et catabolisme, protéasome/ubiquitine, autophagie : **4 cibles**
- ARNm : *processing*, transport : **3 cibles**
- Voies de signalisation de NF κ B, du TNF α , inflammation : **3 cibles**
- Fonctions intracellulaires du calcium et magnésium (signalisation, RE, cytosquelette) : **3 cibles**
- Membrane apicale, jonctions cellulaires : **3 cibles**

Tableau XIX : Fonctions des transcrits déstabilisés par la galectine-3.

Fonctions	% de transcrits
Métabolisme	6,49
Remodelage du cytosquelette, régulation du cytosquelette d'actine par les Rho GTPases	5,19
Fonctions, maintenance et biogenèse des organelles, transport vésiculaire et trafic intracellulaire	12,99
Conformation des protéines, catabolisme protéasome/ubiquitine, autophagie	5,19
ARNm : <i>processing</i> , transport	3,90
UPR, réponse au stress (thermique, oxydatif, calcium, radiations ionisantes, virus osmotique, hypoxie, drogues...)	22,08
Apoptose	10,39
Cycle cellulaire, réparation de l'ADN, remodelage de la chromatine	7,79
Transcription, régulation de la traduction	14,29
Voies de signalisation de NF κ B, du TNF α , inflammation	3,90
Fonctions intracellulaires du calcium et magnésium (signalisation, RE, cytosquelette)	3,90
Membrane apicale, jonctions cellulaires	3,90

Le tableau XX présente la proportion globale de transcrits stabilisés et déstabilisés par la galectine-3 pour chaque catégorie de fonction biologique déterminée.

Tableau XX : Fonctions des transcrits stabilisés et déstabilisés par la galectine-3

Fonctions	% de transcrits
Métabolisme	8,18
Remodelage du cytosquelette, régulation du cytosquelette d'actine par les Rho GTPases	9,12
Fonctions, maintenance et biogenèse des organelles, transport vésiculaire et trafic intracellulaire	20,13
Conformation des protéines, catabolisme protéasome/ubiquitine, autophagie	6,18
ARNm : <i>processing</i> , transport	7,02
UPR, réponse au stress (thermique, oxydatif, calcium, radiations ionisantes, virus osmotique, hypoxie, drogues...)	7,86
Apoptose	6,50
Cycle cellulaire, réparation de l'ADN, remodelage de la chromatine	7,34
Transcription, régulation de la traduction	13,94
Voies de signalisation de NFkB, du TNF α , inflammation	2,62
Fonctions intracellulaires du calcium et magnésium (signalisation, RE, cytosquelette)	2,83
Membrane apicale, jonctions cellulaires	8,28

Ainsi, parmi les rôles des cibles stabilisées et déstabilisées, un mot-clé nouveau par rapport à ce que l'on savait de la galectine-3, et constant dans les différentes approches ressort de l'analyse bibliographique et des annotations GO (*Gene Ontology*) : il s'agit du **stress cellulaire et de la voie UPR (*Unfolded protein response*)**.

IV. Analyse individuelle des 15 transcrits à la fois stabilisés et liés par la galectine-3

Nous avons complété notre analyse en étudiant plus précisément les fonctions des transcrits liés et stabilisés par la galectine-3 afin de déterminer s'ils s'inscrivaient également dans ces registres. Nous avons donc réalisé une analyse bibliographique de ces 15 cibles en nous basant sur les données de la littérature (Genecards, Uniprot, Pubmed, principalement) :

CCDC64 (*BICDR1- Bicaudal D Related 1*) code une protéine, à la fois cytoplasmique et nucléaire, localisée au niveau du cytosquelette. Elle exerce, dans le cytoplasme, une fonction régulatrice du trafic vésiculaire à partir de l'appareil de Golgi et dans le noyau, elle est localisée autour du centrosome et joue un rôle dans le cycle cellulaire. Elle interagit avec des protéines responsables du transport intracellulaire d'organelles et de vésicules (kinésine, dynactine, dynéine). Elle se lie aux Rab GTPases et contrôle la formation, le mouvement des vésicules intracellulaires et leur translocation au travers de la membrane plasmique (Schlager *et al.*, 2010; Valente *et al.*, 2010).

MAPK1IP1L (*MISS protein* ou *Mitogen-Activated Protein Kinase 1 Interacting Protein 1 Like*) code une protéine présente à la fois dans le cytoplasme et le noyau, en fonction du cycle cellulaire. Elle contrôle, par la phosphorylation, l'organisation des microtubules et est impliquée dans l'orientation et le maintien de l'intégrité du fuseau lors de la division cellulaire. Il est intéressant de noter que ce gène est contigu à celui de *LGALS3* (chromosome 14) avec lequel il partage des éléments régulateurs suggérant donc également des fonctions partagées avec la galectine-3.

KRT19 (*Keratin 19*) code la kératine 19, une cytokératine de type 1, protéine constitutive des filaments intermédiaires du cytosquelette responsables de l'intégrité structurale des cellules épithéliales. Elle est présente dans tous les compartiments subcellulaires, et participe au remodelage du cytosquelette qui accompagne la différenciation cellulaire, la prolifération, la migration et le processus métastatique. Elle intervient également dans la signalisation cellulaire notamment par le biais de son interaction avec le complexe β caténine-RAC1, ce qui augmente la translocation nucléaire de la β caténine (Saha *et al.*, 2017). Par ailleurs l'ARNm codant la β caténine (*CTNNB1*) est lui aussi stabilisé, mais non lié par la galectine-3.

EZR (*Ezrin*) code l'ezrine ou villine 2, un membre du groupe des protéines ERM (Ezrine, Radixine et Moésine) qui possèdent un domaine commun FERM (*4.1 protein ezrin radixin moesin domain*) grâce auquel elles interagissent avec de nombreux constituants protéiques ou lipidiques des membranes. Elles lient aussi le cytosquelette d'actine et constituent donc des intermédiaires physiques entre la membrane et le cytosquelette. Ceci leur permet de jouer un rôle fondamental dans les changements morphologiques qui accompagnent les cellules au cours du développement ou de la réparation des lésions tissulaires (transition épithélio-mésenchymateuse) ou encore du processus métastatique (Bosanquet *et al.*, 2014). Ces protéines sont un substrat des tyrosines kinases et jouent le rôle de lien moléculaire entre la membrane plasmique et le cytosquelette d'actine pour la transduction des signaux provenant du pôle apical. L'ezrine est nécessaire à la formation des microvillosités des cellules épithéliales, elle joue un rôle important dans l'adhérence cellule-cellule. Elle interagit avec différents partenaires comme l'actine, CD44 (dont l'ARNm est lié également par la galectine-3), ICAM2, la β 4-intégrine, les protéoglycanes ... (Jiang *et al.*, 2014).

ACTR3 (*Actin-related protein 3*) code une protéine « *actin-like* » à très large répartition subcellulaire, appartenant au complexe ARP2/3 (*Actin-related protein 2/3*) qui est essentiel au maintien de la forme cellulaire et à la motilité cellulaire par assemblage des lamellipodes d'actine. En effet, ce complexe contrôle la polymérisation, l'organisation et le recyclage du réseau de filament d'actine (Goley and Welch, 2006) au cours du remodelage du cytosquelette qui accompagne des processus cellulaires complexes comme le cycle cellulaire, la polarisation, la migration... Cette protéine joue également un rôle dans la régulation de la ciliogenèse (Kim *et al.*, 2010).

FAM83H (*Family With Sequence Similarity 83 Member H*) code une protéine à la fois nucléaire et cytosolique, d'expression tissulaire ubiquitaire, qui possède un domaine de liaison aux filaments de kératine. Elle régule l'assemblage et le désassemblage des filaments de kératine, et donc le remodelage du cytosquelette par le recrutement de la caséine kinase CK-1 α sur les filaments de kératine (kinase dont l'ARNm est lui aussi stabilisé, mais non lié, par la galectine-3). Elle régule la formation des desmosomes ainsi que la migration des cellules épithéliales (Kuga *et al.*, 2016). Une surexpression de FAM83H est associée au cancer colorectal et une perte de fonction à une maladie génétique touchant le développement dentaire (amelogenesis imperfecta) (Kuga *et al.*, 2013).

PSMD3 (*Proteasome 26S Subunit, Non-ATPase 3*) code une protéine à la fois cytosolique et nucléaire, la sous-unité régulatrice S3 du complexe 26S du protéasome. Le complexe 26S du protéasome se forme par assemblage de deux complexes, le complexe 20S qui constitue le cœur de la machinerie de dégradation des protéines ubiquitinylées et le complexe 19S qui a un rôle régulateur et au sein duquel on trouve S3. Cette protéine intervient donc dans la régulation de mécanismes très importants de la vie cellulaire, comme le cycle cellulaire ou l'apoptose et impliquant la cascade de signalisation MAPK (Iio *et al.*, 2015).

UBE2S (*Ubiquitin Conjugating Enzyme E2 S*) code l'enzyme E2S qui appartient au groupe des enzymes de conjugaison E2 impliqués dans l'ubiquitination des protéines et qui donc participe aux étapes initiales de leur marquage en vue de leur dégradation par le protéasome. UBE2S est également un composant du complexe APC/C (*Anaphase Promoting Complex / Cyclosome*), un facteur essentiel du contrôle de la sortie du cycle cellulaire (Garnett *et al.*, 2009). UBE2S contrôle la dégradation de pVHL (*Von Hippel-Lindau*) et par conséquent l'accumulation de HIF1 α (Jung *et al.*, 2006).

STC2 (*Stanniocalcin 2*) code une glycoprotéine sécrétée et exprimée de façon ubiquitaire, la stanniocalcine 2. Elle est abondante dans le réticulum endoplasmique (et l'appareil de Golgi), se lie aux ions métalliques et participe à la régulation de l'homéostasie du calcium et des phosphates au niveau tissulaire (absorption, transport au niveau du rein et de l'intestin grêle) et au niveau cellulaire et subcellulaire, selon des mécanismes peu connus. Elle est surexprimée en réponse à l'hypercalcémie, l'hypoxie ou le stress oxydant et elle est considérée jouer un rôle important dans la réponse cellulaire face à un stress, au même titre que l'hème-oxygénase 1 (HO-1), une protéine de stress avec laquelle elle forme un complexe (Jiang *et al.*, 2012). Elle aurait un rôle cytoprotecteur, grâce à sa capacité à lier de nombreuses molécules, et à ce titre, elle intervient dans la voie UPR (*Unfolded protein response*), son expression étant induite par le facteur ATF4 (*Activating Transcription Factor 4*) via la branche PERK (*PRKR-Like Endoplasmic Reticulum Kinase*). Elle favorise la survie cellulaire et la résistance des cellules au stress, peut-être en régulant la voie de signalisation MAPK (Ito *et al.*, 2004; Joshi *et al.*, 2015; Kim *et al.*, 2015).

SEMA3A (*Semaphorin 3A*) code une protéine de la large famille des sémaphorines (protéines à domaine SEMA) connues initialement pour leur rôle dans le développement du système nerveux. Les sémaphorines de classe 3 modulent des interactions avec toutes sortes de protéines (rôle d'agent chemoattractant ou chemorépulsif) et régulent le cycle cellulaire, l'adhérence cellulaire, la motilité, la survie cellulaire, la progression tumorale. La protéine SEMA3A est sécrétée mais aussi présente au niveau de la membrane plasmique et du cytosquelette (cytosol, noyau) et intervient dans l'apoptose (neurones, macrophages). Cette protéine stimule la signalisation intracellulaire en induisant notamment le recrutement de la sérine/thréonine kinase CDK5 (*Cyclin dependant kinase 5*) dont la galectine-3 stabilise également (sans la lier) une protéine régulatrice CDK5RAP2. Elle exerce des fonctions anti-tumorales dans certaines cellules cancéreuses et pro-tumorales dans d'autres (comme le pancréas par exemple) (Nasarre *et al.*, 2014).

TONSL (*Tonsoku-Like*) code une protéine qui a été appelée IκBR (*Inhibitor Of Kappa B-Related Protein*) ou NFKBIL2 (*NFκB Inhibitor-Like 2*) en raison d'une similarité de structure avec des ligands d'I-κB (*inhibitor of kappa B*), mais ses fonctions sont encore mal connues (Ray *et al.*, 1995). Elle se partage entre le noyau et le cytoplasme. Dans le noyau, elle forme un complexe avec la protéine MMS22L (*Methyl methanesulfonate-sensitivity protein 22-like*) qui est recrutée sur les régions de recombinaison homologue et participe à leur résolution, contribuant ainsi au maintien de l'intégrité du réplisome. Dans le cytosol, son rôle

supposé est celui de régulateur négatif de la transcription médiée par NF- κ B, empêchant ainsi l'entrée du complexe NF- κ B, dans le noyau mais ce rôle n'a pas été démontré (Hayden and Ghosh, 2012).

KPNB1 (*Karyopherin Subunit Beta 1*) code l'importine β 1, une protéine participant au transport de protéines cytoplasmiques jusque dans le noyau *via* les pores nucléaires et grâce à l'énergie des Ran GTPases. Elle est particulièrement sollicitée durant le cycle cellulaire (réplication de l'ADN et mitose). Elle transporte les protéines pourvues de signaux d'import nucléaire (NLS) grâce à sa liaison avec une importine α (hétérodimerisation), ou, de façon autonome, elle permet l'import de certaines protéines des ribosomes ou certaines histones. L'importine β 1 est le transporteur-clé du complexe p50/p65 de la voie canonique NF- κ B et module la prolifération et l'apoptose (Yan *et al.*, 2015). A noter que l'ARNm de *RELA* (codant p65) est également stabilisé mais non lié par la galectine-3. KPNB1 transporte également des ions Zinc dans le noyau où ils sont indispensables à la transcription.

XAB2 (*XPA Binding Protein 2*) code une protéine orthologue du « SYF1 Pre-mRNA-Splicing Factor » de la levure. Il s'agit d'une *protein-binding protein* impliquée dans différents processus nucléaires et cytoplasmiques : la réparation de l'ADN couplée à la transcription (TC-NER : *Transcription-Coupled Nucleotide Excision Repair*), la transcription, l'épissage des pré-ARNm et l'export des ARN. De plus, elle intervient dans la progression du cycle cellulaire (Kuraoka *et al.*, 2008). XAB2 joue aussi un rôle essentiel dans le cadre de l'épissage alternatif des pré-ARNm lors des étapes initiales de l'embryogenèse (Guilgur *et al.*, 2014).

PITX1 (*Pituitary homeobox 1*) code une protéine nucléaire qui est un facteur de transcription jouant un rôle important dans la morphogenèse de nombreux organes. On connaît peu les mécanismes de son action. Parmi les 1440 cibles potentielles de ce facteur de transcription (*The Harmonisome Database, TRANSFAC predicted transcription factor targets dataset*), on trouve de nombreux gènes impliqués dans l'ubiquitination des protéines et le protéasome, le fonctionnement du cytosquelette, le transport nucléaire, la réponse aux stress, l'apoptose, le métabolisme des glucides et glycoconjugués, le métabolisme lipidique. Un certain nombre de ses gènes-cibles figurent dans nos listes de transcrits impactés par la galectine-3. La perte d'expression de PITX1 est observée dans de nombreux types de cancers, permettant de proposer que pour ce facteur de transcription, au cours de la cancérogenèse, des fonctions de gène suppresseur de tumeur. Les tumeurs qui maintiennent les niveaux

d'expression les plus élevés seraient celles qui seraient les plus résistantes à la chimiothérapie suggérant donc un rôle dans la résistance au stress (Takenobu *et al.*, 2016).

GFPT1 code une enzyme du métabolisme glucidique, la Glutamine-Fructose-6-Phosphate Transaminase 1, qui catalyse la formation de glucosamine-6-phosphate à partir du fructose-6-phosphate. Elle appartient à la voie des hexosamines, voie située au carrefour des voies cataboliques (métabolisme énergétique) et anaboliques du glucose. Cette enzyme est l'enzyme limitante de la voie des hexosamines. Elle régule la disponibilité des précurseurs pour la synthèse des N- et O-glycoprotéines au niveau du réticulum endoplasmique. Elle est un acteur important de l'équilibre de l'usage intracellulaire du glucose, partagé entre besoins énergétiques (divisions cellulaires) et synthèses des glycoconjugués, équilibre rompu au cours de la cancérogenèse. De cet équilibre dépendent les O-glycosylations de facteurs nucléaires qui régulent le cycle cellulaire (FOXO1 : *Forkhead box protein O1*) ou les voies de survie en réponse à de nombreux stress. De plus, elle permet le couplage de la voie UPR au métabolisme glucidique ; le facteur « UPR » XBP1 (*X-Box Binding Protein 1*) stimule sa transcription afin de rétablir l'homéostasie protéique (Hazari *et al.*, 2016; Vincenz and Hartl, 2014; Wang *et al.*, 2014).

Nous avons précédemment identifié 12 groupes de fonctions rassemblant les transcrits stabilisés et déstabilisés par la galectine-3. L'analyse bibliographique des 15 cibles liées et stabilisées par cette lectine a montré la participation de nombreux transcrits à plusieurs fonctions, ce qui nous a permis de rassembler certaines d'entre elles, comme par exemple UPR et apoptose. Ainsi, nous proposons de regrouper ces 15 transcrits au sein de 8 grands processus biologiques, certains ARNm étant parfois rattachés à plusieurs processus comme c'est le cas par exemple pour *UBE2S* :

- développement/morphogenèse : *EZR*, *SEMA3A*, *FAM83H*, *KRT19*, *PSMD3*, *KPNB1*, *XAB2*, *PITX1* ;
- cycle cellulaire/mitose : *CCDC64*, *MAPK1IP1L*, *PSMD3*, *UBE2S*, *SEMA3A*, *TONSL*, *KPNB1*, *XAB2* ;
- architecture cellulaire et cytosquelette : *CCDC64*, *KRT19*, *EZR*, *ACTR3*, *SEMA3A*, *FAM83H*, *MAPK1IP1L* ;
- voie UPR, dégradation des protéines, apoptose : *PSMD3*, *UBE2S*, *STC2*, *SEMA3A*, *GFPT1*, *KPNB1* ;

- signalisation cellulaire : *EZR*, *PSMD3*, *UBE2S*, *SEMA3A*, *KPNB1* ;
- métabolisme (ions, glycosylation) : *STC2*, *KPNB1*, *GFPT1* ;
- acides nucléiques (ARN et ADN) : *XAB2*, *PITX1* ;
- trafic intracellulaire : *CCDC64*, *KPNB1*.

En conclusion, ces grandes catégories sont en accord avec celles identifiées lors de l'étude des transcrits stabilisés et déstabilisés par la galectine-3.

Parallèlement, nous avons colligé les rôles déjà identifiés de la galectine-3 dans la littérature envers, soit des acides nucléiques (transcription) soit des protéines (apoptose, inflammation, trafic cellulaire, signalisation...) La figure 104 illustre ces registres connus. Nous pouvons constater qu'une partie des fonctions biologiques de nos 15 cibles s'intègrent également dans ce qui est déjà connu des fonctions de la galectine-3. Ceci souligne la cohérence des résultats issus du transcriptome et nous permet d'envisager d'en faire une interprétation valide pour de nouvelles fonctions de la galectine-3. Une parties des fonctions déjà connues de la galectine-3 s'exercent donc au niveau post-transcriptionnel sur des ARNm codant des protéines en lien avec les fonctions connues de la galectine-3. De plus deux de ces ARNm-cibles apparaissaient, par l'analyse bibliographique directement liés à la voie UPR : *STC2* et *GFPT1*.

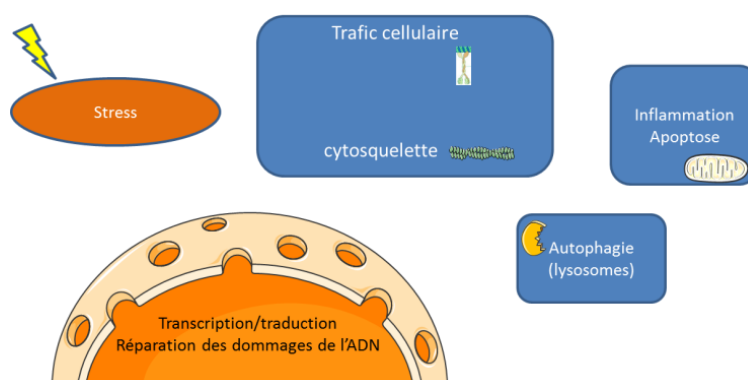


Figure104 : Illustration de fonctions connues pour la galectine-3 retrouvées dans l'analyse bibliographique des 15 transcrits stabilisés et liés par cette lectine.

V. La voie UPR et les voies qu'elle active secondairement peuvent-elles rendre compte de toutes les catégories fonctionnelles préalablement caractérisées ?

Lors de l'étude des fonctions des transcrits stabilisés et déstabilisés par la galectine-3, l'identification de cibles dont les fonctions pouvaient être rattachées directement ou indirectement à la voie UPR telles que *STC2* et *GFPT1* nous ont orientés vers l'hypothèse de l'implication de cette lectine dans ce processus biologique. En effet, l'activation de la voie UPR suit une logique préventive destinée à garantir la survie cellulaire. Ainsi elle est associée à des modifications du métabolisme cellulaire (lipidique, protéique et énergétique principalement), à une modulation de la transcription de nombreux gènes, et à une adaptation à l'accroissement du catabolisme protéique (nécessitant de mettre en place la machinerie de dégradation des protéines mal conformées). De plus, son activation conduit à un arrêt du cycle cellulaire en phase G1 (Brewer and Diehl, 2000), à des mécanismes de transcription couplés à la réparation de l'ADN et enfin à l'apoptose cellulaire si l'homéostasie cellulaire n'est pas rétablie.

Avant de poursuivre nos investigations concernant cette voie UPR, il convient de rappeler quelques aspects à son sujet afin de pouvoir proposer d'y intégrer l'ensemble des processus biologiques que nous avons identifiés en étudiant les ARNm stabilisés et déstabilisés par la galectine-3 :

Le réticulum endoplasmique (RE) joue un rôle important dans l'homéostasie cellulaire générale : rôle dans l'homéostasie calcique, dans la synthèse, la mise en conformation et la maturation des protéines sécrétées et transmembranaires (un tiers des protéines cellulaires concernées) et des glycolipides (c'est-à-dire les biomolécules requérant des étapes de N- et/ou O-glycosylation) ainsi que dans la synthèse des phospholipides. Il communique avec le compartiment golgien via un réseau vésiculaire impliquant le cytosquelette. Un système de contrôle de qualité permet aux protéines bien conformées de sortir du RE pour poursuivre leur maturation. Les protéines anormales (estimées à 30 % des protéines néosynthétisées) sont quant à elles dégradées par le biais de la voie ERAD (*ER associated protein degradation*) faisant intervenir le protéasome, ou par autophagie, lorsque la voie ERAD ne parvient pas à désengorger le RE. Certaines protéines échappent au contrôle qualité du RE et sont quand même adressées vers l'appareil de Golgi. Au niveau golgien, il existe également un système

de contrôle de qualité des protéines et celles qui sont détectées comme étant mal conformées font l'objet d'un transport rétrograde vers le RE puis le cytosol, pour être dégradées *via* la voie ERAD.

Toutes les situations mettant en péril l'intégrité du RE telles que la privation en nutriments, l'hypoxie, la dérégulation de la synthèse protéique, du métabolisme lipidique ou de l'homéostasie calcique, peuvent être à l'origine d'un stress du RE (Kato and Nishitoh, 2015). Un déséquilibre entre la charge de protéines qui entrent dans le RE pour y être conformées et les capacités de prise en charge de ces protéines par le RE engendre un stress du RE qui sera détecté par 3 protéines transmembranaires. Celles-ci vont stimuler une réponse cellulaire appelée UPR dont le but est de restaurer l'homéostasie du RE. La chronologie de la réponse cellulaire à l'activation de la voie UPR est la suivante : mise en place de la machinerie de catabolisme des protéines en excès, cette réponse initiale est la voie ERAD ; puis activation de 3 grands mécanismes (Janssens *et al.*, 2014) qui sollicitent le noyau :

- Diminution de la traduction des protéines pour limiter la charge du RE : mécanisme d'atténuation de la traduction ;
- Augmentation de la biosynthèse (transcription-traduction) des protéines chaperonnes afin d'augmenter les capacités de conformation du RE ;
- Augmentation de la biosynthèse des composants de la voie ERAD.

En cas d'échec du retour à l'homéostasie cellulaire malgré la mise en place de ces processus à visée corrective, deux autres processus cellulaires peuvent intervenir :

- Activation de l'autophagie qui fait intervenir les lysosomes ;
- Apoptose qui fait intervenir également les mitochondries.

Les 3 senseurs du stress du RE (figure 105), points de départ de la réponse UPR, sont les protéines membranaires PERK (*PRKR-Like Endoplasmic Reticulum Kinase*), ATF6 (*activating transcription factor6*) et IRE1 (*inositol-requiring enzyme 1*). Elles sont maintenues à l'état inactif, en dehors de toute situation de stress, par leur interaction avec la protéine chaperonne Bip (GRP78) au niveau de la lumière du RE. En cas de stress du RE, Bip se détache de ces 3 senseurs pour contribuer à l'acquisition de la conformation des protéines. La dissociation de ce facteur est également observée lors de la baisse du contenu en calcium du RE (Kato and Nishitoh, 2015).

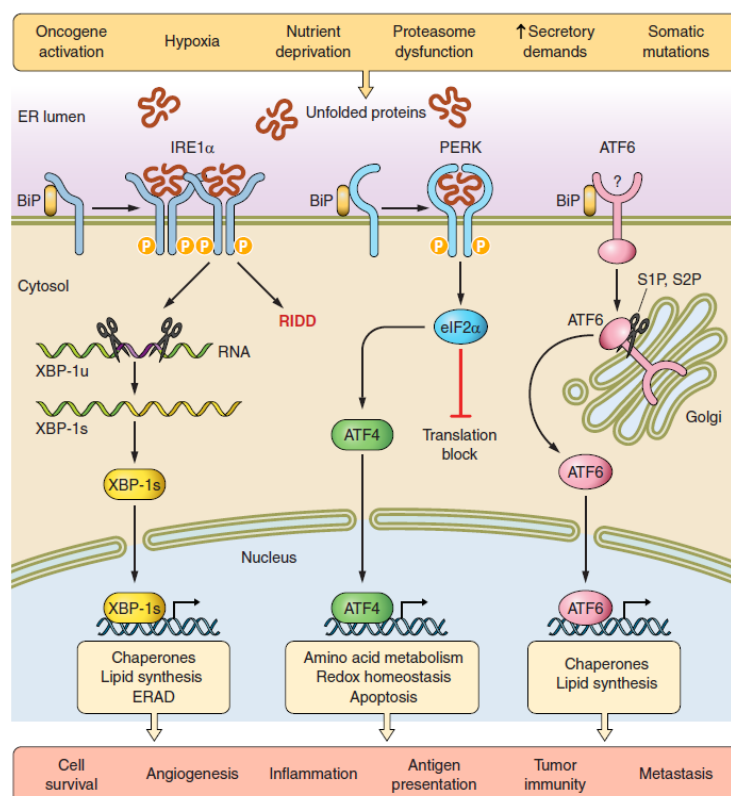


Figure 105 : Acteurs de la voie UPR et les 3 branches qui lui sont associées *via* IRE1, PERK et ATF6 (Oakes, 2017).

Une sollicitation exagérée de la production de nouvelles protéines amène la cellule à voir s'accumuler des protéines mal conformées au sein du RE. Ceci active la voie UPR *via* 3 protéines transmembranaires IRE1, PERK et ATF6 dont la signalisation conduit à une réponse permettant à la cellule de faire face à cette situation : augmentation notamment de la synthèse de protéines chaperones, des lipides, activation de la voie ERAD et éventuellement induction de l'apoptose si le stress perdure. Dans les cellules cancéreuses, une activation de cette voie est souvent observée favorisant la survie des cellules tumorales, l'angiogénèse, le processus métastatique...

La voie UPR, dès qu'elle est enclenchée, se déroule selon deux phases successives : une phase d'adaptation réversible pendant laquelle le métabolisme cellulaire est modifié pour faire face à la situation de stress du RE, puis une phase irréversible associée à l'activation de l'apoptose si la cellule n'arrive pas à restaurer son homéostasie (Hetz, 2012). Ainsi, selon que le stress soit temporaire ou chronique et que la réponse soit correctrice ou non, la voie UPR est en lien avec d'autres voies: inflammation, autophagie, contrôle et maintenance du génome, apoptose...

Après avoir rappelé l'ensemble des acteurs et processus rattachés à la voie UPR, nous pouvons dire que 7,9% de nos cibles (stabilisées et déstabilisées) sont en lien direct avec la voie UPR, 6,2% avec le système de dégradation des protéines, 6,5 % sont en lien avec l'apoptose et 20% sont en rapport avec le trafic cellulaire et les organelles (RE, appareil de Golgi, lysosomes, mitochondries, exosomes). D'autres cibles sont rattachées à ce processus comme le cytosquelette (9,1 %), et le métabolisme calcique (2,8 %). Des cibles participent au processus pour partie, il s'agit de celles associées à la transcription, en particulier TCNER (*Transcription-coupled nucleotide excision repair*), et à la traduction (13,9 %). Enfin, des cibles ne sont pas directement concernées par cette activité (jonctions cellulaires, membrane apicale...).

L'analyse à l'aide de l'outil GSEA a identifié les **17 cibles de MYC** comme étant le principal processus biologique commun aux transcrits stabilisés et déstabilisés par la galectine-3. Ce résultat concorde avec le fait que MYC agit sur l'apoptose, la prolifération et la différenciation par le contrôle l'expression de nombreux gènes cellulaires (environ 15% d'entre eux). En effet, MYC est un facteur de transcription qui se dimérise avec la protéine MAX pour réguler l'expression de gènes impliqués dans la croissance cellulaire. Son activité positive sur la croissance cellulaire se situe dans deux domaines principaux (Stine *et al.*, 2015).

1/ la stimulation du métabolisme, en particulier du métabolisme intermédiaire, avec la stimulation de la production énergétique (catabolisme du glucose et des lipides), et la stimulation de la biosynthèse des macromolécules (via la stimulation des voies de biosynthèse des précurseurs des acides nucléiques, des protéines et des lipides) ;

2/ la stimulation de la biogenèse des organelles, en particulier des ribosomes (stimulation de la traduction et stimulation de la transcription des gènes impliqués dans la machinerie traductionnelle) et des mitochondries (production d'ATP). Le rôle de MYC dans la biogenèse des autres organelles est très mal connu.

MYC participe également, dans le cadre de ces fonctions positives sur la croissance, à la régulation du cycle cellulaire (coordination entre synthèse d'ADN, transcription et traduction des gènes-clés).

Par ailleurs, la protéine MYC agit avec des protéines partenaires dont certaines sont présentes dans les listes des transcrits stabilisés par la galectine-3 : *MTDH* (métadhérine) qui

est une *poly(A)-BP* et une *dsDNA-BP* qui active la transcription de *MYC* et de *NFKB* ; *MXI1* (*MAX interactor 1*), qui est lui un répresseur transcriptionnel de *MYC* et encore *CCNY* qui active la transcription des gènes médiée par *MYC*, par exemple.

La seconde fonction identifiée par l'analyse bioinformatique GSEA est le **fuseau mitotique (15 cibles)**. La plupart des cibles codent des protéines du cytosquelette impliquées dans la régulation de la mitose (*INCEMP*, composant du *Chromosomal Passenger Complex* et régulateur-clé de la mitose, par exemple). Ceci pourrait être en lien avec le fait que la détection de protéines mal conformées dans le noyau s'accompagne d'un arrêt du cycle cellulaire en phase G1 (Brewer *et al.*, 1999).

L'analyse bibliographique des fonctions des cibles stabilisées par la galectine-3 indique donc que cette lectine participerait à l'homéostasie cellulaire globale, c'est-à-dire l'homéostasie des organelles (RE, golgi, lysosomes, mitochondries, exosomes), et celles des protéines qui rendent possibles le trafic entre les différents compartiments subcellulaires via l'équilibre entre leur synthèse et dégradation. Elle aurait un rôle dynamique de fluidification des déplacements qui font intervenir le cytosquelette dans le cadre du trafic du noyau vers la membrane plasmique, vers les organites, des organites vers la membrane plasmique, sachant que ces voies de transport peuvent être à double sens, en particulier le transport antérograde du RE vers le Golgi ou le transport rétrograde du Golgi vers le RE, pour recycler les protéines résidentes du RE ou en cas de stress du Golgi. La galectine-3 jouerait donc un rôle de pivot dans la lutte pour préserver l'homéostasie cellulaire et dont la voie centrale est UPR mais dont le champ d'action s'élargirait aux gènes-clés pour la coordination de l'expression des nombreux acteurs qui composent ces voies associées (figure 106).

Notre analyse permet donc d'unifier tous les rôles connus de la galectine-3 relatés dans la littérature (trafic, inflammation, apoptose, transcription, signalisation cellulaire...). UPR semblerait donc être le chaînon manquant qui relierait la galectine-3 à toutes ses fonctions. La galectine-3 a donc un rôle précoce dans la réponse au stress, elle n'intervient donc pas que dans la phase tardive associée à l'inflammation et à l'apoptose mais pourrait agir en amont en stimulant l'expression de protéines pouvant favoriser le retour à l'homéostasie du RE par le biais de la réponse UPR.

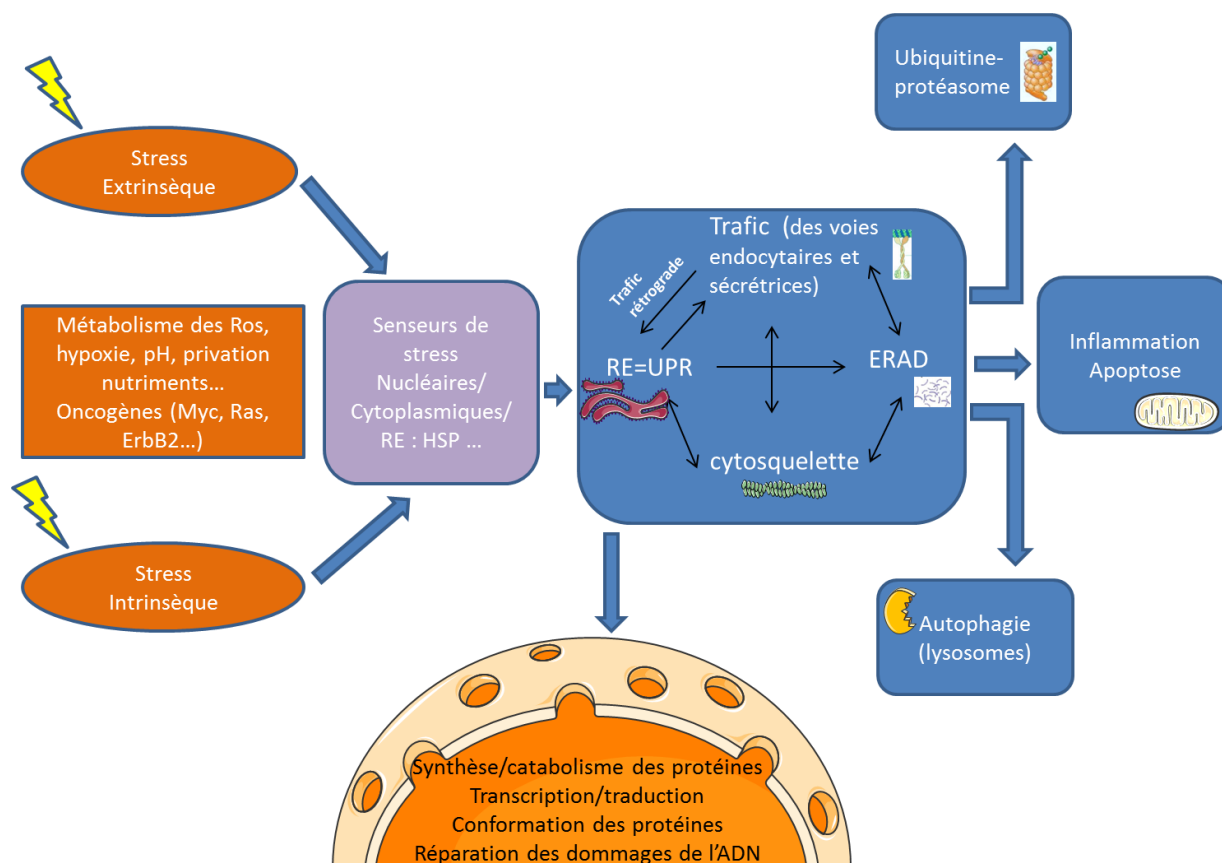


Figure 106 : Processus physiopathologique dans lequel s'intègrent les fonctions des transcrits stabilisés par la galectine-3.

VI. Etude détaillée des transcrits-cibles rattachables directement ou indirectement à la voie UPR

VI.1. Transcrits directement rattachés à la voie UPR

Les transcrits désignés comme « directement rattachés à la voie UPR » sont ceux figurant dans des répertoires de gènes étiquetés UPR, répertoires accessibles par interrogation de bases de données. Il est à noter que certaines cibles sont absentes de certains répertoires mais présentes dans d'autres et que donc l'étude bibliographique est indispensable.

VI.1.1. Voie UPR et transcrits stabilisés par la galectine-3

a. Analyse orientée à l'aide de l'outil GSEA

Afin de déterminer les cibles de la voie UPR parmi les transcrits de la liste des ARNm stabilisés/déstabilisés par la galectine-3, nous avons étudié ceux identifiés à l'aide de l'outil *Gene Set Enrichment Analysis (GSEA, Broad Institute)*. Ce logiciel dont la base de données « UPR » contient 113 gènes définis comme « *genes up-regulated during unfolded protein response* » a rattaché 12 cibles, toutes stabilisées par la galectine-3, à la voie UPR. Parmi elles, *STC2*, cible préalablement identifiée comme étant à la fois liée et stabilisée par la galectine3. Les 11 autres cibles sont :

- ***XPOT*** (*Exportin For TRNA*) codant l'exportine T, une protéine Ran GTPase impliquée dans l'export nucléaire des ARN de transfert (Gupta *et al.*, 2016) ;
- ***EIF4A1*** (*Eukaryotic Translation Initiation Factor 4A1*) codant un facteur d'initiation de la traduction nécessaire à la fixation des ARNm sur les ribosomes. De plus, ce facteur contrôle l'expression du facteur CHOP (*C/EBP homologous protein*) appartenant à la branche PERK de la voie UPR (Fredriksson *et al.*, 2014) ;
- ***IARS*** codant une enzyme, l'*Isoleucyl-ARNt Synthetase*. Cette enzyme est un gène cible de la voie PERK-ATF4 (Cameron *et al.*, 2015) ;
- ***DKC1*** (*Dyskerin Pseudouridine Synthase 1*) codant une protéine ayant un rôle dans la stabilisation et la maintenance des télomérases, l'adhérence cellulaire, la réponse aux dommages de l'ADN. De plus, elle limite le stress du RE (Kropski *et al.*, 2013) ;
- ***MTHFD2*** codant un enzyme bi-fonctionnel mitochondrial (*Methylenetetrahydrofolate Dehydrogenase (NADP+ Dependent) 2, Methenyltetrahydro-folate*). Il s'agit d'un gène cible d'ATF4 (Huggins *et al.*, 2015) ;
- ***PAIP1*** (*Poly(A) Binding Protein Interacting Protein 1*) codant une protéine se liant aux *poly(A) binding proteins* et à eIF4A ainsi qu'eIF3. Elle est impliquée dans l'initiation de la traduction. Via son interaction avec un facteur d'initiation de la traduction, elle est impliquée dans la réponse au stress (Sharma *et al.*, 2016) ;

- **ATF6**, codant une protéine transmembranaire du RE qui, après clivage en réponse à un stress du RE, libère l'un des trois facteurs de transcription-clés activant les gènes-cibles de la voie UPR (branche impliquant l'appareil de Golgi) (figure 105) ;
- **EXOSC1 (Exosome Component 1)**, codant une sous-unité à rôle non catalytique du complexe ARN exosome impliqué dans le *processing* et la dégradation de nombreux ARN cellulaires : ARN de transfert et ARN ribosomiques impliqués dans la traduction et ARN messagers, en particulier ceux possédant des motifs ARE dans leur 3'UTR (Liu *et al.*, 2006) ;
- **PARN (Poly(A)-Specific Ribonuclease)**, codant une 3'exoribonucléase qui agit préférentiellement sur la queue poly(A) des ARN messagers, première étape de la dégradation des ARNm. Cette ribonucléase intervient aussi sur les ARNm intrinsèquement peu stables car comportant des régions ARE dans leur 3' UTR via l'interaction avec des *RNA binding proteins* (Wu *et al.*, 2005) ;
- **ARFGAP1**, codant l'*Arf GTPase activating protein 1* qui est une protéine à fonction de chaperonne et de régulateur de la petite GTPase ARF1, impliquée dans le fonctionnement du cytosquelette et du trafic vésiculaire entre le RE et l'appareil de Golgi et dans l'autophagie (Boon *et al.*, 2014; Drin *et al.*, 2009). Ce gène est rattaché aux 2 branches UPR : IRE1 α et XBP1 ;
- **SHC1 (SHC-transforming protein 1)**, codant une protéine impliquée dans le couplage du complexe GRB2/SOS à la signalisation RAS-MAPK qui régule la durée de vie et les fonctions des espèces réactives de l'oxygène (Miyazawa and Tsuji, 2014).

b. Analyse orientée par les bases de données

Nous nous sommes ensuite intéressés aux cibles qui, parmi les 385 transcrits étaient stabilisés par la galectine-3, et pouvaient avoir un lien avec la voie UPR. Nous les avons donc recherchées dans un certain nombre de répertoires de gènes « UPR » en particulier :

- celui du Weizmann Institute of Science (Version 4.2.1.13 Updated 2016 September 27) qui comporte 88 gènes “UPR” ;
- celui du MD Anderson Cancer Center de l'Université du Texas (date de la dernière version consultée non précisée) qui comporte 91 gènes “UPR”.

7 ARNm stabilisés par la galectine-3 sont codés par des gènes pouvant être rattachés à la voie «UPR»:

- *ATF6*, *ARFGAP1*, *EXOSC1*, *PARN*, et *SHC1* sont présents dans les 3 répertoires (GSEA, Weizmann Institute et MC Anderson) ;
- *STC2* et *GFPT1* (stabilisés et liés) sont présents dans 2 répertoires sur 3 (*STC2* : GSEA et MD Anderson ; *GFPT1* : Weizmann Institute et MC Anderson). Ces deux gènes sont par ailleurs rattachés à « UPR » font l'objet de publications récentes dans lesquelles ils sont considérés comme des biomarqueurs « UPR » d'intérêt important en pathologie (Yang *et al.*, 2016; Yeung *et al.*, 2012).

c. Analyse orientée par la bibliographique

Pour étudier de façon exhaustive les cibles stabilisées par la galectine-3 en lien avec la voie UPR, et peut-être ainsi en découvrir d'autres, nous avons réalisé une étude bibliographique exhaustive et consulté *a posteriori* des répertoires plus larges pour authentifier nos cibles.

Il s'agissait du répertoire « UPR related genes » (Weizmann Institute of Science, 158 gènes) qui a permis d'assigner au répertoire UPR le gène :

- ***DERL1*** (*Derlin 1*) codant une protéine de liaison à la E3 Ubiquitine ligase (voie XBP1) et qui participe à la sortie du RE des protéines mal conformées en vue de leur dégradation via la voie ERAD (You *et al.*, 2017)

et du répertoire AmiGO2 version 2.4.26, amigo2b (185 entrées) qui nous a permis d'assigner au répertoire « UPR » les cibles :

- ***TBL2*** (*Transducin Beta Like 2*) codant la transducine interagissant avec PERK, modulant la signalisation et la survie cellulaire lors du stress du réticulum endoplasmique (Tsukumo *et al.*, 2014) ;
- ***TMEM33*** (*Transmembrane Protein 33*) codant une protéine dont l'expression est induite par le stress du RE et qui agit comme un modulateur de la réponse UPR (voie IRE1 et voie PERK) (Sakabe *et al.*, 2015).

Par ailleurs, la bibliographie et le répertoire « ERAD related genes » (Weizmann Institute of Science, 502 gènes) nous ont permis d'élargir les cibles de la galectine-3 aux

gènes « ERAD » après étude bibliographique (*Pubmed*, *Genecards*). Ainsi, nous en avons mis en évidence les cibles ERAD suivantes :

- **DERL1** (*Derlin 1*). Ce gène est retrouvé à la fois dans le répertoire UPR et dans le répertoire ERAD.
- **JKAMP** (*JNK1/MAPK8-Associated Membrane Protein*) codant un régulateur de MAPK8 et qui intervient dans la voie ERAD dépendante de l'ubiquitine (Tcherpakov *et al.*, 2009) ;
- **SEL1L** (*SEL1L ERAD E3 Ligase Adaptor Subunit*) codant une protéine impliquée dans la rétrotranslocation/dislocation des protéines mal conformées (voie ERAD) (Jeong *et al.*, 2016) ;
- **UBE2K** (*Ubiquitin Conjugating Enzyme E2 K*) codant l'enzyme de conjugaison E2K, régulant la dégradation des protéines à demi-vie courte et des protéines mal conformées *via* la voie ERAD, elle est aussi impliquée dans la régulation du complexe NFκB (Middleton and Day, 2015) ;
- **GET4** (*Golgi To ER Traffic Protein 4*) codant une protéine à fonction de *chaperone-binding protein*, impliquée dans la rétrotranslocation des protéines mal conformées, du Golgi vers le RE, en vue de leur dégradation par le protéasome ou la voie ERAD (Wang *et al.*, 2011) ;
- **USP14** (*Ubiquitin Specific Peptidase 14*) codant une enzyme cytoplasmique de déubiquitination de certaines protéines, elle est aussi l'inhibiteur physiologique de la voie ERAD dans des conditions cellulaires sans stress. Son expression est souvent dérégulée dans les cancers, elle aurait des propriétés anti-apoptotiques (Mines *et al.*, 2009; Zhu *et al.*, 2016) ;
- **TP53** codant un gène suppresseur de tumeur qui est un facteur de transcription. p53 a pour cible des gènes impliqués dans de nombreux processus cellulaires dont la plupart sont partagés par les transcrits cibles de la galectine-3 tels que l'arrêt du cycle cellulaire, la réparation de l'ADN, l'apoptose, le métabolisme, l'autophagie, la traduction des ARNm et les mécanismes de rétrocontrôle (Fischer, 2017).

VI.1.2. Voie UPR et transcrits déstabilisés par la galectine-3

Parmi les transcrits déstabilisés par la galectine-3, aucun n'est présent dans le répertoire « GSEA ». Cependant, deux cibles ont retenu notre attention même si elles ne

figuraient pas dans les répertoires consultés. En effet leur proximité de fonction avec la réponse au stress est apparue dans la bibliographie.

- Il s'agit pour la première de **CARHSP1** (*Calcium Regulated Heat Stable Protein 1*), codant une protéine de choc thermique, dépendante du calcium. Cette protéine est une *RNA-binding protein*, régulateur post-transcriptionnel, qui se fixe sur le 3'UTR d'ARNm-cibles et les stabilise, vraisemblablement dans une voie de stabilisation impliquant HuR. Elle est localisée sur le plan subcellulaire dans les P-bodies mais aussi dans les granules exosomes. Elle a pour cible l'ARNm du TNF α , médiateur central de l'inflammation (Pfeiffer *et al.*, 2011). Dans le contexte "UPR", il apparaît cohérent que l'ARNm de **CARHSP1** soit déstabilisé dans notre modèle.
- La deuxième cible est **EIF2AK1** (*Eukaryotic Translation Initiation Factor 2 Alpha Kinase 1*). Cet ARNm code une kinase qui joue un rôle de répresseur de l'initiation de la traduction au cours de différents types de stress, y compris le stress du RE. La déstabilisation de cet ARNm apparaît cohérente avec l'hypothèse d'un rôle de la galectine-3 qui serait de donner à la cellule la possibilité de restaurer le fonctionnement du RE sans entraîner la cellule vers d'autres étapes modifiant de façon importante le destin cellulaire et incluant l'apoptose mais aussi le changement de répertoire des protéines exprimées.

VI.1. Transcrits ayant des fonctions dans le retour à l'homéostasie du réticulum endoplasmique mais non répertoriés comme appartenant à la voie UPR

Il s'agit de transcrits-cibles stabilisés ou déstabilisés par la galectine-3 dont les fonctions, après analyse bibliographique, nous permettaient de penser qu'ils pourraient éventuellement jouer un rôle dans la réponse UPR, même s'ils n'étaient pas répertoriés en tant que tels dans les bases de données.

VI.1.1. Cibles à fonctions de chaperonnes stabilisées par la galectine-3

Différents transcrits stabilisés par la galectine-3 peuvent être rattachés à la fonction de rétablissement de la conformation des protéines (protéines ayant un **rôle de chaperonne**).

Celle-ci est mise en place en réponse à l'activation de la voie UPR lors de la phase réversible d'adaptation au stress. Parmi celles-ci se trouvent :

- **PSMG3** (*Proteasome Assembly Chaperone 3*) et **POMP** (*Proteasome Maturation Protein*) codant chacun une protéine chaperonne nécessaire à la maturation du complexe 20S du protéasome ;
- **PPIL2** et **PPIL3** (*Peptidylprolyl Isomerase Like 2-3*) codant des membres de la famille des cyclophilines, famille de protéines ayant un rôle important dans l'acquisition de la conformation des protéines. PPIL2 peut également agir comme une E3 ubiquitine ligase et PPIL3 intervient dans l'épissage des pré-ARNm ;
- **DNAJA3**, **DNAJB6** et **DNAJC7** (*DnaJ Heat Shock Protein Family (Hsp40) Member A3, B6 et C7*) codant des protéines de la famille des protéines DNAJ qui est une des deux classes majeures de protéines chaperonnes impliquées, en collaboration avec les protéines HSP70 ou HSP90, dans la conformation et l'assemblage des protéines oligomériques. DNAJB6 agit également comme une co-chaperonne de la protéine HSP70 et atténue l'activation de la caspase-3. DNAJC7 est une co-chaperonne des protéines chaperonnes HSP70 et HSP90 ;
- **HSPA1A** et **HSPA1B** (*Heat Shock Protein Family A (Hsp70) Member 1A et 1B*) codant un membre de la famille des protéines chaperonnes HSP70. Elles agissent notamment en limitant l'agrégation des protéines, dans le cytosol et le noyau ;
- **MKKS** (*Mac Kusik-Kaufman syndrome*), chaperonne liant les protéines mal conformées et qui joue également un rôle dans le transport de protéines dans les organelles.

VI.1.2. Cibles stabilisées par la galectine-3 et impliquées dans les voies secondairement activées à l'issue de la réponse UPR

L'analyse des listes d'ARN messagers stabilisés par la galectine-3 nous a permis d'identifier des cibles en rapport avec des fonctions déjà connues de la galectine-3 et pouvant être connectées à la voie UPR, en particulier **l'apoptose, l'inflammation et la voie NFκB** dont voici quelques exemples :

- **API5**, **DUS2**, **CHP1**, **CIAPIN1**, **DUS2**, **TMBIM4**, **TRAF2** codant des régulateurs négatifs de l'apoptose ;

- **C1D, CASP7, CSNK1D, CSNK2A1, FAM188A** codant des inducteurs/régulateurs positifs de l'apoptose ;
- **HSH2D** codant un modulateur de la réponse apoptotique ;
- **CXCL1, IL1RAP**, codant une protéine induisant la synthèse de facteurs pro-inflammatoires lors d'un stress ;
- **RELA1**, codant un composant de NFκB ;
- **CARD16, ASCC1, BCL2L13, BNIP3, MTDH, TRIM8**, codant des régulateurs positifs de NFκB ;
- **TNFRSF12A** codant un récepteur du TNF ;
- **ARRDC2, BIRC3, PSMD10, UBE2L3** codant des régulateurs négatifs de la voie NFκB.

Par ailleurs, de nombreuses cibles associées au trafic intracellulaire sont présentes dans la liste des transcrits stabilisés par la galectine-3. Un intérêt tout particulier peut être porté sur un certain nombre d'entre elles qui concerne le trafic rétrograde de l'appareil de Golgi vers le RE, qui en situation de stress, permet de réorienter les protéines mal conformées vers le RE puis vers le cytosol en vue de leur dégradation : **ARFGAP1** (déjà citée en qualité d'appartenance à UPR, voie IRE1α et protéine importante du transport et trafic vésiculaire de la voie ARF1), **COG7** (*Component of oligomeric golgi complex 7*), **SCFD1**, qui régule COG4 (*Component of oligomeric golgi complex*), **GET4** (*chaperone-binding protein*, appartenance à la voie ERAD), **KIFAP3** (protéine associée à la kinésine, très concentrée dans le RE et dans le cytosquelette du noyau et du cytoplasme et impliquée dans le transport rétrograde Golgi-RE et dans l'organisation du fuseau mitotique) et enfin **RAB6A** (protéine Rab qui régule globalement le trafic et les sens antérograde et rétrograde du transport golgien). Ceci est donc en adéquation avec le rôle connu de cette lectine dans le trafic intracellulaire, rôle très largement décrit dans la littérature.

VI.1.3. Cibles déstabilisées par la galectine-3

Parmi les ARN messagers-cibles déstabilisés se trouve une forte proportion (22%) de transcrits codant des protéines de réponse à des stress variés, comme par exemple :

- Adaptation au stress hypoxique : **VHL** (protéine impliquée dans le complexe ubiquitine ligase E3, et rôle de répresseur transcriptionnel de certains gènes pour

favoriser le retour à l'homéostasie de l'oxygène), **USP20** (un enzyme de débiquitination stabilisant HIF1 α qui est un régulateur transcriptionnel de l'adaptation au stress hypoxique et un substrat de la protéine VHL) ;

- Réponse au stress microbien : **ADGRG6** ;
- Réponse au stress viral : **SEC14L1** ;
- Réponse au stress thermique : **NUP160** ;
- Réponse aux dommages de l'ADN et remodelage de la chromatine : **DESI2**, **CBX6**, **PDRG1**, **MORF4L2**.

En ce qui concerne l'**apoptose**, on note une majorité d'ARNm-cibles déstabilisés à fonctions pro-apoptotiques, parmi lesquels :

- **STK4** codant la *serine/threonine kinase 4*, une kinase activée à la suite d'un stress et qui intervient après l'activation des caspases, en stimulant la condensation de la chromatine puis la fragmentation de l'ADN. Elle exerce un rôle pro-apoptotique. Sa déstabilisation par la galectine-3 est donc cohérente avec un rôle pour cette lectine qui selon notre hypothèse se situerait en amont, avec la mise en place d'une réponse permettant de résoudre le stress du RE justement pour éviter l'engagement cellulaire dans l'apoptose ;
- **TNFRSF10B** codant le membre 10b de la superfamille des récepteurs du TNF. Sa fonction est essentielle pour l'initiation de la cascade des caspases au cours de l'apoptose induite par un stress d'où la cohérence de sa déstabilisation potentielle par la galectine-3.

En ce qui concerne le **métabolisme**, on peut constater des fonctions des ARNm déstabilisés antagonistes à celles des transcrits stabilisés vus précédemment, par exemple : **GNPDA2** codant la glucosamine-6-Phosphate déaminase. Cet enzyme catalyse la réaction de désamination de la glucosamine-6-phosphate en fructose-6-phosphate (Oikari *et al.*, 2016). Cet enzyme est l'enzyme-clé de synthèse de glucoconjugués comme l'acide hyaluronique. La voie des hexosamines, qui régule l'équilibre des métabolismes des hexoses comme substrats énergétiques ou comme précurseurs de la biosynthèse des glycoconjugués, est régulée par deux groupes de deux enzymes GFPT1 (GAT1) et GFPT2 (GAT2) d'une part et GNPDA1 et GNPDA2 d'autre part qui agissent de façon inverse ou en concurrence (figure

107). De façon cohérente, alors que nous avons vu précédemment l'ARNm de *GFPT1* stabilisé, nous voyons celui de *GNPDA2* qui catalyse la réaction inverse, déstabilisé.

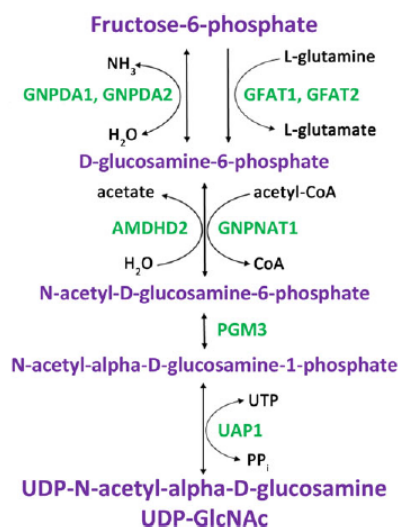


Figure 107 : Intervention opposée des enzymes GFPT1 et GNPDA1 dans le métabolisme des hexoses (Oikari *et al.*, 2016).

Une cible qu'il nous a paru important de mentionner est **DROSHA** codant une ribonucléase qui exécute la phase initiale de maturation des micro-ARN. C'est donc une enzyme-clé de la régulation post-transcriptionnelle et traductionnelle car impliquée dans le mécanisme généraliste de la synthèse des microARN. Par le biais de sa déstabilisation et donc de la diminution de la synthèse de microARN, la galectine-3 pourrait contribuer à inhiber la dégradation des transcrits et tenter de maintenir le mécanisme global de la traduction.

Ainsi, les fonctions « en miroir » des transcrits stabilisés et déstabilisés par la galectine-3, par exemple sur le métabolisme glucidique comme vu précédemment ou sur l'apoptose (anti-apoptotiques pour les ARNm stabilisés par la galectine-3 et pro-apoptotiques pour ceux déstabilisés) sont concordants avec l'hypothèse de l'implication de cette lectine dans la voie UPR et nous le verrons plus loin, avec les fonctions déjà connues de la galectine-3.

VII. Qu'en est-il des transcrits liés à la fois par la galectine-3 et HuR ?

Comme nous l'avons vu précédemment, l'étude des transcrits sélectionnés par RNA-IP a été effectuée avec un anticorps anti-galectine-3 ou un anticorps anti-HuR. Nous avons recoupé ces deux listes afin d'identifier des transcrits qui puissent être liés simultanément par ces deux protéines (mais pas forcément sur la même région du transcrit). Ainsi, nous avons analysé une liste de 261 transcrits associés à ces deux protéines.

Différentes catégories de fonctions pour ces transcrits sont retrouvées et notamment :

- Cytosquelette / jonctions cellulaires / trafic intracellulaire / transporteurs cellulaires ;
- Facteur de traduction, protéines ribosomiques ;
- Métabolisme : glucidique, lipidique, ionique ;
- ADN et ARN : facteur de transcription, histones, facteur d'épissage, snoRNA ;
- Signalisation cellulaire ;
- Stress/ inflammation : protéines de choc thermique, cytokines.

A la différence des listes précédentes d'ARNm « Gal-3 stabilisés/déstabilisés », nous trouvons davantage de facteurs impliqués dans le cycle de vie des ARNm (c'est-à-dire leur synthèse, traduction, transport) dans le répertoire des ARNm « Gal-3-HuR liés ». Ces résultats sont donc concordants avec les fonctions de HuR décrites dans ce processus mais aussi avec le rôle nouveau que nous avons décrit pour la galectine-3 dans le cycle de vie des ARNm.

Cependant, des fonctions communes entre ces deux listes de transcrits concernant le trafic intracellulaire, le métabolisme, l'inflammation, le cytosquelette, entre autres, existent également.

Le recoupement des différentes listes de transcrits (figure 108) a permis d'établir que :

- Sur les 891 ARNm liés par la galectine-3 (et non stabilisés par la galectine-3), 261 sont liés également par HuR ;
- Sur les 370 ARNm stabilisés par la galectine-3 (et non lié par la galectine-3), 48 sont liés également par HuR ;
- Sur les 15 ARNm liés et stabilisés par la galectine-3, 8 sont liés également par HuR (*MAPK1IP1L*, *KRT19*, *EZR*, *FAM83H*, *PSMD3*, *UBE2S*, *TONSL*, *GFPT1*) ;

- Sur les 55 transcrits déstabilisés par la galectine-3, aucun n'est lié à la fois par HuR et Gal3 ;
- Sur les 55 transcrits déstabilisés par la galectine-3, 14 sont liés par HuR.

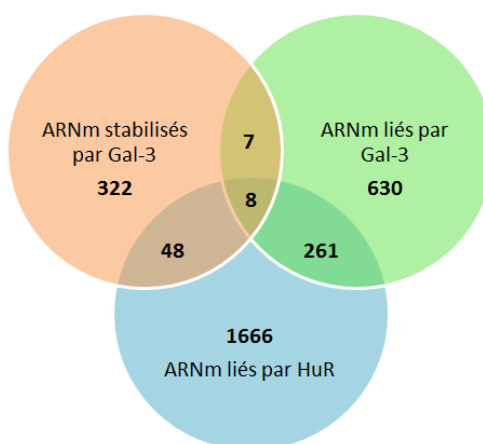


Figure 108 : Recouplement des listes des transcrits liés par la galectine-3, HuR et stabilisés par la galectine-3.

L'existence de cibles liées conjointement par ces deux protéines peut laisser supposer un rôle nouveau de la galectine-3 en lien avec HuR en dehors de la stabilisation de certaines espèces d'ARNm et concernant peut être le trafic intracellulaire des ARNm.

Il est intéressant de noter que peu des transcrits (12 sur 906 soit 1,3%) liés par la galectine-3 (mais non stabilisés ou déstabilisés) concernent la voie UPR ou la voie ERAD, suggérant ainsi que la liaison des ARNm par la galectine-3 ne serait pas impliquée dans ces fonctions. Les 12 transcrits sont les suivants : *HDGF*, *KHSRP*, *SSR1*, *SEC63*, *EXOSC4*, *EXOSC7*, *AARS*, *HSPA5*, *P4HB*, *TPP1*, *PSMC3*, *MARCH6*.

En conclusion, cette partie du travail consacrée à l'interprétation des données de *whole transcriptome stability* et RNA-IP ouvre des pistes quant au rôle de la galectine-3 autour de la réponse au stress et de la voie UPR, en tant **qu'acteur important de la coordination des différentes phases de réponse au stress** et notamment au cours de **la phase adaptative** (figure 109):

- en stabilisant les ARNm impliqués au début de la réponse qui codent des chaperonnes, des protéines participant à l'acquisition d'une bonne conformation protéique (co-chaperonnes), des précurseurs des voies de glycosylation pour restaurer le trafic des glycoprotéines et glycolipides ;
- en stabilisant des composants du système de destruction des protéines mal conformées et marquées par l'ubiquitine (via le système ubiquitine-protéasome) ;
- en stabilisant des facteurs de transcription qui code des protéines rattachées aux grandes fonctions identifiées ou qui stabilisent eux-mêmes des ARNm.

La galectine-3, pour permettre la mise en œuvre de cette phase adaptative, empêcherait de façon coordonnée l'activation des gènes pro-apoptotiques :

- en déstabilisant des gènes codant des facteurs pro-apoptotiques ou pro-inflammatoires ;
- en déstabilisant des facteurs de transcription qui régulent ce type de gènes de façon plus généraliste.

Le tableau XXI résume l'ensemble des cibles majeures présentées dans cette partie des résultats.

Tout ce long travail d'analyse bibliographique et critique de répertoires de transcriptomes réalisé dans le but de révéler des rôles nouveaux de la galectine-3 nous permettra de sélectionner des cibles d'intérêt qui feront l'objet d'investigations moléculaires dans le cadre de la poursuite de ce travail.

	Cibles rattachées directement à la voie UPR-ERAD			Cibles rattachées indirectement à la voie UPR			
	GSEA	Bases de données	Complément par analyse bibliographique	Chaperonne	Apoptose inflammation NFKB	Trafic intra-cellulaire	Métabolisme
Cibles stabilisées par la galectine-3	STC2		DER11	PSMG3	API5	ARFGAP1	GFPT1...
	ATF6		DERL1	POMP	DUS2 CHP1	COG7	
	ARFGAP1		JKAMP	PPIL2	CIAPIN1 DUS2	SCFD1	
	EXOSC1		SEL1L	PPIL3	TMBIM4	GET4	
	PARN		TBL2	DNAJA3	TRAF2 C1D	KIFAP3	
	SHC1		TMEM33	DNAJB6	CASP7 CSNK1D CSNK2A1	RAB6A	
	X POT		GFPT1	DNAJC7	FAM188A HSH2D		
	EIF4A1			HSPA1A	CXCL1, IL1RAP		
	IARS			HSPA1B	RELA1 CARD16, ASCC1		
	DKC1			MKKS	BCL2L13 BNIP3 MTDH TRIM8 TNFRSF12A ARRDC2 BIRC3 PSMD10 UBE2L3		
MTHFD2							
PAIP1							
Cibles déstabilisées par la galectine-3			CARHSP1		STK4		GNPDA2...
			EIF2AK1		TNFRSF10B		

Discussion

L'expression des mucines membranaires *MUC1*, *MUC4*, *MUC16* est dérégulée au cours de la cancérogenèse pancréatique, en particulier celle de *MUC4* qui n'est pas exprimée par le pancréas exocrine adulte sain. Cependant, les mécanismes moléculaires à l'origine de ces modifications d'expression ne sont pas complètement élucidés. La galectine-3, également surexprimée au cours de l'adénocarcinome pancréatique canalaire, est une lectine capable de se lier aux glycanes, aux protéines, ce qui explique ses nombreuses fonctions biologiques. Elle intervient aussi dans le métabolisme des ARNm via son rôle dans l'épissage des pré-ARNm. Les données de la littérature indiquent que les protéines impliquées dans le métabolisme des ARNm participent le plus souvent à plusieurs étapes du cycle de vie des transcrits, ce qui pourrait donc être le cas pour cette galectine. Tel était le contexte scientifique au début de ce travail de thèse.

Nous nous sommes donc intéressés au rôle de la galectine-3 dans la régulation de l'expression des mucines membranaires, avec un intérêt particulier pour *MUC4*. Les différentes stratégies et expériences développées ont donc permis de mettre en évidence et de caractériser plusieurs rôles de la galectine-3, dont un rôle nouveau dans le cycle de vie des transcrits de *MUC4*, fonction qui semble bien pouvoir être élargie à d'autres ARNm.

I. Etude de l'influence de la galectine-3 sur l'expression des mucines membranaires

La lignée cancéreuse pancréatique humaine CAPAN-1 se caractérise par une expression des transcrits des mucines membranaires *MUC1*, *MUC4*, *MUC16* et de la galectine-3. La perte de l'expression du gène *LGALS3* dans ce modèle cellulaire (cellules Sh1) (Merlin *et al.*, 2011) est associée à une diminution de l'expression de *MUC1* et de *MUC4* mais pas de *MUC16* par rapport aux cellules contrôles (Sc). L'expression conjointe de ces trois mucines membranaires et de la galectine-3 a également été caractérisée dans d'autres lignées cellulaires d'origine pancréato-biliaire, comme le montrent les données de transcriptomique obtenues par Barretina (Barretina *et al.*, 2012). Notre modèle peut donc être jugé comme représentatif du modèle de la cancérogenèse épithéliale pancréatique que nous avons choisi d'étudier.

A côté de ces résultats concernant les lignées cellulaires, l'étude de l'expression tissulaire des ARNm de ces mucines et de la galectine-3 (étude effectuée sur 10 couples de tissus d'adénocarcinomes pancréatiques canaux et tissus sains associés) indiquent une corrélation entre l'expression des transcrits *LGALS3* avec ceux de *MUC4* mais pas avec ceux

de *MUC1* et de *MUC16*. Ces données suggèrent donc que, dans notre modèle d'adénocarcinome pancréatique canalaire, le lien le plus direct entre la galectine-3 et les ARNm de mucines concerne principalement *MUC4*. Afin d'expliquer l'influence de la galectine-3 sur l'expression des mucines membranaires, nous avons orienté nos recherches vers la détermination d'une régulation s'exerçant soit au niveau transcriptionnel soit au niveau post-transcriptionnel.

Actuellement, la relation entre l'expression de *MUC1* ou de *MUC4* avec la galectine-3, protéines toutes trois surexprimées au cours de l'adénocarcinome pancréatique n'est pas décrite dans la littérature. Le rôle de la galectine-3 dans la régulation de l'expression de *MUC2* a déjà fait l'objet de plusieurs études. Dudas *et al.* (2002) et Song *et al.* (2005) ont en effet montré que la galectine-3 régule l'expression de *MUC2* dans des cellules cancéreuses coliques humaines au niveau transcriptionnel par le biais de la stabilisation du complexe de transcription AP-1 sur un de ses éléments cis (-1871) présents dans le promoteur du gène *MUC2*. La galectine-3 régule également l'expression du gène codant la cycline D1 en stabilisant la fixation du facteur de transcription Sp1 sur le promoteur de ce gène (Lin *et al.*, 2002). La galectine-3 ayant une localisation nucléaire dans les cellules Sc (Merlin *et al.*, 2011) et les promoteurs de *MUC1* et de *MUC4* possédant des sites de fixation pour les facteurs AP-1 et Sp1 (Perrais *et al.*, 2001), nous avons tout d'abord émis l'hypothèse d'une régulation transcriptionnelle exercée par la galectine-3 sur *MUC1* et de *MUC4*. Les résultats obtenus lors des expériences de co-transfections cellulaires transitoires en présence du vecteur d'expression de la galectine-3 et des différentes constructions du promoteur de *MUC1* ou de *MUC4* dans les cellules Sh1 indiquent que cette lectine ne module pas l'activité transcriptionnelle de *MUC1* et de *MUC4*. Le processus décrit pour *MUC2* ne s'applique donc pas à ces deux mucines *MUC1* et *MUC4*.

Ces résultats nous ont donc conduits à envisager un autre mécanisme permettant d'expliquer la modulation de l'expression de *MUC1* et de *MUC4* par la galectine-3.

L'étude menée par Fischer *et al.* (2003) a mis en évidence l'implication de l'élastase des polynucléaires neutrophiles, produite lors du processus inflammatoire, dans la modification de la stabilité des transcrits de *MUC4*. En effet, leur demi-vie passe de 5 h à 21 h dans ce contexte inflammatoire. En se basant sur ces données, nous avons donc étudié l'hypothèse selon laquelle la galectine-3 module l'expression de *MUC1* et de *MUC4* au niveau post-transcriptionnel, notamment en influençant la stabilité de leurs transcrits. Pour cela, nous

avons déterminé les demi-vies des ARNm de *MUC1* et de *MUC4* dans les cellules Sh1 et Sc en présence d'actinomycine D.

Les résultats obtenus indiquent une diminution statistiquement significative de la demi-vie des transcrits de *MUC4* en absence de cette protéine. La demi-vie des transcrits de *MUC1* n'est pas différente entre les cellules Sc et Sh1, ce qui indique que la galectine-3 ne semble pas modifier la stabilité des ARNm de *MUC1*. Nos données mettent en évidence que la galectine-3 stabilise spécifiquement les transcrits de *MUC4* et exerce donc une régulation au niveau post-transcriptionnel.

Le rôle de stabilisation des ARNm exercées par la galectine-3 sur les ARNm de *MUC4* est novateur puisque non décrit à l'heure actuelle dans la littérature. L'idée d'une stabilisation possible des transcrits de *MUC4* a été développée par Fischer *et al.* (2003) dans le cadre de cellules non épithéliales mais dans un contexte lié à l'inflammation pour lesquels les mécanismes n'ont pas été élucidés mais pourraient potentiellement être attribuables à la galectine-3 exprimées également dans les polynucléaires neutrophiles.

II. Etude des mécanismes de stabilisation des transcrits de *MUC4* par la galectine-3

Afin d'expliquer les mécanismes mis en jeu dans la stabilisation des transcrits de *MUC4* par la galectine-3, nous avons dans un premier temps exploré l'hypothèse d'une stabilisation médiée par des éléments *cis* présents dans la région 3'UTR de *MUC4*. En effet, un élément important dans la stabilité des ARNm est leur région 3'UTR qui comporte notamment des sites de fixation pour des microARN, pour des RBP stabilisatrices ou déstabilisatrices. L'analyse bioinformatique du 3'UTR de *MUC4* a révélé la présence de divers de ces éléments et plus précisément d'ARE présents au nombre de trois. Nous nous sommes dans un premier temps intéressés à ces éléments *cis* puisque la protéine HuR est largement décrite dans la littérature pour stabiliser les ARNm comme par exemple ceux de l'*INFβ* (Takeuchi, 2015). Cependant, les expériences de transfections réalisées en présence d'une construction du 3'UTR de *MUC4* et des vecteurs d'expression de la galectine-3 et de HuR n'ont pas mis en évidence d'effet potentialisateur de HuR sur la stabilisation exercée par la galectine-3 via ces régions.

Nous nous sommes donc intéressés au motif *CA repeat* présent également dans la région 3'UTR de *MUC4* et conservé chez le rat et la souris. Dans la littérature, cet élément *cis*

est décrit comme participant à la stabilisation des ARNm du *VEGF* (Shih and Claffey, 1999) et de *CD154* (Hamilton *et al.*, 2008) via la fixation de la protéine hnRNP-L. Les résultats obtenus lors des co-transfections du 3'UTR de *MUC4*, sous forme native ou dépourvue du CARE, en présence des vecteurs d'expression de la galectine-3 et/ou de hnRNP-L ainsi que l'utilisation de *si*-ARN ciblant hnRNP-L nous indiquent que l'élément CARE et hnRNP-L participent à la régulation de l'expression de *MUC4*. Nous avons également démontré par *RNA pull-down* que hnRNP-L se lie directement sur le CARE comme décrit par Shih et claffey (1999) pour les ARNm du *VEGF*. Cependant nous n'avons pas mis en évidence d'interaction directe du CARE avec la galectine-3. Ce dernier résultat est cohérent avec les données de la littérature qui décrivent que la plupart des actions médiées par la galectine-3 sur un ARNm s'exercent de façon indirecte, que ce soit pour la régulation transcriptionnelle où elle agit par la stabilisation de facteur de transcription, ou pour son rôle dans l'épissage où elle intervient par l'intermédiaire de la ribonucléoprotéine U1.

De façon intéressante, les résultats de *RNA pull down* indiquent que la galectine-3 intensifie la liaison de hnRNP-L sur le CARE. Nous pouvons donc en déduire que la galectine-3 stabilise les transcrits de *MUC4* en potentialisant la fixation de hnRNP-L sur cette région. Afin de déterminer si la galectine-3 agit de façon indirecte en se liant à hnRNP-L, nous avons réalisé des expériences de PLA, co-immunofluorescence et co-immunoprécipitation qui ont mis en évidence une interaction entre ces deux protéines.

Ces résultats sont en adéquation avec les données récemment publiées (Fritsch *et al.*, 2016). En effet, cette équipe a montré la possibilité d'une interaction entre la galectine-3 et la protéine hnRNP-L par spectrométrie de masse sur laquelle nous pouvons nous appuyer pour valider nos résultats. Par ailleurs, il est intéressant de noter que leurs travaux ont également mis en évidence une interaction entre la galectine-3 et une autre protéine de la famille hnRNP : hnRNPA2B1. Ils démontrent que cette interaction est impliquée dans l'assemblage de la machinerie d'épissage ainsi que dans l'export des ARNm. Enfin, une autre étude a décrit également l'interaction entre la galectine-3 et un autre membre de la famille hnRNP qu'est hnRNP Q (Yoo *et al.*, 2009). Elle démontre que la formation de ce complexe stabilise la protéine hnRNP Q. L'inhibition de l'expression de la galectine-3 entraîne la dégradation de cette protéine et s'accompagne d'une diminution de la prolifération des cellules cancéreuses coliques.

Qu'en est-il du mécanisme mis en jeu par hnRNP-L pour stabiliser les transcrits de *MUC4*? Nous pouvons avancer plusieurs hypothèses en fonction des données de la littérature :

- La fixation de hnRNP-L pourrait protéger les transcrits de *MUC4* de la dégradation exercée par certains microARN en empêchant leur fixation. C'est ce mécanisme qui est mis à profit lors de la stabilisation des ARNm du *VEGF* par hnRNP-L. En effet, en conditions hypoxiques, hnRNP-L est transloquée du noyau vers le cytoplasme et se lie sur le CARE présent dans le 3'UTR des ARNm du *VEGF*. Cette interaction empêche l'activité inhibitrice des mir-297 et -299 (qui s'exerce en condition normoxique) en bloquant l'accès au complexe miRISC qui se lie au même élément (Jafarifar *et al.*, 2011, figure 110). Ceci entraîne donc la stabilisation des ARNm du *VEGF*.

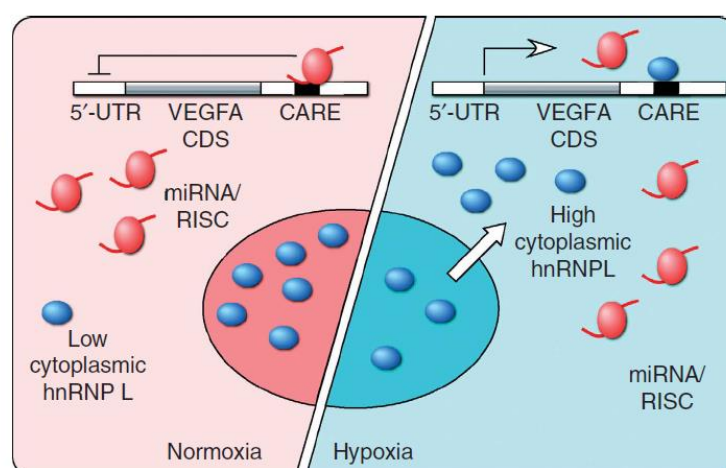


Figure 110 : Mécanisme de stabilisation des transcrits du *VEGF* par hnRNP-L (Jafarifar *et al.*, 2011).

- Il a été démontré dans la littérature que le motif *CA repeat* peut être un site de clivage endonucléolytique (Bracken *et al.*, 2011; Hui *et al.*, 2003). En se fixant sur cette région, hnRNP-L pourrait protéger les transcrits de *MUC4* de cette dégradation. Ainsi, Hui *et al.* ont décrit la présence d'un motif CARE au niveau de l'intron 13 de l'ARNm de *eNOS*. En absence d'hnRNP-L, cet ARNm est clivé en amont du CARE et à l'inverse, l'ajout de hnRNP-L restaure la stabilité de ces transcrits.

Afin d'étudier la relevance de nos résultats au niveau physiologique, nous avons étudié l'expression de MUC4 et hnRNP-L sur des tissus de souris WT et KO pour la galectine-3 et nos résultats confirment la relation entre la diminution de l'expression de la protéine hnRNP-L et des ARNm de *MUC4* au niveau du jéjunum. De façon parallèle, l'absence de variation d'expression de hnRNP-L au niveau du poumon ne s'accompagne d'aucune modification de la quantité des transcrits de *MUC4*. Cette différence de résultats entre tissu digestif et pulmonaire pourrait s'expliquer par l'expression plus intense de la galectine-3 au niveau du jéjunum mais également par l'hétérogénéité du tissu pulmonaire qui a été étudié de façon globale. D'autre part, il est intéressant de noter que la littérature indique que les souris KO pour la galectine-3 ne présentent pas d'anomalies phénotypiques notables et notamment ne développent pas spontanément de cancer (Eude-Le Parco *et al.*, 2009). Les entérocytes des souris KO présentent des anomalies (Delacour *et al.*, 2008) au niveau de la membrane basolatérale avec formation de structures de type microvillosités. Ces souris présentent également différents troubles qui peuvent être reliés aux fonctions de la galectine-3 : anomalies immunologiques, altération du métabolisme glucidique et lipidique avec une inflammation systémique (Pang *et al.*, 2013) et une diminution de la sécrétion de mucus lors de la stimulation des voies respiratoires par des allergènes (Ge *et al.*, 2010). Cette dernière observation est intéressante puisque en lien avec les résultats que nous avons obtenus : la galectine-3 est associée à l'expression des mucines respiratoires comme MUC4.

Au cours des expériences de PLA, nous avons constaté une augmentation de l'interaction entre la galectine-3 et hnRNP-L dans les cellules en mitoses. Afin de confirmer cette observation, nous avons traité nos cellules par le nocodazole, un agent bloquant les cellules en phase G2/M. Les résultats d'imagerie confirment une interaction accrue entre la galectine-3 et hnRNP-L au cours de ce traitement. Ce phénomène pourrait s'expliquer par l'augmentation de l'expression de hnRNP-L lors du passage de la phase S à G2 observée dans les travaux de Lane (Lane *et al.*, 2013). Ceci pourrait être à l'origine d'une augmentation quantitative de l'interaction entre ces deux protéines. Par ailleurs, les cellules mitotiques sont caractérisées par un arrêt des ribosomes sur les ARNm, un blocage de la traduction et une protection des transcrits (Sivan *et al.*, 2007; Tanenbaum *et al.*, 2015). Cette répression de la traduction permet une synthèse sélective de protéines nécessaires à la cellule lors de la division cellulaire. L'interaction augmentée entre la galectine-3 et hnRNP-L observée en mitose est donc en adéquation avec une protection accrue des transcrits lors de la mitose, et

notamment ceux de la volumineuse O-glycoprotéine MUC4, dont la synthèse protéique ne serait pas indispensable à ce moment de la vie cellulaire.

III. Liaison de la galectine-3 à des ARNm matures de mucines dans le cytoplasme

La galectine-3 ne possède pas dans sa structure de domaine de fixation aux ARNm. Comme nous l'avons décrit précédemment, les données de la littérature indiquent qu'elle interagit de façon indirecte avec les pré-ARNm via sa fixation avec la ribonucleoprotéine U1 participant ainsi au processus l'épissage. Cependant, aucune étude de la littérature ne s'est intéressée à l'interaction directe entre la galectine-3 et les transcrits ni même à son interaction avec des ARNm matures. Néanmoins, dans une étude à grande échelle menée par Castello *et al.* en 2012 afin d'identifier de façon exhaustive l'ensemble des RBP par spectrométrie de masse, il s'avère que la galectine-3 figure parmi la liste de ces protéines (Castello *et al.*, 2012).

Nous avons étudié la liaison de la galectine-3 sur les ARNm de mucines par imagerie cellulaire et avons mis en évidence une colocalisation périnucléaire entre les ARNm de *MUC4* et la galectine-3. Ceci suggère une interaction avec ces transcrits après leur export nucléaire donc avec des ARNm matures et épissés. De plus, nous avons effectué une RNA-IP à l'aide d'un anticorps anti-galectine-3 afin de collecter les complexes ARNm-protéines et d'analyser par séquençage à haut débit de l'ensemble des ARNm immunoprécipités. Là encore, les résultats indiquent une interaction de la galectine-3 avec des ARNm des mucines matures. Par ailleurs, il est intéressant de noter que l'interaction se fait au niveau du plus gros exon des transcrits de mucines, par l'intermédiaire d'un motif décrit par Bolisetti comme étant impliqué dans la rétention des grands exons lors du processus d'épissage (Bolisetti and Beemon, 2012). Ceci pourrait suggérer un rôle de la galectine-3 dans l'intégration des gros exons, caractéristiques de certains transcrits de mucines, lors de ce processus. Cette expérience, réalisée avec un *crosslinking* préalable des complexes ARNm-protéines par le formaldéhyde ne permet pas de déterminer si l'interaction de la galectine-3 avec les transcrits via ce motif est directe ou fait intervenir la liaison de la galectine-3 à une autre protéine.

Certes, la galectine-3 ne peut pas être considérée comme une RBP dite conventionnelle car elle ne possède pas de domaines classiques de fixation aux ARNm. Cependant une équipe a décrit la présence dans sa structure d'une région de faible complexité (des aa 29 à 113)

(Cornejo-García *et al.*, 2016). Ces types de domaines sont répertoriés comme des motifs étant présents dans certaines RBP et notamment dans celles ayant un rôle dans la formation de structures cellulaires telles que des granules nucléaires et cytoplasmiques (Helder *et al.*, 2016). Cette région pourrait donc être impliquée dans la liaison de la galectine-3 directement sur les ARNm mais ceci reste à déterminer.

Nous pouvons ensuite nous demander quand a lieu la liaison de la galectine-3 sur les transcrits. Deux hypothèses sont envisageables :

- la galectine-3 serait liée aux pré-ARNm lors de l'épissage dans le noyau, resterait associée aux transcrits lors de l'export nucléaire et les accompagnerait dans le cytoplasme ;
- la galectine-3 nucléaire se dissocierait des transcrits dans le noyau, puis l'ARNm serait exporté et la galectine-3 cytoplasmique se réassocierait avec les transcrits.

Dans tous les cas, la galectine-3 accompagnerait les transcrits à plusieurs stades du cycle de vie des ARNm, tout comme cela a été décrit pour HuR (Srikantan and Gorospe, 2011).

IV. Liens entre mucines et galectine-3 dans le cancer du pancréas

Les mucines membranaires MUC4, MUC1, MUC16 et la galectine-3 sont surexprimées au cours de l'adénocarcinome pancréatique. Notre étude a démontré que la galectine-3 stabilise spécifiquement les transcrits de *MUC4* et pas ceux de *MUC1*. L'interconnexion entre mucines et galectines au cours du cancer est largement décrite dans la littérature. Ces deux catégories de protéines ont leur rôle propre dans la cancérogenèse mais celles-ci interagissent et/ou agissent de façon synergique afin de faciliter le processus métastatique en permettant la redistribution des mucines à la membrane plasmique des cellules cancéreuses ce qui démasque des molécules d'adhésion au niveau des surfaces cellulaires (Senapati *et al.*, 2011; Zhao *et al.*, 2009). De plus, la galectine-3 régule l'expression de *MUC2* dans le cancer colique au niveau transcriptionnel (Song *et al.*, 2005). Le lien que nous avons démontré concernant la régulation de l'expression de *MUC4* par la galectine-3 renforce donc le lien entre ces deux acteurs de la cancérogenèse pancréatique.

Une explication peut être apportée quant au rôle sélectif de la stabilisation des ARNm de *MUC4* par la galectine-3 par rapport à ceux de *MUC1*. En effet, la mucine MUC1 est exprimée à l'état basal dans un grand nombre de cellules épithéliales, dont le pancréas, alors que l'expression de MUC4 en est absente à l'état physiologique. MUC1 pourrait donc plutôt

être considérée comme une mucine « de ménage » à la différence de MUC4. De plus, plusieurs études ont mis en évidence l'implication de MUC4 (mais aussi de MUC16) au cours de l'inflammation qui est un état associé au développement du cancer (Das *et al.*, 2016; Haridas *et al.*, 2014; Lundmark *et al.*, 2015). Comme nous l'avons présenté dans l'introduction, le rôle de la galectine-3 dans l'inflammation est également décrit dans la littérature. Nos résultats suggèrent donc que MUC4 pourrait appartenir à un groupe de gènes dont l'expression est induite en partie par la galectine-3 dans les situations inflammatoires comme celles associées par exemple aux lésions précancéreuses pancréatiques.

Comment expliquer alors la diminution de l'expression des ARNm de *MUC1* en absence de galectine-3 alors que la stabilité des transcrits n'est pas altérée ? L'hypothèse d'une modulation de l'expression de microARN par la galectine-3 pourrait être avancée afin d'expliquer la diminution des ARNm de *MUC1* sans modification de l'activité du promoteur de *MUC1* et de la stabilité de ses transcrits par la galectine-3. Des liens ont été décrits dans la littérature entre MUC1, galectine-3 et microARN. En effet, dans la boucle inverse de celle décrite dans notre étude, Ramasamy *et al.* (2007) ont mis en évidence que MUC1 régule l'expression du mir-322 qui déstabilise les transcrits de *LGALS3*. D'autre part, la perte d'expression de la galectine-3 peut être associée à des modulations de l'expression de microARN. Ainsi, le niveau du mir let-7c, régulateur négatif de la transcription de *KRAS*, est augmenté en absence de cette lectine (Levy *et al.*, 2011). On pourrait donc tout à fait supposer que dans notre modèle, l'expression de la galectine-3 est associée à la perte d'expression d'un mir ciblant *MUC1*.

Le ciblage de la galectine-3 constituerait donc une stratégie thérapeutique intéressante permettant d'une part de réduire l'expression de MUC4 et de MUC1, toutes deux impliquées dans la cancérogenèse pancréatique, et d'autre part d'empêcher son interaction avec les mucines, interaction impliquée dans le processus métastatique.

Actuellement, dans le cas du PDAC, les taux de réponse à la chimiothérapie sont faibles. Ceci serait lié à l'hétérogénéité tumorale ainsi qu'au microenvironnement tumoral caractérisé par un stroma dense limitant l'accès de la tumeur aux agents thérapeutiques et aux cellules immunitaires (Matsuoka and Yashiro, 2016).

Le traitement de référence du PDAC a longtemps reposé sur l'utilisation de la gemcitabine en monothérapie (Saif, 2011). Depuis 2011, un traitement basé sur l'utilisation du protocole FOLFIRINOX (association d'oxaliplatine, d'irinotécan, de leucovorine et de 5-

Fluorouracile) permet d'augmenter la survie des patients par rapport au traitement de référence basé sur la gemcitabine seule (survie de 11,1 mois avec FOLFIRINOX *versus* 6,8 mois avec la gemcitabine) (Hammel *et al.*, 2015; Saif and Chabot, 2011). Cependant, du fait de sa toxicité importante, il ne peut être proposé qu'aux patients dont l'état général le permet. Le Nab-paclitaxel (paclitaxel lié à l'albumine), un poison du fuseau mitotique (molécule de la famille des taxanes), a été autorisé par la FDA en 2013 en association avec la gemcitabine. Celui-ci permet une amélioration de la survie de 8.5 mois *versus* 6.6 mois avec la gemcitabine seule (Von Hoff *et al.*, 2013). Néanmoins, tout comme le Folfirinox, la toxicité de ce traitement est importante et il ne peut donc être proposé qu'aux patients dont l'état général le permet.

De nombreuses stratégies thérapeutiques sont en cours d'évaluation et les mucines constituent des cibles potentielles. Ainsi, l'immunothérapie passive à l'aide d'anticorps dirigés contre des protéines anormalement exprimées par les cellules cancéreuses et notamment MUC1 est développée (Jonckheere *et al.*, 2010). L'immunothérapie active stimulant la production d'anticorps par le patient est également testée, comme par exemple le vaccin peptidique basé sur l'utilisation du peptide K-Ras muté (Koido *et al.*, 2011) ou le gVAX, un vaccin développé à partir de cellules cancéreuses pancréatiques irradiées transfectées pour produire du GMCSF. Ce facteur de croissance augmente l'activité des cellules présentatrices d'antigènes qui vont activer à leur tour les lymphocytes cytotoxiques CD4 et CD8 (Oyasiji and Ma, 2015). Enfin, de nombreuses thérapies ciblées sont en cours d'évaluation tels que les inhibiteurs d'angiogenèse, de tyrosines kinases et de métalloprotéases (Hammel *et al.*, 2015; Izeradjene and Hingorani, 2007; Seicean *et al.*, 2015).

Dans ce contexte, le développement d'inhibiteurs de la galectine-3 serait donc intéressant de par le lien qu'elle entretient avec les mucines surexprimées dans ce cancer. De plus, comme nous l'avons vu dans l'introduction, la galectine-3 est impliquée dans le processus fibrotique. Son inhibition constituerait donc une stratégie potentielle pour améliorer l'accessibilité aux cellules tumorales des agents thérapeutiques utilisés dans le PDAC et dont la distribution est limitée par le stroma intense retrouvé dans ce cancer.

V. Rôle de la galectine-3 dans le métabolisme des ARNm

Ayant démontré que la galectine-3 stabilise les ARNm de *MUC4* et qu'elle interagit avec des transcrits matures dans le cytoplasme, nous avons voulu dans la suite de nos travaux

déterminer si la galectine-3 était impliquée dans le stockage des ARNm dans le cytoplasme. En effet, au niveau cytoplasmique, les ARNm sont en échange constant entre des structures dynamiques, telles que les P-Bodies ou les granules de stress, et le pool d'ARNm en traduction. Par PLA et immunofluorescence, nous avons étudié, à la fois à l'état basal et après un choc thermique, l'interaction entre la galectine-3 et XRN-1 ou DDX6, marqueurs de P-Bodies et TIA-1, eIF2 α ou G3BP, marqueurs de granules de stress. Nous avons démontré que la galectine-3 interagit rarement avec ces marqueurs suggérant qu'elle ne serait pas principalement présente dans ces deux types de granules. Afin de déterminer le devenir des ARNm de *MUC4* lors d'un stress, nous avons étudié l'interaction entre les transcrits de *MUC4* et TIA-1 d'une part et la galectine-3 d'autre part à l'état basal ou après un choc thermique. Les résultats indiquent que le stress n'induit pas une augmentation de l'interaction entre TIA-1 et *MUC4* alors que l'interaction des transcrits de *MUC4* avec la galectine-3 diminue. L'ensemble de ces résultats suggèrent donc que l'interaction entre la galectine-3 et les ARNm de *MUC4* se produit dans d'autres structures que les P-bodies et les granules de stress. Enfin, par FISH, nous avons montré que le stress induit une colocalisation nucléaire entre la galectine-3 et les ARNm de *MUC4*.

Chaque type de stress auquel est soumise la cellule s'accompagne d'une réponse différente (Hetz, 2012). Il est intéressant de noter que dans le cas particulier du stress thermique, nous observons une diminution quantitative de l'interaction entre les transcrits de *MUC4* et la galectine-3 par rapport à l'état basal ainsi qu'une redistribution de la colocalisation au niveau nucléaire et non plus cytosolique comme observé en absence de stress. Plusieurs hypothèses peuvent être avancées pour expliquer ces observations (relocalisation et diminution de l'interaction) :

- l'export des ARNm est habituellement interrompu lors d'un stress. Ce mécanisme serait à l'origine de l'accumulation des ARNm dans le noyau d'où un renforcement de la colocalisation observée au niveau nucléaire où la galectine-3 est associée aux transcrits pour l'épissage. Ceci expliquerait également la diminution quantitative de l'interaction puisque l'interaction cytosolique serait dans ce cas fortement diminuée (Burgess *et al.*, 2011; Zander *et al.*, 2016) ;
- le stress thermique pourrait être associé à des modifications de la phosphorylation de la galectine-3, processus à l'origine d'une perturbation de son trafic nucléocytoplasmique ;

- en situation de stress thermique, les transcrits de *MUC4* pourraient être dégradés (?) car inutiles à la cellule en situation d'urgence vitale, et donc non stockés dans le cytoplasme (?), ce qui serait à l'origine d'une interaction cytosolique diminuée, la colocalisation nucléaire étant associée à l'épissage.

Si la galectine-3 n'interagit pas avec les transcrits dans les granules de stress et les P-bodies, dans quelles structures se produit donc l'interaction ?

Fritsch *et al.*, (2016) ont démontré que la galectine-3 interagit avec la protéine hnRNPA2/B1. Or cette dernière joue un rôle important dans le transport des ARNm neuronaux (Percipalle *et al.*, 2009). De plus, Elvira *et al.*, (2006) ont analysé la composition de granules neuronaux et y ont retrouvés hnRNPA2/B1 mais aussi hnRNP-L. Or nos travaux ont permis de démontrer que la galectine-3 interagit avec la protéine hnRNP-L. On peut donc supposer que la galectine-3, par le biais de son interaction avec hnRNP-L, ou bien directement via son domaine de faible complexité, interagit avec des ARNm présents dans des granules traductionnellement inactifs les protégeant ainsi de la dégradation. Ainsi, Fawal *et al.*, (2011) ont décrit des structures appelées « granules ALK », distinctes des SG et des PB, ayant une composition différente de RBP et contenant des ARNm polyadénylés, les protégeant ainsi de la dégradation. On peut émettre l'hypothèse que des granules de ce type, dans lesquels la galectine-3 serait présente pour protéger les ARNm, pourraient exister dans nos cellules.

Des résultats préliminaires concernant l'étude de l'expression de *MUC4* au niveau protéique par *western blot* semble indiquer une très faible diminution dans les cellules Sh1. Ceci suggère donc qu'en présence de galectine-3 les ARNm de *MUC4* sont protégés de la dégradation sans pour autant que cela ne s'accompagne d'une augmentation de la traduction. Cette observation concorde avec un stockage de transcrits dans des structures où la traduction est inactive. On peut supposer que la galectine-3, dans le contexte d'une production intense de protéines génératrices d'un stress du réticulum, stabilise et protège les ARNm de *MUC4* en les stockant dans des granules dont la nature n'a pas encore été déterminée. Ils ne seraient pas traduits car non indispensables à la cellule, le temps que les mécanismes permettant à la cellule de retrouver l'homéostasie se mettent en place. Ceci permettrait à la cellule de stocker ces transcrits codant pour des protéines de grandes tailles dont la synthèse n'est pas immédiatement requise.

DeGracia *et al.* (2008) ont montré dans leur étude que les granules à HuR permettent de regrouper des ARNm fonctionnellement ou structurellement reliés. Ceux-ci forment ainsi des « opérons à ARN » permettant une régulation concertée de leur traduction et dégradation ce qui assure une réponse cellulaire coordonnée lors d'un stress. De la même façon que celle exercée par HuR, il est possible que la galectine-3 entre dans la composition de granules regroupant des ARNm aux fonctions reliées afin de les stabiliser.

Nous avons précédemment évoqué l'implication de la galectine-3 et de MUC4 lors de l'inflammation associée au développement du cancer (Das *et al.*, 2016; Iacobini *et al.*, 2011; Lundmark *et al.*, 2015). Les résultats de RNA-IP et de *whole transcriptome stability* nous ont permis d'étudier de façon plus large le rôle de la galectine-3 sur l'ensemble des transcrits exprimés dans notre modèle cellulaire. Afin d'identifier des groupes fonctionnels d'ARNm ciblés par la galectine-3, nous avons analysé les fonctions connues de groupes d'ARNm, stabilisés ou déstabilisés par la galectine-3, liés par la galectine-3 ou liés par la galectine-3 et HuR. Ceci nous a permis d'identifier un dénominateur commun à ces fonctions et s'inscrivant dans le contexte d'une cellule cancéreuse : la voie UPR et la voie ERAD qui l'accompagne. Ce répertoire a ensuite été confirmé pour l'étude élargie à l'ensemble de 385 ARNm stabilisés par la galectine-3.

L'hypothèse de la voie UPR s'inscrit parfaitement dans le cadre d'un modèle de cellules cancéreuses car, dans ce contexte, les voies de signalisation associées au stress du RE sont dérégulées et une activation constitutive de la voie UPR est fréquemment observée. En effet, le métabolisme de ces cellules nécessite une synthèse accrue de protéines pour répondre à leur niveau de prolifération élevé (Kato and Nishitoh, 2015). De plus, l'environnement tumoral pauvre en nutriment et hypoxique (par rapport à la demande) stimule l'activation de cette voie. Par ailleurs, les oncogènes comme MYC, en augmentant la synthèse de protéines et par conséquent leur entrée dans le RE, stimulent la voie UPR pour augmenter les capacités d'acquisition de bonne conformation des protéines des cellules cancéreuses. De plus, s'agissant d'un modèle cellulaire épithélial de cancer pancréatique, une surexpression massive des mucines membranaires (*MUC1*, *MUC4* et *MUC16*) mais aussi de mucines sécrétées (*MUC5AC*, *MUC6*, en particulier), peut tout à fait contribuer à l'installation d'un état de stress du réticulum endoplasmique et de l'appareil de Golgi (Hetz, 2012).

Les cibles ARNm que nous avons identifiées sont impliquées dans divers groupes de fonctions tels que le trafic cellulaire et vésiculaire, le métabolisme, le catabolisme des

protéines par le système ubiquitine-protéasome, la signalisation cellulaire, la voie UPR, la transcription et la traduction, le cycle cellulaire, tous ces groupes de fonctions étant reliables entre eux par la seule voie UPR. De manière analogue à DeGracia pour caractériser la fonction de la protéine HuR au sein des granules HuR, nous proposons d'utiliser le concept de « régulon » pour désigner le rôle de la galectine-3 sur ces répertoires de transcrits (DeGracia *et al.*, 2008).

L'identification récente de la galectine-3 à l'intersection d'un réseau de gènes structuré autour de 4 gènes-maîtres (Verardo *et al.*, 2016) est en accord avec le concept de « régulon » que met également en évidence notre étude. Les travaux récents de cette équipe ont placé la galectine-3 au centre des quatre groupes de fonctions suivants (figure 111) :

- Survie cellulaire sous la dépendance du facteur FOXA1 (Forkhead Box A1) ;
- Détermination du devenir cellulaire sous la dépendance du facteur KLF4 (Kruppel-like factor 4) ;
- Métabolisme lipidique et de l'inflammation sous la dépendance du facteur NFκB ;
- Développement sous la dépendance du facteur FOXD1 (Forkhead Box D1).

Il est intéressant de noter dans ce réseau la présence du facteur MAPK1IPL1, membre de la voie des MAPK, aux fonctions mal connues, dont le gène est contiguï sur le génome avec le gène *LGALS3* et avec lequel il partage des éléments de régulation. L'ARNm de ce gène fait partie de la liste restreinte des 15 cibles liées et stabilisées par la galectine-3.

Par ailleurs, ces quatre acteurs majeurs de ces différents réseaux, tout comme la galectine-3, sont impliqués dans le développement de maladies rénales, cardiovasculaires et du diabète.

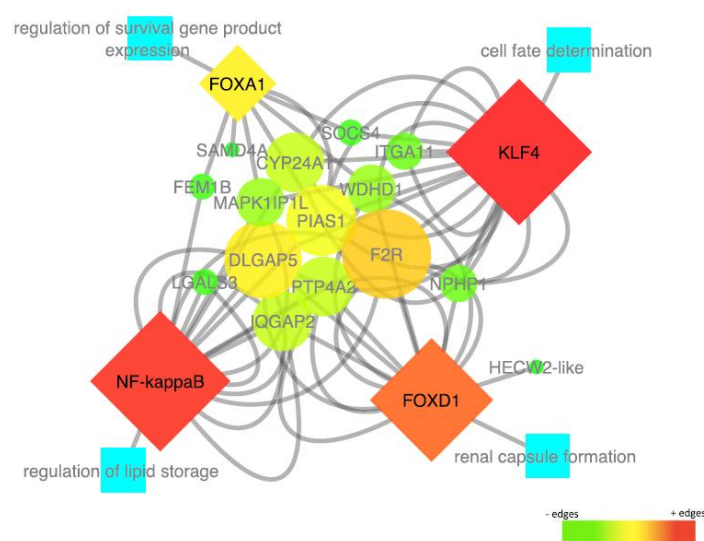


Figure 111 : Le gène *LGALS3* situé à l'intersection des quatre réseaux de fonctions sous la dépendance de NFκB, FOXD1, KLF4 et FOXA1 (Verardo *et al.*, 2016).

Plusieurs études se sont intéressées à l'expression de la galectine-3 dans les cellules exposées à un stress du RE et au lien qu'il peut exister entre galectines et cette situation physiopathologique. Ainsi, (Timoshenko *et al.*, 2016) ont mis en évidence une modification du profil d'expression des galectines dans des cellules HL-60 soumises à un stress du RE (induit par hypoxie ou par un agent chimique) avec notamment une augmentation de l'expression de la galectine-3. De plus, l'inhibition de l'expression de la galectine-1, qui possède des fonctions redondantes avec la galectine-3 (Patterson *et al.*, 2004), s'accompagne d'une altération du stress du RE (Le Mercier *et al.*, 2008). L'ensemble de ces résultats sont donc concordants avec l'hypothèse d'une implication de cette lectine dans la voie UPR qui est particulièrement active dans les cellules cancéreuses.

Ainsi, notre lignée cellulaire CAPAN-1, caractérisée par une production importante de mucines, et notamment MUC4, pourrait être dans un état de stress chronique dû à la présence de ces volumineuses protéines en quantité anormale dans le réticulum endoplasmique. Il est en effet décrit dans la littérature que les cellules mucisécrétrices sont caractérisées par un stress du RE (Hetz, 2012). Il est intéressant de noter qu'au niveau respiratoire, la protéine IRE1 β , composante d'une des branches de la voie UPR, régule la production de mucines secrétées et participe à l'augmentation de leur expression lors de l'inflammation (Martino *et al.*, 2013). Cette donnée renforce donc l'idée d'un lien direct entre mucines et voie UPR, surtout en situation cancéreuse.

L'ensemble de nos résultats semble donc converger vers l'hypothèse d'un rôle de coordination exercé par la galectine-3 pour la mise en place des différentes voies nécessaires à une réponse cellulaire adaptative en situation de stress du RE. Celle-ci serait impliquée dans la voie UPR elle-même, *via* la constitution d'un « **régulon** » composé d'un set d'ARNm déstabilisés et d'un set d'ARNm stabilisés codant globalement les protéines nécessaires à la mise en place, à l'activation de la voie UPR, et à son fonctionnement pendant un temps suffisant pour favoriser le retour à l'homéostasie. Les ARNm stabilisés pourraient donc, pour certains d'entre eux, être traduits en protéines « UPR » qui seraient prioritaires à la traduction et au trafic dans le RE et le Golgi. Certains ARN déstabilisés antagonistes des précédents au plan fonctionnel verraient leur traduction réduite. Mais, en confinant certains autres transcrits « stabilisés », comme ceux de *MUC4*, dans des granules où ils seraient inactifs au plan traductionnel, la galectine-3 permettrait à la cellule de faire face à ce changement temporaire de programme transcriptionnel et traductionnel, en désengorgeant le RE et l'appareil de

Golgi. La galectine-3 ne serait donc pas impliquée que dans la mise en place de la voie UPR. Les ARNm qu'elle stabilise ou déstabilise concernent également des acteurs-clés des voies concernées en aval d'UPR pour orienter la cellule vers ses destins cellulaires possibles, fonction de la durée et de l'issue de la tentative de restauration de l'homéostasie (retour à l'homéostasie, stress chronique, autophagie, apoptose). Elle pourrait donc coordonner les différentes phases de l'évolution de la réponse cellulaire au stress en participant au contrôle du métabolisme cellulaire, des différentes voies de signalisation concernées, au contrôle du cycle cellulaire et de la transcription/traduction ainsi qu'au contrôle de la communication et du trafic entre les organelles concernés (noyau, RE, Golgi, lysosomes, mitochondries).

La galectine-3 interviendrait dans la phase de réponse adaptative réversible (figure 112) par le biais de son implication dans la stabilisation de transcrits impliqués dans le retour de l'homéostasie protéique. Parmi les cibles stabilisées se trouvent des acteurs des voies ATF6 et XBP1 mais pas de la branche PERK. Or cette dernière est la voie principale induisant l'activation de l'apoptose si le stress n'est pas résolu (Tameire *et al.*, 2015). Par ailleurs, parmi les transcrits stabilisés figurent les protéines de la famille DNAJ. Ces protéines chaperonnes exercent des fonctions anti-apoptotiques en s'opposant au processus de mort cellulaire induit par le facteur CHOP (dépendant de la voie PERK) par le biais de l'inhibition de la translocation mitochondrial de Bax (Thacker *et al.*, 2013). De façon parallèle, la déstabilisation des transcrits impliqués dans la voie UPR, qui représentent 22 % des ARNm déstabilisés par la galectine-3, pourrait de façon complémentaire contribuer à réduire la quantité d'ARNm impliqués dans le passage vers la phase apoptotique.

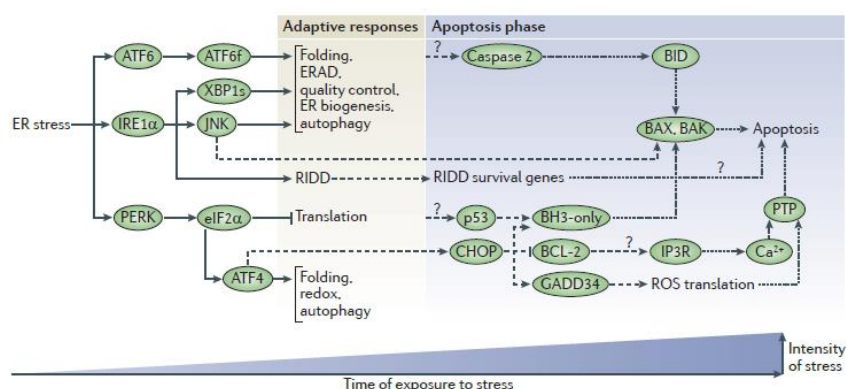


Figure 112 : Les deux phases de la voie UPR en fonction de l'intensité et de la durée d'exposition au stress (Hetz, 2012).

La voie UPR permet donc de restaurer l'homéostasie du RE et de façon plus générale, l'homéostasie cellulaire. Il s'agit d'une période transitoire qui donne à la cellule un délai lui permettant de s'orienter vers le retour à l'équilibre. Néanmoins, lorsque ce délai se prolonge et que de nouveaux signaux de stress arrivent, l'objectif de la voie UPR devient d'engager la cellule vers l'apoptose. Dans ce contexte, la galectine-3 aurait pour rôle de différer cet engagement vers l'apoptose.

Ainsi, notre étude suggère une nouvelle fonction originale pour la galectine-3 consistant en la régulation post-transcriptionnelle d'un ensemble d'ARNm partageant des fonctions communes de type « *RNA regulons* » au sein de la voie UPR mais aussi des voies qui lui sont connectées. Elle permettrait ainsi à la cellule cancéreuse de s'adapter aux stimulations intrinsèques et extrinsèques dont elle fait l'objet pour rétablir son homéostasie protéique.

Conclusions et perspectives

Ce travail nous a permis de mettre en évidence une nouvelle fonction de la galectine-3 dans le cycle de vie des ARNm qu'est la régulation de l'expression de la mucine membranaire MUC4 au niveau post-transcriptionnel. Nous avons démontré, *in vitro*, dans la lignée cancéreuse pancréatique humaine CAPAN-1, que celle-ci stabilise les transcrits de *MUC4* en potentialisant la fixation de la *RNA Binding Protein* hnRNP-L sur le motif *CA repeat* présent dans la région 3'UTR de *MUC4*. Par ailleurs, ce mécanisme est spécifique de *MUC4* et ne concerne pas les transcrits de la mucine MUC1. La régulation de l'expression de *MUC4* via hnRNP-L est également retrouvée *in vivo* dans des tissus épithéliaux digestifs murins. Par ailleurs, nos résultats indiquent que la galectine-3 interagit directement avec la protéine hnRNP-L mais faiblement avec des marqueurs de P-Bodies ou de granules de stress.

Nous avons observé que la galectine-3 se fixe aux ARNm matures de *MUC4* au niveau de la région périnucléaire, suggérant que la galectine-3 pourrait exercer un rôle plus large dans le cycle de vie des ARNm et notamment stabiliser les ARNm via leur stockage dans des granules cytoplasmiques dont le type reste à déterminer, mais autres que les P-Bodies ou les granules de stress. Cette fonction permettrait de protéger et stabiliser tout ou partie d'un répertoire de transcrits identifiés par les expériences de *whole transcriptome stability* et d'immunoprécipitation des ARNm-protéines. La question de la traduction de ces transcrits est ouverte et doit être explorée. Les analyses préliminaires semblent converger vers une fonction commune à l'ensemble de ces ARNm qu'est la mise en place de la voie UPR, réponse induite suite à un stress du réticulum, et particulièrement active dans les cellules cancéreuses et les cellules mucisécrétoires. Ainsi, la galectine-3 coordonnerait l'expression d'un « regulon » des transcrits convergeant vers cette fonction. Elle coordonnerait aussi l'expression d'autres acteurs-clés de la résolution du stress une fois la voie UPR enclenchée et ceci dans le cadre des voies et mécanismes impliqués dans ERAD, l'autophagie, l'apoptose, l'adaptation du trafic entre le réticulum et l'appareil de Golgi, le remodelage du cytosquelette et l'adaptation de la transcription et de la traduction ainsi que du cycle cellulaire et de la réparation des dommages de l'ADN. Les rôles de la galectine-3 avaient été partiellement élucidés dans un certain nombre de ces voies et mécanismes mais jamais reliés entre eux dans le cadre d'une « super-fonction » de coordination moléculaire entre des voies, mais aussi de coordination cellulaire, ces voies et ces événements moléculaires nécessitant des dialogues et des échanges entre tous les organites et compartiments subcellulaires.

Nos travaux mettent donc en évidence un nouveau rôle de la galectine-3 dans la stabilisation des ARNm de *MUC4* mais aussi sa liaison avec des transcrits matures, interaction jusqu'alors décrite dans la littérature avec des pré-ARNm. Ceci attribue donc de nouvelles fonctions à la galectine-3 dans le cycle de vie des ARNm. Par ailleurs, notre étude renforce les interconnexions largement décrites dans la littérature entre mucines et galectine-3, et ce particulièrement en situation cancéreuse.

Ce travail offre différentes perspectives de recherche :

Régulation de l'expression de *MUC1* par la galectine-3

Notre étude a démontré que l'invalidation du gène *LGALS3* s'accompagne d'une diminution de la quantité des transcrits de *MUC1*. Nos résultats indiquent que ceci n'est pas lié à une modification de l'activité du promoteur de *MUC1* sous l'influence de la galectine-3 ni à une modification de la stabilité des transcrits de cette mucine. L'hypothèse d'une modulation du profil de microARN par la galectine-3 peut être avancée. Afin d'étudier celle-ci, nous étudierons le niveau d'expression dans les cellules Sc et Sh1 de microARN décrits dans la littérature pour réguler l'expression de *MUC1* ou pour lesquels des sites de fixation sont présents dans les transcrits de cette mucine. Nous sélectionnerons ceux différentiellement exprimés les cellules Sc et Sh1 et étudierons leur rôle à l'aide la construction contenant le 3'UTR de *MUC1* cloné en vecteur luciférase. Pour cela nous effectuerons des cotransfections transitoires dans les cellules Sh1 de cette construction, du vecteur d'expression de la galectine-3 et du microARN ou de l'antimir correspondant afin d'évaluer sa fonctionnalité sur l'expression de *MUC1* en fonction de la présence ou de l'absence de galectine-3.

Stabilisation des transcrits par la galectine-3

Le croisement des données des ARNm liés (RNA-IP) et stabilisés par la galectine-3 (*whole transcriptome stability*) révèlent que seuls 15 transcrits sont à la fois stabilisés et liés par cette lectine. Ceci suggère donc que la stabilisation des ARNm par la galectine-3 est probablement indirecte, faisant par exemple appel à d'autres RBP. Afin de déterminer celles-ci, les séquences 3'UTR de l'ensemble des transcrits stabilisés par la galectine-3 seront

alignées afin d'identifier un (ou des) motif(s) commun(s). Le niveau d'expression de la RBP (ou des RBP) reconnaissant ce motif sera (seront) étudié(s) dans les cellules Sc et Sh1 et des expériences de co-immunoprécipitation permettront d'apprécier l'interaction entre la galectine-3 et cette (ces) protéine(s).

Détermination de lieu de stockage de stabilisation des ARNm par la galectine-3

Nos résultats suggèrent que la galectine-3 stabilise les transcrits dans des structures cytoplasmiques autres que les granules de stress et les P-Bodies. Afin de déterminer le type de granules impliqués, un isolement et une purification des granules contenant la galectine-3 seront effectués dans les cellules Sc en se basant sur les protocoles décrits dans la littérature (Elvira *et al.*, 2006; Jønson *et al.*, 2007). L'étude de leur protéome sera réalisée afin de déterminer leur composition et de d'évaluer si celle-ci se rapproche de granules déjà décrits dans la littérature. De plus, il sera possible d'étudier, par micro-array, les ARNm présents dans ces granules et de comparer la liste des transcrits obtenus à celle issue de l'expérience d'immuno-précipitation des complexes ARNm-protéines. Cette comparaison nous permettra de distinguer les transcrits stockés dans les granules avec la galectine-3 de ceux qui lui sont associés en dehors de ces structures. De plus, par le biais de l'identification de partenaires protéiques de la galectine-3 dans ces granules, la réalisation de co-immunofluorescence avec un anticorps anti-galectine-3 et des anticorps dirigés contre les autres protéines de ces granules permettra de déterminer la localisation cellulaire de ces structures.

Galectine-3 et réponse UPR

Nos analyses préliminaires semblent indiquer que la galectine-3 stabilise des transcrits dont les fonctions convergent vers la voie UPR. Afin d'évaluer l'impact de la galectine-3 dans cette réponse, nous évaluerons la différence d'activation de la voie UPR entre les cellules Sc et Sh1 en étudiant le niveau d'expression protéique d'acteurs de cette voie (par *Western-blot* ou par étude protéomique). Pour cela, nous étudierons l'expression de ces cibles par *Western-blot* dans les cellules Sc et Sh1 (et de façon plus exhaustive par spectrométrie de masse) afin de déterminer si la stabilisation des transcrits s'accompagne d'une augmentation de la synthèse protéique des acteurs de la voie UPR (ATF6, HSPA1A, HSPA1B, Derlin, SEL1L...). S'il s'avère que celle-ci est activée dans les cellules Sc, nous traiterons ces

cellules avec un inhibiteur de la voie UPR afin de déterminer si le niveau d'activation redevient similaire à celui-ci observé dans les cellules Sh1 après traitement. De façon intéressante, il sera également possible de traiter les cellules Sc avec l'inhibiteur de la galectine-3 et son analogue structural pour évaluer si ce traitement se traduit par une baisse de l'activation de la voie UPR. Si l'inhibition de galectine-3 s'accompagne d'une réduction du stress du réticulum, l'utilisation d'agents thérapeutiques ciblant cette galectine permettrait, en plus de diminuer l'expression de MUC4, d'atténuer cette voie dont l'activation favorise la survie des cellules cancéreuses. La galectine-3 constituerait donc une cible pour inhiber la voie UPR puisque celle-ci stabilise des transcrits potentiellement impliqués dans ce processus.

Par ailleurs, il serait intéressant d'étudier *in vivo*, par analyses immunohistochimiques dans des souris WT et KO pour la galectine-3, les niveaux d'expression d'acteurs de la voie UPR identifiés comme différentiellement exprimés entre les cellules Sc et Sh1 afin de confirmer nos résultats. Des observations préliminaires semblent indiquer des différences de morphologies du RE entre les souris WT et KO (Dr D Delacour). Cet axe nous permettra de poursuivre notre collaboration avec le Dr D Delacour.

Le niveau d'expression de ces cibles sera ensuite étudié sur des tissus d'adénocarcinome pancréatique canalaire par RT-qPCR et immunohistochimie afin d'étudier la relevance de nos données au cancer. L'ensemble de ces résultats nous permettront de confirmer l'implication de cette lectine dans la voie UPR.

Enfin, l'étude du niveau d'expression protéique des cibles stabilisées par la galectine-3 permettrait de faire la différence entre les transcrits inactifs sur le plan de la traduction et ceux dont la stabilisation s'accompagne d'une augmentation de la synthèse protéique. Cette étude permettrait de déterminer les transcrits dont la stabilisation se justifie par un besoin de traduction accrue et qui sont donc indispensables aux cellules cancéreuses dans la situation de stress du réticulum.

Appendice technique

I. Modèles d'étude

I.1. Lignées cellulaires humaines

I.1.1. Lignée cancéreuse humaine pancréatique CAPAN-1

La lignée cellulaire CAPAN-1 (ATCC HTB-79) est issue d'une métastase hépatique d'un adénocarcinome pancréatique canalaire d'un homme caucasien âgé de 40 ans. Ces cellules, qui ont une morphologie de cellules épithéliales prismatiques, poussent en monocouche (figure 113 ; Coppin, 2012). Cette lignée est mutée pour les gènes *K-ras* et *TP53* et n'exprime pas les gènes suppresseurs de tumeur *p16/CDKN2A* et *SMAD4/DPC4*.

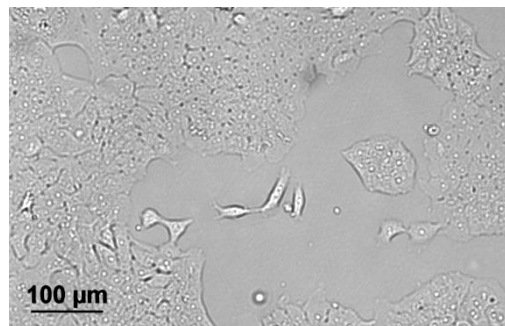


Figure 113 : Observation microscopique de cellules CAPAN-1 (grossissement x 100, Coppin, 2012).

I.1.2. Lignées cellulaires Sh1 et Sc

Les cellules Sh1 dérivent de la lignée cellulaire CAPAN-1 invalidée spécifiquement pour le gène codant la galectine-3 humaine par *short hairpin RNA* (*sh*-RNA) suite à une infection rétro-virale (Merlin *et al.*, 2011). Le *sh*-RNA utilisé est dirigé contre la partie 3' de l'exon 5 (GAAGAAAGACAGTCGGTTT) du gène *LGALS3*. Une construction *Scramble* (Sc), utilisée comme témoin négatif dans notre étude et spécifique d'aucun ARNm humain, a été déterminée après criblage de banques d'ADN génomique et de transcrits humains (AAGTCAATCAACACGGTAACA).

I.2. Tissus humains

Dix échantillons de tissus d'adénocarcinomes pancréatiques canaux et dix tissus pancréatiques sains adjacents ont été obtenus en collaboration avec la tumorotheque du CHRU de Lille (numéro d'autorisation CSTMT220 et consentement obtenu pour chaque

patient). Après ablation chirurgicale, les tissus ont été immédiatement congelés dans l'azote liquide en prévision de l'extraction des ARNm ou inclus en paraffine en vue d'une analyse immunohistochimique.

I.3. Tissus murins

Les tissus murins de jéjunum et de poumon sont issus de souris adultes de fond génétique C57BL/6, sauvages (*wild type*) ou invalidées (*knock-out*) pour le gène *LGALS3* (collaboration avec le Dr Delphine Delacour, Institut Jacques Monod, Paris, numéro d'acceptation de l'expérimentation : A75-13-17). Les tissus ont été immédiatement congelés dans l'azote liquide en prévision de l'extraction des ARNm ou inclus en paraffine en vue d'une analyse immunohistochimique.

II. Biologie cellulaire

II.1. Culture cellulaire

II.1.1. Milieux de culture

Le milieu de culture complet des cellules Sh1 et Sc dérive du milieu de culture des cellules CAPAN-1 et est constitué des éléments suivants :

- RPMI (*Roswell Park Memorial Institute medium*) 1640 (Gibco)
- Sérum de veau fœtal (SVF) (Gibco) 15 % (v/v)
- Pénicilline/streptomycine (Gibco) 1 % (v/v)
- L-glutamine (Gibco) 2 mM
- Puromycine (InvivoGen) 200 ng/ml

II.1.2. Conditions de culture

Les milieux de culture ainsi que les différents réactifs utilisés sont de qualité culture cellulaire et sont stériles. L'ensemble des manipulations est réalisée sous hotte à flux laminaire.

Les cellules sont ensemencées en flacons de culture de 25 cm² (Falcon®) une fois par semaine. Pour cela, après élimination du milieu de culture, les cellules sont rincées avec du

tampon phosphate salin 1X selon Dubelcco (D-PBS 1X, dépourvu de Ca^{2+} et de Mg^{2+}) avant l'ajout de 2 ml de trypsine-EDTA (trypsine 0,05 % (v/v), EDTA 0,5 mM). Les flacons de culture sont incubés pendant 10 min à 37 °C.

Lorsque les cellules sont décollées du support, l'ajout de 8 ml de milieu complet permet la neutralisation de la trypsine.

La suspension obtenue est placée dans un tube à centrifuger de 50 ml et centrifugée à 470 x *g* durant 4 min à température ambiante ($T^{\circ}\text{amb.}$). Après aspiration du surnageant, les cellules sont remises en suspension avec 10 ml de milieu complet.

Un comptage cellulaire est ensuite réalisé en cellule de Malassez avec du bleu de trypan à 0,05 % (m/v), puis le nombre désiré de cellules est déposé dans des flacons de culture neufs ($5,10^5$ cellules par flacon de 25 cm^2 ou $1,5.10^6$ cellules par flacon de 75 cm^2).

A la fin de l'étape d'ensemencement, les flacons sont placés dans une étuve à 37 °C en présence de 5 % de CO_2 . Le milieu de culture est ensuite renouvelé tous les jours.

II.2. Restauration de phénotype : traitement des cellules par la galectine-3 recombinante

La galectine-3 recombinante a été obtenue en collaboration avec le Dr Frédéric Frénois. Pour cela, l'ADNc correspondant aux 251 premiers aa de la galectine-3 a été cloné dans le vecteur pET15b possédant un tag Histidine (Merck Millipore). Le surnageant de culture de bactéries *Escherichia coli* BL21 pLysS (DE3) (Merck Millipore) transformées par ce vecteur a été déposé sur une résine afin de récupérer la galectine-3 recombinante. Celle-ci a ensuite été purifiée par dialyse.

Lors du traitement des cellules par de la galectine-3 recombinante les cellules sont ensemencées en plaque 24 puits ($2,5.10^4$ cellules par puits dans un volume de 500 μl de milieu de culture) une semaine avant le traitement.

Puis, 40 h avant la récupération des lysats cellulaires, les cellules sont privées de sérum de veau fœtal (SVF). Pour cela, elles sont rincées deux fois avec du D-PBS 1X puis placées dans un milieu contenant 0,01 % (m/v) de Sérumbumaine bovine (BSA), 1% (v/v) de pénicilline/streptomycine et 1 % (v/v) de L-Glutamine.

Quatre heures plus tard, 4 μM de galectine-3 recombinante sont ajoutés pour une durée de 36 h. Au cours de cette période, deux renouvellements du milieu contenant la galectine-3 recombinante sont effectués 12 h et 24 h après le début du premier traitement.

II.3. Inhibiteurs de la galectine-3

Une molécule inhibitrice de la galectine-3 ciblant le domaine CRD de la protéine (CG 0713-74, 4-O-[3-O-(3-methoxybenzyl)- β -D-galactopyranosyl]-1-azido-1,2-dideoxy-2-(3-methoxybenzoyl)- β -D-glucopyranosyl, figure 114A) et un analogue structural n'ayant pas d'effet inhibiteur sur la galectine-3 (JD 0214-104, 1,2;4',6'-di-O-(4-methoxybenzylidene-1- α -D-cellobiopyranoside, figure 114B) ont été obtenus sous forme de lyophilisat (collaboration avec le Dr Cyrille Grandjean, Faculté des Sciences et Techniques, Nantes). L'analogue structural contient un résidu glucose à la place du galactose et ses groupements hydroxyl en positions 4 et 6 sont protégés, ce qui entraîne son absence de liaison à la galectine-3.

Avant utilisation, ces molécules sont reconstituées à une concentration de 20 mM dans de l'eau stérile.

Pour effectuer le traitement, les cellules sontensemencées en flacon de culture de 25 cm² et cultivées jusqu'à confluence (7 jours de culture).

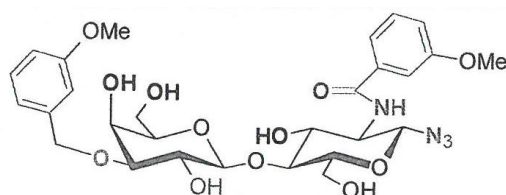
Avant l'ajout des molécules, elles sont sevrées pendant 2 h 30 dans un milieu dépourvu de SVF contenant 0,01% de BSA (m/v).

Un prétraitement par de l'actinomycine D (8 μ M, antibiotique inhibiteur de l'ARN polymérase, Sigma Aldrich) est réalisé 30 min avant l'addition des molécules.

Ensuite, l'inhibiteur (ou l'analogue structural) utilisé à une concentration finale de 200 μ M, est additionné dans le milieu, toujours en présence d'actinomycine D.

Après 6 h d'incubation, les ARN sont extraits afin de quantifier l'expression de *MUC4* par RT-qPCR.

CG 0713-74 : inhibiteur de la galectine-3



JD 0214-104 : analogue structural

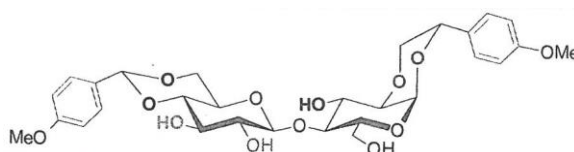


Figure 114 : Structures de l'inhibiteur de la galectine-3 (CG 0713-74) et de l'analogue structural (JD 0214-104) (collaboration avec le Dr Cyrille Grandjean).

II.4. Etude des cellules en phase G2/M et cytométrie en flux

Afin de bloquer les cellules en phase G2/M, celles-ci sont traitées par du nocodazole pendant 16 h (400 ng/ml (Dawkins *et al.*, 2016), agent anti-mitotique stimulant la dépolymérisation des microtubules, Sigma Aldrich).

Puis, les cellules sont décollées du support de culture avec de la trypsine et lavées deux fois par du D-PBS 1X. Pour chaque condition d'analyse, 1 million de cellules est mis en suspension dans 200 µl de D-PBS puis 1 ml d'éthanol à 70 % glacé est ajouté.

Après 10 min d'incubation à 4 °C et une centrifugation de 10 min à 300 x *g*, le culot cellulaire est lavé par 2 ml de D-PBS 1X puis repris par 100 µl de D-PBS 1X. Cette suspension est mise en contact avec 100 µl de RNase A (Promega, 4 mg/ml) pendant 15 min à T° amb.

Puis, après l'ajout d'1 ml de Iodure de propidium-Triton (25 µg/ml-0,1 %, Sigma Aldrich), une incubation de 30 min est effectuée à T° amb. à l'abri de la lumière.

Les cellules sont ensuite analysées par cytométrie en flux sur le *CyAn-ADP*TM (Beckman Coulter) de la Plate-forme BICeL (Lille) en fonction de leur diploïdie (2N ou 4N) à l'aide du logiciel *Summit version 4.3* (Beckman Coulter).

II.5. Transfections transitoires

II.5.1. Transfections transitoires à l'Effectene[®]

Afin de réaliser des transfections cellulaires transitoires avec l'Effectene[®] (Qiagen), les cellules sontensemencées la veille de la manipulation dans des plaques à 6 puits (4.10⁵ cellules par puits dans 2 ml de milieu de culture)

Trois témoins sont utilisés lors de chacune des transfections (Coppin, 2012) :

- le vecteur pGL3 contrôle (contenant l'*enhancer* et le promoteur du virus SV40) est utilisé pour s'assurer de l'efficacité de la transfection ;
- le vecteur pGL3 basique, dépourvu de séquence promotrice ou d'*enhancer* eucaryotes, permet de déterminer l'activité basale du vecteur vide ;
- une transfection sans ADN plasmique est utilisée comme témoin négatif.

Au cours de chacune des manipulations, chaque construction de promoteur est testée trois fois.

Pour ce faire, 1 µg de la construction promotrice et 0,2 µg du (ou des) vecteur(s) d'expression sont mélangés dans un microtube de 1,5 ml avec du tampon EC (tampon de condensation, Qiagen) pour obtenir un volume final de 100 µl.

Puis, la condensation de l'ADN est réalisée grâce à l'ajout de 3,2 µl d'*Enhancer*. L'ensemble est incubé 5 min à T°amb. et centrifugé brièvement avant l'apport de 5 µl d'Effectene®. Ce mélange est homogénéisé puis incubé 15 min à T°amb.

Parallèlement, après aspiration du milieu de culture, les cellules sont rincées avec du D-PBS 1X puis 1,6 ml de milieu neuf complet sont ajoutés dans chaque puits.

A la fin de l'incubation, le mélange ADN/*Enhancer*/Effectene est complété avec 0,6 ml de milieu neuf complet puis homogénéisé et transféré goutte à goutte sur les cellules.

A l'issue de cette étape, les plaques sont inclinées de droite à gauche et d'avant en arrière afin de répartir le mélange sur le tapis cellulaire avant de les incubé pendant 48 h à 37 °C. Les cellules sont ensuite lysées pour permettre l'extraction des ARNm ou des protéines.

II.5.2. Transfections transitoires par des siARN

Les séquences du *pool* de siARN ciblant hnRNP-L (Dharmacon®, GE Healthcare) sont les suivantes (5'→3') : GCGAUGAGCUGGGAGUGAA, UGAAAGUAUUCUCAGGCAA, GUUUGUAGAGGCUUACUUA et UAAGAUGAACUGUGACCGA.

Les conditions contrôles sont réalisées à l'aide d'un *pool* de siARN « non-targeting » (siControl Dharmacon®) constitué des oligonucléotides suivants (5'→3') : UGGUUUACAUGUCGACUAA, UGGUUUACAUGUUGUGUGA, UGGUUUACAUGUUUUCUGA et UGGUUUACAUGUUUUCCUA.

Les siARN sont repris à une concentration de 20 µM (solution mère) avec du tampon *siRNA buffer* 1X (Dharmacon®) et sont dilués au ¼ dans ce même tampon extemporanément pour la réalisation des transfections.

Les transfections sont effectuées dans des plaques 6 puitsensemencées avec 4.10⁵ cellules par puits 24 h avant la manipulation de façon à obtenir une confluence de 60 % le jour des transfections.

Dans un premier tube, 10 µl d'une dilution de siARN (25 nM) et dans un second tube 4 µl de l'agent de transfection DharmaFECT 2® (Dharmacon®) sont respectivement mélangés

à 190 µl et 196 µl de milieu de culture dépourvu de SVF et d'antibiotique (Pénicilline/Streptomycine).

Ces deux solutions sont homogénéisées par tapotement et incubées 5 min à T°amb. avant d'être regroupées et placées 20 min à T°amb.

Par ailleurs, les cellules sont rincées par du D-PBS 1X avant de recevoir le mélange siARN/DharmaFECT 2® qui est au préalable complété par 1,6 ml de milieu de culture contenant 15 % (v/v) de SVF.

Les plaques sont agitées de droite à gauche et d'avant en arrière afin de permettre une répartition homogène du mélange sur les cellules et sont placées pendant 48 h à 37 °C avant de réaliser la lyse des cellules.

II.5.3. Transfections transitoires par des oligonucléotides antisens

Les séquences des oligonucleotides antisens (Eurogentec) ont été choisies afin de cibler la région *CA repeat* présente dans le 3'UTR de *MUC4* (AS 1: 5' TCGTGTGTGTGTGTGTGTGTGCAC 3') ou une région du 3'UTR de *MUC4* dépourvue d'élément régulateur (AS 2: 5' CACAGGCTAGTGTCTTCTGTGGGT 3'). Ces oligonucléotides sont repris à une concentration de 100 µM (solution mère) dans de l'eau stérile et sont utilisés à la concentration finale de 1 µM.

Les transfections sont réalisées avec le réactif « oligofectamine Transfection Reagent » (Thermo Fisher Scientific) dans des plaques 6 puitsensemencées avec $2,5 \cdot 10^5$ cellules par puits 24 h avant les transfections de façon à obtenir une confluence à 40 % le jour de la manipulation.

Dans un premier tube, 15 µl d'oligonucléotides antisens et dans un second tube 4 µl de l'agent de transfection Oligofectamine® sont mélangés respectivement à 170 µl et 11 µl d'Opti-MEM™ I Reduced Serum Media (Thermo Fisher Scientific).

Ces deux solutions sont homogénéisées par tapotement et incubés 10 min à T°amb. avant d'être regroupées et placées 20 min à T°amb.

Les cellules sont rincées par 1 ml de milieu dépourvu de SVF et d'antibiotique puis 800 µl de milieu sans SVF et sans antibiotique sont ajoutés dans chaque puits.

Les 200 µl résultants du mélange des deux tubes sont déposés sur les cellules et les plaques sont agitées d'avant en arrière et de droite à gauche pour permettre une répartition homogène du mélange sur les cellules et sont placées 4 h à 37 °C.

Après cette incubation, 500 µl de milieu de culture sans antibiotique mais contenant 3 fois plus de SVF que le milieu complet sont ajoutés dans chaque puits. Les plaques sont de nouveau placées à 37 °C pour une durée de 48 h avant de réaliser l'extraction des ARN.

II.6. Imagerie

II.6.1. Hybridation in situ couplée à de l'immunofluorescence (FISH)

Les cellules sontensemencées dans des chambres de culture Lab-Tek I à 8 puits (Dutscher) ($3,5 \cdot 10^4$ cellules par lame) et sont maintenues en culture une semaine.

Les lames sont ensuite récupérées et les cellules sont lavées 3 fois sur glace avec du D-PBS 1X contenant du Ca^{2+} et du Mg^{2+} avant d'être fixées pendant 20 min dans du paraformaldéhyde (PFA) à 4 % (m/v) dans du D-PBS 1X (Affymetrix).

A la fin de cette étape, les cellules sont lavées une fois avec du D-PBS 1X et le PFA est neutralisé en ajoutant 200 µl de NH_4Cl 50 mM (dilué dans du D-PBS 1X Ca^{2+} - Mg^{2+}) pendant 20 min à T° amb.

Puis, les cellules sont perméabilisées pendant 20 min dans une solution de Saponine à 0,2 % (m/v), Triton X-100® 0,1 % (v/v) diluée dans du D-PBS 1X, suivi d'un lavage avec du D-PBS 1X Ca^{2+} - Mg^{2+} .

Ensuite, les sites aspécifiques sont saturés pendant 20 min en ajoutant une solution de Saponine-Triton X-100® additionnée de 1% (m/v) de BSA.

Afin de vérifier la spécificité de la sonde ARN (contrôle négatif), un traitement par de la RNase A (Promega, 4 mg/ml) peut être effectué afin de dégrader les ARN. Pour cela, la RNase A est diluée dans du D-PBS 1X Ca^{2+} - Mg^{2+} au 1/40 et ajoutée sur les cellules pendant 30 min à 37 °C. Puis les cellules sont incubées 5 min avec 200 µl de Tris-HCl 1 M pH 8.

Les cellules sont ensuite incubées avec la sonde marquée par le fluorophore cy3.5 (pendant 1 h à 37 °C pour la sonde oligodT ou pendant 4 h à 42 °C pour la sonde spécifique des ARNm de *MUC4*). Celles-ci sont utilisées à une concentration de 1,3 µg/ml dans le tampon suivant :

- Solution saline de citrate de sodium 2X (SSC) (NaCl 0,3M, citrate de sodium 0,03M)
- BSA 0,005 % (m/v) (Sigma Aldrich)
- ARN de levure 1 mg/ml (Sigma Aldrich)
- Sulfate de dextran 10 % (v/v) (Merck Millipore)
- Formamide 25 % (v/v) (Merck Millipore)

Les séquences des sondes sont les suivantes (5'→3') (Vandenhoute *et al.*, 1997) :

MUC4 : Cy3.5-GTCGGTGACAGGAAGAGGGGTGGCGTGACCTGTGGATGCTGAGGAAGT

oligo dT : Cy3.5-TT

A l'issue de l'hybridation, les cellules sont rincées successivement 2 fois par du SSC 4X et 2 fois par du SSC 2X (5 min par lavage).

Elles sont ensuite incubées avec l'anticorps primaire polyclonal anti-galectine-3 pendant une nuit à l'abri de la lumière (Abcam ab31707, dilution au 1/150 dans du tampon SSC 2X avec 0,2 % de saponine (m/v) et 1 % de BSA (m/v)).

Le lendemain, les cellules sont rincées 3 fois 5 min avec du tampon SSC 2X puis elles sont incubées avec l'anticorps secondaire anti-lapin pendant 45 min à T° amb. (Alexa fluor[®] 488, (Thermo Fisher Scientific), dilution au 1/150 dans du SSC 2X contenant 0,2 % (m/v) de saponine et 1 % (m/v) de BSA).

A la fin de cette incubation, 3 lavages de 5 min avec du tampon SSC 2X sont effectués et les lames sont montées avec 70 µl de Vectashield[®] (Vector) contenant du DAPI afin de marquer les noyaux. Les lames sont conservées à 4 °C à l'abri de la lumière (Burgess *et al.*, 2011).

Celles-ci sont analysées en mode séquentiel à l'aide du microscope confocal LSM 710 (Carl Zeiss) (Plate-forme BICeL, Lille) et du logiciel Zeiss Efficent Navigation. Les clichés sont pris au grossissement x40 ou x63.

II.6.2. Immunofluorescence ou Co-immunofluorescence

Pour l'analyse par immunofluorescence (IF) ou co-IF, le protocole est le même que décrit au paragraphe II.6.1 en dehors du fait que le traitement à la RNase A et le marquage par la sonde ne sont pas effectués. De plus, les dilutions des anticorps primaires et secondaires sont réalisées dans du D-PBS 1X Ca²⁺-Mg²⁺ et non pas dans du SSC 2X. Le tableau XXII reprend la liste des anticorps primaires utilisés dans notre étude. Les anticorps secondaires (AlexaFluor[®] 488 poluet anti-lapin, AlexaFluor[®] 488 chèvre anti-souris and AlexaFluor[®] 633 poulet anti-chèvre (Thermo Fisher Scientific)) sont utilisés à une dilution de 1/500 et choisis en fonction de l'espèce de l'anticorps primaire.

Les lames sont analysées en mode séquentiel à l'aide du microscope confocal LSM 710 (Carl Zeiss) (Plate-forme BICeL, Lille) et du logiciel Zeiss Efficent Navigation. Les clichés sont pris au grossissement x40 ou x63.

Tableau XXII : Anticorps primaires et dilutions utilisés dans les expériences d'imagerie.

Anticorps (référence)	Dilution
Galectine-3 (Abcam, ab31707)	1/150
GW182 (Abcam, ab70522)	1/200
TIA-1 (Santa Cruz, sc-1751)	1/200
Xrn-1 (Abcam, ab70259)	1/200
eIF2 α (Enzo, ADI KAP C130D)	1/200
DDX6 (Bethyl A300-461A)	1/200
G3BP1 (BD Biosciences 611126)	1/200
hnRNP-L (Abcam, ab6106)	1/200
MUC4 (Santa Cruz, sc-53945)	1/200
ErbB2 (Santa Cruz, sc-284)	1/200
ZEB-1 (Cell signaling 3396)	1/200

II.6.3. Proximity Ligation Assay

Les réactifs « Duolink[®] In Situ Reagents » (Sigma Aldrich) sont utilisés pour la réalisation des expériences de *Proximity Ligation Assay* (PLA).

Les cellules sontensemencées dans des chambres de culture Lab-Tek I à 8 puits (Dutscher) (3,5. 10⁴ cellules par lame) et sont maintenues en culture sont maintenues en culture une semaine.

Les lames sont ensuite récupérées et les cellules sont lavées 3 fois sur glace avec du D-PBS 1X contenant du Ca²⁺ et du Mg²⁺ avant d'être fixées pendant 20 min dans du paraformaldéhyde (PFA) à 4 % (Affymetrix).

A la fin de cette étape, les cellules sont lavées une fois avec du D-PBS 1X et le PFA est neutralisé en ajoutant 200 μ l de NH₄Cl 50 mM (dilué dans du D-PBS 1X Ca²⁺-Mg²⁺) pendant 20 min à T° amb.

Ensuite, les sites aspécifiques sont saturés pendant 20 min en ajoutant une solution de saponine (0,2 % (m/v))-BSA (3% (m/v)) diluée dans du D-PBS 1X Ca²⁺-Mg²⁺.

Les cellules sont incubées la nuit avec les anticorps primaires dilués dans du D-PBS 1X Ca²⁺-Mg²⁺ contenant 3% (m/v) de BSA et 0,2 % (m/v) de saponine (tableau XXII).

Le lendemain, les lames sont lavées 3 fois avec du D-PBS 1X. Les anticorps secondaires utilisés (dilués au 1/5 dans la solution *Antibody Diluent*[®]) sont couplés à une sonde d'ADN soit positive soit négative. L'anticorps secondaire anti-lapin couplé à la sonde Duolink[®] In Situ PLA[®] Anti-lapin PLUS est incubé soit avec l'anticorps secondaire anti-souris couplé à la sonde Duolink[®] In Situ PLA[®] Probe MINUS, soit avec l'anticorps secondaire anti-chèvre couplé la sonde Duolink[®] In Situ PLA[®] Anti-chèvre MINUS pendant 1 h à 37 °C.

Les lames sont ensuite lavées 2 fois 5 min avec du tampon de lavage A (Tris-HCl 0,01 M, pH 7,4, NaCl 0,15 M, Tween[®]20 0,05 % (v/v)). Puis, la ligase (1 U/μl) diluée au 1/40 dans la solution de ligation est ajoutée pendant 30 min à 37 °. Les lames sont lavées 2 fois 2 min par du tampon de lavage A.

Ensuite, les lames sont incubées 100 min à 37 °C à l'abri de la lumière en présence de la polymérase (10 U/μl, dilution au 1/80 dans la solution d'amplification) puis rincées 2 fois pendant 10 min avec du tampon B (Tris-HCl 0,2 M, pH 7,5, NaCl 0,1 M) et une fois 1 min par le tampon B dilué au 1/100 dans du D-PBS 1X, toujours à l'abri de la lumière.

Un marquage de la F-actine est réalisé par la phalloïdine-FITC (Thermo Fisher Scientific) pendant 20 min à T°amb.

Les lames sont ensuite séchées à température ambiante à l'obscurité puis montées avec le liquide de montage (Vectashield[®]) dans lequel se trouve du DAPI.

Les lames sont analysées en mode séquentiel à l'aide du microscope confocal LSM 710 (Carl Zeiss) (Plate-forme BICeL, Lille) (λ excitation = 598 nm et λ émission = 634 nm) et du logiciel *Zeiss Efficient Navigation*. Les clichés sont pris au grossissement x40 ou x63.

La moyenne du nombre de spots de colocalisation présents dans chaque cellule est calculée sur un minimum de 13 cellules avec le module de comptage de cellules du logiciel ImageJ.

II.7. Immunohistochimie

Les lames sur lesquelles sont fixés les tissus murins sont déparaffinées par incubations successives dans des bains de xylène (2 fois 5 min), d'éthanol 100 % (2 fois 5 min), d'éthanol 95 % (2 fois 5 min) et d'eau (1 fois 5 min).

Les sites antigéniques sont démasqués en chauffant les lames au micro-onde (1 min 30 s pour hnrnp-l, 20 min pour Muc4) dans du tampon citrate pH 6 (Dako).

Les peroxydases endogènes sont ensuite désactivées dans un bain de peroxyde d'hydrogène à 3% (v/v, Sigma Aldrich) pendant 30 min et les lames sont rincées 2 fois 5 min avec du D-PBS 1X.

Une étape de perméabilisation est réalisée avec le tampon Teng-T 1X (Tris-HCl 10 mM, pH 8, EDTA 5 mM, NaCl 150 mM, gélatine 0,25 % (m/v), Tween[®] 20 0,01 % (v/v)) pendant 30 min à T°amb. puis 2 lavages de 5 min sont effectués avec du D-PBS 1X. Les lames sont ensuite placées 1 h à température ambiante dans le réactif de blocage (M.O.M Mouse Ig blocking reagent, Vector Laboratories) afin de réaliser la saturation puis sont rincées 2 fois pendant 5 min avec du D-PBS 1X.

L'anticorps primaire est dilué dans du diluent M.O.M (Vector Laboratories) contenant 0,1 % (v/v) de Triton X-100[®] (hnRNP-L : Abcam ab6106, 1/100 et Muc4 : Santa Cruz sc-33654, 1/200). Les lames sont mises au contact de l'anticorps primaire à T°amb. (1 h pour hnRNP-L, 30 min pour Muc4) puis lavées 4 fois 5 min avec du D-PBS 1X.

L'anticorps secondaire (M.O.M biotinylated anti-Mouse IgG reagent, Vector Laboratories) est ajouté pendant 10 min à T°amb. puis les lames sont lavées 4 fois 5 min avec du D-PBS 1X.

Les lames sont incubées pendant 5 min à T°amb. avec un complexe avidine-biotine couplé à la peroxydase de Raifort permettant l'amplification du signal (M.O.M Vectastain Elite). Les lames sont ensuite lavées 6 fois pendant 5 min avec du D-PBS 1X.

Les lames sont incubées avec le 3,3'diaminobenzidine (DAB, Sigma Aldrich) (3-20 min pour hnRNP-L, 2-3 min pour Muc4) puis une étape de lavage par de l'eau déminéralisée permet d'arrêter la réaction.

La contre-coloration est effectuée en plaçant les lames dans l'hématoxyline pendant 20 s puis dans une solution de carbonate de lithium pendant 2 s et celles-ci sont rincées à l'eau.

Les tissus sont déshydratés avant le montage des lames par une incubation dans l'éthanol à 95 % pendant 5 min, puis 2 fois dans l'éthanol absolu pendant 5 min et enfin 2 fois pendant 5 min dans le xylène.

Les lames sont montées avec une lamelle de verre avec le liquide de montage Pertex (Mounting medium for light microscopy, Histolab).

III. Biologie moléculaire

III.1. Extraction des ARN totaux

III.1.1. Extraction des ARN totaux à partir de culture cellulaire

L'extraction des ARNm est effectuée avec le kit « NucleoSpin® RNA II » (Macherey Nagel) sur des cellules post-confluentes (confluence + 1 jour) cultivées en flacon de culture de 25 cm².

Pour ce faire, les cellules sont lysées par 350 µl du tampon de lyse RA1 (isothiocyanate de guanidine) contenant 3,5 µl de β-mercaptoéthanol (1 %).

Le lysat obtenu est transféré dans une colonne « NucleoSpin® Filter » qui est centrifugée 1 min à 11 000 x g.

Ensuite, 350 µl d'éthanol à 70 % sont ajoutés à l'éluat et le mélange est déposé dans une colonne « NucleoSpin® RNA II » pour être centrifugé 1 min à 11 000 x g.

Après élimination de l'éluat, 350 µl de tampon de lavage MBD sont additionnés et la colonne est centrifugée 1 min à 11 000 x g.

L'éluat est éliminé et 95 µl du mélange constitué de 10 µl de rDNase et de 90 µl du tampon de l'enzyme sont placés sur le filtre pour une durée de 15 min à T°amb.

Après l'ajout de 200 µl de tampon de lavage RA2, la colonne est centrifugée 30 s à 11 000 x g.

Ensuite, 600 µl de tampon de lavage RA3 sont déposés sur la colonne et celle-ci est centrifugée 30 s à 11 000 x g. Un nouveau lavage par 250 µl de tampon RA3 est effectué suivi d'une centrifugation pendant 2 min à 11 000 x g.

La colonne est ensuite centrifugée à vide pendant 1 min à 11 000 x g afin d'éliminer les traces d'éthanol.

Après cette étape, la colonne est placée dans un microtube neuf et 60 µl d'eau *RNase free* sont déposés sur la membrane. Une centrifugation de 1 min à 11 000 x g permet de collecter l'éluat. Les ARN ainsi obtenus sont conservés à -80 °C.

La mesure de la densité optique des échantillons avec le spectrophotomètre NanoDrop® permet de déterminer la concentration des échantillons. La mesure des DO à 230, 260, et 280 nm est réalisée afin d'évaluer la pureté des échantillons par le biais du calcul des rapports suivants : DO 260 nm/DO 280 nm (contamination par des protéines) et DO 260 nm/DO 230 nm (contamination par des sels ou des solvants).

III.1.2. Extraction des ARN totaux à partir de tissus

Les ARN de tissus sont extraits avec kit « NucleoSpin® RNA II » (Macherey Nagel) comme décrit au paragraphe III.2.1. La seule différence réside dans le fait que le morceau de tissu est préalablement broyé manuellement à l'aide d'un petit pilon dans un microtube de 1,5 ml contenant 500 µL de RA1 et 5 µL de β-mercaptoéthanol.

III.2. Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)

III.2.1. Reverse Transcription : synthèse de l'ADNc

Afin de réaliser la synthèse de l'ADNc, 1 µg d'ARN total est mis en présence d'1 µl d'amorce oligo(dT)₁₅ (Promega) dans un volume final d'eau de 14 µl. L'ensemble est incubé 5 min à 70 °C puis refroidit 5 min sur glace.

Ensuite, le mélange réactionnel suivant est additionné :

- dNTPs (20 mM, Eurogentec)	5 µl
- Tampon M-MLV RT 5X (Tris-HCl 50 mM, pH 8,3, KCl 75 mM, MgCl ₂ 3 mM, DTT 10 mM, Promega)	5 µl
- Reverse transcriptase M-MLV (200 U, Promega)	1 µl
- Eau stérile	14 µl

Le tube est placé 1 h à 42 °C puis 5 min à 94 °C et enfin 5 min sur glace. L'ADNc est conservé à - 80°C.

III.2.2. Amplification de l'ADNc par PCR

La réaction d'amplification par PCR semi-quantitative est réalisée avec le mélange suivant :

- Tampon enzyme 10X (Tris-HCl 100 mM, pH 8,3, KCl 500 mM, MgCl ₂ 15 mM, Roche Diagnostic)	2,5 µl
- dNTP (20 mM, Eurogentec)	1 µl
- Amorce sens (5 µM, MWG ou Eurogentec)	1 µl
- Amorce anti-sens (5 µM, MWG ou Eurogentec)	1 µl
- Taq DNA Polymérase (5 U/µl, Roche Diagnostic)	1 µl
- ADNc	2,5 µl
- Eau sterile	13,5 µl

Les transcrits 28S sont utilisés comme contrôle interne. Le tableau XXIII reprend les paires d'oligonucléotides spécifiques de chaque gène étudié.

L'amplification est effectuée dans le thermocycleur « MJ MiniTM » (BioRad) avec les conditions suivantes :

Cycles :

MUC1/MUC4/hnRNP-L/28S : 94 °C, 2 min ; 30 cycles [94 °C, 45 s ; 60 °C, 1 min ; 72 °C, 1 min] ; 72°C, 10 min puis 4 °C à l'infini.

LGALS3 : 94 °C, 4 min ; 33 cycles [94 °C, 1 min ; 60 °C, 1 min ; 72 °C, 2 min] ; 72 °C, 10 min puis 4 °C à l'infini.

Une électrophorèse en gel d'agarose 1 % (m/v) est réalisée dans du tampon TBE 1X (Tris 89 mM, acide borique 89 mM, pH 8,3, EDTA 2 mM) en présence de 0,7 µg/ml de bromure d'éthidium (BET) pour visualiser les amplicons.

L'utilisation du logiciel ImageQuanTL (GE Healthcare) permet la quantification des bandes obtenues. L'intensité du signal détecté est rapporté à celui de l'ARN 28S afin de normaliser la quantité des transcrits d'intérêt présents dans l'échantillon.

Tableau XXIII : Séquences des oligonucléotides utilisés en RT-PCR.

Gène	Taille de l'amplicon (pb)		Amorces (5'→3')	Température d'hybridation (T _m , °C)	Nombre de cycles d'amplification
MUC4	596	sens antisens	CGCGGTGGTGGAGGCGTTCTT GAAGAATCCTGACAGCCTTCA	60	30
MUC1	447	sens antisens	GAACTACGGGCAGCTGGACATC GCTCTCTGGGCCAGTCCTCCTG	60	30
LGALS3	719	sens antisens	ATGGCAGACAATTTTTCGCTCC ATGTCACCAGAAATTCCCAGTT	60	33
hnRNP-L	185	sens antisens	TTCTGCTTATATGGCAATGTGG GACTGACCAGGCATGATGG	58	35
28S	231	sens antisens	GCAGGGCGAAGCAGAAGGAAACT TGAGATCGTTTCGGCCCCAA	60	30

III.1. Extraction des ARN et Reverse transcription spécifique lors de la réversion de phénotype

III.1.1. Récupération des cellules

Les cellules sont décollées du support de culture par trypsinisation. Le culot cellulaire obtenu est lavé deux fois par du D-PBS 1X avant d'être remis en suspension dans 200 µl de D-PBS 1X froid.

Un comptage des cellules en lame de Malassez est effectué et un ajustement du volume de D-PBS 1X est réalisé de façon à obtenir une concentration de $7,5 \cdot 10^3$ cellules par microlitre.

III.1.2. Extraction des ARN

L'extraction des ARNm est effectuée avec le kit « SuperScript III CellsDirect cDNA Synthesis System » (Invitrogen).

Le mélange réactionnel suivant est préparé dans un microtube de 0,2 ml :

- $7,5 \cdot 10^3$ cellules	1 µl
- Tampon de resuspension	10 µl
- RNaseOUT® (40 U/µl)	1 µl

Le mélange est incubé 10 min à 75 °C. Ensuite, 5 µl de DNase I® et 1,6 µl de tampon 10X de la DNase I sont additionnés pour d'éliminer les traces d'ADN et le tube est placé 5 min à T°amb. Puis, 1,2 µl d'EDTA 25 mM sont ajoutés suivi d'une incubation de 5 min à 70 °C pour d'arrêter la réaction.

III.1.3. Synthèse de l'ADNc

Afin d'effectuer la synthèse de l'ADNc, les réactifs suivants sont ajoutés au mélange précédent :

- dNTPs (10 mM)	1 µl
- oligo(dT) ₂₀ (50 mM)	2 µl

Le tube est placé 5 min à 70 °C puis 2 min sur glace.

Ensuite sont ajoutés :

- Tampon RT 5X	6 µl
- RNaseOUT [®] (40 U/µl)	1 µl
- Enzyme SuperScript III RT [®] (200 U/µl)	1 µl
- DTT 0,1 M	1 µl

Le tube est incubé 50 min à 50 °C puis 5 min à 85 °C pour inactiver la réaction. Puis 1 µl de l'enzyme RNase H[®] (2 U/µl) est ajouté et l'ensemble est placé 20 min à 37 °C. L'ADNc obtenu est conservé à -20 °C.

III.2. PCR quantitative en temps réel (qPCR)

III.2.1. Quantification avec la technique TaqMan[®]

a. Quantification absolue

Lors de l'utilisation de la technologie TaqMan, une quantification absolue des transcrits de *MUC1*, *MUC4*, *MUC16* et de *LGALS3* est réalisée par l'établissement d'une droite de calibration à partir de dilutions des plasmides (Coppin, 2012). En effet, un fragment de *MUC4*, *MUC1*, *MUC16* ou de *LGALS3* reconnu par les sondes utilisées dans la technique de qPCR a été sous-cloné en vecteur pCR[®]2.1 (TA cloning Kit, invitrogen). Les amorces et sondes utilisées sont présentées dans le tableau XXIV.

L'ARN 18S est utilisé comme standard interne. Il est co-amplifié au cours de la même réaction de PCR à l'aide des réactifs du kit « Ribosomal RNA Control Reagents » (Applied Biosystems). Les amorces et sondes des gènes cibles à étudier sont utilisées à la concentration de 10 µM (Eurogentec).

Le mélange utilisé est le suivant :

- Amorce sens	1 µl
- Amorce antisens	1 µl
- Sonde	1 µl
- Amorce 18S sens (1 µM)	0,375 µl
- Amorce 18S antisens (1 µM)	0,375 µl
- Sonde 18S (40 µM)	0,125 µl
- Absolute TM QPCR ROX mix (Abgene)	12,5 µl
- Eau stérile	6,625 µl
- ADNc	2 µl

L'amplification est réalisée dans le thermocycleur « ABI 7900 » (Applied Biosystems). Chaque Ct obtenu est converti en nanogramme à l'aide de la droite d'étalonnage. Enfin, pour chacun des échantillons, le ratio entre la quantité de la cible d'intérêt et la quantité de 18S est calculé.

Tableau XXIV : Sondes et amorces utilisées en qPCR TaqMan[®].

Gène		Amorces (5'→3')
MUC4	sens	TCAGCTGAGGCCTTGCCTT
	antisens	TCAGTCACCTTCCCTTTTCCA
	sonde	FAM-TAAGGCGCCATTGCTTTTGGGAGA-Tamra
MUC1	sens	CAGACGTCAGCGTGAGTGAGTGATG
	antisens	CTGACAGACAGCCAAGGCAAT
	sonde	FAM-TGCTGGTCTGTGTTCTGGTTGCGCT-Tamra
MUC16	sens	CAGTCAACTACATGACACATT
	antisens	ACTCTGTCTACTCTCCGAGCC
	sonde	FAM-GTGATGGTGAAATTCAGGTAGA-TAMRA
LGALS3	sens	GCCACTGGCCCCTATGG
	antisens	CCCAGGCAAAGGCAGGTTATA
	sonde	FAM-AGGCACAATCAGTGGC-Tamra
GAPDH	sens	CCACATCGCTCAGACACCAT
	antisens	CCAGGCGCCCAATAC
	sonde	FAM-AAGGTGAAGGTCGGAGTCAACGGATTTG-Tamra

b. Quantification relative

Quantification des transcrits *CCDC64*, *MAPK1IP1L*, *PSMD3*, *KRT19*, *ACTR3*, *GFPT1*, *PITX1*, *STC2*, *EZR*

Ces transcrits ont été quantifiés en utilisant les transcrits *18S* comme référence interne avec une quantification relative à l'aide de la méthode des delta de Ct (ΔCt) puisque nous ne disposons pas des plasmides permettant d'établir la gamme d'étalonnage pour les transcrits d'intérêt. La liste des amorces et sondes utilisées dans cette situation est présentée dans le tableau XXV.

La technique de quantification relative des transcrits fait appel aux calculs suivants :

- la différence ΔCt entre les valeurs de Ct du gène cible (par exemple *MUC4*) et du gène de ménage (*GAPDH*) est calculée pour l'échantillon à analyser (échantillon) et le contrôle (contrôle) :

$$\Delta Ct (\text{échantillon}) = Ct (\text{cible}_{\text{échantillon}}) - Ct (\text{ménage}_{\text{échantillon}})$$

$$\Delta Ct (\text{contrôle}) = Ct (\text{cible}_{\text{contrôle}}) - Ct (\text{ménage}_{\text{contrôle}})$$

- puis la différence entre les delta de Ct entre le contrôle et l'échantillon est déterminée de la façon suivante: $\Delta \Delta Ct = \Delta Ct (\text{échantillon}) - \Delta Ct (\text{contrôle})$

- un ratio d'expression normalisé d'un gène cible est ensuite calculé à l'aide de la formule : $2^{-\Delta \Delta Ct}$.

Tableau XXV: Sondes et amorces utilisées en qPCR TaqMan®.

Gène		Amorces (5'→3')
CCDC64	sens	AGAGGAACCAGGACATGAGC
	antisens	CTCCCCTTCTCGGTTCTCAA
	sonde	FAM-AGCTCGAGCACTTAGAGCAA-Tamra
MAPK1IP1L	sens	GACCAGCACCAACAGGAATG
	antisens	AGTTGGGGCTGGATAAGGAC
	sonde	FAM-CTTTCCTCCTTCCGGACCAT-Tamra
PSMD3	sens	GTGTCATTGAGGCCAGCATC
	antisens	GGCTCTCGGGTGGAAATAGAT
	sonde	FAM-ACCACGAGAAGGGCTATGTC-Tamra
KRT19	sens	GGCGAGCTAGAGGTGAAGAT
	antisens	TGTCGATCTGCAGGACAATC
	sonde	FAM-ACAAGATTCTTGGTGCCACC-Tamra
ACTR3	sens	GTGGCACGGGGTATACAAAA
	antisens	TTCATCACCCCTCCTTTGAGC
	sonde	FAM-TCAGCAAAAGTGGGTGATCA-Tamra
GFPT1	sens	AGCCCTCTGTTGATTGGTGT
	antisens	CACACGAGAGAGATTGCAGC
	sonde	FAM-ACAGAACAGGCAAAGACAAGA-Tamra
PITX1	sens	CCAAGAAGAAGAAGCAGCGG
	antisens	TTCTTGAACCAGACCCGCA
	sonde	FAM-ACTTCACAAGCCAGCAGTTG-Tamra
STC2	sens	GCAGAATACAGCGGAGATCC
	antisens	CAAATCCCATGTAAGCCCCG
	sonde	FAM-GGTGTGGCGTGTTTGAATGT-Tamra
EZR	sens	AAGGGTTCTGCTCTGACTCC
	antisens	TGGTTTCGGCATTTTCGGTT
	sonde	FAM-TGCTGCTGGATAGTCGTGTT-Tamra

Quantification des transcrits de *MUC4* après transfections par siARN

Lors de la quantification des transcrits de *MUC4* par RT-qPCR après traitement par les siARN hnRNP-L, l'utilisation de la *GAPDH* s'est avérée être un standard interne meilleur que

les transcrits *18S*. Les amorces et sondes dirigées contre les gènes cibles et la *GAPDH* sont utilisées à la concentration de 10 μM (Eurogentec).

Le mélange réactionnel utilisé dans ce cas est le suivant :

- Amorce sens	1 μl
- Amorce antisens	1 μl
- Sonde	1 μl
- Amorce <i>GAPDH</i> sens	1 μl
- Amorce <i>GAPDH</i> antisens	1 μl
- Sonde <i>GAPDH</i>	1 μl
- Absolute TM QPCR ROX mix (Abgene)	12,5 μl
- Eau stérile	4,5 μl
- ADNc	2 μl

Par ailleurs, pour cette condition, nous avons réalisé une quantification relative des transcrits de *MUC4* avec la méthode des delta de Ct (ΔCt) puisque nous ne disposons pas de fragment de la *GAPDH* cloné en vecteur pCR2.1 permettant l'établissement de courbes d'étalonnage.

III.2.2. Quantification avec la technique EvaGreen[®]

Cette technique est utilisée pour la quantification des transcrits *luciférase* après transfections cellulaires et des ARNm de *Muc4* et *hnrnp-l* murins. Les transcrits de la *GAPDH* sont utilisés comme standard interne. Les amorces sont présentées dans le tableau XXVI.

Le mélange réactionnel est le suivant :

- Amorce sens (5 μM , Eurogentec)	2 μl
- Amorce antisens (5 μM , Eurogentec)	2 μL
- SsoFast TM EvaGreen [®] supermix 2X (Biorad)	10 μl
- ADNc	1 μl
- Eau stérile	5 μl

L'amplification est réalisée sur le thermocycleur CFX96TM Real-Time System (Bio-Rad) avec le programme suivant : 95 °C 30 s : 1 cycle puis [95 °C, 3s ; 60 °C, 3s] 40 cycles.

Tableau XXVI : Séquences des amorces utilisées en qPCR EvaGreen®.

Gène		Amorces (5' → 3')
luciférase	sens	CCTGAAGTCTCTGATTAAGT
	antisens	GCGACACCTGCGTCGAAG
GAPDH	sens	CTCAACTTTAACTGGAAAGAATGTC
	antisens	TCCTTTTCACCAGCAAGCT
Muc4 souris	sens	CCTTCACTGATAACCGCTGCTT
	antisens	GCGGAGGCATTTTCATCC
Gapdh souris	sens	AGGTCGGTGTGAACGGATTG
	antisens	TGTAGACCATGTAGTTGAGGTCA

La quantification des transcrits utilisée est basée sur une quantification relative utilisant la méthode des delta de Ct.

III.3. Etude de la demi-vie des transcrits

Pour déterminer la demi-vie des transcrits de mucines, 8 µg/ml d'actinomycine D (Sigma Aldrich) sont ajoutés pendant toute la durée de la manipulation dans le milieu de culture de cellules post-confluentes cultivées en flacon de culture de 25 cm² (Debailleul *et al.*, 1998). Les ARN sont extraits avant traitement (temps zéro) puis après différents temps d'incubation avec l'actinomycine D (T3h, T6h, T9h, T12h, T24h et T30h) et sont analysés par RT-qPCR ou séquençage à haut débit.

III.4. Analyse par séquençage à haut débit

III.4.1. Analyse de la demi-vie des transcrits

Afin d'étudier la stabilité des transcrits (*whole transcriptome stability*), les bibliothèques sont préparées à partir des ARNm issus des temps T3h, T6h, T9h, T12h, T24h et T30h de traitement à l'actinomycine D avec le kit « TruSeq Stranded mRNA Sample Preparation Kit » (Illumina) selon les recommandations du fournisseur.

Pour ce faire, les ARN sont fragmentés pendant 2 min à 94 °C. Un premier puis un second brin d'ADN complémentaires des ARNm purifiés sont synthétisés afin de générer de l'ADNc double brin.

Un nucléotide Adénosine est ajouté à leur extrémité 3' pour éviter la ligation des différents brins entre eux. Les codes-barres servant à identifier chaque échantillon sont ensuite ajoutés. Une réaction de PCR permet d'amplifier les ADNc afin de générer les bibliothèques d'ADNc.

En vue de leur passage sur le séquenceur, celles-ci sont « multiplexées » (en 2 groupes contenant 7 bibliothèques par groupe) puis analysées sur le séquenceur HiSeq2500 (Illumina, collaboration avec le Pr Philippe Froguel, UMR CNRS 8199, Lille). Pour cela, une amplification en ponts des fragments d'ADNc est effectuée suivie d'un séquençage des *clusters* d'ADNc avec des nucléotides fluorescents. L'acquisition des couleurs permet de déduire la succession des nucléotides (figure 115).

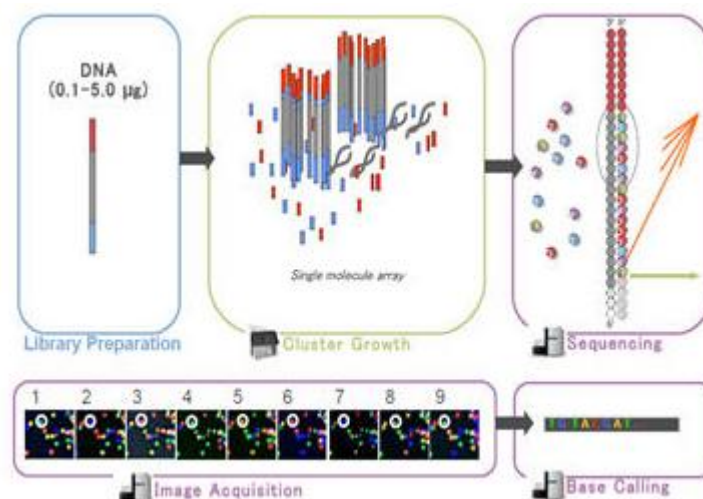


Figure 115 : Principe des différentes étapes du séquençage à haut débit utilisant la technologie Illumina.

Les séquences (*reads*) obtenues sont alignées sur le génome de référence (hg19, Tophat v2.0.12 (Kim *et al.*, 2013). Les transcrits sont assemblés et quantifiés en FPKM (*fragments per kilobase of exon per million reads mapped*) avec l'outil Cufflinks v2.2.1 (Trapnell *et al.*, 2012, 2013).

Trois transcrits, *RPLP0* (NM_001002), *PGK1* (NM_000291) et *TRAP1* (NM_016292)), dont le niveau d'expression est resté stable au cours du traitement par actinomycine D sont choisis pour normaliser le nombre de *reads* pour l'ensemble des transcrits analysés (Tani *et al.*, 2012). Pour ce faire, la moyenne géométrique des FPKM de ces trois transcrits est calculée pour chaque temps. Puis pour chaque temps, la valeur des FPKM de chaque transcrit étudié est divisée par la moyenne géométrique des FPKM des transcrits de référence au temps correspondant. A partir de ces résultats normalisés ensuite sont construites les courbes de décroissance permettant de déterminer la valeur de leur demi-vie des transcrits. Seuls les transcrits pour lesquels la courbe de décroissance théorique se rapproche du modèle observé (coefficient de corrélation supérieur à 0.9) sont retenus dans la suite de l'étude.

III.4.2. Etude des ARNm immunoprécipités lors de la RNA-IP

Les ARNm obtenus lors de l'expérience d'immunoprécipitation des complexes ARNm-protéines (paragraphe IV.8) sont analysés par *RNA-sequencing* (collaboration avec la plateforme de génomique fonctionnelle du Dr Martin Figeac, Lille) sur le séquenceur PGM (Personal Genome Machine[®], Life Technologies).

Les librairies d'ADNc sont préparées à l'aide du kit « Ion total RNA-Seq Kit v2 » (Life Technologies) sans l'étape de fragmentation par la RNase III. Pour ce faire, après une étape de ligation avec des adaptateurs, les ARN servent de matrice à une transcription inverse afin d'obtenir des ADNc. Les ADNc obtenus sont amplifiés par PCR et purifiés afin de générer des librairies d'ADNc.

Celles-ci sont ensuite analysées sur le séquenceur (PGM). La technologie de ce séquenceur est basée sur le changement de pH induit par la libération de protons faisant suite à l'incorporation des nucléotides lors de la polymérisation de l'ADN.

Les signaux obtenus sont convertis sous forme de séquence et les « lectures » (*reads*) sont alignées sur la banque des transcrits de référence (RefSeq) grâce à l'outil Bowtie 2 (Langmead and Salzberg, 2012).

III.5. Sous-clonage de produits PCR en vecteur plasmidique pCR[®]2.1

III.5.1. Ligation

La ligation du produit de PCR en vecteur pCR[®]2.1 est effectuée avec le kit « TA Cloning kit » (Invitrogen). Pour cela, le mélange suivant est incubé une nuit à 14 °C :

- Produit PCR	x µl
- Vecteur pCR2.1 (25 ng/µl)	2 µl
- T4 DNA ligase (5 U/µl)	1 µl
- Tampon de ligation 10X	1 µl
- Eau stérile	qsp 10 µl

x est le volume correspondant à la quantité du produit de PCR (y) en nanogramme calculé avec l'équation :

$y = (\text{taille en paire de base du produit PCR}) \times (50 \text{ ng du vecteur pCR}^{\text{®}}2.1) / (\text{taille en paire de base du vecteur pCR}^{\text{®}}2.1).$

III.5.2. Transformation bactérienne par choc thermique

La transformation de bactéries compétentes INVαF' (One Shot competent cells, Invitrogen) est effectuée de la façon suivante : 2 µl de produit de ligation sont ajoutés à 50 µl de bactéries et l'ensemble est placé 30 min dans la glace.

L'insertion du plasmide dans les bactéries est réalisée par un choc thermique en incubant ce mélange 30 s à 42 °C.

Puis, après avoir posé les bactéries 2 min sur la glace, 250 µl de milieu SOC (*Super Optimal broth with Catabolite repression* composé de tryptone 2 % (m/v), extrait de levure 0,5 % (m/v), NaCl 10 mM, KCl 2,5 mM, MgCl₂ 10 mM, glucose 20 mM) sont ajoutés et la suspension est agitée doucement à 37 °C pendant 1 h.

Enfin, 50 µl de bactéries transformées sont étalées dans une boîte de Pétri contenant du milieu LB-Agar (tryptone 1 % (m/v), extrait de levure 0,5 % (m/v), NaCl 0,5 % (m/v), agar 1,5 % (m/v)) additionné de 50 µg/ml de l'antibiotique kanamycine et sur lequel 40 µl d'une solution de X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) à 40 mg/ml ont été préalablement répartis. La boîte de Pétri est ensuite placée une nuit à 37 °C.

III.5.3. Sélection des clones

Le principe de sélection des clones ayant intégré le produit de PCR est basé sur le changement de coloration du substrat X-Gal. En effet, celui-ci devient bleu à la suite de l'hydrolyse réalisée par la β -galactosidase. Cependant, l'insertion du produit de PCR a lieu au sein du gène codant cette enzyme (LacZ), ce qui empêche sa production. Ainsi, les colonies ayant intégrées le produit PCR sont incapables d'hydrolyser le substrat X-Gal et restent donc blanches.

Les clones blancs sélectionnés sont ensuite repiqués dans 5 ml de milieu LB contenant 50 μ g/ml de kanamycine puis la suspension est mise sous agitation douce pendant une nuit à 37 °C.

L'ADN plasmidique est extrait avec le kit « NucleoSpin 8 Plasmid » (Macherey Nagel) comme décrit au paragraphe III.10 puis il est analysé en séquençage nucléotidique (GATC) afin de s'assurer de la présence de l'insert et l'absence d'erreur générée lors de la réaction de PCR. L'amorce servant au séquençage est la suivante : 5' CAGGAAACAGCTATGAC 3'.

III.6. Mutagenèse dirigée

PCR

La mutagenèse dirigée est effectuée avec le kit « QuickChange Site-Directed Mutagenesis » (Stratagène). La première étape consiste en une amplification PCR avec des amorces spécifiques contenant la ou les mutations désirée(s). Celles-ci sont choisies de façon à respecter les critères suivants :

- Taille comprise entre 25 et 45 nucléotides ;
- Température d'hybridation ≥ 78 °C calculée avec la formule : $81.5 + 0.41 (\%GC) - 675/N$ - % mésappariement où N représente le nombre de bases de l'amorce ;
- Pourcentage de GC ≥ 40 % ;
- Terminaison de l'amorce par un ou plusieurs C ou G ;
- Mutation désirée située au milieu de l'amorce avec 10 à 15 nucléotides de chaque côté.

Les amorces utilisées sont les suivantes :

Del CA Sens : 5' TGTGCACGCGCGCGTGGAGTTCATAATGTGGTGATGG 3'

Del CA Antisens : 5' CCATCACACATTATGAACTCCACGCGCGCGTGCACA 3'

L'amplification par PCR est effectuée avec le mélange ci-dessous :

- Tampon de l'enzyme 10X (Stratagène)	5 µl
- Amorce sens (Eurogentec)	125 ng
- Amorce antisens (Eurogentec)	125 ng
- dNTP mix (Stratagène)	1 µl
- ADN plasmidique	20 ng
- Eau stérile	qsp 50 µl

L'enzyme ADN polymérase *pfu Turbo* (1 µl : 2,5 U) est ensuite ajoutée à ce mélange.

L'amplification est réalisée dans le thermocycleur « MJ Mini™ » (BioRad) avec le programme suivant : 95 °C 30s ; 18 cycles [95 °C 30s ; 55°C 1 min ; 68 °C 11 min] puis 4 °C à l'infini.

A la fin de la réaction de PCR, le tube est refroidi 2 min sur la glace et 1 µl de l'enzyme *DpnI* (10 U/µl) est ajouté au mélange pendant 1 h 30 min à 37°C.

Transformation bactérienne, purification de l'ADN plasmide et séquençage

Pour réaliser la transformation bactérienne, 1 µl du mélange précédent est mis en présence de 50 µl de bactéries compétentes «XL1-Blue Supercompetent cells » (Stratagène) et l'ensemble est placé 30 min dans la glace.

Le choc thermique est effectué en plaçant le tube 45 sec à 42 °C puis 2 min dans la glace. Ensuite 500 µl de SOC sont ajoutés et la suspension bactérienne ainsi obtenue est placée 1 h à 37°C sous agitation douce.

Enfin, 50 µl de cette suspension sont étalés dans une boîte de pétri contenant du LB agar avec 50 µg/ml d'ampicilline sur lequel a été préalablement étalé du X-Gal (40 mg/ml). La boîte est incubée une nuit à 37 °C.

Les clones blancs obtenus sont mis en culture pendant une nuit sous agitation douce dans 5 ml de LB liquide avec 50 µg/ml d'ampicilline et l'ADN plasmidique est ensuite extrait

avec le kit « NucleoSpin 8 Plasmid » (Macherey Nagel) (paragraphe III.10). Puis, la séquence de celui-ci est vérifiée par séquençage nucléotidique.

III.7. Clonage du motif exon2-exon3 de MUC4

Le clonage nous a permis d'insérer une séquence nucléotidique supplémentaire dans le vecteur plasmidique « pGL3 promoter » (Promega) contenant le 3'UTR de *MUC4* au niveau des sites de restriction de l'enzyme HindIII.

PCR

Une amplification par PCR est réalisée sur un ADNc préparé à partir de lignée cellulaire avec des amorces situées autour de la région à insérer. Les amorces choisies respectent les critères ci-dessous avec un site de restriction pour l'enzyme de HindIII dont le site de coupure est A/AGCTT :

- les amorces débutent par CGC en 5' ;
- ce motif est suivi par le motif de restriction reconnu par l'enzyme (pour HindIII : AAGCTT) ;
- puis sont ajoutés une vingtaine de nucléotides contenant la séquence à insérer.

Les amorces utilisées sont les suivantes :

hind3 sens ex2-3 : 5' CGCAAGCTTGGACTCCCCTCTGAAGAT 3'

hind3 reverse ex2-3 : 5' CGCAAGCTTAGGTTGTGGGGGGTGGT 3'

L'amplification par PCR est réalisée avec le mélange suivant :

- | | |
|---|-----------|
| - Tampon de l'enzyme 10X (Roche Diagnostic) | 5 µl |
| - Amorce sens 5 µM (Eurogentec) | 4 µl |
| - Amorce antisens 5 µM (Eurogentec) | 4 µl |
| - dNTP 20 mM (Eurogentec) | 4 µl |
| - Taq polymérase (5 U/µl, Roche Diagnostic) | 4 µl |
| - ADN matrice | 7 µl |
| - Eau stérile | qsp 50 µl |

La réaction d'amplification est réalisée dans le thermocycleur « MJ Mini™ » (BioRad) avec le programme suivant : 94 °C, 2 min ; 35 cycles [94 °C, 45 s ; 60 °C, 1 min ; 72 °C, 1 min] ; 72 °C, 10 min puis 4 °C à l'infini.

Purification de l'insert

Le produit PCR est purifié avec le kit « Wizard SV gel and PCR clean-up system » (Promega).

Pour cela 45 µl du produit de PCR sont transférés dans un microtube de 1,5 ml et le même volume de « Membrane binding solution » est ajouté.

Après homogénéisation par aspiration-refoulement, le mélange est déposé sur une colonne et est centrifugé 1 min à 16 000 x g à 4 °C.

Ensuite, 700 µl de tampon « Membrane wash solution » sont déposés sur la membrane et la colonne est centrifugée 1 min à 16 000 x g à 4 °C. De la même manière, un second lavage est effectué avec 500 µl de « Membrane wash solution » suivi d'une centrifugation de 5 min.

Ensuite, une centrifugation à vide est réalisée pendant 2 min à 16 000 x g à 4°C. A l'issue de cette étape, la colonne est transférée sur un microtube de 1,5 ml et 42 µl d'eau sont ajoutés sur la membrane, suivi d'une incubation de 1 min puis d'une centrifugation de 1 min à 16 000 x g à 4 °C afin de collecter l'éluat contenant le produit PCR.

Digestion de l'insert purifié

L'insert est digéré par l'enzyme de restriction choisie au niveau des sites de restrictions insérés dans le produit PCR. Pour cela, le mélange suivant est placé 24 h à 37 °C :

- | | |
|---|-------|
| - Insert | 27 µl |
| - Tampon 10 X de l'enzyme (New England Biolabs) | 3 µl |
| - Enzyme de restriction (HindIII, 20 000 U/ml, New England Biolabs) | 2 µl |

Après digestion de l'insert par l'enzyme de restriction, le produit est à nouveau purifié et l'élution est réalisée avec 17 µl d'eau stérile.

Digestion du vecteur

Le vecteur dans lequel l'insert sera introduit est clivé par l'enzyme de restriction HindIII en incubant le mélange suivant 2 h à 37 °C :

- Vecteur	2 µl
- Tampon 10X de l'enzyme (New England Biolabs)	2 µl
- Enzyme de restriction (HindIII, New England Biolabs)	2 µl
- Eau stérile	14 µl

Traitement à la phosphatase et purification du vecteur digéré déphosphorylé

Le vecteur digéré est ensuite déphosphorylé afin d'éviter sa re-circularisation en ajoutant 2 µl de phosphatase (1 U/µl, Rapid DNA Dephos and ligation Kit, Roche Diagnostic) et 2 µl du tampon de l'enzyme. Le mélange est placé 1 h à 37 °C.

Le vecteur digéré déphosphorylé est à nouveau purifié avec le kit « Wizard SV gel and PCR clean-up system » et l'élution est réalisée par 17 µl d'eau stérile.

Ligation du vecteur et de l'insert

La ligation de l'insert dans le vecteur est réalisée à l'aide du kit « Rapid DNA Dephos and ligation Kit » (Roche Diagnostic). Un contrôle négatif est réalisé sans insert et la ligation est réalisée avec 3 fois plus d'insert que de vecteur. Le mélange réactionnel suivant est placé une nuit à 20 °C :

	Contrôle	Ratio 1/3
- Insert	0 µl	3 µl
- Vecteur	1 µl	1 µl
- T4 DNA Ligase (5 U/µl, Roche Diagnostic)	1 µl	1 µl
- T4 DNA Ligation Buffer 2X (Roche Diagnostic)	5 µl	5 µl
- Eau stérile	3 µl	0 µl

Transformation bactérienne, extraction de l'ADN plasmidique et séquençage

Le lendemain, 3 µl du produit de ligation sont ajoutés à 50 µl de bactéries compétentes *Escherichia coli* (*E. Coli*) JM109 (Promega). Le mélange est incubé 30 min dans la glace et le choc thermique est réalisé par une incubation du tube à 42 °C pendant 45 s.

Après avoir posé le tube 2 min dans la glace, 250 µl de milieu SOC sont additionnés et la suspension est incubée à 37 °C pendant 1 h sous agitation douce.

Puis, 50 µl de bactéries transformées sont étalés dans une boîte de Pétri contenant du milieu LB-Agar additionné d'un antibiotique (50 µg/ml d'ampicilline) qui est ensuite placée à 37°C pendant une nuit.

Les clones sont repiqués dans 5 ml de milieu LB contenant 50 µg/ml d'ampicilline puis placés sous agitation douce pendant une nuit à 37 °C.

L'ADN plasmidique est extrait à l'aide du kit « NucleoSpin 8 Plasmid » (Macherey Nagel) comme décrit au paragraphe III.10 et sa séquence est vérifiée par séquençage nucléotidique.

III.8. Préparation d'ADN plasmidique en petit volume (Mini-Prep)

Après mise en culture des bactéries en petit volume (5 ml de milieu LB contenant l'antibiotique adapté), l'ADN plasmidique est extrait avec le kit « NucleoSpin 8 Plasmid » (Macherey Nagel).

Pour cela, 1,4 ml de la suspension bactérienne précédemment obtenue sont placés dans un microtube de 1,5 ml et centrifugés 10 min à 1 000 x *g*.

Le culot bactérien est remis en suspension avec 250 µl de tampon de resuspension A1 puis 250 µl de tampon de lyse A2 sont ajoutés. L'ensemble est homogénéisé par retournement du tube et laissé 5 min à T°amb.

Ensuite, 350 µl de tampon de neutralisation sont additionnés au mélange et le lysat est placé dans une barrette filtrante « NucleoSpin Plasmid *Filter Strips* » placée dans la partie supérieure du collecteur « NucleoVac ». Le vide appliqué permet de récupérer le filtrat dans la barrette « NucleoSpin Plasmid *Binding Strips* », qui a été préalablement disposée dans la zone inférieure du « NucleoVac ».

Cette barrette est transférée dans la partie supérieure du « NucleoVac ». Une nouvelle étape d'application du vide permet l'adsorption de l'ADN plasmidique sur la membrane de silice de la colonne.

Les différents lavages suivants sont réalisés avec application du vide entre chacun d'entre eux : 600 µl de tampon AW (un lavage), 900 µl de tampon A4 (2 lavages).

Le vide est maintenu pendant 10 min puis 100 µl de tampon d'élution NE sont déposés sur la membrane pendant 1 min à T°amb., l'élution de l'ADN plasmidique est réalisée par application du vide pendant 5 min.

La concentration et la qualité de l'ADN plasmidique extrait sont ensuite déterminées à l'aide d'un spectrophotomètre (NanoDrop®).

III.9. Préparation d'ADN plasmidique en grande quantité (Maxi-Prep)

Pour disposer des constructions plasmidiques et des différents vecteurs d'expression en grande quantité, le kit « NucleoBond® Xtra Maxi Plus EF » (Macherey-Nagel) est utilisé après transformation de bactéries *E. coli* (JM109) (Promega).

III.9.1. Transformation des bactéries *E. coli* JM109

Dans un tube à centrifuger de 15 ml sont mélangés 100 µl de bactéries *E. coli* JM109 (Promega) et 100 ng du plasmide d'intérêt.

Le tube est ensuite placé 30 min dans la glace et le choc thermique est réalisé par une incubation à 42 °C pendant 45 s.

Après avoir laissé le tube 2 min dans la glace, 900 µl de milieu SOC sont additionnés et la suspension est mise à 37 °C pendant 1 h sous agitation douce.

Ensuite, 50 µl de bactéries transformées sont étalés dans une boîte de Pétri contenant du milieu LB-Agar additionné d'un antibiotique (50 µg/ml d'ampicilline) qui est ensuite placée à 37 °C pendant une nuit.

Un clone isolé est prélevé et placé dans 5 ml de milieu liquide LB contenant 50 µg/ml d'ampicilline. Le mélange est incubé pendant 8 h à 37 °C sous agitation douce.

Ensuite, 300 µl de la suspension bactérienne sont transférées dans 100 ml de milieu liquide LB additionné de 50 µg/ml d'ampicilline. L'ensemble est incubé pendant une nuit à 37 °C sous agitation douce.

III.9.2. Extraction plasmidique

La purification de l'ADN plasmidique est réalisée à l'aide du kit « NucleoBond® Xtra Maxi Plus EF » (Macherey Nagel).

Pour cela, la suspension bactérienne obtenue est centrifugée pendant 10 min à 5 000 x *g* à 4 °C et le culot bactérien est mis en suspension dans 12 ml de tampon de resuspension (RES-EF).

Après transfert du mélange dans un tube à centrifuger de 50 ml, 12 ml de tampon de lyse sont ajoutés (LYS-EF) et le tube est laissé 5 min à T°amb.

Parallèlement, une étape d'équilibration de la colonne est effectuée avec 35 ml de tampon EQU-EF.

A la fin de l'incubation, le lysat est additionné de 12 ml de tampon de neutralisation (NEU-EF) et l'ensemble bien homogénéisé est incubé dans la glace pendant 5 min.

Ensuite, la suspension est déversée dans la colonne et cette dernière est lavée avec 10 ml de tampon FIL-EF.

A la suite de cette étape, le filtre est éliminé et différents lavages sont réalisés par l'ajout dans la colonne de 90 ml de tampon de lavage ENDO-EF puis de 45 ml de tampon de lavage WASH-EF.

L'étape d'éluion de l'ADN plasmidique est effectuée par l'addition dans la colonne de 15 ml de tampon ELU-EF. L'éluat est collecté dans un tube à centrifuger de 50 ml puis 10,5 ml d'isopropanol lui sont ajoutés pour précipiter l'ADN.

L'ensemble est transféré dans une seringue sur laquelle est placé un filtre où va se fixer le matériel plasmidique. Puis, 5 ml d'éthanol à 70 % sont ajoutés afin de laver le précipité et l'ADN plasmidique est élué par 500 µl d'H₂O-EF.

La concentration et la qualité de l'ADN plasmidique extrait sont ensuite déterminées à l'aide d'un spectrophotomètre (NanoDrop[®]).

IV. Biochimie

IV.1. Mesure de l'activité luciférase

Le dosage de l'activité luciférase est effectué sur des extraits cellulaires totaux préparés à la suite des transfections.

Pour se faire, les cellules sont rincées par du D-PBS 1X puis 110 µl de « Reporter Lysis Buffer » 1X (RLB) (Promega) sont déposés sur les cellules pendant 15 min.

Ensuite, le tapis cellulaire est gratté et le lysat obtenu est placé dans un microtube de 1,5 ml préalablement refroidi dans la glace.

Après agitation vive du tube pendant 15 s, celui-ci est centrifugé durant 3 min à 4 °C à 15 500 x g. Le surnageant est transféré dans un microtube neuf.

La mesure de l'activité enzymatique de la luciférase est réalisée avec le kit « Luciferase assay System » (Promega). Le principe repose sur la mise en contact des extraits cellulaires avec le substrat de l'enzyme, la luciférine. La quantité de lumière relative (RLU) émise en 15 s par le mélange comprenant 20 µl d'extrait cellulaire et 100 µl de luciférine est mesurée par le luminomètre Mithras LB 940 (Berthold) de la Plate-forme BICeL (Lille).

Pour normaliser les résultats, les valeurs d'activité luciférase sont rapportées à la concentration de protéines de l'échantillon (RLU/mg de protéines) ou à l'activité β -galactosidase. Ils sont ensuite exprimés en facteur d'induction par rapport aux valeurs obtenues avec le vecteur d'expression vide dont le niveau d'activité luciférase normalisé est arbitrairement fixé à 1.

IV.2. Dosage de l'activité β -galactosidase

Le dosage s'effectue en plaque 96 puits sur les extraits cellulaires préparés avec le tampon RLB 1X (Promega) à l'aide du kit « β -galactosidase Enzyme Assay System » (Promega). Les points de gamme et les échantillons sont déposés à chaque fois en double.

Une gamme d'étalonnage est préparée en diluant la solution stock de β -galactosidase au 1/10 000 de la façon décrite dans le tableau XXVII :

Tableau XXVII : Gamme d'étalonnage pour le dosage de l'activité β -galactosidase.

Standard β -galactosidase (milliunités)	Volume de la solution diluée au 1/10 000 (μ l)	Volume de RLB 1X (μ l)
0	0	50
1,0	10	40
2,0	20	30
3,0	30	20
4,0	40	10
5,0	50	0

Ensuite, 25 μ l de RLB 1X et 25 μ l d'échantillon sont déposés dans chaque puits. Puis 50 μ l d'assay buffer 2X sont ajoutés et la plaque est placée 15 min sous agitation douce à 37 °C jusqu'à développement d'une coloration jaune.

La réaction est arrêtée par l'ajout de 150 μ l de carbonate de sodium 1M et les densités optiques obtenues sont mesurées sur le lecteur de microplaque *Model 680* (Biorad) à 420 nm. La droite d'étalonnage permet de déterminer l'activité β -galactosidase de chaque échantillon.

IV.3. Préparation d'extraits cellulaires totaux

Les extraits cellulaires totaux sont préparés à partir de cellules post-confluentes (confluence + 2 jours) ensemencées en flacon de culture de 75 cm².

Pour se faire, les cellules sont grattées dans leur milieu de culture et la suspension obtenue est placée dans un tube à centrifuger de 50 ml. Une étape de rinçage du flacon de culture est réalisée avec du D-PBS 1X froid.

Après une centrifugation de 5 min à 535 x *g* à 4 °C et élimination du surnageant, le culot est remis en suspension dans 0,5 ml de D-PBS 1X. Le mélange est placé dans un microtube de 1,5 ml et centrifugé pendant 1 min à 9 000 x *g*. Une nouvelle étape de lavage par du D-PBS 1X est effectuée et le culot cellulaire est resuspendu dans 500 µl de tampon de lyse dont la composition est la suivante :

- Tris-HCl 50 mM, pH 7,4
- NaCl 150 mM
- EDTA 1 mM
- Aprotinine 1 µg/ml
- Leupeptine hémi-sulfate 1 µg/ml
- PMSF 1 mM
- NP-40 1 % (v/v)
- Fluorure de sodium 5 mM
- Orthovanadate de sodium 5 mM
- Désoxycholate de sodium 0,25 % (m/v)

Le mélange est placé pendant 30 min dans de la glace en frottant les tubes toutes les 10 min sur un portoir métallique.

Après une étape de centrifugation de 10 min à 16 000 x *g* à 4 °C, le surnageant est récupéré et conservé à -80 °C.

Une étape préalable de *crosslinking* peut être réalisée avant l'extraction des protéines afin de fixer les interactions protéine-protéines (Mouchel-Vielh *et al.*, 2011). Pour cela, les cellules sont mises en présence de formaldéhyde à 1 % (v/v) pendant 10 min à T°amb.

La réaction est arrêtée par l'ajout de 10 ml d'une solution de glycine 0,25 M préparée dans du D-PBS 1X pendant 5 min.

Les cellules sont rincées par 5 ml de D-PBS 1X froid puis par du D-PBS 1X froid contenant des inhibiteurs de protéases (leupeptine 10 µg/ml, aprotinine 10 µg/ml, EDTA 0,2 mM et PMSF 0,5 mM).

Les cellules sont grattées dans 10 ml de D-PBS 1X froid contenant les inhibiteurs de protéases. Les extraits cellulaires sont ensuite préparés comme décrit au début du paragraphe.

IV.4. Dosage protéique

Le dosage protéique est basé sur une méthode colorimétrique utilisant l'acide bicinchoninique (BCATM protein Assay Reagent, Pierce). Une gamme d'étalonnage est effectuée parallèlement au dosage avec des concentrations croissantes de sérumalbumine bovine.

Pour cela, dans une microplaque de 96 puits, 2 µl d'échantillon sont déposés par puits et 200 µl du réactif BCA (reconstitué extemporanément : 5 ml de solution A pour 0,1 ml de solution B, Pierce) sont additionnés dans chacun des puits. La plaque est placée sous agitation douce pendant 1 h à 37 °C et les densités optiques sont ensuite mesurées à 570 nm avec le lecteur de microplaque *Model 680* (Biorad). La droite d'étalonnage permet de déduire les concentrations protéiques à partir des densités optiques.

IV.5. Western-blot

Pour réaliser un *Western-blot*, les protéines (25 µg) sont dénaturées 5 min à 100 °C après l'ajout (v/v) d'une solution dénaturante 2X faisant office également de tampon de charge dont la composition est la suivante :

- Tris-HCl 0,5 M, pH 6,8
- Glycérol 20 % (v/v)
- β-mercaptoéthanol 10 % (v/v)
- Bleu de bromophénol 0,05 % (m/v)
- SDS 0,1 % (m/v)

Puis les échantillons sont déposés sur un gel de polyacrylamide dénaturant dont la réticulation est de 10 %. L'électrophorèse est effectuée dans un tampon Tris-HCl 25 mM, pH 8,3, glycine 192 mM, SDS 0,05 % (v/v) sous un courant constant de 35 mA par gel pendant 1 h 30.

A la suite de cette électrophorèse, les protéines sont transférées sur une membrane de nitrocellulose (0,22 µm, Protran Amersham) par électrotransfert sous une tension constante de 100 V pendant 1 h dans du tampon de transfert (Tris-HCl 25 mM, pH 8,3, glycine 192 mM, SDS 0,005 % (v/v), méthanol 20 % (v/v)).

La membrane est ensuite lavée avec du tampon Tris-HCl 50 mM, pH 8,3, NaCl 150 mM, Tween® 20 0,05 % (v/v) (T-BST).

Afin de saturer les sites de fixation aspécifiques, la membrane est placée pendant 3 h sous agitation à T °amb. dans du tampon TBS-T contenant 5 % (m/v) de lait en poudre écrémé. Ensuite, trois lavages de 5 min de la membrane sont réalisés dans du TBS-T.

La membrane est incubée avec l'anticorps primaire pendant une nuit à 4 °C sous agitation douce (tableau XXVIII).

Le lendemain, quatre lavages de la membrane dans le TBS-T sont effectués (une fois 15 min puis trois fois 5 min). Celle-ci est ensuite incubée avec l'anticorps secondaire couplé à la peroxydase de Raifort pendant 1 h 30.

La membrane est lavée 3 fois avec du TBS-T et placée pendant 5 min dans le mélange luminol-H₂O₂ du kit « SuperSignal® West PICO » (Pierce). Les bandes sont ainsi détectées par chimioluminescence et leur acquisition est réalisée sur l'appareil LAS 4000 (Fujifilm).

Le logiciel ImageQuantTL (GE Healthcare) permet la quantification des bandes visualisées par densitométrie. La normalisation est réalisée en rapportant l'intensité du signal obtenu pour la protéine d'intérêt sur celui de la référence interne utilisée pour chaque échantillon (β-actine).

Tableau XXVIII : Anticorps utilisés pour les *Western-Blot*.

Anticorps primaire	Référence	Dilutions
Galectine-3	Polyclonal de lapin	Primaire : 1/300
	Abcam ab31707	Secondaire : 1/4000
hnRNP-L	Monoclonal de souris	Primaire : 1/2000
	Abcam ab6106	Secondaire : 1/5000
β-actine	Monoclonal de souris	Primaire : 1/5000
	Sigma A5441	Secondaire : 1/20000

IV.6. Western-blot MUC4

Du fait du haut poids moléculaire de MUC4 (>900 kDa), la migration des protéines est effectuée en gel d'agarose 2 % (m/v) dans le même tampon de migration que celui utilisé pour les *Western-blot* classiques.

Après dénaturation de 20 µg de protéines pendant 5 min à 100 °C dans le tampon de charge dénaturant, l'électrophorèse est réalisée pendant 6 h sous une tension constante de 100 V.

Puis le transfert des protéines est effectué passivement sur une membrane de nitrocellulose de 0,22 µm (Protran Amersham) de la manière suivante : le gel est placé sur un pont de papier Whatman® 3 MM qui trempe dans du tampon de transfert, la membrane est déposée au-dessus du gel et recouverte par une épaisseur de papier Whatman® 3 MM sur laquelle est posé un poids. Le transfert des protéines est alors permis par le pouvoir absorbant du papier Whatman® et se déroule toute la nuit.

La membrane est ensuite placée dans une solution de lait écrémé à 5 % (m/v) dans du TBS-T pendant 2 h afin d'être saturée.

Puis, après trois lavages de 5 min dans TBS-T, la membrane est incubée 2 h à T°amb avec l'anticorps primaire anti-MUC4 (Santa Cruz sc-53945, dilution 1/200).

Trois lavages dans le TBS-T sont à nouveau réalisés puis la membrane est incubée avec l'anticorps secondaire anti-souris (1/10 000) couplé à la peroxydase de Raifort pendant 1 h à T°amb.

La membrane est lavée trois fois avec du TBS-T puis la détection de l'activité enzymatique couplée à l'anticorps secondaire est effectuée avec le kit « SuperSignal® West PICO » (Pierce) (paragraphe IV.5).

IV.7. Co-immunoprécipitation

Une étape préalable de *crosslinking* est réalisée avant l'extraction des protéines afin de fixer les interactions protéine-protéines. Pour cela, les cellules sont mises en présence de formaldéhyde à 1 % (v/v) pendant 10 min à T°amb.

La réaction est arrêtée par l'ajout de 10 ml d'une solution de glycine 0,25 M préparée dans du D-PBS 1X pendant 5 min.

Les cellules sont rincées par 5 ml de D-PBS 1X froid puis par du D-PBS 1X froid contenant des inhibiteurs de protéases (leupeptine 10 µg/ml, aprotinine 10 µg/ml, EDTA 0,2 mM et PMSF 0,5 mM).

Les cellules sont grattées dans 10 ml de D-PBS 1X froid contenant les inhibiteurs de protéases. Les extraits cellulaires sont ensuite préparés comme décrit au paragraphe IV.3.

Pour effectuer la co-immunoprecipitation, 300 µg d'extraits cellulaires totaux sont mis en présence 2 µg de l'anticorps anti-galectine-3 (Santa Cruz, sc-20157) ou de l'IgG irrelevante (Millipore, 12-370) dans un microtube de 1,5 ml qui est placé une nuit à 4 °C sous agitation circulaire.

Le lendemain, le volume global de billes (EZview Red protein A Affinity Gel, Sigma Aldrich) correspondant à 40 µl de billes pour chaque condition est prélevé et 750 µl de tampon RIPA, dont la composition est la suivante, sont ajoutés :

- Tris-HCl 20 mM pH 7,5
- NaCl 150 mM
- EDTA 1 mM
- EGTA 1 mM
- NP-40 10 % (v/v)
- Désoxycholate de sodium 1 mM
- Glycerophosphate de sodium 1 mM
- Pyrophosphate de sodium 2.5 mM
- Orthovanadate de sodium 1 mM
- Leupeptine 1 µg/ml
- Aprotinine 1 µg/ml
- PMSF 1 mM

Le mélange billes-RIPA est vortexé, centrifugé 30 s à 8200 x g et le surnageant est éliminé. Cette étape est répétée une fois.

Les billes sont ensuite remises en suspension dans 375 µl de RIPA. Puis 40 µl de ces billes sont ajoutées à chaque tube contenant les complexes protéiques-anticorps et celui-ci est placé pendant 1 h à 4 °C sur un agitateur circulaire.

Les tubes sont ensuite centrifugés 30 s à 8200 x g et le surnageant est éliminé.

Ensuite, 750 µl de tampon de lyse sont ajoutés sur les billes. Les tubes sont placés 5 min sur l'agitateur circulaire à 4 °C, centrifugés 30 s à 8200 x *g* et les surnageants sont éliminés. Cette étape est répétée 4 fois au total.

Après le dernier lavage, les billes sont resuspendues dans 20 µl de tampon de lyse et 20 µl de tampon dénaturant SDS 2X et sont ébouillantées 5 min à 99 °C.

Les surnageants sont ensuite étudiés par *Western-blot*.

IV.8. Immunoprécipitation des complexes ARNm-protéines (RNA- IP)

Les cellules sontensemencées dans des boîtes de culture circulaire de 100 mm de diamètre (1,5 million par boîte) pendant 4 jours afin d'obtenir la confluence du tapis cellulaire.

Les complexes ARN-protéines sont fixés par un traitement de 10 min à T°amb. avec du formaldéhyde à 1% (v/v, Sigma Aldrich).

La réaction est arrêtée par l'ajout pendant 5 min de 5 ml d'une solution 0,25 M de glycine préparée dans du D-PBS 1X.

Les cellules sont ensuite rincées avec 5 ml de D-PBS 1X froid puis par 2 ml de D-PBS 1X froid contenant les inhibiteurs de protéases suivants :

- Leupeptine 10 µg/ml
- Aprotinine 10 µg/ml
- EDTA 0,2 mM
- PMSF 0,5 mM

Le tapis cellulaire est ensuite gratté dans 3 ml de D-PBS 1X froid contenant les inhibiteurs de protéases. Le lysat obtenu est transféré dans un tube à centrifuger de 50 ml. Un nouveau grattage de la boîte de pétri est réalisé afin de collecter les cellules restantes avec 2 ml de D-PBS 1X froid contenant les inhibiteurs de protéases.

La suspension est ensuite centrifugée à 4°C pendant 5 min à 685 x *g* et le surnageant est éliminé.

Le culot est mis en suspension dans 1 ml du tampon de lyse suivant :

- Tris-HCl 50 mM pH 7,5
- NP40 1% (v/v)
- Désoxycholate de sodium 0,5 % (m/v)
- SDS 0,05 % (m/v)

- EDTA 1 mM
- NaCl 150 mM
- Solution de vanadyl ribonucleoside 2 mM (Sigma-Aldrich)
- RNAsin 100 U/ml (Promega)
- Cocktail d'inhibiteurs de proteases 25 µl/ml (Sigma-Aldrich)

Puis, les cellules sont soniquées pendant 2 cycles de 5 min (30 s ON, 30 s OFF) avec le Bioruptor (Diagenode).

Le matériel insoluble est éliminé au cours d'une centrifugation de 10 min à 16 000 x *g* à 4°C. Le surnageant est recueilli dans un microtube de 1,5 ml.

Une étape de lavage est réalisée en ajoutant 75 µl de billes de protéine A, préalablement lavées à l'aide du tampon de lyse comme décrit au paragraphe IV.7, en plaçant le tube contenant les billes et le surnageant 1 h sur un agitateur circulaire à 4 °C.

A la fin de cette étape, le tube est centrifugé pendant 5 min à 1 300 x *g* à 4°C et le surnageant est transféré dans un microtube de 1,5 ml et 15 µg d'anticorps primaire sont ajoutés (anti-galectine-3 Santa Cruz sc-20157 ; anti-HuR Santa Cruz sc-20694 ; anti-TIA-1 Santa Cruz sc-1751 ; IgG irrelevante Millipore 12-370 ; IgG irrelevante Santa Cruz sc-2028). Le mélange est placé une nuit à 4 °C en rotation.

Le lendemain, 75 µl de billes de protéine A lavées sont ajoutées dans le tube qui est placé en rotation à 4 °C pendant 4 h.

Après centrifugation pendant 45 s à 3300 x *g* à 4 °C, le surnageant est éliminé et les billes sont lavées par 1 ml de tampon de lyse contenant soit 1 M (spécifiquement pour l'anticorps anti-galectine-3 et l'isotype correspondant) ou 3 M d'urée (pour les autres anticorps et les isotypes correspondants).

Le tube est placé en rotation 10 min à 4 °C puis centrifugé à 4 °C à 3300 x *g* et le surnageant est éliminé afin de pouvoir effectuer un nouveau lavage des billes. Au total 5 lavages sont réalisés par échantillon.

A l'issue du dernier lavage, les billes sont mises en suspension dans 100 µl de tampon de resuspension :

- Tris HCl 50 mM pH 7,0
- EDTA 5 mM
- DTT 10 mM
- SDS 1 % (m/v)

La réversion du *crosslinking* est effectuée en incubant le tube 45 min à 70°C au bain-marie.

Les ARN sont ensuite extraits par l'ajout de 2 volumes de phénol-chloroforme-alcool isoamylique pH 4,5 (Thermo Fisher Scientific) par volume de lysat. Le tube est vortexé puis incubé 5 min à T°amb et centrifugé 5 min à 11 700 x g pendant 5 min. La phase supérieure aqueuse est transférée dans un microtube de 1,5 ml siliconé. Son volume est noté Vi.

De l'éthanol absolu (2,5 x Vi), de l'acétate de sodium 3M (1/10 x Vi), du glycoblue (20 µg/µl, Life Technologies) sont ensuite ajoutés afin de précipiter les ARN en plaçant le tube une nuit à -20°C.

Le lendemain, celui-ci est centrifugé à 4°C pendant 20 min à 16 000 x g, le surnageant est éliminé. Le culot est séché au Speed Vac pendant 10 min à chaud, puis resuspendu dans 20 µl d'eau.

Afin d'éliminer toute trace d'ADN, un traitement par la DNase (Macherey Nagel) est effectué en ajoutant 2 µl de tampon de l'enzyme et 0,2 µl de DNase dans le tube qui est placé 10 min à 37 °C puis centrifugé 30 s à 8 200 x g. A l'issue de cette étape, une nouvelle précipitation des ARN est réalisée comme décrite précédemment.

Le lendemain, le tube est centrifugé à 4 °C pendant 20 min à 16000 x g, le surnageant est éliminé et le culot est séché au Speed Vac pendant 10 min à chaud. Celui-ci est ensuite resuspendu dans 12 µl d'eau stérile (Niranjanakumari *et al.*, 2002; Peritz *et al.*, 2006).

Les échantillons sont ensuite analysés soit par RT-qPCR (*MUC4*) soit par RNA-sequencing (collaboration avec la plateforme de génomique fonctionnelle du Dr Martin Figeac, Lille) sur le séquenceur PGM (Life Technologies) afin de déterminer l'ensemble des ARNm immunoprécipités.

IV.9. RNA-pull down

IV.9.1. Synthèse de la sonde ARN et marquage avec la biotine

La préparation de la sonde débute par une étape de PCR sur de l'ADNc de cellules Sc à l'aide d'amorces spécifiques entourant le site de liaison à étudier et possédant, pour l'amorce sens, un site reconnu par la T7 Polymérase (Abdelmohsen *et al.*, 2011) :

- Amorce sens : 5' CCAAGCTTCTAATACGACTCACTATAGGGAGA-séquence 3' avec séquence = AGGCTGTCAGGATTCTTCA afin d'amplifier la région correspondant au *CA repeat* présent dans le 3'UTR de *MUC4*.
- Amorce antisens : 5' CAAAAAGTCAGAGAGACTT 3'.

L'amplification par PCR est réalisée avec le programme suivant sur le thermocycleur « MJ Mini™ » (BioRad) : 94 °C, 2 min ; 36 cycles [94 °C, 45 s ; 51 °C, 1 min ; 72 °C, 1 min] ; 72°C, 10 min puis 4 °C à l'infini.

Cette étape est suivie d'une électrophorèse en gel d'agarose du produit de PCR et d'une purification sur gel avec le kit « Wizard SV gel and PCR clean-up system » (Promega) comme au paragraphe III.8. Le produit PCR est quantifié à l'aide d'un spectrophotomètre (Nanodrop®).

La conversion de l'ADNc en ARN et le marquage de la sonde avec la biotine est réalisée avec le kit « Biotin labeling mix » (Roche). Pour se faire, le mélange suivant est incubé 2 h à 37 °C:

- Produit PCR	100 ng (volume max 8,5 µl)
- Biotine	2 µl
- Tampon transcription 5X	4 µl
- DTT 100 mM	2 µl
- RNAsine	0,5 µl
- T7 polymérase (15-20 U)	1 µl
- Eau stérile	qsp 18 µl

Ensuite, un traitement par 2 µl de DNase I diluée dans son tampon (Machery Nagel) est réalisé pendant 10 min à 37 °C. La réaction est arrêtée par l'ajout de 2 µl d'EDTA 0,2 M pH 8,0.

Les ARN obtenus sont précipités une nuit à -20 °C avec 3,3 µl d'acétate de sodium 2 M, 56 µl d'éthanol absolu et 0,3 µl de glycobule.

Le mélange est centrifugé pendant 20 min à 16 000 x *g* à 4 °C puis le culot obtenu est séché à l'aide du Speed Vac pendant 5 min et repris par 20 µl d'eau stérile.

La sonde obtenue est dosée à l'aide d'un spectrophotomètre (Nanodrop®).

IV.9.2. Interaction ARN-protéines

Pour permettre la fixation des protéines sur la sonde, 100 µg d'extraits cellulaires totaux sont dilués dans du tampon de lyse (composition décrite au paragraphe IV.3) dépourvu de NaCl afin de ramener la concentration saline (NaCl) à 60 mM. Ceux-ci sont incubés avec 3 µg de sonde ARN marquée à la biotine pendant 30 min à 25 °C.

IV.9.3. Purification des complexes ARN-protéines

La quantité totale de billes de streptavidine souhaitée (30 µl par condition) (Dynabeads M280 Streptavidin, Life Technologies) est prélevée et celles-ci sont lavées 2 fois avec 1 ml de tampon A (NaCl 0,05 M, NaOH 0,1 M) et une fois avec 1 ml de tampon B (NaCl 100 mM). Pour cela, 1 ml de tampon est ajouté dans le tube et celui-ci est placé 2 min sur l'agitateur circulaire puis 1 min sur le portoir magnétique et le surnageant est éliminé.

Après les lavages, les billes sont reprises avec le même volume de tampon B' (NaCl 60 mM) que celui correspondant au volume de billes prélevé initialement. Puis 30 µl de billes sont ajoutés dans chaque tube contenant les complexes ARN-protéines. Ceux-ci sont placés sur l'agitateur circulaire pendant 30 min puis sur le portoir magnétique pendant 2 min et le surnageant est éliminé.

Puis, 1 ml de tampon B' est ajouté aux billes. Le mélange est placé 5 min sur l'agitateur circulaire puis 2 min sur le portoir magnétique et le surnageant est éliminé. Cette étape est répétée une fois.

Le culot de billes est ensuite repris par 15 µl de tampon de charge SDS loading 2X utilisé pour le *Western-blot* et les échantillons sont placés 5 min à 99 °C afin de dissocier les protéines des billes. Les surnageants obtenus sont ensuite étudiés par *Western Blot*.

V. Statistiques

Les données ont été analysées à l'aide du test de Student. Les résultats avec un p inférieur à 0,05 sont considérés comme statistiquement significatifs.

Bibliographie

A

Abdelmohsen, K., Tominaga, K., Lee, E.K., Srikantan, S., Kang, M.-J., Kim, M.M., Selimyan, R., Martindale, J.L., Yang, X., Carrier, F., *et al.* (2011). Enhanced translation by Nucleolin via G-rich elements in coding and non-coding regions of target mRNAs. *Nucleic Acids Res.* 39, 8513–8530.

Adeli, K. (2011). Translational control mechanisms in metabolic regulation: critical role of RNA binding proteins, microRNAs, and cytoplasmic RNA granules. *Am. J. Physiol. Endocrinol. Metab.* 301, E1051-1064.

Advedissian, T., Deshayes, F., Poirier, F., Grandjean, C., and Viguier, M. (2015). Galectins, a class of unconventional lectins. *Med. Sci. MS* 31, 499–505.

Aizer, A., Kalo, A., Kafri, P., Shraga, A., Ben-Yishay, R., Jacob, A., Kinor, N., and Shav-Tal, Y. (2014). Quantifying mRNA targeting to P-bodies in living human cells reveals their dual role in mRNA decay and storage. *J. Cell Sci.* 127, 4443–4456.

Albrecht, H., and Carraway, K.L. (2011). MUC1 and MUC4: switching the emphasis from large to small. *Cancer Biother. Radiopharm.* 26, 261–271.

Anderson, P., Kedersha, N., and Ivanov, P. (2015). Stress granules, P-bodies and cancer. *Biochim. Biophys. Acta* 1849, 861–870.

Andrianifahanana, M., Moniaux, N., and Batra, S.K. (2006). Regulation of mucin expression: mechanistic aspects and implications for cancer and inflammatory diseases. *Biochim. Biophys. Acta* 1765, 189–222.

Apostolopoulos, V., Stojanovska, L., and Gargosky, S.E. (2015). MUC1 (CD227): a multi-tasked molecule. *Cell. Mol. Life Sci. CMLS* 72, 4475–4500.

Argüeso, P., Guzman-Aranguez, A., Mantelli, F., Cao, Z., Ricciuto, J., and Panjwani, N. (2009). Association of cell surface mucins with galectin-3 contributes to the ocular surface epithelial barrier. *J. Biol. Chem.* 284, 23037–23045.

Argüeso, P., Mauris, J., and Uchino, Y. (2015). Galectin-3 as a regulator of the epithelial junction: Implications to wound repair and cancer. *Tissue Barriers* 3.

Audie, J.P., Janin, A., Porchet, N., Copin, M.C., Gosselin, B., and Aubert, J.P. (1993). Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by in situ hybridization. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* 41, 1479–1485.

B

Baek, J.-H., Kim, S.-J., Kang, H.G., Lee, H.-W., Kim, J.-H., Hwang, K.-A., Song, J., and Chun, K.-H. (2015). Galectin-3 activates PPAR γ and supports white adipose tissue formation and high-fat diet-induced obesity. *Endocrinology* 156, 147–156.

Bailey, P., Chang, D.K., Nones, K., Johns, A.L., Patch, A.-M., Gingras, M.-C., Miller, D.K., Christ, A.N., Bruxner, T.J.C., Quinn, M.C., *et al.* (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 531, 47–52.

- Balagopal, V., and Parker, R. (2009). Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. *Curr. Opin. Cell Biol.* 21, 403–408.
- Bardeesy, N., and DePinho, R.A. (2002). Pancreatic cancer biology and genetics. *Nat. Rev. Cancer* 2, 897–909.
- Barondes, S.H., Castronovo, V., Cooper, D.N., Cummings, R.D., Drickamer, K., Feizi, T., Gitt, M.A., Hirabayashi, J., Hughes, C., and Kasai, K. (1994). Galectins: a family of animal beta-galactoside-binding lectins. *Cell* 76, 597–598.
- Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res.* 33, 7138–7150.
- Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., *et al.* (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 603–607.
- Bensaude, O. (2011). Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription* 2, 103–108.
- Berberat, P.O., Friess, H., Wang, L., Zhu, Z., Bley, T., Frigeri, L., Zimmermann, A., and Büchler, M.W. (2001). Comparative analysis of galectins in primary tumors and tumor metastasis in human pancreatic cancer. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* 49, 539–549.
- Bichara, M., Attmane-Elakeb, A., Brown, D., Essig, M., Karim, Z., Muffat-Joly, M., Micheli, L., Eude-Le Parco, I., Cluzeaud, F., Peuchmaur, M., *et al.* (2006). Exploring the role of galectin 3 in kidney function: a genetic approach. *Glycobiology* 16, 36–45.
- Blanchard, H., Yu, X., Collins, P.M., and Bum-Erdene, K. (2014). Galectin-3 inhibitors: a patent review (2008-present). *Expert Opin. Ther. Pat.* 24, 1053–1065.
- Bolisetty, M.T., and Beemon, K.L. (2012). Splicing of internal large exons is defined by novel cis-acting sequence elements. *Nucleic Acids Res.* 40, 9244–9254.
- Boon, J.Y., Dusonchet, J., Trengrove, C., and Wolozin, B. (2014). Interaction of LRRK2 with kinase and GTPase signaling cascades. *Front. Mol. Neurosci.* 7, 64.
- Bosanquet, D.C., Ye, L., Harding, K.G., and Jiang, W.G. (2014). FERM family proteins and their importance in cellular movements and wound healing (review). *Int. J. Mol. Med.* 34, 3–12.
- Boundedjah, O., Desforges, B., Wu, T.-D., Pioche-Durieu, C., Marco, S., Hamon, L., Curmi, P.A., Guerquin-Kern, J.-L., Piétrement, O., and Pastré, D. (2014). Free mRNA in excess upon polysome dissociation is a scaffold for protein multimerization to form stress granules. *Nucleic Acids Res.* 42, 8678–8691.
- Bracken, C.P., Szubert, J.M., Mercer, T.R., Dinger, M.E., Thomson, D.W., Mattick, J.S., Michael, M.Z., and Goodall, G.J. (2011). Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic cleavage. *Nucleic Acids Res.* 39, 5658–5668.

Braeuer, R.R., Shoshan, E., Kamiya, T., and Bar-Eli, M. (2012). The sweet and bitter sides of galectins in melanoma progression. *Pigment Cell Melanoma Res.* 25, 592–601.

Brewer, J.W., and Diehl, J.A. (2000). PERK mediates cell-cycle exit during the mammalian unfolded protein response. *Proc. Natl. Acad. Sci. U. S. A.* 97, 12625–12630.

Brewer, J.W., Hendershot, L.M., Sherr, C.J., and Diehl, J.A. (1999). Mammalian unfolded protein response inhibits cyclin D1 translation and cell-cycle progression. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8505–8510.

Brook, M., and Gray, N.K. (2012). The role of mammalian poly(A)-binding proteins in coordinating mRNA turnover. *Biochem. Soc. Trans.* 40, 856–864.

Buisine, M.P., Devisme, L., Savidge, T.C., Gespach, C., Gosselin, B., Porchet, N., and Aubert, J.P. (1998). Mucin gene expression in human embryonic and fetal intestine. *Gut* 43, 519–524.

Buisine, M.P., Devisme, L., Copin, M.C., Durand-Réville, M., Gosselin, B., Aubert, J.P., and Porchet, N. (1999). Developmental mucin gene expression in the human respiratory tract. *Am. J. Respir. Cell Mol. Biol.* 20, 209–218.

Buisine MP, Porchet N, Van Seuningen I. (2008) Mucin expression and regulation during development and cell differentiation. In Van Seuningen, I. (sous la direction de). *Mucin book: The epithelial mucins, Structure/function. Roles in cancer and inflammatory diseases.* Trivandrum (IND), Research Signpost, p.:75-94.

Burgess, H.M., Richardson, W.A., Anderson, R.C., Salaun, C., Graham, S.V., and Gray, N.K. (2011). Nuclear relocalisation of cytoplasmic poly(A)-binding proteins PABP1 and PABP4 in response to UV irradiation reveals mRNA-dependent export of metazoan PABPs. *J. Cell Sci.* 124, 3344–3355.

C

Caffery, B., Joyce, E., Heynen, M.L., Jones, L., Ritter, R., Gamache, D.A., and Senchyna, M. (2008). MUC16 expression in Sjogren's syndrome, KCS, and control subjects. *Mol. Vis.* 14, 2547–2555.

Calabretta, S., and Richard, S. (2015). Emerging Roles of Disordered Sequences in RNA-Binding Proteins. *Trends Biochem. Sci.* 40, 662–672.

Cameron, T.L., Bell, K.M., Gresshoff, I.L., Sampurno, L., Mullan, L., Ermann, J., Glimcher, L.H., Boot-Handford, R.P., and Bateman, J.F. (2015). XBP1-Independent UPR Pathways Suppress C/EBP- β Mediated Chondrocyte Differentiation in ER-Stress Related Skeletal Disease. *PLoS Genet.* 11, e1005505.

Carmody, S.R., and Wente, S.R. (2009). mRNA nuclear export at a glance. *J. Cell Sci.* 122, 1933–1937.

- Carraway, K.L., Rossi, E.A., Komatsu, M., Price-Schiavi, S.A., Huang, D., Guy, P.M., Carvajal, M.E., Fregien, N., Carraway, C.A., and Carraway, K.L. (1999). An intramembrane modulator of the ErbB2 receptor tyrosine kinase that potentiates neuregulin signaling. *J. Biol. Chem.* 274, 5263–5266.
- Carraway, K.L., Ramsauer, V.P., Haq, B., and Carothers Carraway, C.A. (2003). Cell signaling through membrane mucins. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 25, 66–71.
- Carraway, K.L., Theodoropoulos, G., Kozloski, G.A., and Carothers Carraway, C.A. (2009). Muc4/MUC4 functions and regulation in cancer. *Future Oncol. Lond. Engl.* 5, 1631–1640.
- Carvalho, R.S., Fernandes, V.C., Nepomuceno, T.C., Rodrigues, D.C., Woods, N.T., Suarez-Kurtz, G., Chammas, R., Monteiro, A.N., and Carvalho, M.A. (2014). Characterization of LGALS3 (galectin-3) as a player in DNA damage response. *Cancer Biol. Ther.* 15, 840–850.
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., *et al.* (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406.
- Chaturvedi, P., Singh, A.P., and Batra, S.K. (2008). Structure, evolution, and biology of the MUC4 mucin. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 22, 966–981.
- Choudhury, A., Moniaux, N., Winpenny, J.P., Hollingsworth, M.A., Aubert, J.P., and Batra, S.K. (2000). Human MUC4 mucin cDNA and its variants in pancreatic carcinoma. *J. Biochem. (Tokyo)* 128, 233–243.
- Chugh, S., Gnanapragassam, V.S., Jain, M., Rachagani, S., Ponnusamy, M.P., and Batra, S.K. (2015). Pathobiological implications of mucin glycans in cancer: Sweet poison and novel targets. *Biochim. Biophys. Acta BBA - Rev. Cancer* 1856, 211–225.
- Ciafrè, S.A., and Galardi, S. (2013). microRNAs and RNA-binding proteins: a complex network of interactions and reciprocal regulations in cancer. *RNA Biol.* 10, 935–942.
- Cléry, A., and Allain, F.H.-T. (2013). From structure to function of RNA binding domains (Landes Bioscience).
- Colnot, C., Sidhu, S.S., Balmain, N., and Poirier, F. (2001). Uncoupling of chondrocyte death and vascular invasion in mouse galectin 3 null mutant bones. *Dev. Biol.* 229, 203–214.
- Cook, K.B., Kazan, H., Zuberi, K., Morris, Q., and Hughes, T.R. (2011). RBPDB: a database of RNA-binding specificities. *Nucleic Acids Res.* 39, D301-308.
- Cooper, D.N.W. (2002). Galectinomics: finding themes in complexity. *Biochim. Biophys. Acta* 1572, 209–231.
- Coppin, L. (2012) Rôle de la galectine-3 dans la régulation de l'expression des mucines membranaires MUC1 et MUC4 dans un modèle d'adénocarcinome pancréatique canalaire humain. Mémoire pour le diplôme d'études spécialisées de Biologie Médicale et de Docteur en Pharmacie dirigé par le Pr Pascal Pigny

Coppin, L., Benomar, K., Corfiotti, F., Cattan, S., Renaud, F., Lapere, C., Leteurtre, E., Vantyghem, M.-C., Truant, S., and Pigny, P. (2016). CA-125, but not galectin-3, complements CA 19-9 for discriminating ductal adenocarcinoma versus non-malignant pancreatic diseases. *Pancreatol. Off. J. Int. Assoc. Pancreatol. IAP AI* 16, 115–120.

Corfield, A.P. (2015). Mucins: a biologically relevant glycan barrier in mucosal protection. *Biochim. Biophys. Acta* 1850, 236–252.

Cornejo-García, J.A., Romano, A., Guéant-Rodríguez, R.M., Oussalah, A., Blanca-López, N., Gaeta, F., Tramoy, D., Josse, T., Doña, I., Torres, M.J., *et al.* (2016). A non-synonymous polymorphism in galectin-3 lectin domain is associated with allergic reactions to beta-lactam antibiotics. *Pharmacogenomics J.* 16, 79–82.

D

Dagher, S.F., Wang, J.L., and Patterson, R.J. (1995). Identification of galectin-3 as a factor in pre-mRNA splicing. *Proc. Natl. Acad. Sci. U. S. A.* 92, 1213–1217.

Danguy, A., Camby, I., and Kiss, R. (2002). Galectins and cancer. *Biochim. Biophys. Acta* 1572, 285–293.

Das, S., and Batra, S.K. (2015). Understanding the Unique Attributes of MUC16 (CA125): Potential Implications in Targeted Therapy. *Cancer Res.* 75, 4669–4674.

Das, S., Rachagani, S., Sheinin, Y., Smith, L.M., Gurumurthy, C.B., Roy, H.K., and Batra, S.K. (2016). Mice deficient in Muc4 are resistant to experimental colitis and colitis-associated colorectal cancer. *Oncogene* 35, 2645–2654.

Dawkins, J.B.N., Wang, J., Maniati, E., Heward, J.A., Koniali, L., Kocher, H.M., Martin, S.A., Chelala, C., Balkwill, F.R., Fitzgibbon, J., *et al.* (2016). Reduced Expression of Histone Methyltransferases KMT2C and KMT2D Correlates with Improved Outcome in Pancreatic Ductal Adenocarcinoma. *Cancer Res.* 76, 4861–4871.

Debailleul, V., Laine, A., Huet, G., Mathon, P., d’Hooghe, M.C., Aubert, J.P., and Porchet, N. (1998). Human mucin genes MUC2, MUC3, MUC4, MUC5AC, MUC5B, and MUC6 express stable and extremely large mRNAs and exhibit a variable length polymorphism. An improved method to analyze large mRNAs. *J. Biol. Chem.* 273, 881–890.

Decker, C.J., and Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb. Perspect. Biol.* 4, a012286.

DeGracia, D.J., Jamison, J.T., Szymanski, J.J., and Lewis, M.K. (2008). Translation arrest and ribonemics in post-ischemic brain: layers and layers of players. *J. Neurochem.* 106, 2288–2301.

Dekker, J., Rossen, J.W.A., Büller, H.A., and Einerhand, A.W.C. (2002). The MUC family: an obituary. *Trends Biochem. Sci.* 27, 126–131.

Delacour, D., Koch, A., Ackermann, W., Eude-Le Parco, I., Elsässer, H.-P., Poirier, F., and Jacob, R. (2008). Loss of galectin-3 impairs membrane polarisation of mouse enterocytes in vivo. *J. Cell Sci.* 121, 458–465.

Delacour, D., Koch, A., and Jacob, R. (2009). The Role of Galectins in Protein Trafficking. *Traffic* 10, 1405–1413.

Distler, M., Aust, D., Weitz, J., Pilarsky, C., Grü, and Tzmann, R. (2014). Precursor Lesions for Sporadic Pancreatic Cancer: PanIN, IPMN, and MCN. *BioMed Res. Int.* 2014, e474905.

Drin, G., Bigay, J., and Antonny, B. (2009). Regulation of vesicular transport by membrane curvature. *Med. Sci. MS* 25, 483–488.

Dudas, S.P., Yunker, C.K., Sternberg, L.R., Byrd, J.C. and Bresalier, R.S. (2002). Expression of human intestinal mucin is modulated by the beta-galactoside binding protein galectin-3 in colon cancer. *Gastroenterology*. 123, 817-826.

Dumic, J., Dabelic, S., and Flögel, M. (2006). Galectin-3: An open-ended story. *Biochim. Biophys. Acta BBA - Gen. Subj.* 1760, 616–635.

E

Eisenstein, M. (2015). Startups use short-read data to expand long-read sequencing market. *Nat. Biotechnol.* 33, 433–435.

Elad-Sfadia, G., Haklai, R., Balan, E., and Kloog, Y. (2004). Galectin-3 augments K-Ras activation and triggers a Ras signal that attenuates ERK but not phosphoinositide 3-kinase activity. *J. Biol. Chem.* 279, 34922–34930.

Elola, M.T., Blidner, A.G., Ferragut, F., Bracalente, C., and Rabinovich, G.A. (2015). Assembly, organization and regulation of cell-surface receptors by lectin-glycan complexes. *Biochem. J.* 469, 1–16.

Elvira, G., Wasiak, S., Blandford, V., Tong, X.-K., Serrano, A., Fan, X., Sánchez-Carbente, M. del R., Servant, F., Bell, A.W., Boismenu, D., *et al.* (2006). Characterization of an RNA Granule from Developing Brain. *Mol. Cell. Proteomics* 5, 635–651.

Escande, F., Lemaitre, L., Moniaux, N., Batra, S.K., Aubert, J.-P., and Buisine, M.-P. (2002). Genomic organization of MUC4 mucin gene. Towards the characterization of splice variants. *Eur. J. Biochem.* 269, 3637–3644.

Eude-Le Parco, I., Gendronneau, G., Dang, T., Delacour, D., Thijssen, V.L., Edelmann, W., Peuchmaur, M., and Poirier, F. (2009). Genetic assessment of the importance of galectin-3 in cancer initiation, progression, and dissemination in mice. *Glycobiology* 19, 68–75.

Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007). P bodies: at the crossroads of post-transcriptional pathways. *Nat. Rev. Mol. Cell Biol.* 8, 9–22.

F

Fabian, M.R., Sonenberg, N., and Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* 79, 351–379.

Fauquette, V., Aubert, S., Groux-Degroote, S., Hemon, B., Porchet, N., Van Seuningen, I., and Pigny, P. (2007). Transcription factor AP-2alpha represses both the mucin MUC4 expression and pancreatic cancer cell proliferation. *Carcinogenesis* 28, 2305–2312.

Fawal, M., Jean-Jean, O., Vanzo, N., and Morello, D. (2011). Novel mRNA-containing cytoplasmic granules in ALK-transformed cells. *Mol. Biol. Cell* 22, 726–735.

Fei, F., Joo, E.J., Tarighat, S.S., Schiffer, I., Paz, H., Fabbri, M., Abdel-Azim, H., Groffen, J., and Heisterkamp, N. (2015). B-cell precursor acute lymphoblastic leukemia and stromal cells communicate through Galectin-3. *Oncotarget* 6, 11378–11394.

Ferlay, J., Partensky, C., and Bray, F. (2016). More deaths from pancreatic cancer than breast cancer in the EU by 2017. *Acta Oncol. Stockh. Swed.* 55, 1158–1160.

Fischer, M. (2017). Census and evaluation of p53 target genes. *Oncogene*.

Fischer, B.M., Cuellar, J.G., Diehl, M.L., deFreytas, A.M., Zhang, J., Carraway, K.L., and Voynow, J.A. (2003). Neutrophil elastase increases MUC4 expression in normal human bronchial epithelial cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 284, L671-679.

Fortuna-Costa, A., Gomes, A.M., Kozlowski, E.O., Stelling, M.P., and Pavão, M.S.G. (2014). Extracellular galectin-3 in tumor progression and metastasis. *Mol. Cell. Oncol.* 4, 138.

Fredriksson, L., Wink, S., Herpers, B., Benedetti, G., Hadi, M., de Bont, H., Groothuis, G., Luijten, M., Danen, E., de Graauw, M., *et al.* (2014). Drug-induced endoplasmic reticulum and oxidative stress responses independently sensitize toward TNF α -mediated hepatotoxicity. *Toxicol. Sci. Off. J. Soc. Toxicol.* 140, 144–159.

Fritsch, K., Mernberger, M., Nist, A., Stiewe, T., Brehm, A., and Jacob, R. (2016). Galectin-3 interacts with components of the nuclear ribonucleoprotein complex. *BMC Cancer* 16, 502.

Fu, X.-D., and Ares, M. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.* 15, 689–701.

Fu, X., Gao, X., Ge, L., Cui, X., Su, C., Yang, W., Sun, X., Zhang, W., Yao, Z., Yang, X., *et al.* (2016). Malonate induces the assembly of cytoplasmic stress granules. *FEBS Lett.* 590, 22–33.

Fukumori, T., Kanayama, H.-O., and Raz, A. (2007). The role of galectin-3 in cancer drug resistance. *Drug Resist. Updat. Rev. Comment. Antimicrob. Anticancer Chemother.* 10, 101–108.

Funasaka, T., Raz, A., and Nangia-Makker, P. (2014). Nuclear transport of galectin-3 and its therapeutic implications. *Semin. Cancer Biol.* 27, 30–38.

G

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* 8, 113–126.

- Garnett, M.J., Mansfeld, J., Godwin, C., Matsusaka, T., Wu, J., Russell, P., Pines, J., and Venkitaraman, A.R. (2009). UBE2S elongates ubiquitin chains on APC/C substrates to promote mitotic exit. *Nat. Cell Biol.* *11*, 1363–1369.
- Garzon, R., Marcucci, G., and Croce, C.M. (2010). Targeting microRNAs in cancer: rationale, strategies and challenges. *Nat. Rev. Drug Discov.* *9*, 775–789.
- Ge, X.N., Bahaie, N.S., Kang, B.N., Hosseinkhani, M.R., Ha, S.G., Frenzel, E.M., Liu, F.-T., Rao, S.P., and Sriramaraio, P. (2010). Allergen-induced airway remodeling is impaired in galectin-3-deficient mice. *J. Immunol. Baltim. Md 1950* *185*, 1205–1214.
- Gerstberger, S., Hafner, M., and Tuschl, T. (2014a). A census of human RNA-binding proteins. *Nat. Rev. Genet.* *15*, 829–845.
- Gerstberger, S., Hafner, M., Ascano, M., and Tuschl, T. (2014b). Evolutionary conservation and expression of human RNA-binding proteins and their role in human genetic disease. *Adv. Exp. Med. Biol.* *825*, 1–55.
- Geuens, T., Bouhy, D., and Timmerman, V. (2016). The hnRNP family: insights into their role in health and disease. *Hum. Genet.* *135*, 851–867.
- Glavey, S.V., Huynh, D., Reagan, M.R., Manier, S., Moschetta, M., Kawano, Y., Roccaro, A.M., Ghobrial, I.M., Joshi, L., and O'Dwyer, M.E. (2015). The cancer glycome: carbohydrates as mediators of metastasis. *Blood Rev.* *29*, 269–279.
- Glinskii, O.V., Sud, S., Mossine, V.V., Mawhinney, T.P., Anthony, D.C., Glinsky, G.V., Pienta, K.J., and Glinsky, V.V. (2012). Inhibition of prostate cancer bone metastasis by synthetic TF antigen mimic/galectin-3 inhibitor lactulose-L-leucine. *Neoplasia N. Y. N* *14*, 65–73.
- Goley, E.D., and Welch, M.D. (2006). The ARP2/3 complex: an actin nucleator comes of age. *Nat. Rev. Mol. Cell Biol.* *7*, 713–726.
- Gorgoni, B., and Gray, N.K. (2004). The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: a developmental perspective. *Brief. Funct. Genomic. Proteomic.* *3*, 125–141.
- Goss, D.J. and Kleiman, F.E. (2013). Poly(A) binding proteins: are they all created equal? *Wiley Interdiscip Rev RNA.* *4*, 167-179.
- Graille, M. and Séraphin, B. (2012). Surveillance pathways rescuing eukaryotic ribosomes lost in translation. *Nat Rev Mol Cell Biol.* *13*, 727-735.
- Groux-Degroote, S., Harduin-Lepers, A. and Delannoy, P. (2008) Biosynthesis of mucin O-glycan chains in normal and pathological states. In Van Seuning, I. (sous la direction de). *Mucin book: The epithelial mucins, Structure/function. Roles in cancer and inflammatory diseases*. Trivandrum (IND), Research Signpost, p 39-54.

Guha, P., Kaptan, E., Bandyopadhyaya, G., Kaczanowska, S., Davila, E., Thompson, K., Martin, S.S., Kalvakolanu, D.V., Vasta, G.R., and Ahmed, H. (2013). Cod glycopeptide with picomolar affinity to galectin-3 suppresses T-cell apoptosis and prostate cancer metastasis. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 5052–5057.

Guilgur, L.G., Prudêncio, P., Sobral, D., Liszekova, D., Rosa, A., and Martinho, R.G. (2014). Requirement for highly efficient pre-mRNA splicing during *Drosophila* early embryonic development. *eLife* *3*, e02181.

Guo, X., Zheng, S., Dang, H., Pace, R.G., Stonebraker, J.R., Jones, C.D., Boellmann, F., Yuan, G., Haridass, P., Fedrigo, O., *et al.* (2014). Genome reference and sequence variation in the large repetitive central exon of human MUC5AC. *Am. J. Respir. Cell Mol. Biol.* *50*, 223–232.

Gupta, A., Kailasam, S., and Bansal, M. (2016). Insights into the Structural Dynamics of Nucleocytoplasmic Transport of tRNA by Exportin-t. *Biophys. J.* *110*, 1264–1279.

H

Hamilton, B.J., Wang, X.-W., Collins, J., Bloch, D., Bergeron, A., Henry, B., Terry, B.M., Zan, M., Mouland, A.J., and Rigby, W.F.C. (2008). Separate cis-trans Pathways Post-transcriptionally Regulate Murine CD154 (CD40 Ligand) Expression : a novel function for ca repeats in the 3'-untranslated region. *J. Biol. Chem.* *283*, 25606–25616.

Hammel, P. (2006). Pancreatic cancer: what is new in 2005 ?. *Bull. Cancer (Paris)* *93*, 67–72.

Hammel, P., Neuzillet, C., Bendaoud, S., Lekhal, C., de Mestier, L., and Hentic, O. (2015). Pancreatic adenocarcinoma: 10 years of progress. *Bull. Cancer (Paris)* *102*, S62-71.

Hanisch, F.G., and Müller, S. (2000). MUC1: the polymorphic appearance of a human mucin. *Glycobiology* *10*, 439–449.

Hanson, R.L., and Hollingsworth, M.A. (2016). Functional Consequences of Differential O-glycosylation of MUC1, MUC4, and MUC16 (Downstream Effects on Signaling). *Biomolecules* *6*.

Haridas, D., Ponnusamy, M.P., Chugh, S., Lakshmanan, I., Seshacharyulu, P., and Batra, S.K. (2014). MUC16: molecular analysis and its functional implications in benign and malignant conditions. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* *28*, 4183–4199.

Hatrup, C.L., and Gendler, S.J. (2008). Structure and function of the cell surface (tethered) mucins. *Annu. Rev. Physiol.* *70*, 431–457.

Haudek, K.C., Voss, P.G., Locascio, L.E., Wang, J.L., and Patterson, R.J. (2009). A mechanism for incorporation of galectin-3 into the spliceosome through its association with U1 snRNP. *Biochemistry (Mosc.)* *48*, 7705–7712.

Haudek, K.C., Spronk, K.J., Voss, P.G., Patterson, R.J., Wang, J.L., and Arnoys, E.J. (2010a). Dynamics of galectin-3 in the nucleus and cytoplasm. *Biochim. Biophys. Acta* *1800*, 181–189.

Hughes, R.C. (1997). The galectin family of mammalian carbohydrate-binding molecules. *Biochem. Soc. Trans.* 25, 1194–1198.

Hughes, R.C. (1999). Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochim. Biophys. Acta* 1473, 172–185.

Hui, J., Reither, G., and Bindereif, A. (2003). Novel functional role of CA repeats and hnRNP L in RNA stability. *RNA N. Y.* 9, 931–936.

I

Iacobini, C., Menini, S., Ricci, C., Blasetti Fantauzzi, C., Scipioni, A., Salvi, L., Cordone, S., Delucchi, F., Serino, M., Federici, M., *et al.* (2011). Galectin-3 ablation protects mice from diet-induced NASH: a major scavenging role for galectin-3 in liver. *J. Hepatol.* 54, 975–983.

Iio, E., Matsuura, K., Nishida, N., Maekawa, S., Enomoto, N., Nakagawa, M., Sakamoto, N., Yatsuhashi, H., Kurosaki, M., Izumi, N., *et al.* (2015). Genome-wide association study identifies a PSMD3 variant associated with neutropenia in interferon-based therapy for chronic hepatitis C. *Hum. Genet.* 134, 279–289.

Ito, D., Walker, J.R., Thompson, C.S., Moroz, I., Lin, W., Veselits, M.L., Hakim, A.M., Fienberg, A.A., and Thinakaran, G. (2004). Characterization of stanniocalcin 2, a novel target of the mammalian unfolded protein response with cytoprotective properties. *Mol. Cell. Biol.* 24, 9456–9469.

Ivanov, P., Kedersha, N., and Anderson, P. (2011). Stress puts TIA on TOP. *Genes Dev.* 25, 2119–2124.

Izeradjene, K., and Hingorani, S.R. (2007). Targets, trials, and travails in pancreas cancer. *J. Natl. Compr. Cancer Netw. JNCCN* 5, 1042–1053.

J

Jackson, R.J., Hellen, C.U. and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol.* 11 ; 113–127.

Jafarifar, F., Yao, P., Eswarappa, S.M., and Fox, P.L. (2011). Repression of VEGFA by CA-rich element-binding microRNAs is modulated by hnRNP L. *EMBO J.* 30, 1324–1334.

Janssens, S., Pulendran, B., and Lambrecht, B.N. (2014). Emerging functions of the unfolded protein response in immunity. *Nat. Immunol.* 15, 910–919.

Jensen, K.B., and Darnell, R.B. (2008). CLIP: crosslinking and immunoprecipitation of in vivo RNA targets of RNA-binding proteins. *Methods Mol. Biol. Clifton NJ* 488, 85–98.

Jeong, H., Sim, H.J., Song, E.K., Lee, H., Ha, S.C., Jun, Y., Park, T.J., and Lee, C. (2016). Crystal structure of SEL1L: Insight into the roles of SLR motifs in ERAD pathway. *Sci. Rep.* 6, 20261.

Jiang, H.-B., Xu, M., and Wang, X.-P. (2008). Pancreatic stellate cells promote proliferation and invasiveness of human pancreatic cancer cells via galectin-3. *World J. Gastroenterol.* 14, 2023–2028.

- Jiang, J., Westberg, J.A., and Andersson, L.C. (2012). Stanniocalcin 2, forms a complex with heme oxygenase 1, binds hemin and is a heat shock protein. *Biochem. Biophys. Res. Commun.* 421, 274–279.
- Jiang, L., Phang, J.M., Yu, J., Harrop, S.J., Sokolova, A.V., Duff, A.P., Wilk, K.E., Alkhamici, H., Breit, S.N., Valenzuela, S.M., *et al.* (2014). CLIC proteins, ezrin, radixin, moesin and the coupling of membranes to the actin cytoskeleton: a smoking gun? *Biochim. Biophys. Acta* 1838, 643–657.
- Jin, C., Rajabi, H., and Kufe, D. (2010). miR-1226 targets expression of the mucin 1 oncoprotein and induces cell death. *Int. J. Oncol.* 37, 61–69.
- Jonckheere, N., and Van Seuningen, I. (2008). The membrane-bound mucins: how large O-glycoproteins play key roles in epithelial cancers and hold promise as biological tools for gene-based and immunotherapies. *Crit. Rev. Oncog.* 14, 177–196.
- Jonckheere, N., and Van Seuningen, I. (2010). The membrane-bound mucins: From cell signalling to transcriptional regulation and expression in epithelial cancers. *Biochimie* 92, 1–11.
- Jonckheere, N., Skrypek, N., and Van Seuningen, I. (2010). Mucins and pancreatic cancer. *Cancers* 2, 1794–1812.
- Jonckheere, N., Skrypek, N., Merlin, J., Dessein, A.F., Dumont, P., Leteurtre, E., Harris, A., Desseyn, J.-L., Susini, C., Frénois, F., *et al.* (2012). The mucin MUC4 and its membrane partner ErbB2 regulate biological properties of human CAPAN-2 pancreatic cancer cells via different signalling pathways. *PloS One* 7, e32232.
- Jonckheere, N., Skrypek, N., Frénois, F., and Van Seuningen, I. (2013). Membrane-bound mucin modular domains: from structure to function. *Biochimie* 95, 1077–1086.
- Jonckheere, N., Skrypek, N., and Van Seuningen, I. (2014). Mucins and tumor resistance to chemotherapeutic drugs. *Biochim. Biophys. Acta* 1846, 142–151.
- Jonckheere, N., Lahdaoui, F., and Van Seuningen, I. (2015). Targeting MUC4 in pancreatic cancer: miRNAs. *Oncoscience* 2, 799–800.
- Jønson, L., Vikesaa, J., Krogh, A., Nielsen, L.K., Hansen, T. vO, Borup, R., Johnsen, A.H., Christiansen, J., and Nielsen, F.C. (2007). Molecular composition of IMP1 ribonucleoprotein granules. *Mol. Cell. Proteomics MCP* 6, 798–811.
- Joshi, A.D., Carter, D.E., Harper, T.A., and Elferink, C.J. (2015). Aryl hydrocarbon receptor-dependent stanniocalcin 2 induction by cinnabaric acid provides cytoprotection against endoplasmic reticulum and oxidative stress. *J. Pharmacol. Exp. Ther.* 353, 201–212.
- Ju, T., Wang, Y., Aryal, R.P., Lehoux, S.D., Ding, X., Kudelka, M.R., Cutler, C., Zeng, J., Wang, J., Sun, X., *et al.* (2013). Tn and sialyl-Tn antigens, aberrant O-glycomics as human disease markers. *Proteomics Clin. Appl.* 7, 618–631.
- Jung, C.-R., Hwang, K.-S., Yoo, J., Cho, W.-K., Kim, J.-M., Kim, W.H., and Im, D.-S. (2006). E2-EPF UCP targets pVHL for degradation and associates with tumor growth and metastasis. *Nat. Med.* 12, 809–816.

K

- Kaplan, J.C. and Delpech, M. (1989) *Biologie moléculaire et médecine*, Paris, (1ère édition) Médecine Sciences, Flammarion, 310 p.
- Kato, H., and Nishitoh, H. (2015). Stress Responses from the Endoplasmic Reticulum in Cancer. *Front. Oncol.* 5.
- Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., *et al.* (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149, 753–767.
- Kaur, S., Kumar, S., Momi, N., Sasson, A.R., and Batra, S.K. (2013). Mucins in pancreatic cancer and its microenvironment. *Nat. Rev. Gastroenterol. Hepatol.* 10, 607–620.
- Kedersha, N., and Anderson, P. (2007). Mammalian stress granules and processing bodies. *Methods Enzymol.* 431, 61–81.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36.
- Kim, J., Lee, J.E., Heynen-Genel, S., Suyama, E., Ono, K., Lee, K., Ideker, T., Aza-Blanc, P., and Gleeson, J.G. (2010). Functional genomic screen for modulators of ciliogenesis and cilium length. *Nature* 464, 1048–1051.
- Kim, P.-H., Na, S.-S., Lee, B., Kim, J.-H., and Cho, J.-Y. (2015). Stanniocalcin 2 enhances mesenchymal stem cell survival by suppressing oxidative stress. *BMB Rep.* 48, 702–707.
- Kim, S.-J., Lee, H.-W., Gu Kang, H., La, S.-H., Choi, I.J., Ro, J.Y., Bresalier, R.S., Song, J., and Chun, K.-H. (2014). Ablation of galectin-3 induces p27(KIP1)-dependent premature senescence without oncogenic stress. *Cell Death Differ.* 21, 1769–1779.
- Koch, A., Poirier, F., Jacob, R., and Delacour, D. (2010). Galectin-3, a novel centrosome-associated protein, required for epithelial morphogenesis. *Mol. Biol. Cell* 21, 219–231.
- Koido, S., Homma, S., Takahara, A., Namiki, Y., Tsukinaga, S., Mitobe, J., Odahara, S., Yukawa, T., Matsudaira, H., Nagatsuma, K., *et al.* (2011). Current immunotherapeutic approaches in pancreatic cancer. *Clin. Dev. Immunol.* 2011, 267539.
- Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J., and Nogues, G. (2004). Multiple links between transcription and splicing. *RNA N. Y. N* 10, 1489–1498.
- Kornblihtt, A.R., Schor, I.E., Alló, M., Dujardin, G., Petrillo, E., and Muñoz, M.J. (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat. Rev. Mol. Cell Biol.* 14, 153–165.
- Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* 11, 597–610.

Kropski, J.A., Lawson, W.E., Young, L.R., and Blackwell, T.S. (2013). Genetic studies provide clues on the pathogenesis of idiopathic pulmonary fibrosis. *Dis. Model. Mech.* 6, 9–17.

Krześlak, A., and Lipińska, A. (2004). Galectin-3 as a multifunctional protein. *Cell. Mol. Biol. Lett.* 9, 305–328.

Kufe, D.W. (2009). Mucins in cancer: function, prognosis and therapy. *Nat. Rev. Cancer* 9, 874–885.

Kuga, T., Kume, H., Kawasaki, N., Sato, M., Adachi, J., Shiromizu, T., Hoshino, I., Nishimori, T., Matsubara, H., and Tomonaga, T. (2013). A novel mechanism of keratin cytoskeleton organization through casein kinase I α and FAM83H in colorectal cancer. *J. Cell Sci.* 126, 4721–4731.

Kuga, T., Sasaki, M., Mikami, T., Miake, Y., Adachi, J., Shimizu, M., Saito, Y., Koura, M., Takeda, Y., Matsuda, J., *et al.* (2016). FAM83H and casein kinase I regulate the organization of the keratin cytoskeleton and formation of desmosomes. *Sci. Rep.* 6, 26557.

Kumar, S., Cruz, E., Joshi, S., Patel, A., Jahan, R., Batra, S.K., and Jain, M. (2016). Genetic Variants of Mucins: Unexplored Conundrum. *Carcinogenesis* bgw120.

Kuraoka, I., Ito, S., Wada, T., Hayashida, M., Lee, L., Saijo, M., Nakatsu, Y., Matsumoto, M., Matsunaga, T., Handa, H., *et al.* (2008). Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. *J. Biol. Chem.* 283, 940–950.

L

Lafarga, V., Cuadrado, A., Lopez de Silanes, I., Bengoechea, R., Fernandez-Capetillo, O., and Nebreda, A.R. (2009). p38 Mitogen-activated protein kinase- and HuR-dependent stabilization of p21(Cip1) mRNA mediates the G(1)/S checkpoint. *Mol. Cell. Biol.* 29, 4341–4351.

Lahdaoui, F., Delpu, Y., Vincent, A., Renaud, F., Messenger, M., Duchêne, B., Leteurtre, E., Mariette, C., Torrisani, J., Jonckheere, N., *et al.* (2015). miR-219-1-3p is a negative regulator of the mucin MUC4 expression and is a tumor suppressor in pancreatic cancer. *Oncogene* 34, 780–788.

Lamblin, G., Lhermitte, M., Klein, A., Houdret, N., Scharfman, A., Ramphal, R., and Roussel, P. (1991). The carbohydrate diversity of human respiratory mucins: a protection of the underlying mucosa? *Am. Rev. Respir. Dis.* 144, S19-24.

Lane, K.R., Yu, Y., Lackey, P.E., Chen, X., Marzluff, W.F., and Cook, J.G. (2013). Cell cycle-regulated protein abundance changes in synchronously proliferating HeLa cells include regulation of pre-mRNA splicing proteins. *PloS One* 8, e58456.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.

Le, N., Sund, M., Vinci, A., and GEMS collaborating group of Pancreas 2000 (2016). Prognostic and predictive markers in pancreatic adenocarcinoma. *Dig. Liver Dis. Off. J. Ital. Soc. Gastroenterol. Ital. Assoc. Study Liver* 48, 223–230.

Le Hir, H., Saulière, J., and Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. *Nat. Rev. Mol. Cell Biol.* 17, 41–54.

Le Mercier, M., Mathieu, V., Haibe-Kains, B., Bontempi, G., Mijatovic, T., Decaestecker, C., Kiss, R., and Lefranc, F. (2008). Knocking down galectin 1 in human hs683 glioblastoma cells impairs both angiogenesis and endoplasmic reticulum stress responses. *J. Neuropathol. Exp. Neurol.* 67, 456–469.

Leroy, X., Buisine, M.-P., Leteurtre, E., Aubert, S., Buob, D., Porchet, N., and Copin, M.-C. (2006). MUC1 (EMA): A key molecule of carcinogenesis?. *Ann. Pathol.* 26, 257–266.

Lévy, P. (sous la direction de) (2009) "Item 155 : tumeurs du pancréas" *In Abrégé d'hépatogastroentérologie*. Paris, Masson, p. 1-9.

Levy, R., Biran, A., Poirier, F., Raz, A., and Kloog, Y. (2011). Galectin-3 mediates cross-talk between K-Ras and Let-7c tumor suppressor microRNA. *PloS One* 6, e27490.

Li, L., Li, J., and Gao, J. (2014). Functions of galectin-3 and its role in fibrotic diseases. *J. Pharmacol. Exp. Ther.* 351, 336–343.

Lin, H.-M., Pestell, R.G., Raz, A., and Kim, H.-R.C. (2002). Galectin-3 enhances cyclin D(1) promoter activity through SP1 and a cAMP-responsive element in human breast epithelial cells. *Oncogene* 21, 8001–8010.

Liu, F.-T., and Rabinovich, G.A. (2005). Galectins as modulators of tumour progression. *Nat. Rev. Cancer* 5, 29–41.

Liu, F.-T., and Rabinovich, G.A. (2010). Galectins: regulators of acute and chronic inflammation. *Ann. N. Y. Acad. Sci.* 1183, 158–182.

Liu, F.-T., Patterson, R.J., and Wang, J.L. (2002). Intracellular functions of galectins. *Biochim. Biophys. Acta* 1572, 263–273.

Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* 127, 1223–1237.

Lui, J., Castelli, L.M., Pizzinga, M., Simpson, C.E., Hoyle, N.P., Bailey, K.L., Campbell, S.G., and Ashe, M.P. (2014). Granules harboring translationally active mRNAs provide a platform for P-body formation following stress. *Cell Rep.* 9, 944–954.

Lundmark, A., Davanian, H., Båge, T., Johannsen, G., Koro, C., Lundeberg, J., and Yucel-Lindberg, T. (2015). Transcriptome analysis reveals mucin 4 to be highly associated with periodontitis and identifies pleckstrin as a link to systemic diseases. *Sci. Rep.* 5, 18475.

M

Madsen, C.B., Lavrsen, K., Steentoft, C., Vester-Christensen, M.B., Clausen, H., Wandall, H.H., and Pedersen, A.E. (2013). Glycan elongation beyond the mucin associated Tn antigen protects tumor cells from immune-mediated killing. *PloS One* 8, e72413.

Majumder, S., Chari, S.T., and Ahlquist, D.A. (2015). Molecular detection of pancreatic neoplasia: Current status and future promise. *World J. Gastroenterol.* 21, 11387–11395.

- Mangus, D.A., Evans, M.C., and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* 4, 223.
- Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell* 136, 719–730.
- Martino, M.B., Jones, L., Brighton, B., Ehre, C., Abdulah, L., Davis, C.W., Ron, D., O’Neal, W.K., and Ribeiro, C.M.P. (2013). The ER stress transducer IRE1 β is required for airway epithelial mucin production. *Mucosal Immunol.* 6, 639–654.
- Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* 15, 108–121.
- Matsuoka, T., and Yashiro, M. (2016). Molecular targets for the treatment of pancreatic cancer: Clinical and experimental studies. *World J. Gastroenterol.* 22, 776–789.
- Mazurek, N., Byrd, J.C., Sun, Y., Hafley, M., Ramirez, K., Burks, J., and Bresalier, R.S. (2012). Cell-surface galectin-3 confers resistance to TRAIL by impeding trafficking of death receptors in metastatic colon adenocarcinoma cells. *Cell Death Differ.* 19, 523–533.
- Menini, S., Iacobini, C., Blasetti Fantauzzi, C., Pesce, C.M., and Pugliese, G. (2016). Role of Galectin-3 in Obesity and Impaired Glucose Homeostasis. *Oxid. Med. Cell. Longev.* 2016, 9618092.
- Merlin, J., Stechly, L., de Beaucé, S., Monté, D., Leteurtre, E., van Seuningen, I., Huet, G., and Pigny, P. (2011). Galectin-3 regulates MUC1 and EGFR cellular distribution and EGFR downstream pathways in pancreatic cancer cells. *Oncogene* 30, 2514–2525.
- Middleton, A.J., and Day, C.L. (2015). The molecular basis of lysine 48 ubiquitin chain synthesis by Ube2K. *Sci. Rep.* 5, 16793.
- Millevoi, S. and Vagner S (2010). Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic acids Res.* 38, 2757-2774.
- Mines, M.A., Goodwin, J.S., Limbird, L.E., Cui, F.-F., and Fan, G.-H. (2009). Deubiquitination of CXCR4 by USP14 is critical for both CXCL12-induced CXCR4 degradation and chemotaxis but not ERK activation. *J. Biol. Chem.* 284, 5742–5752.
- Miyazawa, M., and Tsuji, Y. (2014). Evidence for a novel antioxidant function and isoform-specific regulation of the human p66Shc gene. *Mol. Biol. Cell* 25, 2116–2127.
- 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.
- Moniaux, N., Nollet, S., Porchet, N., Degand, P., Laine, A., and Aubert, J.P. (1999). Complete sequence of the human mucin MUC4: a putative cell membrane-associated mucin. *Biochem. J.* 338 (Pt 2), 325–333.
- Moniaux, N., Andrianifahanana, M., Brand, R.E., and Batra, S.K. (2004). Multiple roles of mucins in pancreatic cancer, a lethal and challenging malignancy. *Br. J. Cancer* 91, 1633–1638.

Morgado, M., and Carson, D.D. (2017). PPAR γ Modulation of Cytokine-Stimulated MUC16 (CA125) Expression in Breast and Ovarian Cancer-Derived Cells. *J. Cell. Biochem.* *118*, 163–171.

Morgado, M., Sutton, M.N., Simmons, M., Warren, C.R., Lu, Z., Constantinou, P.E., Liu, J., Francis, L.L.W., Conlan, R.S., Bast, R.C., *et al.* (2016). Tumor necrosis factor- α and interferon- γ stimulate MUC16 (CA125) expression in breast, endometrial and ovarian cancers through NF κ B. *Oncotarget* *7*, 14871–14884.

Mouchel-Vielh, E., Rougeot, J., Decoville, M., and Peronnet, F. (2011). The MAP kinase ERK and its scaffold protein MP1 interact with the chromatin regulator Corto during *Drosophila* wing tissue development. *BMC Dev. Biol.* *11*, 17.

Müller-McNicoll, M., and Neugebauer, K.M. (2013). How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nat. Rev. Genet.* *14*, 275–287.

N

Nabi, I.R., Shankar, J., and Dennis, J.W. (2015). The galectin lattice at a glance. *J. Cell Sci.* *128*, 2213–2219.

Nakahara, S., and Raz, A. (2007). Regulation of cancer-related gene expression by galectin-3 and the molecular mechanism of its nuclear import pathway. *Cancer Metastasis Rev.* *26*, 605–610.

Nangia-Makker, P., Balan, V., and Raz, A. (2008). Regulation of tumor progression by extracellular galectin-3. *Cancer Microenviron. Off. J. Int. Cancer Microenviron. Soc.* *1*, 43–51.

Nasarre, P., Gemmill, R.M., and Drabkin, H.A. (2014). The emerging role of class-3 semaphorins and their neuropilin receptors in oncology. *OncoTargets Ther.* *7*, 1663–1687.

Nath, S., and Mukherjee, P. (2014). MUC1: a multifaceted oncoprotein with a key role in cancer progression. *Trends Mol. Med.* *20*, 332–342.

Newlaczyl, A.U., and Yu, L.-G. (2011). Galectin-3--a jack-of-all-trades in cancer. *Cancer Lett.* *313*, 123–128.

Nieminen, J., St-Pierre, C., Bhaumik, P., Poirier, F., and Sato, S. (2008). Role of galectin-3 in leukocyte recruitment in a murine model of lung infection by *Streptococcus pneumoniae*. *J. Immunol. Baltim. Md 1950* *180*, 2466–2473.

Niranjanakumari, S., Lasda, E., Brazas, R., and Garcia-Blanco, M.A. (2002). Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. *Methods San Diego Calif* *26*, 182–190.

Norbury, C.J. (2013). Cytoplasmic RNA: a case of the tail wagging the dog. *Nat. Rev. Mol. Cell Biol.* *14*, 643–653.

O

Oakes, S.A. (2017). Endoplasmic reticulum proteostasis: a key checkpoint in cancer. *Am. J. Physiol. Cell Physiol.* 312, C93–C102.

Obermair, A., Schmid, B.C., Packer, L.M., Leodolter, S., Birner, P., Ward, B.G., Crandon, A.J., McGuckin, M.A., and Zeillinger, R. (2002). Expression of MUC1 splice variants in benign and malignant ovarian tumours. *Int. J. Cancer* 100, 166–171.

Oikari, S., Makkonen, K., Deen, A.J., Tyni, I., Kärnä, R., Tammi, R.H., and Tammi, M.I. (2016). Hexosamine biosynthesis in keratinocytes: roles of GFAT and GNPDA enzymes in the maintenance of UDP-GlcNAc content and hyaluronan synthesis. *Glycobiology* 26, 710–722.

Oyasiji, T., and Ma, W.W. (2015). Novel adjuvant therapies for pancreatic adenocarcinoma. *J. Gastrointest. Oncol.* 6, 430–435.

P

Panas, M.D., Ivanov, P., and Anderson, P. (2016). Mechanistic insights into mammalian stress granule dynamics. *J. Cell Biol.* 215, 313–323.

Pang, J., Rhodes, D.H., Pini, M., Akasheh, R.T., Castellanos, K.J., Cabay, R.J., Cooper, D., Perretti, M., and Fantuzzi, G. (2013). Increased adiposity, dysregulated glucose metabolism and systemic inflammation in Galectin-3 KO mice. *PloS One* 8, e57915.

Parent, A., and Bisailon, M. (2006). Synergy between transcription and mRNA processing events. *Med. Sci. MS* 22, 626–632.

Paron, I., Scaloni, A., Pines, A., Bachi, A., Liu, F.-T., Puppini, C., Pandolfi, M., Ledda, L., Di Loreto, C., Damante, G., *et al.* (2003). Nuclear localization of Galectin-3 in transformed thyroid cells: a role in transcriptional regulation. *Biochem. Biophys. Res. Commun.* 302, 545–553.

Parry, S., Hanisch, F.G., Leir, S.-H., Sutton-Smith, M., Morris, H.R., Dell, A., and Harris, A. (2006). N-Glycosylation of the MUC1 mucin in epithelial cells and secretions. *Glycobiology* 16, 623–634.

Patterson, R.J., Wang, W., and Wang, J.L. (2004). Understanding the biochemical activities of galectin-1 and galectin-3 in the nucleus. *Glycoconj. J.* 19, 499–506.

Pejnovic, N.N., Pantic, J.M., Jovanovic, I.P., Radosavljevic, G.D., Djukic, A.L., Arsenijevic, N.N., and Lukic, M.L. (2013). Galectin-3 is a regulator of metaflammation in adipose tissue and pancreatic islets. *Adipocyte* 2, 266–271.

Percipalle, P., Raju, C.S., and Fukuda, N. (2009). Actin-associated hnRNP proteins as transacting factors in the control of mRNA transport and localization. *RNA Biol.* 6, 171–174.

Peritz, T., Zeng, F., Kannanayakal, T.J., Kilk, K., Eiríksdóttir, E., Langel, U., and Eberwine, J. (2006). Immunoprecipitation of mRNA-protein complexes. *Nat. Protoc.* 1, 577–580.

Perrais, M., Pigny, P., Ducourouble, M.P., Petitprez, D., Porchet, N., Aubert, J.P., and Van Seuning, I. (2001). Characterization of human mucin gene MUC4 promoter: importance of growth factors and proinflammatory cytokines for its regulation in pancreatic cancer cells. *J. Biol. Chem.* 276, 30923–30933.

Pfaff, J., and Meister, G. (2013). Argonaute and GW182 proteins: an effective alliance in gene silencing. *Biochem. Soc. Trans.* 41, 855–860.

Pfeiffer, J.R., McAvoy, B.L., Fecteau, R.E., Deleault, K.M., and Brooks, S.A. (2011). CARHSP1 is required for effective tumor necrosis factor alpha mRNA stabilization and localizes to processing bodies and exosomes. *Mol. Cell. Biol.* 31, 277–286.

Pigny, P., Guyonnet-Duperat, V., Hill, A.S., Pratt, W.S., Galiege-Zouitina, S., d'Hooze, M.C., Laine, A., Van-Seuning, I., Degand, P., Gum, J.R., *et al.* (1996). Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. *Genomics* 38, 340–352.

Pinho, S.S., and Reis, C.A. (2015). Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* 15, 540–555.

Pollard, T.D. and Earnshaw, W.C. (2004) *Biologie cellulaire*. Paris, Elsevier, Collection Campus reference, 880 p.

Porchet, N., and Aubert, J.-P. (2004). MUC genes: mucin or not mucin? That is the question. *Med. Sci. MS* 20, 569–574.

Porchet, N., Nguyen, V.C., Dufosse, J., Audie, J.P., Guyonnet-Duperat, V., Gross, M.S., Denis, C., Degand, P., Bernheim, A., and Aubert, J.P. (1991). Molecular cloning and chromosomal localization of a novel human tracheo-bronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs. *Biochem. Biophys. Res. Commun.* 175, 414–422.

Pour, P.M., Pandey, K.K., and Batra, S.K. (2003). What is the origin of pancreatic adenocarcinoma? *Mol. Cancer* 2, 13.

Price-Schiavi, S.A., Jepson, S., Li, P., Arango, M., Rudland, P.S., Yee, L., and Carraway, K.L. (2002). Rat Muc4 (sialomucin complex) reduces binding of anti-ErbB2 antibodies to tumor cell surfaces, a potential mechanism for herceptin resistance. *Int. J. Cancer* 99, 783–791.

Protter, D.S.W., and Parker, R. (2016). Principles and Properties of Stress Granules. *Trends Cell Biol.* 26, 668–679.

R

Rabinovich, G.A. (1999). Galectins: an evolutionarily conserved family of animal lectins with multifunctional properties; a trip from the gene to clinical therapy. *Cell Death Differ.* 6, 711–721.

Rachagani, S., Torres, M.P., Moniaux, N., and Batra, S.K. (2009). Current status of mucins in the diagnosis and therapy of cancer. *BioFactors Oxf. Engl.* 35, 509–527.

Radhakrishnan, P., Mohr, A.M., Grandgenett, P.M., Steele, M.M., Batra, S.K., and Hollingsworth, M.A. (2013). MicroRNA-200c modulates the expression of MUC4 and MUC16 by directly targeting their coding sequences in human pancreatic cancer. *PloS One* 8, e73356.

Rajabi, H., Jin, C., Ahmad, R., McClary, C., Joshi, M.D., and Kufe, D. (2010). Mucin 1 oncoprotein expression is suppressed by the mir-125b oncomir. *Genes Cancer* 1, 62–68.

Rajgor, D., and Shanahan, C.M. (2014). RNA granules and cytoskeletal links. *Biochem. Soc. Trans.* 42, 1206–1210.

Ramasamy, S., Duraisamy, S., Barbashov, S., Kawano, T., Kharbanda, S., and Kufe, D. (2007). The MUC1 and galectin-3 oncoproteins function in a microRNA-dependent regulatory loop. *Mol. Cell* 27, 992–1004.

Ray, P., Zhang, D.H., Elias, J.A., and Ray, A. (1995). Cloning of a differentially expressed I kappa B-related protein. *J. Biol. Chem.* 270, 10680–10685.

Remmers, N., Anderson, J.M., Linde, E.M., DiMaio, D.J., Lazenby, A.J., Wandall, H.H., Mandel, U., Clausen, H., Yu, F., and Hollingsworth, M.A. (2013). Aberrant expression of mucin core proteins and o-linked glycans associated with progression of pancreatic cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 19, 1981–1993.

Ritchie, D.B., Schellenberg, M.J., and MacMillan, A.M. (2009). Spliceosome structure: piece by piece. *Biochim. Biophys. Acta* 1789, 624–633.

Robinson, S.M., Fan, L., White, S.A., Charnley, R.M., and Mann, J. (2016). The role of exosomes in the pathogenesis of pancreatic ductal adenocarcinoma. *Int. J. Biochem. Cell Biol.* 75, 131–139.

Rosty, C. (2004). Towards the identification of new markers of pancreatic cancer by gene expression analysis. *Bull. Cancer (Paris)* 91, 593–598.

Ruvolo, P.P. (2016). Galectin 3 as a guardian of the tumor microenvironment. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1863, 427–437.

S

Sachdeva, M., and Mo, Y.-Y. (2010). MicroRNA-145 suppresses cell invasion and metastasis by directly targeting mucin 1. *Cancer Res.* 70, 378–387.

Saha, S.K., Choi, H.Y., Kim, B.W., Dayem, A.A., Yang, G.-M., Kim, K.S., Yin, Y.F., and Cho, S.-G. (2017). KRT19 directly interacts with β -catenin/RAC1 complex to regulate NUMB-dependent NOTCH signaling pathway and breast cancer properties. *Oncogene* 36, 332–349.

Saif, M.W. (2011). Pancreatic neoplasm in 2011: an update. *JOP J. Pancreas* 12, 316–321.

Saif, M.W. and Chabot, J. (2011) Chemotherapy: Metastatic pancreatic cancer--is FOLFIRINOX the new standard? *Nat Rev Clin Oncol.* 8, 452-453.

- Sakabe, I., Hu, R., Jin, L., Clarke, R., and Kasid, U.N. (2015). TMEM33: a new stress-inducible endoplasmic reticulum transmembrane protein and modulator of the unfolded protein response signaling. *Breast Cancer Res. Treat.* 153, 285–297.
- Schlager, M.A., Kapitein, L.C., Grigoriev, I., Burzynski, G.M., Wulf, P.S., Keijzer, N., de Graaff, E., Fukuda, M., Shepherd, I.T., Akhmanova, A., *et al.* (2010). Pericentrosomal targeting of Rab6 secretory vesicles by Bicaudal-D-related protein 1 (BICDR-1) regulates neuritogenesis. *EMBO J.* 29, 1637–1651.
- Schnaar, R.L. (2016). Glycobiology simplified: diverse roles of glycan recognition in inflammation. *J. Leukoc. Biol.* 99, 825–838.
- Schroeder, J.A., Masri, A.A., Adriance, M.C., Tessier, J.C., Kotlarczyk, K.L., Thompson, M.C., and Gendler, S.J. (2004). MUC1 overexpression results in mammary gland tumorigenesis and prolonged alveolar differentiation. *Oncogene* 23, 5739–5747.
- Seicean, A., Petrusel, L., and Seicean, R. (2015). New targeted therapies in pancreatic cancer. *World J. Gastroenterol.* 21, 6127–6145.
- Senapati, S., Das, S., and Batra, S.K. (2010). Mucin-interacting proteins: from function to therapeutics. *Trends Biochem. Sci.* 35, 236–245.
- Senapati, S., Chaturvedi, P., Chaney, W.G., Chakraborty, S., Gnanapragassam, V.S., Sasson, A.R., and Batra, S.K. (2011). Novel INTERaction of MUC4 and galectin: potential pathobiological implications for metastasis in lethal pancreatic cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 17, 267–274.
- Senapati, S., Gnanapragassam, V.S., Moniaux, N., Momi, N., and Batra, S.K. (2012). Role of MUC4-NIDO domain in the MUC4-mediated metastasis of pancreatic cancer cells. *Oncogene* 31, 3346–3356.
- Sharma, D.K., Bressler, K., Patel, H., Balasingam, N., and Thakor, N. (2016). Role of Eukaryotic Initiation Factors during Cellular Stress and Cancer Progression. *J. Nucleic Acids* 2016, 8235121.
- Shih, S.C., and Claffey, K.P. (1999). Regulation of human vascular endothelial growth factor mRNA stability in hypoxia by heterogeneous nuclear ribonucleoprotein L. *J. Biol. Chem.* 274, 1359–1365.
- Shyu, A.-B., Wilkinson, M.F., and van Hoof, A. (2008). Messenger RNA regulation: to translate or to degrade. *EMBO J.* 27, 471–481.
- Sibley, C.R. (2014). Regulation of gene expression through production of unstable mRNA isoforms. *Biochem. Soc. Trans.* 42, 1196–1205.
- Singh, P.K., and Hollingsworth, M.A. (2006). Cell surface-associated mucins in signal transduction. *Trends Cell Biol.* 16, 467–476.

- Singh, A.P., Chauhan, S.C., Andrianifahanana, M., Moniaux, N., Meza, J.L., Copin, M.C., van Seuningen, I., Hollingsworth, M.A., Aubert, J.P., and Batra, S.K. (2007). MUC4 expression is regulated by cystic fibrosis transmembrane conductance regulator in pancreatic adenocarcinoma cells via transcriptional and post-translational mechanisms. *Oncogene* 26, 30–41.
- Singh, G., Pratt, G., Yeo, G.W., and Moore, M.J. (2015). The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annu. Rev. Biochem.* 84, 325–354.
- Sivan, G., Kedersha, N., and Elroy-Stein, O. (2007). Ribosomal slowdown mediates translational arrest during cellular division. *Mol. Cell. Biol.* 27, 6639–6646.
- Siwicki, M., Engblom, C., and Pittet, M.J. (2016). Gal3 Links Inflammation and Insulin Resistance. *Cell Metab.* 24, 655–656.
- Skrypek, N., Vasseur, R., Vincent, A., Duchêne, B., Van Seuningen, I., and Jonckheere, N. (2015). The oncogenic receptor ErbB2 modulates gemcitabine and irinotecan/SN-38 chemoresistance of human pancreatic cancer cells via hCNT1 transporter and multidrug-resistance associated protein MRP-2. *Oncotarget* 6, 10853–10867.
- Smith, R.W., Blee, T.K. and Gray, N.K. (2014). Poly(A)-binding proteins are required for diverse biological processes in metazoans. *Biochem Soc Trans.* 42, 1229–1237.
- Sonenberg, N., and Dever, T.E. (2003). Eukaryotic translation initiation factors and regulators. *Curr. Opin. Struct. Biol.* 13, 56–63.
- Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136, 731–745.
- Song, S., Byrd, J.C., Mazurek, N., Liu, K., Koo, J.S., and Bresalier, R.S. (2005). Galectin-3 modulates MUC2 mucin expression in human colon cancer cells at the level of transcription via AP-1 activation. *Gastroenterology* 129, 1581–1591.
- Souquere, S., Mollet, S., Kress, M., Dautry, F., Pierron, G., and Weil, D. (2009). Unravelling the ultrastructure of stress granules and associated P-bodies in human cells. *J. Cell Sci.* 122, 3619–3626.
- Srikantan, S., and Gorospe, M. (2011). UneCLIPsing HuR nuclear function. *Mol. Cell* 43, 319–321.
- Srikantan, S., Tominaga, K., and Gorospe, M. (2012). Functional interplay between RNA-binding protein HuR and microRNAs. *Curr. Protein Pept. Sci.* 13, 372–379.
- Srivastava, S.K., Bhardwaj, A., Singh, S., Arora, S., Wang, B., Grizzle, W.E., and Singh, A.P. (2011). MicroRNA-150 directly targets MUC4 and suppresses growth and malignant behavior of pancreatic cancer cells. *Carcinogenesis* 32, 1832–1839.
- Staton, J.M., Thomson, A.M., and Leedman, P.J. (2000). Hormonal regulation of mRNA stability and RNA-protein interactions in the pituitary. *J. Mol. Endocrinol.* 25, 17–34.
- Stewart BW and Kleihues P. (sous la direction de) (2005) Le cancer du pancréas in *Le cancer dans le monde*. Lyon, IARC Press p254-258.

Stine, Z.E., Walton, Z.E., Altman, B.J., Hsieh, A.L., and Dang, C.V. (2015). MYC, Metabolism, and Cancer. *Cancer Discov.* 5, 1024–1039.

T

Takenobu, M., Osaki, M., Fujiwara, K., Fukuhara, T., Kitano, H., Kugoh, H., and Okada, F. (2016). PITX1 is a novel predictor of the response to chemotherapy in head and neck squamous cell carcinoma. *Mol. Clin. Oncol.* 5, 89–94.

Takeuchi, O. (2015). HuR keeps interferon- β mRNA stable. *Eur. J. Immunol.* 45, 1296–1299.

Tameire, F., Verginadis, I.I., and Koumenis, C. (2015). Cell intrinsic and extrinsic activators of the unfolded protein response in cancer: Mechanisms and targets for therapy. *Semin. Cancer Biol.* 33, 3–15.

Tanenbaum, M.E., Stern-Ginossar, N., Weissman, J.S., and Vale, R.D. (2015). Regulation of mRNA translation during mitosis. *eLife* 4.

Tani, H., Imamachi, N., Salam, K.A., Mizutani, R., Ijiri, K., Irie, T., Yada, T., Suzuki, Y., and Akimitsu, N. (2012). Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. *RNA Biol.* 9, 1370–1379.

Tanida, S., Mori, Y., Ishida, A., Akita, K., and Nakada, H. (2014). Galectin-3 binds to MUC1-N-terminal domain and triggers recruitment of β -catenin in MUC1-expressing mouse 3T3 cells. *Biochim. Biophys. Acta* 1840, 1790–1797.

Tcherpakov, M., Delaunay, A., Toth, J., Kadoya, T., Petroski, M.D., and Ronai, Z.A. (2009). Regulation of endoplasmic reticulum-associated degradation by RNF5-dependent ubiquitination of JNK-associated membrane protein (JAMP). *J. Biol. Chem.* 284, 12099–12109.

Thacker, S.A., Robinson, P., Abel, A., and Tweardy, D.J. (2013). Modulation of the unfolded protein response during hepatocyte and cardiomyocyte apoptosis in trauma/hemorrhagic shock. *Sci. Rep.* 3, 1187.

Thijssen, V.L., Heusschen, R., Caers, J., and Griffioen, A.W. (2015). Galectin expression in cancer diagnosis and prognosis: A systematic review. *Biochim. Biophys. Acta BBA - Rev. Cancer* 1855, 235–247.

Timoshenko, A.V., Lanteigne, J., and Kozak, K. (2016). Extracellular stress stimuli alter galectin expression profiles and adhesion characteristics of HL-60 cells. *Mol. Cell. Biochem.* 413, 137–143.

Tourrière, H., Chebli, K., and Tazi, J. (2002). mRNA degradation machines in eukaryotic cells. *Biochimie* 84, 821–837.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31, 46–53.

Tréhoux, S., Lahdaoui, F., Delpu, Y., Renaud, F., Leteurtre, E., Torrisani, J., Jonckheere, N., and Van Seuningen, I. (2015). Micro-RNAs miR-29a and miR-330-5p function as tumor suppressors by targeting the MUC1 mucin in pancreatic cancer cells. *Biochim. Biophys. Acta* 1853, 2392–2403.

Tsukumo, Y., Tsukahara, S., Furuno, A., Iemura, S., Natsume, T., and Tomida, A. (2014). TBL2 is a novel PERK-binding protein that modulates stress-signaling and cell survival during endoplasmic reticulum stress. *PloS One* 9, e112761.

Twyffels, L., Gueydan, C., and Kruys, V. (2011). Shuttling SR proteins: more than splicing factors. *FEBS J.* 278, 3246–3255.

U

Uren, P.J., Burns, S.C., Ruan, J., Singh, K.K., Smith, A.D., and Penalva, L.O.F. (2011). Genomic analyses of the RNA-binding protein Hu antigen R (HuR) identify a complex network of target genes and novel characteristics of its binding sites. *J. Biol. Chem.* 286, 37063–37066.

V

Valente, C., Polishchuk, R., and De Matteis, M.A. (2010). Rab6 and myosin II at the cutting edge of membrane fission. *Nat. Cell Biol.* 12, 635–638.

Vandenhaute, B., Buisine, M.P., Debailleul, V., Clément, B., Moniaux, N., Dieu, M.C., Degand, P., Porchet, N., and Aubert, J.P. (1997). Mucin gene expression in biliary epithelial cells. *J. Hepatol.* 27, 1057–1066.

Verardo, L.L., Silva, F.F., Lopes, M.S., Madsen, O., Bastiaansen, J.W.M., Knol, E.F., Kelly, M., Varona, L., Lopes, P.S., and Guimarães, S.E.F. (2016). Revealing new candidate genes for reproductive traits in pigs: combining Bayesian GWAS and functional pathways. *Genet. Sel. Evol. GSE* 48, 9.

Viguier, M., Advedissian, T., Delacour, D., Poirier, F., and Deshayes, F. (2014). Galectins in epithelial functions. *Tissue Barriers* 2, e29103.

Vincent A (2007) Régulation des gènes MUCs 11P15 et MUC4 par les mécanismes épigénétiques et par les facteurs de différenciation cellulaire. Thèse sous la direction de Isabelle Van Seuningen

Vincenz, L., and Hartl, F.U. (2014). Sugarcoating ER Stress. *Cell* 156, 1125–1127.

Von Hoff, D.D., Ervin, T., Arena, F.P., Chiorean, E.G., Infante, J., Moore, M., Seay, T., Tjulandin, S.A., Ma, W.W., Saleh, M.N., *et al.* (2013). Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *N. Engl. J. Med.* 369, 1691–1703.

W

- Wang, L., and Guo, X.-L. (2016). Molecular regulation of galectin-3 expression and therapeutic implication in cancer progression. *Biomed. Pharmacother.* 78, 165–171.
- Wang, J.L., Gray, R.M., Haudek, K.C., and Patterson, R.J. (2004). Nucleocytoplasmic lectins. *Biochim. Biophys. Acta* 1673, 75–93.
- Wang, L., Friess, H., Zhu, Z., Frigeri, L., Zimmermann, A., Korc, M., Berberat, P.O., and Büchler, M.W. (2000). Galectin-1 and galectin-3 in chronic pancreatitis. *Lab. Investig. J. Tech. Methods Pathol.* 80, 1233–1241.
- Wang, Q., Liu, Y., Soetandyo, N., Baek, K., Hegde, R., and Ye, Y. (2011). A ubiquitin ligase-associated chaperone holdase maintains polypeptides in soluble states for proteasome degradation. *Mol. Cell* 42, 758–770.
- Wang, W., Park, J.W., Wang, J.L., and Patterson, R.J. (2006). Immunoprecipitation of spliceosomal RNAs by antisera to galectin-1 and galectin-3. *Nucleic Acids Res.* 34, 5166–5174.
- Wang, Z.V., Deng, Y., Gao, N., Pedrozo, Z., Li, D.L., Morales, C.R., Criollo, A., Luo, X., Tan, W., Jiang, N., *et al.* (2014). Spliced X-box binding protein 1 couples the unfolded protein response to hexosamine biosynthetic pathway. *Cell* 156, 1179–1192.
- Watson, J., Baker, T., Bell, S., Gann, A., Levine, M. and Losick, R (2009) *Biologie moléculaire du gène*, Paris, (6ème édition) Pearson Education, 688 p.
- Weed, D.T., Gomez-Fernandez, C., Yasin, M., Hamilton-Nelson, K., Rodriguez, M., Zhang, J., and Carraway, K.L. (2004). MUC4 and ErbB2 expression in squamous cell carcinoma of the upper aerodigestive tract: correlation with clinical outcomes. *The Laryngoscope* 114, 1–32.
- Wickramasinghe, V.O., and Laskey, R.A. (2015). Control of mammalian gene expression by selective mRNA export. *Nat. Rev. Mol. Cell Biol.* 16, 431–442.
- Wilkie, G.S., Dickson, K.S., and Gray, N.K. (2003). Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. *Trends Biochem. Sci.* 28, 182–188.
- Wilusz, C.J., and Wilusz, J. (2004). Bringing the role of mRNA decay in the control of gene expression into focus. *Trends Genet. TIG* 20, 491–497.
- Wright, R.D., Souza, P.R., Flak, M.B., Thedchanamoorthy, P., Norling, L.V., and Cooper, D. (2016). Galectin-3-null mice display defective neutrophil clearance during acute inflammation. *J. Leukoc. Biol.*
- Wu, M., Reuter, M., Lilie, H., Liu, Y., Wahle, E., and Song, H. (2005). Structural insight into poly(A) binding and catalytic mechanism of human PARN. *EMBO J.* 24, 4082–4093.

X

Xie, K., Zhi, X., Tang, J., Zhu, Y., Zhang, J., Li, Z., Tao, J., and Xu, Z. (2014). Upregulation of the splice variant MUC4/Y in the pancreatic cancer cell line MIA PaCa-2 potentiates proliferation and suppresses apoptosis: new insight into the presence of the transcript variant of MUC4. *Oncol. Rep.* 31, 2187–2194.

Xie, L., Ni, W.-K., Chen, X.-D., Xiao, M.-B., Chen, B.-Y., He, S., Lu, C.-H., Li, X.-Y., Jiang, F., and Ni, R.-Z. (2012). The expressions and clinical significances of tissue and serum galectin-3 in pancreatic carcinoma. *J. Cancer Res. Clin. Oncol.* 138, 1035–1043.

Xu, D., Liu, S., Zhang, L., and Song, L. (2016). MiR-211 inhibits invasion and epithelial-to-mesenchymal transition (EMT) of cervical cancer cells via targeting MUC4. *Biochem. Biophys. Res. Commun.*

Y

Yan, W., Li, R., He, J., Du, J., and Hou, J. (2015). Importin β 1 mediates nuclear factor- κ B signal transduction into the nuclei of myeloma cells and affects their proliferation and apoptosis. *Cell. Signal.* 27, 851–859.

Yang, C., Peng, P., Li, L., Shao, M., Zhao, J., Wang, L., Duan, F., Song, S., Wu, H., Zhang, J., *et al.* (2016). High expression of GFAT1 predicts poor prognosis in patients with pancreatic cancer. *Sci. Rep.* 6, 39044.

Yeung, B.H.Y., Law, A.Y.S., and Wong, C.K.C. (2012). Evolution and roles of stanniocalcin. *Mol. Cell. Endocrinol.* 349, 272–280.

Yonezawa, S., Higashi, M., Yamada, N., Yokoyama, S., Kitamoto, S., Kitajima, S., and Goto, M. (2011). Mucins in human neoplasms: clinical pathology, gene expression and diagnostic application. *Pathol. Int.* 61, 697–716.

Yoo, B.C., Hong, S.-H., Ku, J.-L., Kim, Y.-H., Shin, Y.-K., Jang, S.-G., Kim, I.-J., Jeong, S.-Y., and Park, J.-G. (2009). Galectin-3 stabilizes heterogeneous nuclear ribonucleoprotein Q to maintain proliferation of human colon cancer cells. *Cell. Mol. Life Sci. CMLS* 66, 350–364.

You, H., Ge, Y., Zhang, J., Cao, Y., Xing, J., Su, D., Huang, Y., Li, M., Qu, S., Sun, F., *et al.* (2017). Derlin-1 promotes ubiquitylation and degradation of the epithelial Na⁽⁺⁾ channel, ENaC. *J. Cell Sci.* 130, 1027–1036.

Z

Zander, G., Hackmann, A., Bender, L., Becker, D., Lingner, T., Salinas, G., and Krebber, H. (2016). mRNA quality control is bypassed for immediate export of stress-responsive transcripts. *Nature.*

Zhang, L., Wang, P., Qin, Y., Cong, Q., Shao, C., Du, Z., Ni, X., Li, P., and Ding, K. (2016). RN1, a novel galectin-3 inhibitor, inhibits pancreatic cancer cell growth in vitro and in vivo via blocking galectin-3 associated signaling pathways. *Oncogene.*

Zhao, Q., Guo, X., Nash, G.B., Stone, P.C., Hilkens, J., Rhodes, J.M., and Yu, L.-G. (2009). Circulating galectin-3 promotes metastasis by modifying MUC1 localization on cancer cell surface. *Cancer Res.* 69, 6799–6806.

Zhao, Q., Barclay, M., Hilkens, J., Guo, X., Barrow, H., Rhodes, J.M., and Yu, L.-G. (2010). Interaction between circulating galectin-3 and cancer-associated MUC1 enhances tumour cell homotypic aggregation and prevents anoikis. *Mol. Cancer* 9, 154.

Zhu, Y., Zhang, J.-J., Xie, K.-L., Tang, J., Liang, W.-B., Zhu, R., Zhu, Y., Wang, B., Tao, J.-Q., Zhi, X.-F., *et al.* (2014). Specific-detection of clinical samples, systematic functional investigations, and transcriptome analysis reveals that splice variant MUC4/Y contributes to the malignant progression of pancreatic cancer by triggering malignancy-related positive feedback loops signaling. *J. Transl. Med.* 12, 309.

Zhu, Y., Zhang, C., Gu, C., Li, Q., and Wu, N. (2016). Function of Deubiquitinating Enzyme USP14 as Oncogene in Different Types of Cancer. *Cell. Physiol. Biochem. Int. J. Exp. Cell. Physiol. Biochem. Pharmacol.* 38, 993–1002.