



**HAL**  
open science

# Caractérisation de l'interaction entre FUS et TDP-43, deux protéines de liaison à l'ARN impliquées dans les maladies neurodégénératives

Clément Demongin

► **To cite this version:**

Clément Demongin. Caractérisation de l'interaction entre FUS et TDP-43, deux protéines de liaison à l'ARN impliquées dans les maladies neurodégénératives. Biochimie [q-bio.BM]. Université Paris-Saclay, 2022. Français. NNT : 2022UPASL039 . tel-03783244

**HAL Id: tel-03783244**

**<https://theses.hal.science/tel-03783244>**

Submitted on 22 Sep 2022

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

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

# Caractérisation de l'interaction entre FUS et TDP-43, deux protéines de liaison à l'ARN impliquées dans les maladies neurodégénératives

*Characterization of the interaction between FUS and TDP-43, two RNA-binding proteins involved in neurodegenerative diseases*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n°577 : structure et dynamique des systèmes vivants (SDSV)

Spécialité de doctorat : Science de la vie et de la santé

Graduate School: Life Science and Health

Référent : Université d'Évry Val d'Essonne

Thèse préparée dans l'unité de recherche **Structure-Activité des Biomolécules Normales et Pathologiques (Université Paris-Saclay, Inserm, Univ Evry)**,

Sous la direction de **Loïc HAMON**, Professeur

Thèse présentée et soutenue à Evry, le 11 juillet 2022, par

**Clément DEMONGIN**

### Composition du Jury

<b>Philippe SAVARIN</b> Professeur, Université Paris 13	Président
<b>Caroline SMET-NOCCA</b> Maître de conférences, Université de Lille	Rapporteuse & Examinatrice
<b>Séverine BOILLEE</b> Chargée de recherche, INSERM, Institut du Cerveau et de la Moelle Epinière	Rapporteuse & Examinatrice
<b>Samuel TRANIER</b> Maître de conférences, IPBS Toulouse	Examineur
<b>David PASTRE</b> Professeur, Université d'Évry/Université Paris-Saclay	Examineur
<b>Loïc HAMON</b> Professeur, Université d'Évry/Université Paris-Saclay	Directeur de thèse



# Remerciements

---

*Avant de rentrer dans le vif du sujet, je tenais à prendre un peu de temps pour remercier vivement l'ensemble des personnes m'ayant accompagné dans cette grande aventure. En commençant par mes collègues et l'ensemble des personnes avec lesquelles j'ai partagé mes journées au laboratoire même si le covid a mis un peu de distance entre nous.*

*Je remercie Vandana Joshi, et Bénédicte Desforges pour m'avoir formé aux techniques de biologie moléculaire et aux manipulations cellulaires. Merci à Jean Salone, et à Alexandre Maucuer pour leur aide et les réponses qu'ils ont pu apporter à mes interrogations. Je remercie aussi Emilie Bertrand pour sa présence et nos échanges qui ont rendu moins longs les jours passés devant l'AFM. Je remercie aussi Krystel El Hage avec qui j'ai partagé mon bureau. Merci également à Lydia Lebouil, secrétaire du laboratoire qui a toujours été disponible en cas de difficultés.*

*Je remercie les autres doctorants (maintenant docteurs) qui ont partagé avec moi leur expérience, et ont aussi su me donner des conseils, en particulier Juan-Carlos Rengifo, Anastasiia Samsonova, Serhii Pankivskyi, ainsi qu'Asaki Kobayashi.*

*Je remercie également mon directeur de thèse Loïc Hamon, qui m'a guidé tout au long de ce projet tout en me laissant assez d'autonomie dans sa réalisation. Je remercie aussi David Pastré de m'avoir accueilli au sein du laboratoire SABNP.*

*Je remercie les stagiaires qui m'ont accompagné quelques mois, et ont pu contribuer à l'avancement de ce projet : Léa Ceschi et Loïc Herbinet.*

*J'adresse une pensée toute particulière au regretté Guillaume Lambert qui était présent au début de ma thèse. Il m'a, non seulement formé, aux techniques de biologie moléculaire et de biochimie, mais surtout nous avons pu avoir de nombreux échanges amicaux et enrichissants.*

*Merci aux membres de mon comité de thèse Michel Malo et Samuel Tranier qui ont suivi l'avancée de mon travail et m'ont apporté des conseils avisés.*

*Je souhaite également remercier l'ensemble des membres du jury, Philippe Savarin qui a présidé ce jury, les rapporteuses, Caroline Smet-Nocca, et Séverine Boillée, ainsi que les examinateurs Samuel Tranier, et David Pastré, qui ont consacré du temps à la relecture de ce travail, se sont assurés de la cohérence des travaux présentés, et ont contribué à la finalisation de ce manuscrit. Ils m'ont ainsi suggéré des corrections*

*pertinentes dans leurs domaines respectifs. Je les remercie aussi pour la discussion qui a fait émerger des points intéressants sur la portée scientifique de ce travail.*

*Je souhaite aussi remercier Jérémie Cossette qui m'a accueilli à la plateforme ImCy du Généthon. Au-delà de son appui pour l'acquisition de mes échantillons, son aide et ses conseils m'ont permis de faire avancer mon projet.*

*J'aimerais aussi remercier les personnes qui s'impliquent pour rendre la science accessible au plus grand nombre. En particulier les auteurs qui m'ont permis d'accéder directement à leurs travaux par le biais de ResearchGate, mais aussi le site Sci-hub qui milite pour une science ouverte, et m'a permis d'accéder à de nombreuses ressources pour la rédaction de ce travail de thèse, alors que le manque d'accès à certains travaux représente un frein au progrès des connaissances.*

*Je remercie plus largement le personnel de l'université que j'ai pu côtoyer, ainsi que les amis que j'ai rencontrés à l'université, en particulier au cours de mes activités associatives. Ensemble nous avons organisé quelques événements autour de la transition écologique, et nous avons l'espoir de rendre l'université plus engagée dans la démarche environnementale. Mais, au-delà de nos convictions, et notre volonté commune de faire notre part, nous avons partagé de très bons moments : Victoria, Tom, Katia, et Guillaume. Nos moments, nos débats et nos repas partagés resteront parmi les meilleurs souvenirs de ces dernières années.*

*Enfin, je pense fortement à ma famille. En premier lieu à mon fils Enzo, qui est né lors de ce travail de thèse. Il a parfois été sur mes genoux lors de la relecture de ce travail. Son enthousiasme, son énergie et son sourire m'ont permis de garder la motivation jusqu'au bout. Ma compagne Gwenaëlle qui m'a toujours apporté son soutien, son amour et m'a énormément aidé. Sans elle, ce travail n'aurait pu aboutir. Je suis conscient des tâches et de la charge mentale dont elle m'a délesté lors de la dernière ligne droite de ce projet, et je promets de me rattraper par la suite.*

*J'adresse un grand merci à mes parents, qui ont toujours été là pour me soutenir et m'encourager. A tous ceux qui ont contribué à la relecture de ce rapport : mes parents, mes beaux-parents et ma grand-tante. Et enfin, mes grands-parents paternels qui ont été présents et m'ont facilité les choses en gardant mon enfant pour que je puisse toujours plus travailler.*

*Je dédie ce travail à mes grands-parents maternels*

# Table des matières

---

<b>REMERCIEMENTS</b> .....	<b>I</b>
----------------------------	----------

<b>TABLE DES MATIERES</b> .....	<b>III</b>
---------------------------------	------------

<b>ABREVIATIONS</b> .....	<b>VII</b>
---------------------------	------------

<b>INTRODUCTION</b> .....	<b>1</b>
---------------------------	----------

<b>BIBLIOGRAPHIE</b> .....	<b>5</b>
----------------------------	----------

<b>I-</b>	<b>LES PROTEINES DE LIAISON A L'ARN</b> .....	<b>5</b>
1.	CARACTERISTIQUES GENERALES.....	5
a)	Différents types de domaines de liaison à l'ARN.....	6
b)	Différents modes d'interaction.....	12
2.	CAS PARTICULIER DES RBP AVEC DES LCD.....	16
a)	Présentation des LCD et des caractéristiques de différents domaines fréquemment rencontrés .....	16
b)	Le domaine de faible complexité se retrouve fréquemment chez les RBP liées aux maladies neurodégénératives .....	18
c)	Fonctions des LCD.....	21
3.	FORMATION DE CONDENSATS PAR LE BIAIS DE SEPARATION DE PHASES LIQUIDE-LIQUIDE.....	25
a)	Mécanisme général de LLPS.....	25
b)	Formation de condensats spécifiques à FUS et TDP-43 .....	27
c)	Granules de stress .....	29
d)	De la séparation de phases à l'agrégation .....	30
4.	LES MALADIES NEURODEGENERATIVES EN LIEN AVEC TDP-43 ET FUS.....	38
a)	Trois maladies qui s'entrecroisent : La SLA, la DFT et la LATE.....	38
b)	Lien entre ces maladies.....	42
<b>II.</b>	<b>TDP-43 ET FUS, DEUX PROTEINES SEMBLABLES QUI INTERAGISSENT</b> .....	<b>45</b>
1.	LA PROTEINE TDP-43.....	45
a)	Caractéristiques, domaines, et structure.....	45
b)	Mode d'interaction et ARN pris en charge par TDP-43 .....	49
2.	LA PROTEINE FUS.....	50
a)	Caractéristiques, domaines et structure.....	51
b)	Mode d'interaction et ARN pris en charge par FUS .....	53
3.	FONCTIONS DE FUS ET TDP-43 .....	55
a)	Fonctions en lien avec l'ADN .....	56
b)	Fonctions en lien avec le métabolisme de l'ARNm.....	58

c)	Prise en charge des LncARN et miARN .....	61
d)	Autres fonctions cellulaires.....	62
4.	MODIFICATIONS ET MUTATIONS DE CES PROTEINES.....	63
a)	Modifications post-traductionnelles .....	63
b)	MPT et mutations en lien avec les maladies.....	67
c)	Comparaison et mise en évidence des points communs entre les deux protéines .....	73
5.	TDP-43 ET FUS INTERAGISSENT .....	77
a)	Mise en évidence de leur interaction.....	77
b)	Conséquences sur leurs fonctions physiologiques ?.....	78
c)	En conditions pathologiques .....	81

## **OBJECTIFS ..... 85**

## **MATERIEL ET METHODES ..... 87**

<b>I.</b>	<b>COMPARAISON DES METHODES POUR L'IDENTIFICATION DE L'INTERACTION PROTEINE- PROTEINE ET OUVERTURE SUR LE <i>MICROTUBULE BENCH</i>.....</b>	<b>87</b>
1.	METHODES CONVENTIONNELLES .....	88
a)	Méthode <i>In vitro</i> .....	88
b)	Méthode <i>In vivo</i> ou <i>In situ</i> .....	92
2.	UNE NOUVELLE METHODE : <i>LE MICROTUBULE BENCH</i> .....	96
a)	Description de la technique.....	96
b)	Avantages et inconvénients.....	97
3.	COMPARAISON DES METHODES .....	98
<b>II.</b>	<b>METHODES UTILISEES AU COURS DE CE TRAVAIL .....</b>	<b>100</b>
1.	EXPERIMENTATIONS AVEC MODELE CELLULAIRE.....	100
2.	EXPERIMENTATIONS <i>IN VITRO</i> .....	107

## **RESULTATS..... 114**

<b>I.</b>	<b>ETUDE DU COMPORTEMENT DES FORMES PLEINE LONGUEUR .....</b>	<b>114</b>
1.	DETECTION DE L'INTERACTION ENTRE TDP-43 ET FUS ET COMPARAISON AVEC D'AUTRES PROTEINES DE LIAISON A L'ARN .....	114
a)	L'interaction entre TDP-43 et FUS est plus importante qu'avec d'autres RBP. ....	115
b)	Confirmation de l'interaction à l'aide de la PLA.....	119
c)	FUS et TDP-43 présentent des comportements différents .....	120
2.	LA MISCIBILITE ENTRE TDP-43 ET FUS DEPEND DU RATIO DE CES PROTEINES.....	123
3.	ETUDE DU COMPORTEMENT DES PROTEINES <i>IN VITRO</i> .....	128
a)	Première approche, les gels de sédimentation .....	128
b)	Observation des agrégats et condensats par immunofluorescence <i>In vitro</i> ....	130
c)	Première approche pour le mélange des protéines <i>In vitro</i> .....	132
d)	Etude du mélange <i>In vitro</i> avec différents ratios .....	133
4.	ANALYSE DE LA STRUCTURE DES CONDENSATS DE TDP-43 ET FUS EN PRESENCE D'ARN .....	135

a)	Effet de l'ARN sur les protéines seules .....	135
b)	Effet de l'ARN sur le mélange FUS et TDP-43.....	137
c)	L'ARN améliore le mélange des protéines .....	138
d)	Observations des structures formées par les protéines en 3 dimensions .....	140
5.	TDP-43 ET FUS COOPERENT POUR SE LIER A L'ARN.....	141
a)	Première approche par gel retard .....	141
b)	Influence de FUS sur l'interaction de TDP-43 avec l'ARN révélée par AFM.....	144
<b>II.</b>	<b>ETUDE DES FORMES TRONQUEES.....</b>	<b>146</b>
1.	MESURE DE L'INTERACTION DES FORMES TRONQUEES AVEC LES FORMES PL GRACE AU MTB.....	146
a)	Les formes tronquées induisent une séparation avec les formes PL.....	146
b)	L'interaction entre FUS et TDP-43 est favorisée par le RRM mais est rompue par les formes tronquées pathologiques de TDP-43.....	148
2.	ETUDE DES FORMES TRONQUEES <i>IN VITRO</i> .....	151
a)	Contrôle des formes tronquées produites .....	151
b)	Aperçu du comportement des formes tronquées seules par immunofluorescence en microscopie optique.....	152
c)	Comportement des formes tronquées en mélange avec l'ARN et les formes PL.. .....	154
3.	ETUDE DES FORMES TRONQUEES PAR AFM.....	158
a)	Le comportement des formes tronquées pathologiques diffère de celui des formes PL .....	158
b)	Mise en évidence de certaines structures formées par les formes tronquées <i>in vitro</i> .....	162
	<b>DISCUSSION.....</b>	<b>165</b>
	<b>CONCLUSION ET PERSPECTIVES.....</b>	<b>173</b>
	<b>TABLE DES ILLUSTRATIONS .....</b>	<b>177</b>
	<b>LISTE DES TABLEAUX.....</b>	<b>179</b>
	<b>LISTE DES BOITES DE TEXTE .....</b>	<b>179</b>
	<b>REFERENCES.....</b>	<b>180</b>

---





# Abréviations

---

Un nombre important d'abréviations listées ci-dessous sont tirées de la littérature scientifique anglophone. Le choix a été fait de conserver la terminologie technique dans le texte, la traduction associée est présentée ici.

<b>AA</b>	Acides aminés
<b>Ac</b>	Anticorps
<b>AFM</b>	<i>Atomic Force Microscope</i> -Microscope à Force Atomique
<b>ARN</b>	Acide ribonucléique
<b>ARNm</b>	ARN messenger
<b>Lnc/ncARN</b>	Long ARN non-codant
<b>miARN</b>	Micro-ARN
<b>Pré-ARNm</b>	ARN pré-messenger
<b>pri-miARN</b>	Transcrit primaire des micro-ARN
<b>BET</b>	Bromure d'éthidium
<b>BSA</b>	<i>Bovine serum albumin</i> -Albumine de sérum bovin
<b>CFP</b>	<i>Cyan fluorescent protein</i> -protéine fluorescente cyan
<b>CoIP</b>	Co-immuno-Précipitation
<b>CSD</b>	<i>Cold shock domain</i> -Domaine de choc froid
<b>Cter</b>	Partie C-terminale de la protéine
<b>DFT</b>	Dégénérescence fronto-temporale
<b>DFT-U</b>	Dégénérescence fronto-temporale avec ubiquitine
<b>dsRBM</b>	<i>Double-stranded RNA binding motif</i> -Motif de liaison à l'ARN double brin
<b>DTT</b>	Dithiothréitol
<b>FUS (TLS)</b>	<i>Fused in sarcoma (Translated In sarcoma)</i>
<b>GFP</b>	<i>Green fluorescent protein</i> -Protéine fluorescente verte
<b>GS</b>	Granules de stress
<b>GST pull-down</b>	Technique de détection des interactions par le biais d'une étiquette GST (Glutathion S-transférase)
<b>GU</b>	Guanine-Uracil (motif)

<b>HA</b>	<i>Human influenza hemagglutinin</i> (Etiquette pour marquage)
<b>His<sub>6</sub></b>	Etiquette de 6 histidines (pour la purification de protéine)
<b>hnRNP</b>	<i>Human heterogeneous nuclear ribonucleoprotein</i> -Ribonucleoprotéine nucléaire hétérogène humaine
<b>HSPG</b>	<i>Heparan sulfate proteoglycans</i> -Sulfate d'héparane protéoglycanes
<b>iCLIP</b>	<i>Individual-nucleotide resolution cross-linking immunoprecipitation</i> (Technique d'identification des interactions protéine-ARN)
<b>IDR</b>	<i>Intrinsically disordered region</i> -Région intrinsèquement désordonnée
<b>iPSC</b>	<i>Induced pluripotent stem cells</i> -Cellules souches pluripotentes induites
<b>ITC</b>	<i>Isothermal titration calorimetry</i> -Calorimétrie de titrage isotherme
<b>KH</b>	<i>K homology</i> -Homologue au K
<b>LARKs</b>	<i>Low complexity aromatic-rich kinked segment/Low-complexity amyloid-like reversible kinked segment</i> -Segment de faible complexité riche en aromatiques/de type amyloïde réversible
<b>LATE</b>	<i>Limbic-predominant Age-related TDP-43 Encephalopathy</i> -Encéphalopathie TDP-43 à prédominance limbique liée à l'âge
<b>LCD</b>	<i>Low complexity domain</i> -Domaine de faible complexité
<b>LLPS</b>	<i>Liquid-liquid Phase separation</i> -transition de phase liquide-liquide
<b>MBD</b>	<i>Microtubule Binding Domain</i> -Domaine de liaison aux microtubules
<b>MND</b>	Maladies neurodégénératives
<b>MPT</b>	Modifications post-traductionnelles
<b>MS</b>	<i>Mouse</i> -Souris (espèce de provenance de l'anticorps)
<b>MtB</b>	<i>Microtubule Bench</i>
<b>NES</b>	<i>Nuclear exportation signal</i> -Signal d'exportation nucléaire
<b>NLS</b>	<i>Nuclear localization signal</i> -Signal de localisation nucléaire
<b>NS</b>	Non significatif
<b>Nter</b>	Partie N-terminale de la protéine
<b>PBS</b>	<i>Phosphate buffered saline</i> -Tampon phosphate salin

<b>PFA</b>	Paraformaldéhyde
<b>PL</b>	Pleine longueur
<b>PLA</b>	<i>Proximity ligation assay</i> -Test de ligature de proximité
<b>PrLD</b>	<i>Prion-like domain</i> -Domaine de type Prion
<b>RB</b>	<i>Rabbit</i> -Lapin (espèce de provenance de l'anticorps)
<b>RBD</b>	<i>RNA binding domain</i> -Domaine de liaison à l'ARN
<b>RBP</b>	<i>RNA binding protein</i> -Protéine de liaison à l'ARN
<b>RFP</b>	<i>Red fluorescent protein</i> -Protéine fluorescente rouge
<b>RGG</b>	Domaine contenant des répétitions RGG : riches en Arginines (R) et en glycines (G)
<b>RMN</b>	Résonance magnétique nucléaire
<b>RNAseq</b>	Technique de séquençage de l'ARN
<b>RNP</b>	<i>Ribonucleoprotein</i> -Protéine Ribonucléique
<b>RRM</b>	<i>RNA Recognition Motif</i> -Motif de reconnaissance de l'ARN
<b>qRRM</b>	<i>quasi-RNA Recognition Motif</i>
<b>SDS</b>	<i>Sodium dodecyl sulfate</i> -Dodécylsulfate de sodium
<b>SLA</b>	Sclérose latérale amyotrophique
<b>SLAf</b>	Forme familiale de la SLA
<b>SLAs</b>	Forme sporadique de la SLA
<b>SPR</b>	<i>Surface plasmon resonance</i> -Résonance plasmonique de surface
<b>TDP-43</b>	<i>Transactive response DNA-binding protein of 43 kDa</i>
<b>UTR</b>	<i>Untranslated region</i> -Région non traduite
<b>YFP</b>	<i>Yellow fluorescent protein</i> -Protéine fluorescente jaune
<b>Y2H</b>	<i>Yeast two-hybrid</i> -Essai à deux hybrides chez la levure
<b>ZnF</b>	<i>Zinc Finger</i> -Doigt de zinc



# Introduction

---

A travers le monde, nous observons une forte croissance de personnes touchées par les maladies neurodégénératives (MND), notamment du fait du vieillissement de la population, ce qui en fait un problème de santé publique majeur[1]. Parmi celles-ci, la sclérose latérale amyotrophie (SLA), la dégénérescence fronto-temporale lobaire (DFT), ou encore l'encéphalopathie TDP-43 à prédominance limbique liée à l'âge (LATE), se manifestent par des symptômes différents, mais présentent des caractéristiques communes, comme la présence de protéines agrégées dans le cytoplasme des neurones affectés. Bien que les mécanismes qui conduisent à la dégénérescence des neurones et au développement de ces pathologies soient encore largement mal compris, des pistes se dessinent. Les inclusions dans les neurones de patients contiennent des protéines qui ont la capacité de former des condensats par le mécanisme de séparation de phases liquide-liquide (LLPS). On suspecte donc que ce mécanisme pourrait conduire à l'agrégation, ou encore des dysfonctionnements dans le métabolisme de l'ARN pourraient induire la mort des cellules[2]. C'est ainsi, qu'un lien avec les protéines de liaison à l'ARN, à la fois impliquées dans le processus de séparation de phases et la prise en charge de l'ARN s'esquisse.

Les protéines de liaison à l'ARN sont caractérisées par la présence d'au moins un domaine de liaison à l'ARN (RBD) au sein de leur séquence. Ces domaines peuvent avoir des structures variées, associées à une spécificité d'interaction plus ou moins marquée, tout en permettant une reconnaissance de l'ARN avec une bonne affinité. Un autre type de domaine est fréquemment rencontré chez cette famille de protéines, il s'agit du domaine de faible complexité (LCD). Ce dernier est connu pour être impliqué dans la mise en place d'interactions multivalentes. Les RBP participent à des mécanismes complexes qui nécessitent des contacts protéine-ARN, et protéine-protéine formant des complexes dynamiques, assurant une large palette de fonctions.

TDP-43 et FUS sont deux protéines de liaison à l'ARN, de la famille des hnRNP qui ont un domaine de reconnaissance de l'ARN (RRM) canonique, et un domaine de faible complexité de type Prion (PrLD). Elles se ressemblent d'un point de vue structural, et participent à de nombreuses fonctions analogues. Les deux sont ainsi impliquées dans la transcription, l'épissage, la stabilisation, et le transport de l'ARN mais aussi la réparation de l'ADN[3]. De plus, elles sont localisées dans les mêmes compartiments ou complexes au sein des cellules[3]. Elles sont toutes deux principalement nucléaires, mais font la navette vers le cytoplasme pour le transport de l'ARN, on les retrouve au niveau des synapses pour la traduction locale, ou encore

dans les complexes de Drosha. Si elles ont de nombreux points communs, leurs modes de liaison à l'ARN sont différents. TDP-43 se lie grâce à ses deux RRM avec une grande affinité et une grande spécificité pour les motifs GU. FUS utilise à la fois son RRM, ses RGG et son doigt de zinc (ZnF) pour se lier à l'ARN davantage par reconnaissance de structure plutôt que par spécificité de séquence[4]. Pour finir, TDP-43 et FUS se lient à environ 30% du transcriptome dans le cerveau humain[5], elles partagent ainsi des cibles communes, notamment des ARNm (ARN messenger) essentiels au bon développement des neurones[6], mais leurs cibles sont largement distinctes[6].

TDP-43 et FUS sont des protéines qui sont intrinsèquement prédisposées à l'agrégation[7, 8] et leur liaison à l'ARN est essentielle pour les maintenir solubles[9]. Ce dernier joue, en effet, un rôle prépondérant dans la formation et la dissolution des complexes. A forte concentration, il induit leur solubilisation, tandis qu'en faible quantité, il favorise les interactions multivalentes entre protéines via le LCD, ce qui peut conduire à la formation d'agrégats. Ce phénomène peut s'observer dans le cytoplasme en conditions de stress, où TDP-43 et FUS s'assemblent en granules avec d'autres RBP pourvues d'un LCD[3, 10]. De cette manière, les granules de stress sont suspectées de favoriser la formation d'inclusions que l'on retrouve dans les cas des MND[11, 12].

L'interaction d'une protéine avec son environnement est primordiale pour l'exercice de ces fonctions, une protéine interagit avec de multiples partenaires. Le comportement des protéines est régi par différents modes : la compétition, la coopération ou la régulation mutuelle[13]. Des processus comme les modifications post-traductionnelles (MPT) permettent de réguler le comportement de ces protéines. Caractériser les interactions protéine-protéine permet, alors, de mieux comprendre leurs fonctions, et les conséquences que peut entraîner un dysfonctionnement, en particulier dans le contexte des LLPS, qui dans certaines conditions, peuvent évoluer vers l'agrégation de protéines, et conduire à l'apparition de troubles pathologiques. Plusieurs études ont déjà mis en évidence une interaction entre TDP-43 et FUS[14–16]. Par ailleurs, des études au sein de modèles animaux indiquent que ces deux protéines sont impliquées dans des processus communs de maturation de l'ARN[17–19]. De plus, ces protéines peuvent co-localiser dans les agrégats de patients atteints de MND[20]. Partant de ce constat, nous cherchons à mieux comprendre cette interaction, et plus précisément à étudier l'influence que ces protéines peuvent avoir l'une sur l'autre, et quel est leur comportement vis-à-vis de l'ARN, en abordant des notions de miscibilité et de séparation de phases.

Ainsi, dans ce manuscrit de thèse nous tâcherons de mieux comprendre la relation entre ces deux RBP, TDP-43 et FUS. Dans une première partie, nous prendrons le temps de définir les protéines de liaison à l'ARN, nous décrirons ce qui les caractérise, en nous focalisant en particulier sur le cas des domaines de faible complexité. Nous étudierons ensuite le mécanisme de séparation de phases, puis les mécanismes qui peuvent conduire au développement des maladies neurodégénératives. Dans une 2<sup>ème</sup> partie nous présenterons plus en détail les deux protéines qui sont au cœur de ce rapport, TDP-43 et FUS, nous présenterons leurs caractéristiques, leurs fonctions, ainsi que leurs ARN cibles, pour mieux cerner leurs ressemblances, mais aussi leurs divergences. Enfin, nous finirons par montrer qu'elles interagissent au sein des cellules, et nous verrons les implications de cette interaction en conditions physiologiques et pathologiques. Nous définirons ensuite les objectifs de ce travail de thèse, avant de présenter les principaux résultats obtenus, en particulier que TDP-43 et FUS interagissent et coopèrent dans leur prise en charge de l'ARN. Mais que cette interaction dépend fortement de la proportion respective de ces protéines. De plus, cette coopération peut être perturbée dans le cas de tronctions pathologiques. Cette interaction sera confirmée par le biais de techniques diverses et complémentaires, comme la *PLA*, le *microtubule bench*, et le gel retard. De plus, l'immunofluorescence *in vitro* nous permettra de mieux cerner les séparations de phases et les notions de miscibilité. Enfin, une étude à l'échelle de la molécule unique (AFM, Microscope à Force Atomique) a permis de mieux comprendre comment s'organisaient les protéines sur l'ARN. Dans une dernière partie, nous discuterons de la conséquence de ce résultat en le confrontant à d'autres travaux, avant de conclure et d'ouvrir sur les perspectives amenées par ce travail.



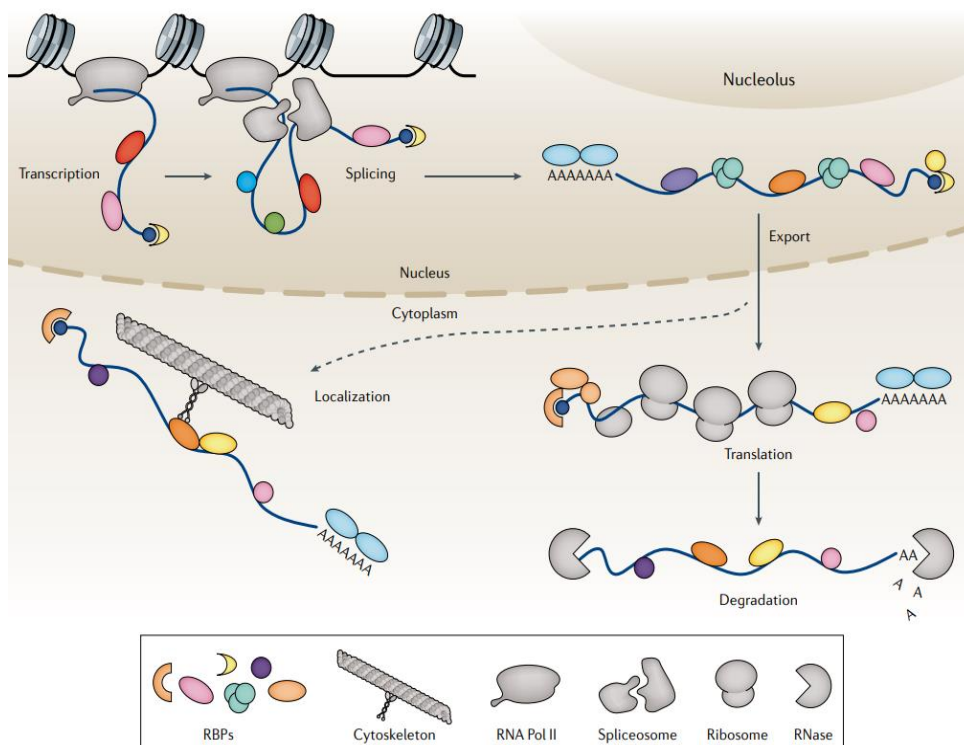


# Bibliographie

## I- Les protéines de liaison à l'ARN

### 1. Caractéristiques générales

Au sein des cellules, un nombre important de protéines ont les capacités de se lier à l'ARN, et de former des complexes qui peuvent être stables ou transitoires, ce sont les protéines de liaison à l'ARN (RBP), impliquées dans la formation de complexes ARN/protéine, les ribonucléoprotéines (RNP). Les RBP sont souvent des protéines très conservées à travers les espèces. Chez l'humain environ 1500 RBP, soit 7,5 % du protéome, ont été répertoriées. Leur détection et leur classement passent par des analyses computationnelles, qui repèrent les caractéristiques typiques des domaines de liaison à l'ARN (RBD) au sein de leur séquence[21]. Les modes d'interaction et les domaines impliqués dans l'interaction sont variés mais présentent des caractéristiques communes. Ces RBP occupent des fonctions importantes pour le métabolisme cellulaire, et très diversifiées comme la régulation de la synthèse des protéines, l'épissage, la maturation de l'ARN, ou son transport (**voir figure 1**). Par



**Figure 1- Les protéines de liaisons à l'ARN contrôlent la vie de l'ARN.**

Les RBP sont impliquées dans un ensemble de fonctions très variées incluant des fonctions nucléaires (transcription, épissage, polyadénylation, exportation), et cytoplasmiques (transport, localisation, traduction, dégradation). RBP : Protéine de liaison à l'ARN ; RNA pol II : Polymérase ARN II. **D'après** Gebauer F, et al. RNA-binding proteins in human genetic disease. *Nat Rev Genet.* 2021 Mar;22(3):185-198. doi: 10.1038/s41576-020-00302-y. Epub 2020 Nov 24. PMID: 33235359.

ailleurs, les RBP peuvent être impliquées dans diverses pathologies en particulier les maladies neurodégénératives[22], le développement de cancers[23] ou encore certaines maladies génétiques[24]. L'interaction avec l'ARN peut être vue de deux façons : la RBP se lie à l'ARN pour assurer certaines fonctions, ou inversement, l'ARN peut interagir avec une RBP pour affecter sa régulation, sa stabilité, ses interactions[21]. Dans cette partie, nous allons passer en revue les différents types de domaines de liaison à l'ARN, avant de nous intéresser aux modes d'interaction qui régissent les comportements de ces protéines, et finalement nous pencher sur une famille de RBP pourvue de domaines de faible complexité.

#### a) Différents types de domaines de liaison à l'ARN

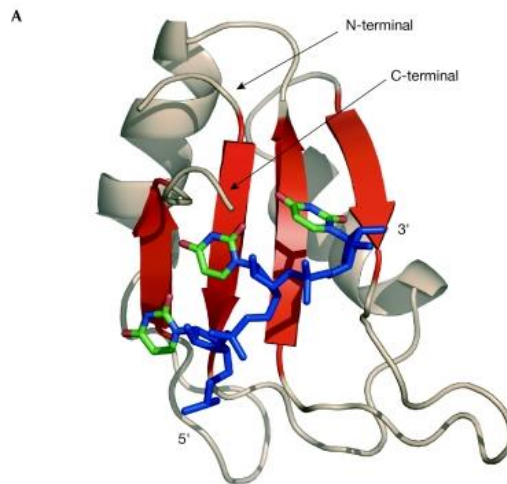
Les RBP peuvent être catégorisées selon leurs domaines de liaison à l'ARN, qui présentent des particularités en termes de spécificités, et de fonctions. Les RBP possèdent souvent des répétitions d'un même domaine, et peuvent aussi en posséder plusieurs différents. Classiquement, l'interaction entre une protéine et l'ARN passe par un domaine structuré et bien défini du type RRM, mais ce n'est pas le seul mode d'interaction possible. Ici nous allons détailler quelques domaines d'interactions fréquemment rencontrés, les plus conventionnels étant les RRM, les KH, les boîtes DEAD et les Znf, mais il existe aussi d'autres domaines moins abondants (environ 30 différents[24]), certains sont plus délicats puisqu'ils n'impliquent pas de structures particulières, nous les développerons ici sous le nom de régions intrinsèquement désordonnées (IDR).

##### ▪ Les domaines de reconnaissance de l'ARN : RRM

Le RRM est le domaine le plus communément rencontré au sein de la famille des protéines de liaison à l'ARN. Il se retrouve chez l'humain dans 0,5 à 1 % des gènes[25] ce qui représente plus de 226 RBP. Il est généralement composé de 75 à 85 acides aminés (AA), formant une succession de 4 brins  $\beta$  et 2 hélices  $\alpha$  suivant l'alternance suivante :  $\beta\alpha\beta\alpha\beta$ [26]. Le RRM est souvent présent en plusieurs copies dans la séquence des protéines. L'interaction du RRM avec l'ARN passe par le contact des feuilletts  $\beta$  à 4 brins qui peut prendre en charge 2 ou 3 nucléotides. L'interaction se fait grâce aux acides aminés contenant des chaînes latérales hydrophobes aromatiques et les bases des acides nucléiques orientées en parallèle aux brins. De cette manière, la reconnaissance des nucléotides est spécifique. Les boucles non structurées, qui varient en taille comme en composition selon les protéines, sont aussi importantes pour la liaison à l'ARN (**voir figure 2**). Le RRM est un domaine assez plastique qui permet d'assurer de nombreuses fonctions (transport, épissage, stabilisation de l'ARN), cette plasticité structurale lui offre une large capacité d'adaptation pour une interaction avec l'ARN à haute affinité et très spécifique[27].

On distingue généralement plusieurs types de RRM. Le **RRM canonique** est le plus couramment rencontré. Il contient deux séquences d'acides aminés très caractéristiques impliquées dans la formation de RNP, l'un se situe dans le brin  $\beta 3$  et est de la forme ([RK]-G-[FY]-GA-[FY]-[ILV]-X-[FY]), l'autre se situe dans le  $\beta 1$  et a pour séquence ([ILV]-[FY]-[ILV]-X-N-L) (ici X peut être remplacé par n'importe AA). La manière de se lier aux acides nucléiques est très semblable pour toutes les protéines contenant ces motifs[28].

Le **quasi-RRM** ou qRRM se distingue du RRM canonique par la présence de trois séquences d'acides aminés conservés dans les boucles reliant les brins  $\beta$ . La première dans la boucle 1 consiste en une séquence du type R-G-L-P-(W/F/Y), la seconde, dans la boucle 5 est de la forme (R/K)-(5AA)-R-Y-(V/I/L)-E-(V-I-L)-F, enfin la dernière contient des résidus chargés positivement au niveau de la boucle 3[29]. Ces boucles sont ainsi impliquées dans la liaison à l'ARN, davantage que les feuillettes, et reconnaissent en particulier les séquences G-tract de l'ARN (répétition d'au moins trois nucléotides guanines). C'est un domaine que l'on retrouve en particulier chez certaines hnRNP comme la hnRNP F[29].



**Figure 2-Le RRM reconnaît spécifiquement la séquence de l'ARN.**

Structure d'un RRM canonique (protéine sex-lethal) qui interagit avec un triplet de nucléotides UUU. Les 4 brins  $\beta$ , mis en évidence en rouge, reconnaissent les nucléotides, en vert, de manière spécifique. **D'après** Stefl R, et al. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep.* 2005 Jan;6(1):33-8. doi: 10.1038/sj.embor.7400325. PMID: 15643449; PMCID: PMC1299235.

Le **RRM hélice** ou xRRM est un domaine qui a la particularité de reconnaître l'ARN double brin. Il diffère du RRM canonique de par sa structure secondaire dans laquelle une troisième hélice  $\alpha$  est présente, ainsi qu'un court feuillet supplémentaire ( $\beta 4'$ ), positionné de cette manière :  $\beta 1$ - $\alpha 1$ - $\beta 2$ - $\beta 3$ - $\alpha 2$ - $\beta 4'$ - $\beta 4$ - $\alpha 3$ . De plus, les séquences d'acides aminés conservés, retrouvées dans les feuillettes  $\beta 1$  et  $\beta 3$  du RRM

canonique sont absentes. C'est alors cette hélice supplémentaire ainsi que 2 acides aminés conservés dans le feuillet  $\beta 2$  qui assurent la liaison au squelette de l'ARN et non directement à des nucléotides spécifiques. Ce domaine se retrouve chez la télomérase p65[30].

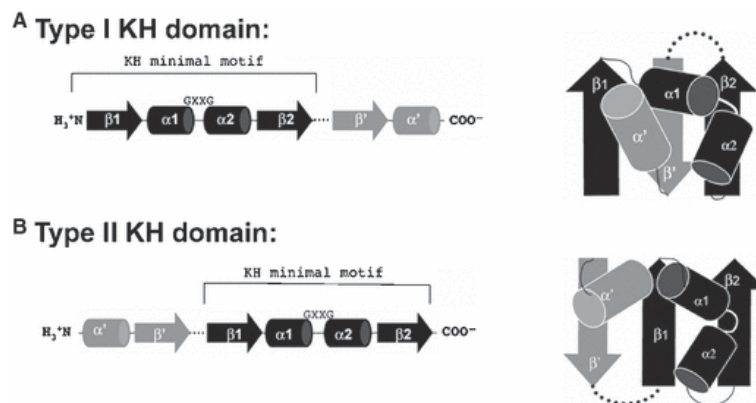
- **DEAD box hélicase**

Certains RBD peuvent être retrouvés dans des protéines impliquées dans des activités enzymatiques. Dans ce cas la liaison de l'ARN à certains sites peut activer ou réprimer une activité catalytique. L'interaction ARN-RBD induit un changement structural favorisant la transformation d'un substrat. C'est le cas des hélicases[31].

Les protéines DEAD box forment la plus importante famille des hélicases qui sont caractérisées par la présence d'un motif Asp-Glu-Ala-Asp d'où l'appellation DEAD, cette famille compte 37 membres chez l'humain. Un domaine similaire est présent chez d'autres hélicases, appelées DEAH, dans lequel l'asparagine est remplacée par une histidine. Les protéines appartenant à la famille des DEAD/H box ont en commun un cœur d'hélicase hautement conservé, avec en particulier des séquences de liaison à l'ATP et à l'ARN[32]. Deux domaines hélicase forment une poche qui abrite le site de liaison à l'ATP, à l'opposé de ce site on trouve le site de liaison à l'ARN. La poche doit être fermée pour se lier et hydrolyser l'ATP, qui est pris en charge par une dizaine d'acides aminés principalement polaires, dont ceux constituant le motif DEAD, formant alors un réseau complexe d'interaction[33]. L'ARN est pris en charge par un ensemble d'acides aminés qui couvrent 5 nucléotides tout en impliquant le squelette de sucre phosphate de l'ARN[34]. Un grand nombre de fonctions est associé à ces RBP, cela va de la régulation de la transcription, à la maturation de l'ARN et à sa dégradation[32].

- **Le domaine KH**

Le domaine KH, pour *K homology* est caractéristique de la famille des HnRNP (*Human heterogeneous nuclear ribonucleoprotein*), parmi laquelle la hnRNP K a été la première à être identifiée. Ce domaine constitué d'environ 70 acides aminés se retrouve dans différentes protéines issues d'espèces diverses y compris chez les archées et les bactéries. C'est un motif qui se retrouve la plupart du temps en plusieurs répétitions au sein d'une séquence. Ces domaines peuvent fonctionner indépendamment ou avec coopérativité[35].



**Figure 3-Structure des domaines KH.**

On distingue 2 types de domaines KH dont la séquence et la structure secondaire sont présentées ici. Le type I (A) se trouve majoritairement chez les eucaryotes, et consiste en 3 feuilles  $\beta$  tête-bêche, le type II plutôt chez les procaryotes, consiste en deux brins parallèles et un brin en antiparallèle. Les lignes en pointillés représentent les boucles variables. **D'après** Valverde R, et al. Structure and function of KH domains. FEBS J. 2008 Jun;275(11):2712-26. doi: 10.1111/j.1742-4658.2008.06411.x. Epub 2008 Apr 15. PMID: 18422648.

Il existe deux types de domaines KH dont les séquences sont semblables, mais les conformations sont différentes (Type I et II). Les deux types sont composés de trois brins  $\beta$  reliés par 3 hélices  $\alpha$ . Dans un cas les 3 brins  $\beta$  sont antiparallèles, tandis que dans le second cas, deux sont parallèles et un est en antiparallèle (**voir figure 3**). Les boucles qui relient les différentes structures ensemble varient selon les protéines considérées. En général, on rencontre le type I chez les eucaryotes[35]. La liaison avec l'ADN se fait au niveau des hélices  $\alpha_1$ ,  $\alpha_2$ , du brin  $\beta_2$  ainsi qu'une boucle nommée GXXG qui peuvent prendre en charge 5 acides nucléiques. La poche centrale de liaison a une tendance plutôt hydrophobe avec des interactions supplémentaires variées qui stabilisent le complexe. La différence avec d'autres domaines de liaison à l'ARN est que dans le cas du KH, il n'existe pas de chaîne aromatique cyclique qui favorise les liaisons d'empilement avec les acides nucléiques ( $\pi$ -stacking)[27].

#### ▪ Les domaines CSD

Le CSD pour *Cold-shock domain*, est appelé ainsi car il a été découvert avec la protéine CspA, chez la bactérie, chez qui une chute soudaine de température provoque une augmentation de l'expression de cette protéine. Ce type de domaine est caractérisé par un site de liaison formant une structure en feuillet  $\beta$  capable de reconnaître l'ARN et l'ADN simple brin. C'est un domaine très conservé qui est présent chez de nombreux organismes. Les protéines qui contiennent des CSD sont impliquées dans des fonctions variées qui vont de la prise en charge de l'ARN, la transcription des gènes à la traduction. Les protéines YB-1 (*Y-box binding protein 1*) ainsi que la protéine Lin28 sont deux exemples de protéines que l'on retrouve dans

cette famille des protéines à CSD. Elles sont toutes deux impliquées dans le développement de cancers[36].

- **Les domaines à doigt de zinc**

Le domaine à doigt de zinc (ZnF, *Zinc Finger*), du type CCHH, est une structure formée d'un assemblage  $\beta\beta\alpha$  qui forme une épingle ( $\beta$ -*hairpin*) liée à un ion  $Zn^{2+}$ . Il est plus généralement connu pour sa reconnaissance spécifique à l'ADN, néanmoins, il est aussi capable de reconnaître l'ARN[37, 38]. La formation d'un complexe du doigt de Zinc avec l'ARN a été mise en évidence par cristallographie[39] et montre que le ZnF reconnaît spécifiquement les bases de l'ARN double brin, comme il le ferait à peu près avec l'ADN, par contact de la chaîne latérale de la partie Nter de l'hélice qui s'insère dans le sillon principal de la séquence de nucléotides. Mais la reconnaissance de l'ARN double brin peut aussi être non spécifique par contact de l'hélice  $\alpha$  avec le squelette sucre-phosphate de l'ARN.

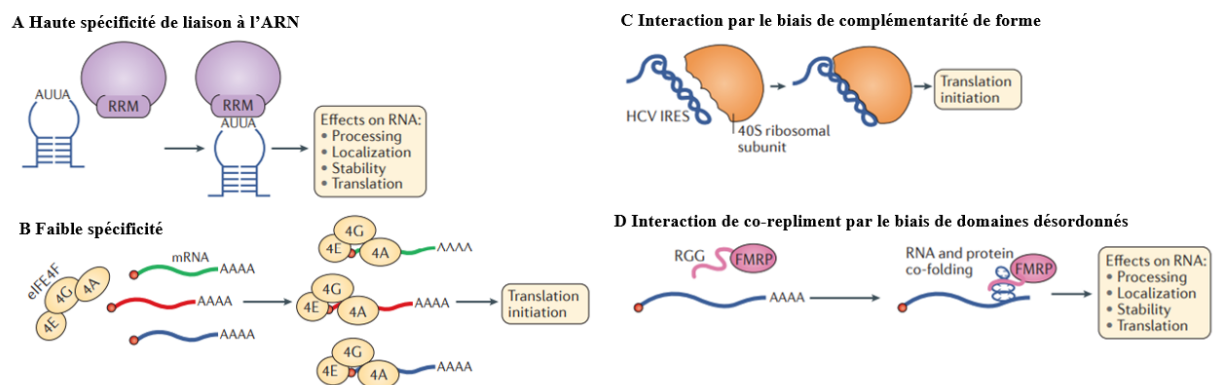
Un autre type de ZnF, le CCCH est capable de reconnaître spécifiquement l'ARN simple brin à travers un réseau intermoléculaire de liaisons hydrogènes entre la chaîne principale de la protéine et les bases nucléotidiques. Ce type de liaison est spécifique et est différent du mode de liaison avec l'ADN[40]. Ainsi les protéines possédant des ZnF sont capables de reconnaître l'ADN de façon directe, l'ARN de façon indirecte par simple reconnaissance de la structure et non de la séquence, mais aussi d'une lecture de l'ARN directe avec une spécificité de reconnaissance des bases.

- **Le domaine de reconnaissance de l'ARN double brin ou dsRBM**

Le dsRBM (*double-stranded RNA binding motif*) est un domaine non spécifique qui reconnaît avant tout la forme de l'ARN, ainsi il permet la liaison à divers ARN structurés en hélice. Ce domaine est généralement présent en mult copie chez une protéine. C'est un domaine composé d'environ 70 acides aminés formant 2 hélices  $\alpha$  opposées à 3 brins  $\beta$  suivant la séquence  $\alpha\beta\beta\beta\alpha$ [27]. Le contact avec l'ARN recouvre 15 bases, la boucle comprise entre les brins  $\beta_1$  et  $\beta_2$  et l'hélice  $\alpha_2$  entrent en contact avec le squelette du sucre phosphate du sillon majeur de l'ARN, ce qui est parfaitement adapté à la structure de l'ARN double brin et non à l'ADN[41]. Dans cette catégorie, on retrouve par exemple les DROSHA et DICER[42] pour la prise en charge des miRNA (micro-ARN), ou encore les kinases PKR qui sont activées par l'interaction avec l'ARN double brin[43] (**voir figure 4, C**).

### ▪ La liaison à l'ARN par le biais de domaines intrinsèquement désordonnés (IDR)

Les RBP contenant des domaines de faible complexité (LCD), qui contiennent une moindre proportion d'acides aminés hydrophobes, au profit d'une importante proportion d'acides aminés polaires ou chargés (Gly, Ser, Arg, Lys, Gln Glu et Asp)[44], peuvent être engagés directement dans la liaison avec l'ARN. C'est le cas des RGG (domaines riches en arginine et en glycine) et des SR (répétition de Serine et Arginine) connu pour se lier à l'ARN[45]. Plus de la moitié des domaines de liaison à l'ARN répertoriés (RBDmap) ont les caractéristiques des IDR, ce sont donc des domaines prépondérants pour l'interaction à l'ARN[46]. De plus, un certain nombre de protéines qui arborent ce type de domaine, semblent interagir uniquement par le biais de ces IDR, ce qui montre que ces régions suffisent pour la liaison à l'ARN[21]. La position des arginines est importante pour la sélectivité de l'interaction avec les séquences d'acides nucléiques riches en guanine, c'est ce qui apparait dans le cas de l'étude de la protéine FMRP[47]. La répétition des acides aminés glycines apporte une grande flexibilité au domaine, qui peut adapter sa conformation et sa forme à l'ARN cible, permettant ainsi sa bonne orientation pour une interaction des paires de bases



**Figure 4-Les protéines de liaison à l'ARN possèdent des modes de liaison variés.**

(A) La reconnaissance de l'ARN peut être spécifique de la séquence comme dans le cas du RRM. Cela se traduit par divers effets sur l'ARN, prise en charge, localisation, stabilisation, traduction. (B) L'interaction peut nécessiter peu de spécificité comme dans le cas des hélicases, ici, l'hélicase eIF4A (4A), participant au complexe du facteur de transcription eIF4F (4F), en association avec eIF4E (4E) et eIF4G (4G). Le complexe s'associe à l'ARN coiffé de manière indépendante à la séquence pour initier la traduction. (C) La RBP peut reconnaître l'ARN par complémentarité de forme, par exemple dans le cas du site d'entrée du ribosome (IRES) du virus de l'hépatite C (HCV). (D) Les régions intrinsèquement désordonnées du type RGG forment un co-repliment avec l'ARN par le biais d'interactions électrostatiques axées sur la complémentarité de forme. Ici le cas de la protéine FMRP (*fragile X mental retardation protein*). **D'après** Hentze MW, et al. *A brave new world of RNA-binding proteins*. *Nat Rev Mol Cell Biol*. 2018 May;19(5):327-341. doi: 10.1038/nrm.2017.130. Epub 2018 Jan 17. PMID: 29339797.



de type Watson-Crick avec l'ARN, cet empilement des acides aminés et des acides nucléiques forme un co-repliment protéine-ARN (**voir figure 4, D**).

Il existe aussi un autre type d'IDR se liant à l'ARN, ce dernier contient des acides aminés aromatiques comme les tyrosines combinés aux glycines ou serines pour former des motifs du type [G/S]Y[G/S] qui peuvent interagir avec les nucléotides par le biais de superposition ( $\pi$ -stacking) ou de liaisons hydrogène[48]. Ainsi, les IDR forment un ensemble malléable, potentiellement multifonction, qui peuvent être impliqués parfois de manière spécifique, et dans d'autres cas de manière non-spécifique.

Plusieurs autres domaines ont été caractérisés, mais n'ont pas été développés ici, comme les LSM, les PUF, les PAZ ou encore les PIWI qui concernent d'autres familles de protéines, et sont moins pertinents dans le cadre de ce rapport. Pour résumer, il existe des modes d'interactions divers entre les protéines et l'ARN, qui sont rendus possibles par des domaines caractéristiques. Ces modes de liaison d'une protéine avec des acides nucléiques, présentent des affinités variables, et peuvent être très spécifiques ou au contraire non spécifiques. Finalement, l'interaction avec l'ARN peut être modulée, notamment par des modifications post-traductionnelles, mais aussi par les partenaires d'interactions.

### b) Différents modes d'interaction

Nous avons vu que la prise en charge de l'ARN, impliquait une liaison entre protéines et acides nucléiques par le biais de domaines particuliers. Mais la liaison protéine-ARN ne suffit pas toujours à effectuer la fonction assurée par une RBP, elle résulte plus globalement de la formation de complexes impliquant aussi des contacts protéine-protéine. Dans cette partie, nous allons étudier comment les protéines peuvent interagir autour de la prise en charge de l'ARN et quels sont les processus qui les régissent. Nous distinguerons trois modes d'interaction : la coopération, la compétition et la régulation mutuelle.

#### ▪ Coopération

La liaison de plusieurs biomolécules à l'ARN est souvent nécessaire pour la formation d'un complexe actif et répondant à une fonction donnée. C'est par exemple le cas de l'épissage qui implique plusieurs protéines distinctes. On parle de coopération lors de l'action conjointe de deux RBP qui ont un but commun et agissent en synergie. Elles peuvent alors interagir physiquement l'une et l'autre ou alors simplement partager la même cible sans contact direct (**voir figure 5, A**)[13].

Il est courant d'observer ce mode d'interaction à différentes étapes de la prise en charge de l'ARN. Les protéines de la famille des hnRNP, notamment connues pour

leur implication dans l'épissage alternatif, se comportent généralement de cette manière. Par exemple, la protéine hnRNPL est connue pour coopérer avec PTBP1 par interaction protéine-protéine pour exclure un exon (P3A) d'un pré-ARNm (ARN pré-messager)[49]. Plusieurs protéines HRNPA1 sont capables de se lier de manière coopérative en se répartissant sur l'ARN pour dérouler sa structure secondaire et faciliter l'épissage[50]. Dans le cas de la polyadénylation de l'ARN on retrouve aussi des mécanismes de coopération comme avec les protéines HuR et Hzf qui co-régulent l'exportation de p53[51]. De la même manière, le facteur d'épissage SF1 et le facteur auxiliaire U2AF sont connus pour interagir par le biais de leur RBD, la liaison à l'ARN lors de l'épissage alternatif est renforcée par cette interaction protéine-protéine qui implique le Nter et le RRM3 de U2AF[31]. Enfin, la protéine de liaison à l'ARN, YB-1, se lie à ses ARN cibles de manière coopérative de façon à effectuer une répression de leur traduction[52].

Il a aussi été montré par des expériences *In vitro* sur différents types d'ARN, que FUS fait preuve de coopérativité positive par le biais de son RRM. FUS ne nécessite pas de structure ou de séquence particulière, mais forme des multimères pour se lier d'une manière dépendante à la longueur. On parle de coopérativité positive lorsque la liaison d'une protéine à l'ARN augmente l'affinité des autres protéines par le biais d'interaction protéine-protéine, ce qui est le cas pour FUS puisqu'il a été montré que l'addition d'une protéine favorise l'association de plusieurs autres protéines à un ARN. Dans le cas de l'ARN prD (un ARN de 48 nucléotides qui ne présente pas de séquence spécifique à FUS), par exemple, au moins quatre protéines sont impliquées[53, 54]. D'un autre côté, TDP-43 se lie coopérativement avec les protéines hnRNP A/B grâce à son LCD pour effectuer l'épissage[55]. De plus, la protéine TDP-43 est capable de s'associer en multimères sur l'ARN. L'étude structurale de la protéine par une combinaison des techniques techniques de RMN et de SAXS (*small-angle X-ray scattering*) et de dynamique moléculaire a permis d'identifier l'interface entre l'ARN et les RRM et ainsi de mettre en évidence les acides aminés impliqués dans la dimérisation de la protéine, en particulier, il existe une interface entre la poche du RRM2 autour de la valine 220 et la boucle 3 du RRM1[56].

#### ▪ **Compétition**

A l'opposé de la coopération, les protéines peuvent être en concurrence pour une même cible. Elles ont des buts différents mais un site de liaison identique, c'est le mode compétitif (**voir figure 5, B**). On retrouve ici aussi la famille des hnRNP qui est régit par ce mode d'interaction dans le cas de la prise en charge des ARN pour l'épissage alternatif. Au sein même de cette famille, hnRNPL et hnRNPLL sont en concurrence puisque la première favorise l'exclusion de l'exon P3A, tandis que la

seconde favorise son inclusion[49]. Mais les hnRNP sont aussi en compétition avec d'autres familles, en particulier les protéines SR. Par exemple, la SRSF1 bloque la liaison coopérative de la HNRNPA1 sur les exons[57]. Les processus de stabilisation de l'ARN et du contrôle de la traduction sont aussi régis par des mécanismes de compétition. Par exemple, dans un cas, la protéine AUF1 est en compétition avec HuR (ou ELAVL1) car leurs ARNm cibles (CDKN1A et CCND1) présentent une superposition de leur site de liaison, et elles ont des effets inverses sur ce dernier (amélioration de la stabilisation contre déstabilisation)[58]. Tandis que dans un autre cas HuR et TIA1 sont en concurrence pour exercer la même fonction en se liant au 3'UTR (Région non traduite) de l'ARNm PDCD4[59].

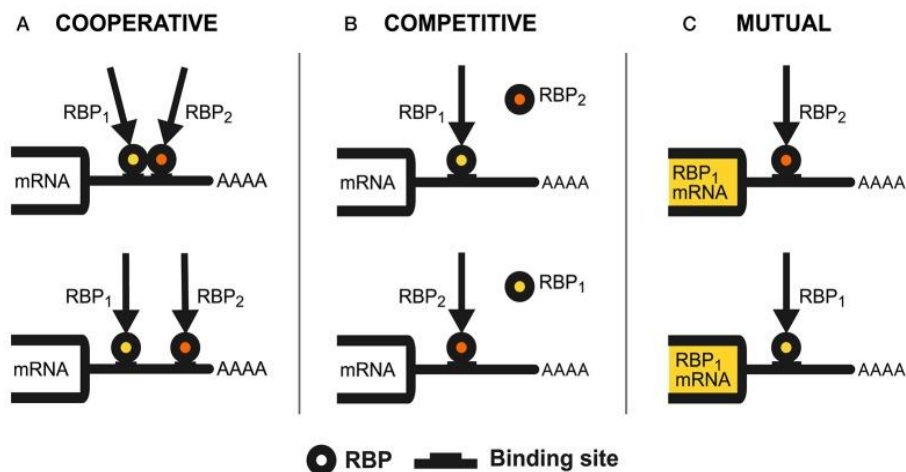
Dans le cas de TDP-43 et FUS, il a été montré que des RBP comme G3BP ou YB-1 pouvaient entrer en compétition avec elles dans l'interaction avec l'ARN, et que cela pouvait favoriser le relargage d'agrégats formés au contact de l'ARN[60].

- **La régulation mutuelle**

Un troisième mode d'interaction qui ne correspond ni à la coopération, ni à la compétition a été identifié, il s'agit du mode mutuel. Le mode mutuel correspond au schéma d'interaction dans lequel l'expression d'une protéine régule l'expression de l'autre (**voir figure 5, C**). En fait, il peut s'agir d'une unique RBP qui s'autorégule en se liant à son propre ARNm, comme de deux RBP distinctes. Dans le premier cas on parle d'une régulation autogène, dans le second cas d'une interaction hétérogène. Ce processus est assez fréquent, en particulier, la régulation autogène est fondamentale pour le maintien de l'homéostasie cellulaire.

Le processus autogène est à l'œuvre pour au moins 57 RBP chez l'homme[13, 61]. C'est un processus qui concerne tout particulièrement les protéines avec des domaines désordonnés qui ont tendance à l'agrégation[62]. Ce qui suggère que ce mode de régulation est tout particulièrement important pour contrôler le niveau d'expression de ces protéines et donc prévenir leur accumulation sous forme d'agrégats. Il est d'ailleurs à l'œuvre pour la protéine TDP-43 qui se lie à son propre ARNm au niveau du 3'UTR pour réprimer sa traduction[63]. Plus largement, c'est un ensemble de mécanismes coordonnés combinant transcription, épissage, et liaison au 3'UTR de l'ARN, impliquant divers facteurs, qui conduit à l'autorégulation de TDP-43 et le maintien d'une concentration adaptée dans la cellule[64]. Un mécanisme similaire a été mis en évidence chez FUS, dont la répression de l'exon 7 de son ARNm lors de l'épissage conduit à son autorégulation[65].

Le processus de régulation hétérogène quant à lui apparaît comme un processus qui régit les RBP de manière hiérarchique. Par exemple, HuR contrôle la régulation de 23 autres RBP[66], mais est elle-même contrôlée par les protéines Hu (HuB, HuC et HuD) qui suppriment son ARNm lors de la différenciation neuronale[67]. En particulier, il a notamment été mis en évidence que HuR agit à la fois sur TDP-43 et sur FUS en régulant leur niveau d'expression[68]. L'épissage est une nouvelle fois au cœur de ce mécanisme de régulation, car c'est lors de ce processus que les RBP interagissent pour contrôler l'expression des autres protéines, comme c'est le cas, par exemple, pour RBFOX2 qui régule 70 RBP en éliminant leur ARN lors d'un mécanisme particulier d'épissage[69]. Enfin, ce mécanisme de régulation est potentiellement à l'œuvre entre FUS et TDP-43, puisqu'il a été détecté des séquences spécifiques de TDP-43 sur le pré-ARNm de FUS, ce qui suggère que la protéine TDP-43 peut le prendre en charge[72,73].



**Figure 5-Différents modes d'interaction entre les protéines de liaison à l'ARN.**

Trois modes de régulation des interactions sont reconnus : (A) le mode coopératif dans lequel deux protéines peuvent interagir physiquement sur un site de liaison commun (haut), mais elles peuvent aussi coopérer sur des sites de liaison distants, sans interaction directe entre elles (bas), il en résulte un effet synergique sur leur cible commune. (B) Le mode compétitif, où les deux RBP partagent un même site de liaison et sont donc en concurrence pour ce site. Leur effet peut être identique ou contraire et le résultat de la régulation est déterminé par une balance entre les deux protéines. (C) La régulation mutuelle, pour laquelle les RBP peuvent réguler l'expression d'une autre RBP ou d'elles-mêmes en se liant à leur propre ARNm. **D'après** Dassi E. *Handshakes and Fights: The Regulatory Interplay of RNA-Binding Proteins*. *Front Mol Biosci*. 2017;4:67. Published 2017 Sep 29. doi:10.3389/fmolb.2017.00067

## 2. Cas particulier des RBP avec des LCD

### a) Présentation des LCD et des caractéristiques de différents domaines fréquemment rencontrés

#### ▪ Qu'est-ce que le LCD

Une protéine a généralement une séquence propre qui permet la formation d'une structure secondaire assurant une fonction précise. Cependant, un certain nombre de protéines ne présentent pas de structure mais assurent des fonctions biologiques complexes[72]. Les domaines de faible complexité sont des séquences avec un biais dans leur composition et notamment une sur-représentation de certains acides aminés. De manière générale, il y a une sous-représentation des acides aminés C, W, Y, F, I, V, et L au profit d'un enrichissement en M, K, R, S, Q, P, et E[73]. On peut ainsi retrouver des répétitions d'un seul ou plusieurs acides aminés, ou alors une distribution apériodique de quelques acides aminés. En général, les LCD ont un caractère désordonné et ne présentent pas de structure secondaire évidente, ce qui les fait apparaître comme des séquences non structurées (*random coil*) avec un aspect très flexible. Toutefois, il est possible de retrouver certaines structures notamment des petites hélices  $\alpha$  au sein de ce domaine. Plus précisément, il apparaît que bien que globalement désordonnés la plupart des LCD peuvent posséder des séquences en partie structurées, c'est ce qui a été mis en évidence par cette étude, qui a calculé que 86 % de ces régions contiennent des informations structurales[74], mais il reste globalement des domaines très flexibles et adaptables, avec une grande capacité à créer des contacts protéine-protéine.

Il existe différents types de LCD, ils peuvent se distinguer selon les acides aminés qui les composent, et les répétitions qu'ils présentent (Région riche en X, où X est un acide aminé quelconque, homorépétition, di-répétition...) ou alors par la structure qui les caractérise (Intrinsèquement désordonnées, bobine, simple hélice  $\alpha$ , flexibilité protéique, fibres amyloïdes)[75]. Ici, nous donnerons des exemples de LCD à travers le domaine de type prion, les RGG et le LARKs.

#### ▪ Le domaine de type Prion

Le domaine Prion a été mis en évidence en premier lieu chez la levure (*Saccharomyces cerevisiae*). Il consiste en un domaine enrichi en acides aminés asparagines, glutamines, tyrosines et glycines qui sont non chargés et favorisent les interactions hydrophobes. Une protéine possède une conformation qui lui est propre et qui peut évoluer dans le cadre de ses interactions. Le domaine prion favorise les mécanismes d'auto-réplication conformationnelle, c'est-à-dire qu'une protéine peut adopter une conformation alternative et la transmettre à d'autres protéines. Ce

phénomène induit des changements dans sa fonction et plus largement dans le phénotype de l'individu. Il peut se perpétuer et même se transmettre entre les individus. Dans le cas de levure, c'est un mécanisme courant qui peut conduire à des phénotypes avantageux, neutres ou délétères, et qui participe à la transmission de l'information et sont donc régis par la sélection naturelle. Par ailleurs, chez l'homme c'est par exemple le cas de la maladie de la vache folle, ou l'encéphalopathie spongiforme bovine qui se transmet par le biais de la protéine Prion (PrP)[76]. En dehors de ce cas pathologique extrême, on retrouve des protéines possédant des caractéristiques d'autoréplication similaires chez l'homme puisqu'elles contiennent des domaines riches en acides aminés non chargés apparentés à celui du Prion chez la levure que l'on nomme donc PrLD pour *Prion-like Domain*. TDP-43 et FUS font partie de cette catégorie de protéine. Les domaines PrLD permettent des interactions de type cation- $\pi$  ou d'empilement  $\pi$ - $\pi$  par le biais d'acides aminés aromatiques. Cela fait des protéines contenant des PrLD des acteurs importants des interactions protéine-protéine.

- **Les motifs RGG ou GAR**

Le RGG est un motif riche en acides aminés arginines et en glycines qui forment un motif qui se répète, il est aussi appelé GAR pour *glycine-arginine-rich region*. Plus de 1000 protéines chez l'homme possèdent un domaine assimilé au RGG. A l'origine, il a été mis en évidence dans le domaine C-terminal de la nucléoline (RGGGFGGRGGFGDRGGRGGGRGG). C'est un site très sensible à la méthylation des arginines par les méthyltransférases[77]. On distingue les motifs tri-RGG qui consistent en 3 répétitions des acides aminés RGG avec de 0 à 4 résidus au milieu : RGG(X0-4)RGG(X0-4)RGG. C'est le cas de FUS par exemple. Il y a les di-RGG de la forme RGG(X0-4)RGG, c'est le cas de Sam68, les di-RG (RG(X0-4)RG) et enfin les tri-RG (RG(X0-4)RG(X0-4)RG)[78]. Ce domaine offre une grande flexibilité aux protéines, et leur permet d'assurer des fonctions diverses, comme l'interaction protéine-protéine, la localisation cellulaire ou la liaison à l'ARN par le biais d'interactions électrostatiques non-spécifiques. Ainsi, parmi les protéines qui possèdent des RGG, la capacité à reconnaître l'ARN est la fonction la plus représentée. On retrouve ce domaine chez les hnRNP comme la hnRNP U qui lui confère la capacité de se lier à l'ARN. Les résidus hydrophobes du RGG permettent les liaisons d'empilement avec l'ARN (**voir chapitre I,1,a, IDR**)[78].

- **Les segments LARKs**

Le domaine LARKs, pour segment de faible complexité riche en aromatiques (*Low complexity aromatic-rich kinded segment* ou *Low-complexity amyloid-like reversible kinded segment*), est un petit segment qui contient des acides aminés

aromatiques de la forme SYSGYS, GYNGFG, ou STGGYG. Il est souvent inclus en plusieurs répétitions dans un LCD plus large. Il est capable de former des interactions de faible énergie par le biais de liaisons hydrogènes ou de Van der Waals, et stabiliser les interactions entre les feuillettes  $\beta$ [79]. Ainsi, ce mécanisme fonctionne comme un scratch (velcro) qui permet l'adhésion des LCD entre eux pour favoriser la formation de condensats (décrit par la suite, **voir I.3**). Ces contacts sont réversibles contrairement à la formation de structures amyloïdogéniques qui fonctionnent comme des fermetures éclairées stériques (*steric-zippers*)[3]. Il y a plus de 400 protéines chez l'homme qui contiennent ce type de domaine avec en moyenne une dizaine de segments LARKs par protéines. Parmi celles-ci 16 %, ont une fonction relative à la liaison à l'ADN et 17 % à l'ARN[79]. Un grand nombre de protéines possédant ce type de segment sont aussi impliquées dans les LLPS. FUS comme TDP-43 arborent ce type de segment au sein de leur LCD[80].

### **b) Le domaine de faible complexité se retrouve fréquemment chez les RBP liées aux maladies neurodégénératives**

#### **▪ Surreprésentation du LCD chez les RBP**

Certaines RBP peuvent être pourvues d'un domaine de faible complexité (LCD), et étrangement une proportion importante des protéines qui sont pourvues d'un domaine de liaison à l'ARN contiennent aussi des LCD. Il y a, en effet, un fort enrichissement du domaine de faible complexité (en particulier pour les domaines riches en Q/N), chez les protéines qui ont la capacité de se lier à l'ARN et à l'ADN[81]. Tous types de domaines de liaison à l'ARN confondus, on compte 40 RBP dans le protéome humain qui possèdent un domaine de type prion (PrLD)[12]. Chez les RBP qui possèdent spécifiquement un RRM, il y a un enrichissement 12 fois plus important du domaine de type prion par rapport aux autres protéines. En effet, sur 210 protéines contenant un RRM canonique, il y en a 26 qui possèdent aussi un domaine de type prion, soit presque 14 %. Inversement, parmi les protéines chez qui on détecte un domaine prion 11,7 % sont des RBP avec un RRM[76]. Il y a, en effet, 246 gènes, soit 1% des gènes chez l'homme codants des protéines contenant un domaine de type prion, cela représente 30 % des protéines de liaison à l'ARN[82] tous types confondus qui possèdent un domaine PrLD. L'ensemble de ces chiffres ont été mis en évidence par des techniques de bioinformatiques[76]. Plus intéressant encore, FUS et TDP-43 se trouvent parmi les protéines contenant un RRM qui ont le comportement de type prion le plus marqué (respectivement 1 et 10 du classement établi[83]). Pour finir, lorsque l'on regarde uniquement les RBP impliquées dans les maladies neurodégénératives une grande majorité contient un LCD.

### Prédire les LCD

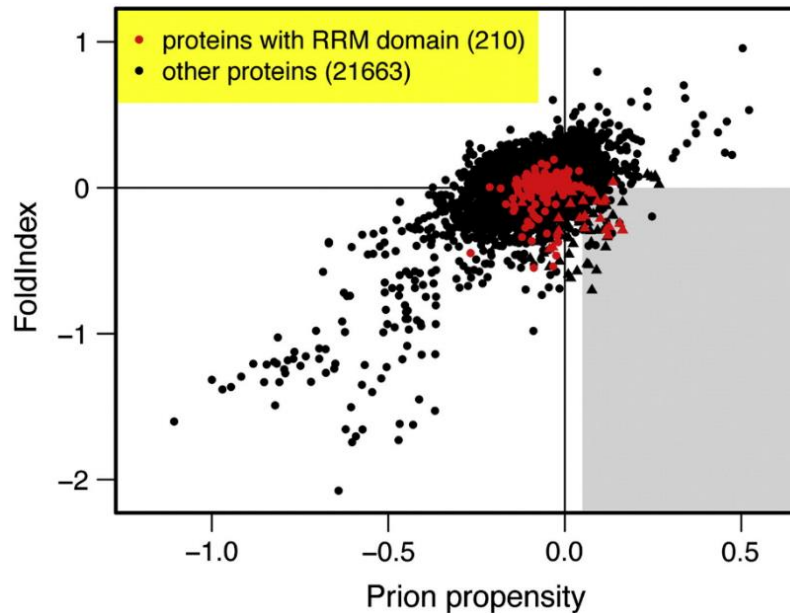
L'identification de domaines de faible complexité et leur catégorisation passent par des analyses informatiques par le biais d'algorithmes spécifiquement développés pour cela. Ils se basent sur la séquence et la composition en acides aminés d'une protéine pour déterminer sa capacité à former des structures ou au contraire l'absence de structure. Le premier à avoir été développé est SEG (1993) qui se base sur la complexité des séquences, ensuite CAST (2000) pour détecter les biais de composition des séquences, puis SIMPLE (2002) qui repère la répétition de motifs. Aujourd'hui, il existe plus de 60 algorithmes de prédiction, qui pour certains utilisent des méthodes comme l'apprentissage machine (*machine learning*) et les réseaux de neurones profonds pour une prédiction plus précise. Un exemple de ces algorithmes est le PONDR (*Predictor of Natural Disordered Regions*) qui atteint une précision de 87 %, il consiste en un réseau de neurones qui a été entraîné en utilisant 162 régions désordonnées (de plus de 30 résidus) provenant de 152 protéines comparées à 290 protéines complètement ordonnées.

A côté des outils informatiques, des techniques comme la cristallographie des rayons X et la RMN (reconnaissance de la flexibilité des chaînes par analyse des déplacements chimiques) permettent d'accéder à la structure et donc aussi de repérer les domaines non structurés. La spectroscopie UV de dichroïsme circulaire permet aussi de détecter un manque de structure secondaire car les domaines désordonnés sont associés à un signal autour d'une longueur d'onde spécifique (environ 200 nm). Pour finir, la base de données *Protein disorder* (DisProt) recense l'ensemble des protéines possédant des domaines désordonnés, soit plus de 800 protéines pour environ 2150 régions intrinsèquement désordonnées.

#### Boîte de texte 1-Prédire les domaines de faible complexité

**D'après** Mier et al., « *Disentangling the Complexity of Low Complexity Proteins* », 2020 ; Vinterhalter et al., « *Bioinformatics Analysis of Correlation between Protein Function and Intrinsic Disorder* », 2021; Tompa, « *Intrinsically Unstructured Proteins* », 2002; Piovesan et al., « *DisProt 7.0* », 2017.





**Figure 6-Un ensemble de protéines de liaison à l'ARN avec des caractéristiques de type prion**  
 78928 protéines humaines ont été scannées et répertoriées sur ce graphe en fonction de leur indice de conformation (FoldIndex : algorithme déterminant les IDR, basé sur la charge moyenne des résidus et la charge nette de la séquence) et leur caractère prion (*Prion propensity*, score moyen calculé par rapport au biais de composition d'une séquence sur une série de 41 AA). Chaque point noir représente une protéine, les points rouges mettent en évidence les protéines contenant un RRM. Les protéines dont la séquence contient des régions de type PrLD sont indiquées par des triangles. Les protéines à caractère prion doivent satisfaire les deux paramètres et se retrouvent dans la zone grise en bas à droite. **D'après** King OD, et al. *The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res.* 2012;1462:61-80. doi:10.1016/j.brainres.2012.01.016 ; Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, Silman I, Sussman JL. *FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics.* 2005 Aug 15;21(16):3435-8. doi: 10.1093/bioinformatics/bti537. Epub 2005 Jun 14. PMID: 15955783.

#### ▪ Lien avec les maladies neurodégénératives

De nombreuses RBP pourvues de LCD ont été reliées à diverses maladies neurodégénératives. C'est le cas de TDP-43 et FUS pour lesquelles on retrouve des agrégats cytoplasmiques dans les neurones de patients atteints de SLA ou de DFT, mais aussi, de plusieurs autres protéines de liaison à l'ARN[84]. L'ataxine 1 comme l'ataxine 2 se retrouvent en cause dans le développement, respectivement, de l'ataxie spinocérébelleuse de type 1 et de type 2 en formant des agrégats à cause d'une expansion de répétition d'une séquence de nucléotides (CAG) dans la séquence codant le LCD[85, 86]. ESWR1 et TAF15, deux protéines qui avec FUS forment la famille FET (FUS/EWS/TAF15) se retrouvent aussi sous forme d'inclusions cytoplasmiques dans les neurones de patients atteint de DFT et de SLA[87]. Les protéines hnRNPA1 et hnRNPA2 ont Un lien avec la SLA, puisqu'une mutation a été mise en évidence pour la forme sporadique de la SLA, mais sont aussi impliquées dans le cas de la protéinopathie multi-système (MSP). Des mutations pathologiques

au sein du domaine PrLD de ces deux protéines exacerbent la capacité de ces protéines à se rassembler en granules de stress puis à évoluer vers des structures insolubles par le biais de processus d'auto-ensemencement conduisant à la formation d'inclusions cytoplasmiques[88]. Enfin, la protéine MATR3 est retrouvée dans les inclusions cytoplasmiques de neurones à la fois des formes sporadiques et familiales de la SLA[89]. De leur côté, TDP-43 et FUS sont connues pour leur lien avec la SLA, la DFT et la LATE (**voir chapitre I. 4.a**).

De nombreux exemples, qui suggèrent un lien entre les protéines de liaison à l'ARN et le développement des maladies neurodégénératives, en pointant un élément commun entre ces protéines, le LCD. Néanmoins, les mécanismes de dégénérescence étant encore mystérieux, rien ne permet de trancher si les RBP sont la cause de ces maladies ou si leur agrégation est simplement la conséquence d'un dérèglement plus large. Ici, nous allons nous pencher plus en détail sur ces domaines de faible complexité, de façon à mieux cerner leurs fonctions, notamment dans le cas de l'interaction protéine-protéine, puis de l'agrégation.

### c) Fonctions des LCD

Les protéines aux domaines non structurés ont l'avantage d'être assez plastiques, ce qui leur confère des capacités particulières. Elles sont impliquées dans des fonctions variées, elles participent notamment aux mécanismes de signalisation ou de régulation, par le biais d'interactions protéine-protéine, protéine-ARN ou protéine-ligand.

Les domaines de faibles complexités ont en premier lieu été classés parmi les domaines d'activation des facteurs de transcription. Ce sont, par exemple, les domaines riches en glutamine du facteur de transcription SP1[90] ou encore la protéine VP16[91] riches en AA acides pour lesquels le lien entre faible complexité et domaine d'activation a été souligné. Il a ensuite été mis en évidence que ces domaines étaient bien plus répandus, et qu'ils pouvaient avoir un rôle plus large dans l'interaction protéine-protéine. En particulier, la formation d'hydrogels, comme c'est le cas avec le LCD de la nucléoporeine constitué de répétition phénylalanine/glycine qui a été le premier à démontrer la capacité de ce type de domaines à mettre en place des auto-interactions[92]. Aujourd'hui, on sait que les protéines possédant des LCD sont impliquées dans une grande variété de fonctions dans les processus cellulaires, comme la régulation de la transcription, la traduction, et le cycle cellulaire, mais aussi dans les fonctions de régulation des assemblages moléculaires comme les ribosomes, l'organisation de la chromatine, les assemblages de microtubules, le transport avec les pores nucléaires ou encore le transport de petites molécules[93]. Ainsi, les protéines contenant des domaines désordonnés peuvent être classés dans

différentes catégories selon les fonctions qu'elles occupent : les chaînes entropiques, les chaperonnes, les effecteurs, les assembleurs, ou les charognards[94]. Les chaînes entropiques regroupent les fonctions qui bénéficient directement du caractère plastique et désordonnés des LCD selon le modèle des espaceurs-autocollants (**voir, boîte 2, ci-contre**). Cette alternance entre régions semi-structurées et régions flexibles confère à la protéine une flexibilité conformationnelle pour interagir avec les molécules environnantes[3]. Dans ce type de catégorie on retrouve par exemple la protéine Titine, une protéine géante et largement non structurée, responsable de la plasticité musculaire, grâce à son domaine PEVK[95]. Les chaperonnes mettent en place des contacts avec de nombreux partenaires capables de s'associer et se dissocier rapidement, c'est par exemple le cas de la famille des Hsp[96]. Le rôle des effecteurs est de modifier l'activité d'une protéine partenaire en s'y liant, par exemple la PKIs qui inhibe l'activité de certaines kinases. Les assembleurs ont des interfaces de liaison importantes de façon à servir de support pour un grand nombre de partenaires pour permettre la formation de gros complexes, c'est le cas par exemple de Ciboulot qui prend part aux complexes de polymérisation de l'actine[97]. Enfin, les charognards peuvent capturer et neutraliser de petits ligands, ce sont le cas, par exemple, des Dsp (*desiccation stress protein*) qui capturent les molécules d'eau[98, 99].

En matière de mécanismes de signalisation cellulaire, les LCD sont particulièrement intéressants puisqu'ils permettent de multiples sites d'interaction qui interviennent dans la liaison à différentes cibles, c'est pourquoi les protéines contenant des LCD fonctionnent comme des plaques tournantes centrales des

### Comment fonctionnent les LCD ?

Concernant le mode de liaison des LCD, ils ont d'abord été décrits sur la base des matériaux plastiques et polymères à l'aide d'un modèle espaceurs et auto-collants (*stickers and spacers*) pour expliquer leur implication dans le contact protéine-protéine. La chaîne latérale des acides aminés (en particulier tyrosine, phénylalanine) forme la partie « auto-collant » qui est responsable des interactions protéine-protéine sans structure apparente, tandis que la partie espaceurs est constituée d'acides aminés inertes qui font le lien entre les motifs auto-collants. Une seconde description, sur le modèle de la feuille de papier a été proposée. Dans celle-ci, les LCD peuvent s'assembler en formant des structures transitoires en feuillets  $\beta$  croisés labiles.

#### Boîte de texte 2-Modes de liaison des LCD

Kato M, Zhou X, McKnight SL. How do protein domains of low sequence complexity work? RNA. 2022 Jan;28(1):3-15. doi: 10.1261/rna.078990.121. Epub 2021 Oct 20. PMID: 34670847; PMCID: PMC8675291.

réseaux de signalisation[100]. Le LCD permet la mise en place d'assemblages dynamiques au cœur des processus de signalisation. Enfin, dans le cas plus spécifique des protéines contenant des RGG, leurs fonctions englobent la répression de la traduction, la régulation de l'apoptose, la régulation de la transcription, l'épissage des ARN pré-messager et le signalement des dommages à l'ADN[78].

Nous allons nous pencher plus en détail sur l'aspect du contact protéine-protéine qui est la fonction prédominante pour ces protéines contenant des LCD. Par ailleurs, le mécanisme récemment mis en évidence, de compartimentalisation ou séparation de phases repose sur ces interactions protéines-protéine, nous allons le détailler par la suite.

- **Interactions multivalentes protéine-protéine**

Une protéine peut établir de nombreuses interactions avec son environnement dans le cadre de ses fonctions, les domaines de faible complexité favorisent ces contacts. Il a été montré que les LCD sont plus fréquemment placés à l'extrémité des séquences (au début ou à la fin de la chaîne), ce qui favorise les contacts avec les autres protéines. Ainsi, les protéines qui ont un LCD en bout de séquence ont plus de partenaires d'interaction par rapport à celles dont le LCD est situé au centre. La longueur du LCD est aussi corrélée aux nombres d'interactions quand il se trouve aux extrémités[101]. De plus, les protéines qui ont un nombre important de partenaires d'interaction, surnommé hub, sont plus fréquemment pourvues de LCD que les autres[102]. Les protéines qui arborent des domaines LARKs qui contiennent une forte proportion de tyrosines qui peuvent interagir entre elles par le biais d'interactions de faible énergie en formant des structures particulières en réseau de feuilletts  $\beta$  croisés[79]. Par ailleurs, les domaines riches en acides aminés polaires (Q/N) ont la capacité de moduler les interactions protéine-protéine, on parle de fermeture éclair polaire (*Polar Zippers*) du fait de la capacité de formation d'un réseau de liaison hydrogène par les chaînes latérales[103]. Les contacts protéines-protéines peuvent aussi se faire par le biais d'interaction d'empilement  $\pi$ - $\pi$  ou cation- $\pi$ [104]. C'est le cas des domaines désordonnés comme le PrLD qui peuvent interagir avec eux même ou ceux d'autres protéines. Ces interactions se produisent dans de nombreux types de domaines désordonnés et peuvent même se produire avec des acides aminés à la surface de domaines structurés[105]. Le LCD de TDP-43 s'assemble avec le LCD d'une autre protéine TDP-43 par le biais d'interactions de faible énergie, ce qui permet de stabiliser les homodimères de la protéine[106]. FUS est capable de mettre en place des interactions entre son RGG et son domaine PrLD ce qui favorise les séparations de phases[107].

- **Régulation des interactions protéines-protéines**

Comme évoqué précédemment, les domaines de faible complexité sont des sites plutôt enclins aux modifications post-traductionnelles ce qui permet notamment de réguler les contacts protéine-protéine. En particulier, la méthylation des arginines, est bien connue pour réguler ces interactions[108]. Chez Sam68 par exemple, les méthylarginines situées près des séquences riches en prolines influencent négativement la liaison au domaine SH3 des partenaires[109]. Inversement, l'introduction de groupements méthyles peut favoriser les interactions en augmentant l'hydrophobicité de certains domaines, comme avec les domaines Tudor contenant des acides aminés aromatiques formant une poche qui favorise les interactions cation- $\pi$  avec le groupement électropositif des méthylarginines[110].

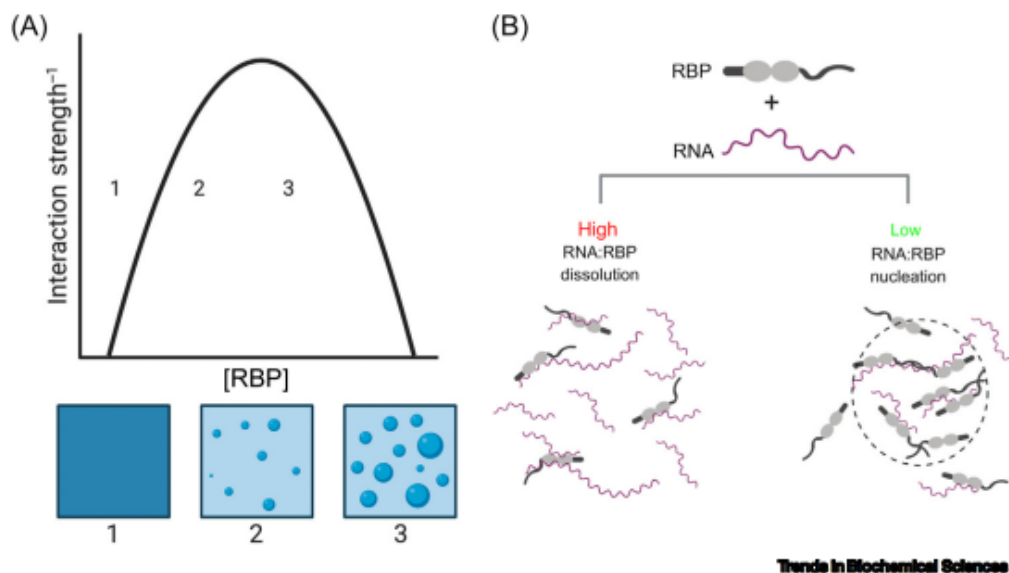
La phosphorylation consiste en l'addition d'un groupement phosphate chargé sur certains acides aminés non chargés comme les sérines, thréonines et tyrosines, par le biais de kinases, ce qui induit un changement dans les interactions d'une protéine avec ses partenaires. Les sites non structurés sont plus susceptibles d'être phosphorylés, comme l'a mis en évidence cette étude qui montre que les régions désordonnées ont une fraction plus importante de sites phosphorylables que les domaines structurés[111]. Par exemple, chez la protéine Bcl-2 ayant un rôle anti-apoptotique, la phosphorylation de deux sérines dans son domaine non structuré, par le complexe  $\nu$ -cyclin-CDK6, a pour conséquence d'annuler son potentiel d'activité[112]. La phosphorylation du domaine PrLD chez FUS régule les interactions protéine-protéine, en limitant sa capacité à former des séparations de phases, et plus largement à s'agréger[113]. Nous étudierons plus en détail les MPT relatives à TDP-43 et FUS par la suite (**voir chapitre II.3.a**).

Nous avons vu que les domaines de faibles complexités étaient très présents chez les RBP, et qu'ils leur permettaient, notamment, de mettre en place des interactions protéine-protéine. De ce fait les protéines peuvent s'assembler en larges complexes et effectuer des fonctions variées. L'augmentation de la concentration locale de protéines peut favoriser la formation d'assemblages complexes, jusqu'à la formation d'organelles sans membrane ou de condensats, opérant ainsi une séparation de phases.

### 3. Formation de condensats par le biais de séparation de phases liquide-liquide

#### a) Mécanisme général de LLPS

Entre les complexes de protéines d'un côté, et les organites cellulaires présentant une membrane de l'autre, un vaste monde a été mis au jour de manière relativement récente[114], il s'agit des condensats, ou compartiments sans membrane issus du processus de séparation de phases liquide-liquide (LLPS). Ces condensats consistent la plupart du temps en assemblages de protéines, d'acides nucléiques, et souvent de protéines ribonucléiques (RNP) qui se concentrent localement pour former des compartiments distincts du milieu cellulaire. Les LLPS sont des phénomènes spontanés et naturels qui font partie de l'homéostasie cellulaire, elles sont le plus souvent liées à la biochimie de l'ARN, c'est aussi le cas des granules de stress par exemple[115]. Le comportement de ces condensats suit un diagramme de phase[116] (**voir figure 7, A**), ce sont des gouttes concentrées de biomolécules avec les propriétés d'un liquide (comme la fusion de gouttes, l'échange des constituants avec le milieu, et le mouillage de surface). D'ailleurs, *In vitro* ces LLPS s'apparentent à des gouttelettes issues d'une séparation de phases de l'huile dans l'eau. Au sein des cellules, les condensats sont souvent un mélange de différents constituants, on parle de condensats hétérotypiques. Bien qu'ils répondent à des propriétés liquides, ces assemblages hétérogènes ne semblent pas avoir de concentration de saturation (concentration en dessous de laquelle les constituants vont se dissocier) comme on peut l'observer pour des condensats homotypiques (formés d'un unique constituant). Ces phases hétérotypiques présentent donc un profil complexe qui résulte d'interactions hétérogènes, leur comportement est donc sensible à divers facteurs, en premier lieu leur composition[116]. L'ARN est particulièrement important dans ces processus de séparation de phases, sa concentration permet leur régulation, à faible ratio, il induit une nucléation des protéines favorisant la formation de condensats, tandis qu'à forte concentration il les déstabilise (**voir figure 7, B**)[116].



**Figure 7-Les transitions de phases liquide-liquide.**

La formation de condensats par le processus de LLPS suit un diagramme de phase (A), lors d'une séparation homotypique, le dépassement d'une concentration limite (concentration de saturation) induit la formation de gouttelettes liquides : on passe d'un milieu homogène (1) à un milieu biphasique (2 et 3). Augmenter la concentration augmente la taille des gouttelettes. Les condensats générés par LLPS au sein des cellules résultent souvent d'interactions hétérotypiques, par exemple entre une ou plusieurs RBP, et l'ARN (B). Dans ce cas, la transition peut être régulée par la quantité locale d'ARN. A haut ratio RNA:RBP les condensats sont dissous. A faible ratio, les protéines se concentrent sur l'ARN et forment des assemblages très denses qui favorisent les contacts protéine-protéine, ce qui permet la nucléation des complexes sous forme d'organelles sans membrane. **D'après** Portz B, Lee BL, Shorter J. *FUS and TDP-43 Phases in Health and Disease. Trends Biochem Sci. 2021;46(7):550-563. doi:10.1016/j.tibs.2020.12.005*

Pour assurer la prise en charge de l'ARN, le phénomène de LLPS est fréquent, l'ARN et les RBD peuvent s'assembler en complexes, comme les granules de stress, les corps de traitement (*processing body*), les *paraspeckles* nucléaires ou encore les granules de transport (tRNP). La classification de ces granules dépend de leurs compositions, de leurs fonctions, de leurs dynamiques et de leurs localisations. Les condensats sont dynamiques et sont en échange constant avec le milieu environnant[117]. Le LCD est important dans le processus de LLPS pour la formation de condensats, par exemple dans le cas de formation de P-body (Les *processing body* sont des condensats formés d'enzymes impliquées dans le renouvellement de l'ARNm) chez la levure, lorsque l'on supprime les LCD des protéines Lsm4 et Dhh1 cela conduit à la suppression de ces granules. De plus, remplacer ce domaine de faible complexité par celui d'une autre protéine permet de rétablir la capacité à former des LLPS[118]. Les protéines TDP-43 et FUS de par leurs domaines de faible complexité contribuent à la formation de LLPS[119, 120], nous allons détailler leurs spécificités dans la partie suivante.

### b) Formation de condensats spécifiques à FUS et TDP-43

Nous avons vu précédemment le cas des RBP avec des domaines de faible complexité. FUS et TDP-43 sont deux protéines de liaison à l'ARN qui possèdent des domaines PrLD, ce qui leur permet de former des condensats. Les domaines PrLD étant capable de former des interactions homotypiques comme hétérotypiques peuvent conduire à la formation de condensats complexes formés de plusieurs

#### Protéines chaperonnes

Les chaperonnes sont des protéines indispensables pour la bonne structuration des protéines dans la cellule, prévenant ainsi l'agrégation. La famille des protéines de choc thermique (*Heat shock proteins*) est impliquée dans la régulation des séparations de phases des RBP. Par exemple, la protéine HspB8 est recrutée dans les séparations de phases de FUS pour prévenir le mauvais repliement de son RRM et donc la transition vers des séparations de phases aberrantes. La protéine TDP-43 quant à elle est prise en charge par Hsp70 et Hsp90, qui semblent jouer un rôle plus large dans la prévention de la neurodégénérescence.

#### Boîte de texte 3-Les protéines chaperonnes

Boczek EE, et al. "HspB8 prevents aberrant phase transitions of FUS by chaperoning its folded RNA-binding domain". *Elife*. 2021 Sep 6;10:e69377. doi: 10.7554/eLife.69377. PMID: 34487489; PMCID: PMC8510580.

Carman A, et al. Chaperone-dependent Neurodegeneration: A Molecular Perspective on Therapeutic Intervention. *J Alzheimers Dis Parkinsonism*. 2013;2013(Suppl 10):007. doi:10.4172/2161-0460.S10-007

protéines. Les tyrosines présentes dans les domaines de type prion de FUS forment des interactions  $\pi$ - $\pi$  ou  $\pi$ -cations avec les arginines présentes dans les RRG[107]. De plus, les liaisons de faible énergie, comme les liaisons hydrogène, ou des interactions hydrophobes participent aussi aux LLPS. Les gouttelettes de FUS formées *In vitro* respectent les critères du comportement liquide, elles ont une forme ronde, elles sont constituées de molécules qui peuvent se déplacer librement et peuvent se réarranger, et enfin, elles peuvent fusionner[121].

La protéine TDP-43 met en place de multiples contacts intermoléculaires par le biais de régions distinctes, notamment grâce à ses domaines PrLD et Nter qui lui permettent de former des condensats. Le Nter permet la dimérisation de la protéine, de plus les contacts Nter-Nter permettent une configuration tête-bêche (*head-to-tail*) des protéines au sein d'oligomères, cette configuration entraîne aussi le contact avec le PrLD, ce qui favorise la nucléation des protéines autour de l'ARN, phase importante dans la formation de condensats[116]. En particulier, l'hélice  $\alpha$  contenu dans le domaine PrLD semble jouer un rôle important dans l'interaction protéine-protéine, et donc plus largement dans les séparations de phases. La conformation en hélice est favorisée dans le contexte



des interactions TDP-43/TDP-43. Et la présence de mutations à ce niveau associées à la SLA, perturbe les séparations de phases et contribue à la transition de l'état liquide à l'état solide[106]. De la même manière l'oxydation de la méthionine située dans ce segment contribue à la transition de phases vers l'état de gel par le biais d'une polymérisation en  $\beta$  croisés[122]. Parmi les exemples de condensats dans lesquels TDP-43 est impliqué, on trouve divers granules, comme les *paraspeckles*[123], les granules de stress[124], les nucléoles[125], les *Cajal bodies*[126], les granules de transport (tRNP), ou les myo-granules. Les tRNP sont des ribonucleoprotéines impliquées dans le transport et le stockage de l'ARNm qui contiennent aussi des miARN. FUS et TDP-43 participent à ces complexes, ce qui les amène dans les neurones à se localiser au niveau des synapses et des dendrites[117]. TDP-43 peut former des myo-granules, un type de condensat qui intervient dans le développement et la régénération des tissus musculaires. *In vitro*, ces structures évoluent vers un processus de nucléation formant des fibrilles[127]. Les *paraspeckles* sont des organelles sans membrane nucléaires caractérisés par la présence du long ARN non-codant NEAT1, TDP-43 comme FUS sont nécessaires à la formation de ces condensats, mais FUS est localisée dans le cœur tandis que TDP-43 se trouve au niveau externe[128]. Enfin, la formation de condensats semble être primordiale pour le bon fonctionnement de la protéine TDP-43, ainsi son rôle dans l'épissage semble être amélioré par le processus de LLPS[129], tandis que chez la souris, une protéine présentant une déficience dans la capacité à former des condensats implique une défaillance des fonctions neuronales et de la régulation de la transcription[130].

FUS a la capacité de mettre en place des contacts inter-domaines de manière intra- et intermoléculaire grâce à ses RGG et son domaine PrLD, ce qui lui offre de nombreuses possibilités pour contrôler et réguler le processus de séparation de phases. De nombreuses protéines FUS peuvent se rassembler au sein d'un même ARN ce qui permet de favoriser les contacts PrLD-PrLD. De plus, le PrLD peut interagir avec le RBD d'autres molécules FUS non liées à l'ARN, ce qui stabilise la présence d'une forte concentration de FUS au même endroit[116]. FUS est impliquée dans de nombreux condensats, tout comme TDP-43 à l'exception des *Cajal Bodies*. Un exemple de condensat spécifique à FUS, est celui mis en place par l'interaction du RGG avec PAR (**voir chapitre II.3.a**) qui favorise la nucléation de FUS au niveau des sites endommagés de l'ADN[131]. Le récepteur d'importation nucléaire Karyopherin- $\beta$ 2 (Kap $\beta$ 2) auquel FUS se lie (au niveau du PY-NLS) joue un rôle de chaperonne, et est capable de prévenir et de renverser la formation de condensats de FUS et notamment de réguler l'interaction protéine-protéine[116].

Nous venons de voir le principe de formation des condensats, tout en montrant que FUS et TDP-43 y participaient activement. C'est un processus important de l'homéostasie cellulaire mais dans certaines conditions, les condensats peuvent évoluer vers un stade de liquide-gel, puis vers la formation d'agrégats insolubles, voir favoriser la croissance de structures fibrillaires. On peut donc passer d'un mécanisme normal et sain à un dysfonctionnement pouvant conduire au développement de pathologies. Ce mécanisme est suspecté de jouer un rôle dans la dégénérescence, c'est ce que nous allons voir par la suite, mais avant cela nous allons décrire les granules de stress, susceptibles d'être une étape intermédiaire entre ces deux états.

### c) Granules de stress

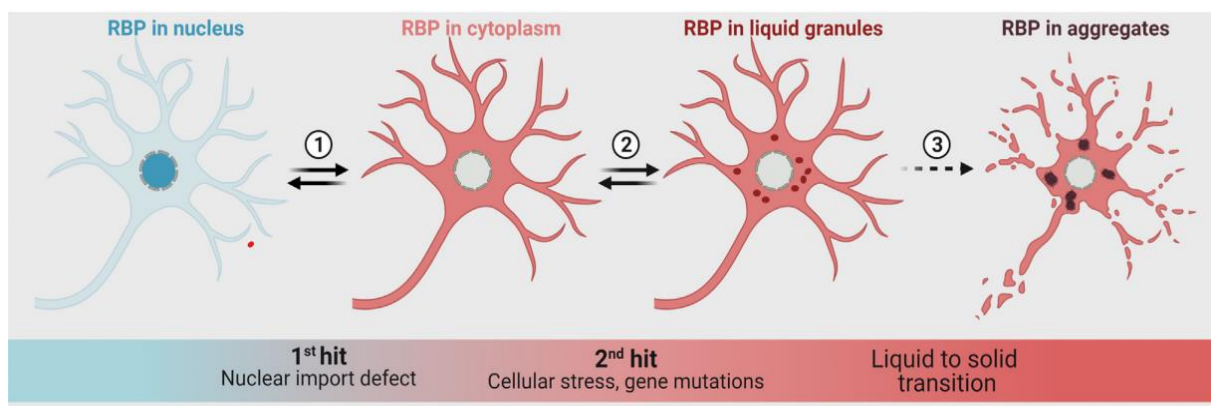
Les granules de stress (GS) sont des complexes rassemblant les RBP et l'ARNm non traduit, formés dans le cytoplasme en réponse à un stress (agents toxiques, infection virale, température, oxydation...). Un jeu complexe d'interactions entre les RBP et l'ARN par des contacts protéine-protéine, protéine-ARN et ARN-ARN conduit à la formation de ce processus transitoire et réversible. La concentration d'ARN est forte dans ces condensats, et la présence de RBP peut s'expliquer par le fait qu'elles stabilisent le complexe à la fois grâce à leurs interactions avec l'ARN mais aussi par les contacts protéine-protéine que leurs LCD permettent. Par exemple, la protéine TIA1 est capable de déclencher la formation de GS lorsqu'elle est surexprimée, tandis qu'un variant de cette protéine dépourvue de son domaine désordonné n'en est pas capable[132]. De même, la protéine G3BP, une RBP qui contient aussi un LCD, est impliquée dans la formation de ces granules, en particulier par le biais de ce domaine qui lui permet de former des dimères[133]. Fait intéressant, Il a été montré que la plupart des RBP impliquées dans les maladies neurodégénératives citées plus tôt (hnRNP, EWSR1, TAF15, ATXN2) jouent un rôle dans la formation des GS[134]. Enfin, il a été montré que TDP-43 participe à la formation de granules de stress par le biais de son domaine de faible complexité[135], de plus les mutations pathologiques observées dans le domaine PrLD affectent la dynamique des granules de stress comme cela a pu être montré sur des lignées cellulaires (NSC-34)[136]. De plus, TDP-43 semble jouer un rôle dans leur régulation[137]. De son côté, FUS aussi peut être localisé dans les granules de stress, et il apparaît que les mutations de la protéine renforcent cette capacité[12].

Les granules de stress ont une fonction protectrice pour la cellule. En cas de stress, elles stockent l'ARNm, la traduction des ARNm est donc stoppée ce qui permet de conserver de l'énergie et de la réserver au profit d'autres processus essentiels à la survie[12]. La formation de ces granules est un processus réversible, ainsi les granules se dissolvent quand le stress s'apaise, et les RBP associées

retournent dans le noyau. Néanmoins, ces dernières sont suspectées d'être un intermédiaire pour la formation d'agrégats en concentrant des protéines sensibles à l'agrégation. De cette manière, la présence de marqueurs des granules de stress (TIA-1, Eif3) dans les inclusions de la protéine TDP-43 chez les patients SLA et DFT-U, suggère un lien entre l'accumulation d'agrégats de RBP, les GS et les MND[138]. Ainsi, dans certaines conditions, ce processus est suspecté de favoriser l'agrégation des protéines[11].

#### d) De la séparation de phases à l'agrégation

Comment peut-on passer d'un mécanisme physiologique essentiel au bon fonctionnement des cellules à un mécanisme d'agrégation toxique en lien avec le développement de pathologies sévères ? Bien que les mécanismes soient encore mal compris, les LCD sont particulièrement ciblés pour expliquer ce phénomène d'agrégation. En effet, si les LCD confèrent une grande plasticité aux RBP, ce qui leur permet de s'adapter à leurs partenaires d'interactions comme à leur environnement pour assurer des fonctions variées, cela leur permet aussi de s'auto-assembler et peut mener à l'agrégation[2]. FUS et TDP-43 ont, par ailleurs, une forte tendance à l'agrégation y compris suivant une faible augmentation de leur niveau d'expression endogène[7, 8].



**Figure 8-La transition de phases peut conduire à l'agrégation des protéines.**

Ce modèle tente d'expliquer comment les RBP se retrouvent agrégées dans les maladies neurodégénératives comme la SLA en proposant deux facteurs qui peuvent s'influencer. Le premier est la délocalisation des protéines nucléaires vers le cytoplasme, cela peut être dû à un problème d'importation nucléaire, par exemple à cause d'une mutation ou d'une défaillance de la machinerie de transport. Le deuxième est la formation de condensats par LLPS, en particulier les granules de stress qui peuvent avec le temps, un défaut des chaperonnes, ou une altération de la dynamique des granules, conduire à l'agrégation. Les protéines opèrent une transition de la phase liquide à solide provoquant une agrégation pathologique. **D'après** Sternburg et al. *Post-translational modifications on RNA-binding proteins: accelerators, brakes, or passengers in neurodegeneration?* Trends Biochem Sci. 2022 Jan;47(1):6-22. doi: 10.1016/j.tibs.2021.07.004. Epub 2021 Aug 5. PMID: 34366183.

La formation d'hydrogels réversibles a été observée pour FUS, par le biais de ses domaines de faible complexité[139]. De plus, ces hydrogels avaient une morphologie de type amyloïde au microscope électronique[121]. Plus intéressant encore, ces structures sont capables d'entraîner d'autres RBP (piégeage hétérotypique)[139]. De manière similaire, la maturation de gouttelettes constituées du domaine PrLD de TDP-43 évolue en structure de type gel ou solide assez semblables à celles observées dans les tissus de patients. La maturation de ces fibres est favorisée par les LLPS[140].

Le maintien et la régulation de TDP-43 en conditions physiologiques dépend à la fois de son autorégulation, sa répartition nucléocytoplasmique et ses capacités à former des condensats, ces trois éléments sont liés et s'influencent, une dysfonction de l'un d'eux peut conduire à l'apparition de comportements pathologiques[141] qui peut permettre d'expliquer la dégénérescence des cellules. L'ARN joue un rôle central dans ces processus et c'est ce que nous allons voir maintenant.

#### ▪ **Rôle de l'ARN dans la nucléation**

Si jusqu'à présent nous avons surtout évoqué le rôle des contacts protéine-protéine, avec notamment le domaine PrLD, l'ARN en lui-même, joue un rôle dans les processus de séparation de phases. Il a un effet ambivalent sur les RBP, à haute concentration (comme au sein du noyau), il permet la solubilisation de ces protéines qui peuvent avoir tendance à l'agrégation (effet de *buffering*), d'un autre côté en faible concentration (comme dans le cytoplasme), il sert d'agent de nucléation pour la formation de condensats mais peut aussi favoriser la maturation d'agrégats. Ici, nous allons donc nous pencher plus particulièrement sur le rôle de l'ARN dans la régulation des LLPS et dans les mécanismes qui peuvent conduire à l'agrégation.

Dans un premier temps, l'ARN régule le comportement des séparations de phases, et prévient la transition de phases de liquide à solide en gardant solubles les RBP, comme cela a été montré avec les RBP possédant des LCD : EWSR1, TAF1, hnRNPA1, FUS et TDP-43[142]. Plus largement, une étude de protéomique à large échelle, a mis en évidence que le phénomène de solubilisation des protéines n'était pas restreint à un petit ensemble de RBP lié au MND comme FUS ou TDP-43 mais qu'il concernait plus largement 1300 protéines, qui devenaient insolubles à la suite de la dégradation de l'ARN dans la cellule[143]. L'ARN permet la formation et la régulation de la taille des LLPS c'est ce qui a été montré directement au sein des cellules avec un modèle de granule ArtiG[144]. Dans le cas de FUS, l'ARN amplifie la séparation de phases en fonction du ratio ARN-protéine[145]. De plus, l'introduction d'un excès d'ARN dans la cellule peut conduire à la formation de multiples compartiments sans membrane[146]. Les propriétés de l'ARN, comme sa taille ou sa structure secondaire, semblent jouer un rôle important dans la formation de

condensats : les ARN plus longs ont une plus grande capacité à conduire la séparation de phases, tandis que les plus petits les inhibent[146]. De manière générale, les ARN longs, flexibles et non structurés, qui contiennent différents motifs de liaison, forment une base pour favoriser les séparations de phases.

D'un autre côté, l'ARN forme un noyau qui permet à FUS de venir s'agréger et favorise la formation de structures fibrillaires, ce qui a été montré *in vitro* à l'aide de concentration substoechiométrique d'ARN prD[53]. La formation d'agrégats semble être dépendante de la quantité d'ARN, ainsi à haute concentration la formation d'agrégats est inhibée. L'ajout d'oligonucléotides spécifiques à TDP-43, permet par ailleurs de dissoudre les transitions de phases aberrantes[9]. De la même façon, les condensats de TDP-43 sont régulés par leur interaction à l'ARN, ou encore stabilisés par des interactions avec d'autres protéines comme la chaperonne Hsp70. Ces complexes peuvent cependant évoluer d'un état liquide à solide quand cette interaction est perturbée[147].

Pour conclure, nous avons vu précédemment que le contact protéine-protéine était important pour répondre à un grand nombre de fonctions. Mais que dans certains cas, il pouvait aussi conduire à des mécanismes pathologiques comme avec la formation d'agrégats fibrillaires. Néanmoins, il apparaît aussi que les interactions ARN-ARN jouent un rôle dans la formation de complexes. L'ARN peut servir de support pour favoriser les interactions protéine-protéine d'un côté, mais d'un autre côté, il peut aussi former des interactions directes par le biais d'appariement de bases de type Watson-Crick, d'appariement de bases non canonique ou d'empilement d'hélices, ce qui favorise la croissance de gros complexes[148]. Ainsi, les contacts ARN-ARN hétérotypiques pourraient jouer un rôle déterminant dans la formation et la régulation des LLPS[146].

### Fibres amyloïdes

Les fibres de type amyloïdes sont des structures fréquemment rencontrées dans le cas de diverses maladies, en particulier les maladies neurodégénératives. Elles résultent de l'assemblage de protéines mal conformées en feuillet  $\beta$  croisés.

Elles sont mises en évidence par les réactifs chimiques spécifiques tels que le *Congo Red* ou la thioflavine S qui s'intercale entre les fibres. Les fibres amyloïdes sont insolubles et toxiques. Elles sont polymorphes, leur structure dépend des conditions d'agrégations.

### Boîte de texte 4-Fibres amyloïdes

**D'après** Naeimi WR, Serio TR. *Beyond Amyloid Fibers: Accumulation, Biological Relevance, and Regulation of Higher-Order Prion Architectures. Viruses.* 2022 Jul 27;14(8):1635. doi: 10.3390/v14081635. PMID: 35893700; PMCID: PMC9332770.

### ▪ La formation de fibres : caractéristique des MND

Les inclusions que l'on rencontre typiquement dans les neurones de patients sont sous forme de fibres de type amyloïde, ce sont des structures formées d'un assemblage de protéines en feuillets  $\beta$  croisés, comme celles qui ont été très bien décrites dans le cas de la maladie d'Alzheimer. Ces structures peuvent être mises en évidence par le biais de marqueurs spécifiques comme le *Congo Red*, et la thioflavine S ou T[149]. Le LCD de TDP-43 a une forte propension à l'agrégation et lui permet notamment de former des structures de type fibre amyloïde[150]. De même pour FUS, une région de son domaine PrLD peut former des structures de type feuillets  $\beta$  croisés qui sont impliquées dans la fibrillation[151]. Bien que ces deux protéines soient capables de former des structures de type amyloïdes la nature des agrégats qu'elles forment dans le cas des MND reste discutée[152–155].

#### Propagation des fragments

La fragmentation est une étape déterminante dans la propagation de maladies avec des agrégats de type Prion. Le système de nettoyage des cellules et l'autophagie pourraient être à la base de la fragmentation des agrégats, qui continueraient de se propager et ainsi de répandre la maladie. En effet, les fibres pourraient être décomposées en fragments, qui joueraient le rôle de noyaux favorisant la nucléation de nouvelles fibres par recrutement et conversion de protéines saines.

#### Boîte de texte 5-Propagation des fragments

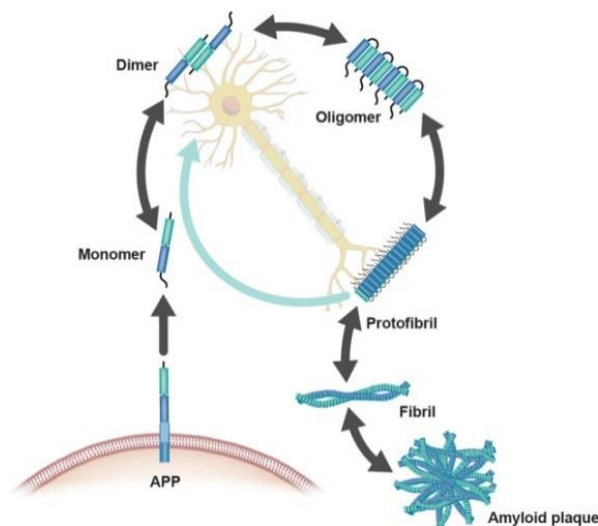
D'après Marrero-Winkens C, Sankaran C, Schätzl HM. From Seeds to Fibrils and Back: Fragmentation as an Overlooked Step in the Propagation of Prions and Prion-Like Proteins. *Biomolecules*. 2020;10(9):1305. Published 2020 Sep 10. doi:10.3390/biom10091305

Une hypothèse permet d'expliquer comment des protéines fonctionnelles peuvent être piégées dans de mauvaises conformations, on parle d'entonnoir de repliement (*protéine-folding funnel*). Une protéine se structure dans un état natif stabilisé par des interactions intramoléculaires qui correspondent au niveau d'énergie minimum (minimum global), cependant dans certain cas une protéine peut se retrouver piégée dans une conformation qui correspond à un mauvais repliement (*protein misfolding*), de plus haute énergie correspondant à un minimum local. Ce mauvais repliement conduit à des contacts intramoléculaires non natifs, et davantage de contacts intermoléculaires, dont la formation d'assemblages intermoléculaires qui peuvent prendre des profils différents : oligomères solubles, agrégats amorphes ou structures amyloïdes riches en feuillets  $\beta$ . Le plus bas niveau d'énergie, et donc le plus stable, correspond à la fibre amyloïde[156].

### ▪ Comment se forment les fibres ?

Une meilleure connaissance sur la formation de fibres à partir des peptides amyloïdes (A $\beta$ 40 ou A $\beta$ 42) dans le cas de la maladie d'Alzheimer nous éclaire sur le

processus général de formation de fibres amyloïdes qui peuvent se rencontrer dans d'autres maladies neurodégénératives. La protéine APP est une protéine transmembranaire, elle peut se retrouver clivée par protéolyse sous forme de fragments de 40 ou 42 résidus (A $\beta$ 40 ou A $\beta$ 42). Ces fragments initialement sous forme de monomères solubles, vont s'assembler en dimères qui vont ensuite évoluer en oligomères. Ce processus résulte d'une première étape de nucléation au cours de laquelle les fragments s'agrègent à partir d'un noyau conduisant à des agrégats désordonnés. A ce stade le processus est encore réversible et l'assemblage est soluble, mais ces premiers agrégats n'en sont pas moins toxiques. Une seconde étape d'élongation va conduire au grossissement des noyaux précédemment formés, à la formation d'assemblages fibrillaires, tout en catalysant la formation de nouveaux noyaux (nucléation secondaire) qui participent à la prolifération des agrégats par le biais d'une boucle de rétroaction positive. Au cours de cette étape, les objets vont évoluer en proto-fibrilles qui peuvent s'associer de manière hélicoïdale et parallèle, puis au cours d'une dernière étape de maturation, les objets vont se stabiliser sous forme de fibres à la structure caractéristique, où les structures en feuillets  $\beta$  s'associent aux brins  $\beta$  dans une orientation perpendiculaire à l'axe de la fibre, et stabilisés par des liaisons hydrogène. Ce processus s'amplifie par accumulation des fibres, jusqu'à la formation de plaques (**voir figure 9**)[157].



**Figure 9-Etapes de formation de fibres : exemple de la maladie d'Alzheimer.**

La formation de fibres amyloïdes à partir de fragments de la protéine APP dans le cas de la maladie d'Alzheimer est bien documenté, ce modèle peut servir de base pour mieux comprendre les processus de nucléation comme ça peut être le cas pour d'autres protéines comme TDP-43 ou FUS. Dans le cas de l'APP, qui est une protéine transmembranaire, elle est clivée et va se retrouver sous forme de peptides (A $\beta$ 40 ou A $\beta$ 42) qui vont s'associer entre eux par contact protéine-protéine pour former un dimère. Ensuite, le dimère évolue en oligomère, dont l'association est favorisée par la mise en place de contacts de feuillets  $\beta$  croisés, ce dernier va grossir pour prendre la forme de proto-fibrille, puis de fibres où les structures en feuillets  $\beta$  s'associent aux brins  $\beta$  dans une orientation perpendiculaire à l'axe de la fibre et stabilisées par des liaisons hydrogène. Ce processus s'amplifie par accumulation des fibres, jusqu'à la formation de plaques amyloïdes **D'après** Hampel H, Hardy J, Blennow K, et al. *The Amyloid- $\beta$  Pathway in Alzheimer's Disease. Mol Psychiatry. 2021;26(10):5481-5503. doi:10.1038/s41380-021-01249-0*

- **Mécanismes de neurodégénérescence : la Propagation**

La nucléation (ou ensemencement, *seeding*) est la capacité d'un agrégat de protéines à servir de support pour la croissance de fibres de type amyloïdes homotypiques. La dissémination des agrégats dans les tissus est mal connue mais peut résulter d'un mécanisme de propagation à la manière de la protéine Prion chez la levure. Quelques travaux ont étudié ce mécanisme de propagation chez des modèles animaux en mettant en avant le côté toxique des fibres de type amyloïdes. Par exemple, la formation de plaques amyloïdes peut être induite par transfert de fibres (issues de cerveaux malades) chez des souris transgéniques sur-exprimant la protéine précurseur de l'amyloïde (APP)[158]. Chez l'homme, un mécanisme similaire a été mis en évidence dans le cas de la maladie de Parkinson. Dans l'espoir d'améliorer leurs symptômes des patients ont reçu des transplantations de cellules dopaminergiques, et à leur mort, 14 ans après, des agrégats d' $\alpha$ -synucléine ont été retrouvés dans ces cellules, ce qui suggère la propagation des agrégats de cellules en cellules[159]. Un phénomène de propagation aussi observé lors des cultures cellulaires[160].

La propagation des agrégats d'une cellule à l'autre est un processus qui permet d'expliquer la diffusion de la maladie. L'agrégat est relâché par la cellule, puis va se répandre de cellule en cellule, aggravant le processus de dégénérescence. Le relargage peut se faire de différentes manières : par le biais d'une rupture de membrane, les agrégats sont simplement relâchés à travers une membrane fragilisée, ce qui permet un transfert passif. Les agrégats peuvent être relâchés par le biais de vésicules comme lors de l'exocytose, où du relargage d'exosomes. Et enfin, les agrégats peuvent être transférés par contact direct par le biais de nanotubes. La capture de ces agrégats par d'autres cellules peut aussi passer par différents phénomènes comme l'endocytose (L'endocytose par le biais des HSPG, *heparan sulfate proteoglycans*, la fusion d'exome, l'endocytose par le biais de récepteur ou encore la phagocytose) qui consiste en l'absorption de l'agrégat par la cellule au niveau de la membrane par le biais de vésicules, ou de récepteurs membranaires[161].

Quelques travaux ont étudié la propagation de la protéine TDP-43. L'introduction de la protéine TDP-43 issue d'agrégats insolubles (provenant de cerveaux malades) dans des cellules (SH-SY5Y) qui expriment la protéine TDP-43 sauvage ou avec une étiquette HA conduit à la formation d'agrégats cytoplasmiques phosphorylés et ubiquitinylés[162]. Un modèle de souris a aussi été développé spécifiquement pour étudier la propagation de TDP-43[163]. Aussi, des agrégats pathologiques provenant de patients DFT peuvent se répandre de cellule en cellule



de manière dépendante au temps à travers différentes régions du système nerveux chez des souris transgéniques qui sur-expriment la protéine TDP-43 sans séquence de translocation nucléaire[164]. De plus, une étude récente a montré par cryo-EM que les filaments de TDP-43, extraits de neurones de patients SLA-DFT, partagent la même morphologie en double spirale, qu'ils viennent du cortex frontal ou moteur, ce qui est compatible avec une dispersion spatiale des structures de type amyloïde[165]. Par conséquent, ces études suggèrent que TDP-43 est capable de suivre le mode de propagation prion.

- **Gain ou perte de fonction**

Le lien entre les RBP et plus largement le métabolisme de l'ARN avec le développement de maladies neurodégénératives est de plus en plus évident. Bien que le mécanisme ne soit pas encore clairement établi, une question fait l'objet d'un grand débat, la dégénérescence peut-elle s'expliquer par une perte ou un gain de fonction des RBP ?

La protéine TDP-43 peut être clivée, phosphorylée, agrégée et délocalisée dans le cytoplasme (**voir chapitre II.4.b**) ce qui peut induire des conséquences sur le fonctionnement des cellules. D'un côté, le gain de fonction résulte de la présence de la protéine (ou de fragments) dans le cytoplasme qui conserverait sa capacité de liaison à l'ARN, ce qui aurait pour conséquence de le piéger ou d'occuper des sites de liaison et donc d'empêcher d'autres protéines de le prendre en charge. D'un autre côté, la délocalisation de la protéine, et donc son absence dans le noyau, peut conduire à de nombreux dysfonctionnements, puisque les fonctions dont la protéine est en charge ne sont plus effectuées. Etant donné l'ensemble des fonctions qu'elle occupe (**voir chapitre II.3**) cela suffirait à induire la neurodégénérescence. Si l'on considère le cas de l'autorégulation par exemple, si ce processus s'opère dans le noyau et que la protéine est dans le cytoplasme, une perte de l'autorégulation conduirait à un cercle vicieux avec une production non contrôlée de la protéine et une formation d'agrégats accrue conduisant à la mort cellulaire. Au contraire, dans le cas où la régulation se fait dans le cytoplasme et que TDP-43 conserve sa liaison à l'ARN, l'autorégulation est amplifiée, il y a donc de moins en moins de protéines dans le noyau, il en résulte alors un nettoyage du noyau (*nuclear clearance phenomom*), ce qui pourrait conduire à un stress cellulaire et finalement à la mort des cellules[166]. Finalement, la question reste débattue, et il est tout à fait possible que la dégénérescence résulte d'une combinaison de ces deux mécanismes, ils ne s'excluent pas mais peuvent fonctionner ensemble pour affecter la viabilité des cellules. Cela met en avant la complexité de la compréhension des mécanismes impliqués dans les MND.

Maintenant que nous avons vu que les RBP pouvaient former des condensats, et que ces derniers sont suspectés d'évoluer vers l'agrégation de ces protéines, dans la prochaine partie nous allons prendre de la hauteur en regardant les conséquences des processus de neurodégénérescence, en développant les maladies en lien avec ces protéines de liaison à l'ARN.

## 4. Les maladies neurodégénératives en lien avec TDP-43 et FUS

### a) Trois maladies qui s'entrecroisent : La SLA, la DFT et la LATE

Les maladies neurodégénératives (MND) représentent un véritable enjeu majeur en matière de santé public. En effet, avec le vieillissement de la population, de plus en plus de cas de ces maladies, qui affectent le système nerveux, sont déclarées. Pour la maladie d'Alzheimer, qui est la MND la plus répandue, on compte par exemple plus d'un million de personnes atteintes en France[1] et on prévoit que ce chiffre va quasiment doubler d'ici 2030 avec le vieillissement de la population. Les MND représentent la 5<sup>ème</sup> cause de mortalité dans le monde après notamment les arrêts cardiaques, et les cancers. Les MND se caractérisent par une mort prématurée des neurones induisant des troubles très variés comme la perte de mémoire, la perte de contrôle musculaire ou des troubles cognitifs et comportementaux. Dans la plupart des cas, la maladie se déclenche après 60 ans et la prévalence augmente avec l'âge. Les différentes MND sont caractérisées, d'une part, par la nature des neurones affectés, et d'autre part, par les dysfonctionnements qui en résultent. Aujourd'hui, un grand effort de recherche au niveau mondial est engagé pour mieux comprendre ces maladies, les mécanismes sous-jacents mais aussi trouver de futurs traitements. Comme nous l'avons vu précédemment les

#### Facteurs de risques

Dans une petite commune de Savoie, une prévalence importante de cas de SLA a été détectée (65 fois plus important que le taux d'incidence en France). Après une longue enquête, ces cas ont été reliés à la consommation d'un champignon, pourtant interdit, mais resté dans la tradition locale : la gyromitre (fausse morille). Ce champignon contient une molécule neurotoxique d'hydrazine, qui bien que le mécanisme soit mal connu, semble augmenter le risque de neurodégénérescence. Des cas similaires d'intoxications alimentaires responsables de la SLA ont aussi été identifiés à l'île de Guam et l'étang de Thau.

A côté de ces cas atypiques liés à une intoxication alimentaire, la cigarette, la pratique sportive exposant à des chocs à la tête, des infections virales, l'exposition aux champs magnétiques, mais aussi les métaux lourds et pesticides, font partie des nombreuses causes environnementales exposées, avec des corrélations souvent faibles.

#### Boite de texte 6-Facteurs de risque pour la SLA

*D'après Lagrange et al. An amyotrophic lateral sclerosis hot spot in the French Alps associated with genotoxic fungi. J Neurol Sci. 2021 Aug 15;427:117558. doi: 10.1016/j.jns.2021.117558. Epub 2021 Jun 25. PMID: 34216974. Et Ingre et al. Risk factors for amyotrophic lateral sclerosis. Clin Epidemiol. 2015;7:181-193. Published 2015 Feb 12. doi:10.2147/CLEP.S37505*

mécanismes en cause dans ses maladies sont encore largement mal compris mais la présence de protéines agrégées est un marqueur commun. Ici nous traiterons en particulier de trois maladies interconnectées, la sclérose latérale amyotrophique (SLA), la dégénérescence fronto-temporale lobaire (DFT), et l'encéphalopathie TDP-43 liée à l'âge à prédominance limbique (LATE) qui ont en commun les mêmes biomarqueurs TDP-43 et FUS.

- **La sclérose latérale amyotrophique (SLA)**

La SLA est une MND caractérisée par une dégénérescence des neurones moteurs. Le terme de sclérose latérale amyotrophique réfère au durcissement des

**Traitements existants et perspectives**

A ce jour, il existe très peu de traitements efficaces pour ces maladies. Seuls le Riluzole et l'Edaravone, permettent de prolonger la durée de vie des patients atteints de la SLA. Cependant, de nombreux candidats médicaments sont testés chaque année mais ne sont pas toujours acceptés faute d'efficacité prouvée. Toutefois, diverses pistes thérapeutiques très prometteuses sont envisagées. Les molécules antioxydantes pour lutter contre le stress oxydatif. La thérapie antisens, par le biais d'oligonucléotides ou de vecteurs viraux, pour réduire l'expression et donc la toxicité de protéines mutées. Enfin, la thérapie cellulaire, à base de cellules souches pluripotentes induites, pour restaurer les fonctions perdues.

tissus et à la fonte musculaire provoqués par la perte de signaux des motoneurones. Elle affecte à la fois les neurones moteurs centraux (cerveau) et périphériques (tronc et moelle épinière), en induisant une perte progressive de la mobilité, une paralysie des membres qui conduit à la mort par arrêt respiratoire[167]. Elle a été décrite pour la première fois par le professeur Jean-Martin Charcot de la pitié Salpêtrière en 1874, c'est pourquoi elle est aussi appelée maladie de Charcot. Aux Etats-Unis, elle porte, en revanche, le nom de la maladie de Lou Gehrig en référence au joueur de baseball atteint de cette pathologie. Le diagnostic est posé entre 40 et 80 ans selon la forme de la maladie. En général, la paralysie totale, puis la mort surviennent 3 à 5 ans après le diagnostic. En France, on compte environ 2300 nouveaux cas par an[168], soit une prévalence de 3 cas pour 100 000 habitants en Europe.

La SLA est une maladie complexe dont les formes peuvent être très variées selon les neurones qui sont affectés. De manière générale, ce sont les neurones responsables de la communication entre

**Boite de texte 7-Traitements existants et perspectives pour la SLA**

*D'après Inserm (en ligne) :*

<https://www.inserm.fr/dossier/sclerose-laterale-amyotrophique-sla-maladie-charcot/>

les muscles et le système nerveux central qui sont touchés, ils peuvent se trouver dans le cortex moteur, le tronc cérébral ou la moelle épinière. Selon les cas, la maladie peut débuter dans une zone localisée où certains neurones vont commencer à dégénérer, avant de s'étendre à d'autres zones plus larges, avec en parallèle une aggravation des symptômes, on parle alors d'une maladie progressive. La SLA est une maladie très hétérogène et les différents variants associés sont classés selon la distribution anatomique de la pathologie. La forme bulbaire qui affecte en particulier la région du visage, de la bouche et de la gorge, mais peut aussi s'étendre à d'autres régions, comprend 33 % des cas, tandis que la forme spinale compte pour 66 % des cas et comprend tous les syndromes associés à la paralysie des membres. Différents phénotypes coexistent dérivant de ces deux catégories principales. De plus, on distingue les formes familiales (fSLA), qui représentent 5 à 10 % des cas, et sont attribuées à des facteurs héréditaires, et les formes sporadiques (sSLA), qui regroupent l'ensemble des cas d'origines inconnues[169].

Différents gènes ont été identifiés et mis en cause pour expliquer le développement de la maladie. Par exemple, 40 % des cas familiaux s'expliquent par la présence d'une répétition étendue d'une séquence d'héxanucléotides sur le chromosome 9 (*C9ORF72*, *Chromosome 9 open reading frame 72*)[170]. Il s'agit d'une répétition de la séquence non-codante de 6 nucléotides GGGGCC (plus de 30 fois, contrairement à moins de 24 dans les cas non pathologiques) sur la séquence du cadre de lecture du chromosome 9. Les répétitions comme celles-ci sont connues pour favoriser l'accumulation de transcript, et la formation d'agrégats piègeant des protéines de liaison à l'ARN (foci d'ARN)[171]. Les mutations du gène codant la protéine superoxyde dismutase 1 (SOD1) ont constitué le premier marqueur mis en évidence[172]. SOD1 est une enzyme qui protège les cellules du stress oxydatif, en participant aux mécanismes d'élimination des radicaux libres. Plus d'une centaine de mutations provoquant la maladie sont associées à ce gène. Ensuite, les deux protéines de liaison à l'ARN, TDP-43 et FUS, ont été découvertes dans les agrégats neuronaux, elles comptent, l'une comme l'autre pour 4 % des formes familiales[173]. En plus, de ces éléments majeurs d'autres gènes, comptant pour moins de 1 % des cas chacun, peuvent expliquer la maladie.

Les causes de la forme sporadique (soit 90 % des cas) sont plus difficiles à expliquer et peuvent s'expliquer par une convergence de facteurs. Néanmoins, la protéine TDP-43 est considérée comme une caractéristique principale des formes sporadiques[174]. En effet, on la retrouve sous forme d'agrégats de protéines mal-conformées dans les tissus de patients. L'altération du système immunitaire est aussi mis en avant, à la fois dans le cas des formes familiales, comme dans les formes

sporadiques, puisque des cellules immunitaires réactives telles que les cellules microgliales sont détectées dans les tissus post-mortem. Ainsi la perturbation du système immunitaire périphérique et la neuro-inflammation peuvent jouer un rôle dans le développement de la maladie[175].

- **La dégénérescence fronto-temporale (DFT)**

La dégénérescence fronto-temporale (DFT) regroupe un ensemble de troubles neurocognitifs qui affectent la partie antérieure des lobes frontaux et/ou temporaux du cerveau. La mort des neurones dans ces régions cérébrales induit des déficits cognitifs variés et des troubles du comportement, ces zones cérébrales étant impliquées chez l'Homme dans la régulation du comportement social, de la prise d'initiative et du traitement des émotions, mais aussi le centre du langage[176]. Dans environ 60 % des cas, les patients ayant une atteinte DFT présentent des troubles comportementaux[177]. Ces patients peuvent présenter des comportements inadaptes, une désinhibition, de l'apathie, des troubles obsessionnels, mais la plupart du temps la mémoire est conservée. D'un autre côté, les patients présentant des troubles du langage, aussi appelés aphasie primaire progressive, représente 40 % des cas[178]. La forme familiale de cette maladie représente environ 40 % des cas. Cette démence est la plus répandue après Alzheimer, elle concerne environ 15 à 22 cas pour 100 000 habitants[179] et touche autant les hommes que les femmes.

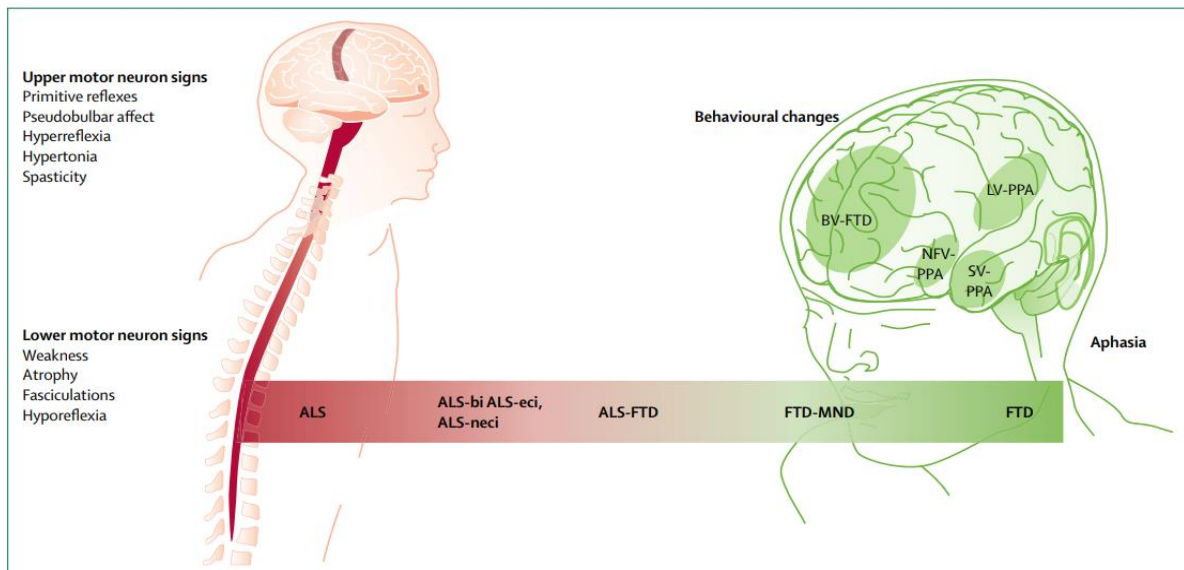
- **L'encephalopathie TDP-43 à prédominance limbique liée à l'âge (LATE)**

La LATE, pour encephalopathie TDP-43 à prédominance limbique liée à l'âge, est une MND qui a été décrite très récemment (en 2019)[180], qui regroupe l'ensemble des pathologies relatives à la protéine TDP-43. C'est un trouble neurologique qui s'observe généralement sur des sujets assez âgés, souvent d'un âge supérieur à 80 ans, et qui provoque des symptômes proches de la maladie d'Alzheimer, comme des pertes de mémoire. Cette pathologie s'étend progressivement sur de larges zones du cerveau où des agrégats cytoplasmiques de TDP-43 se déposent affectant les neurones. Elle se développe en premier lieu dans l'amygdale, elle s'étend ensuite à l'hippocampe, et enfin au gyrus frontal. La LATE englobe les protéinopathies TDP-43 qui sont associées à des troubles cognitifs, et se réfèrent à une perte de localisation de la protéine TDP-43 nucléaire avec la formation de corps d'inclusions dans le cytoplasme des neurones. C'est donc la caractéristique majeure de la LATE qui se vérifie souvent par marquage immunologique après autopsie. Elle est aussi fréquemment associée à une sclérose de l'hippocampe. On estime que la prévalence de la maladie serait 100 fois plus importante que la DFT. En effet, 20 % des cerveaux autopsiés après 80 ans présentent des caractéristiques de la

LATE[180]. De plus, il apparaît qu'environ 15 à 20 % des cas diagnostiqués pour la maladie d'Alzheimer seraient attribuables à la LATE.

### b) Lien entre ces maladies

Ces trois MND présentent des caractéristiques communes et s'entrecroisent. D'abord, en termes de caractéristiques cliniques, elles apparaissent aujourd'hui davantage comme un continuum que comme des maladies bien distinctes. D'un côté, environ 50 % des patients SLA présentent exclusivement des symptômes moteurs, de l'autre 70 % des patients DFT présentent exclusivement des symptômes exécutifs. Entre les deux, un ensemble de cas présentent à la fois des symptômes moteurs, cognitifs et comportementaux à différents degrés (**voir figure 10**)[181]. De même, de nombreux cas de DFT sont liés à la présence d'inclusions cytoplasmiques de TDP-43 et peuvent donc être apparentés à la LATE[180]. Inversement des cas de LATE peuvent présenter des symptômes très proches de la DFT.

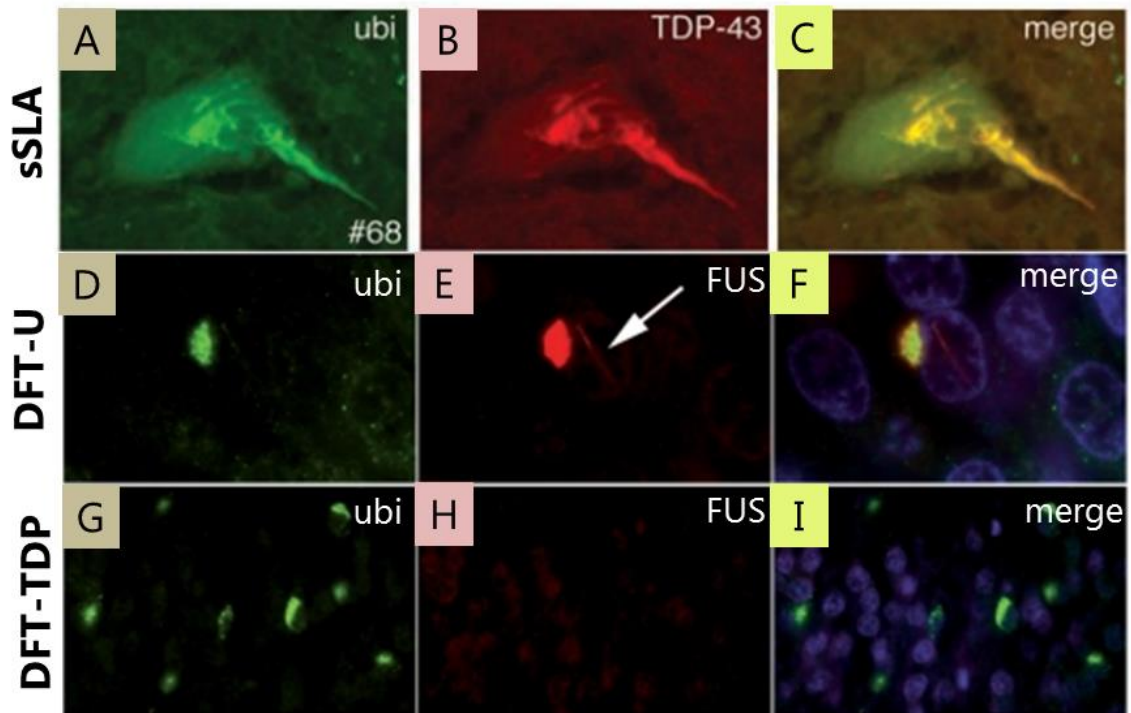


**Figure 10-Le spectre des MND est continu.**

La SLA et la DFT peuvent être considérées comme les extrêmes d'un même spectre. La SLA se caractérise par un dysfonctionnement des neurones moteurs hauts et bas. Environ, la moitié des patients atteints de SLA ont uniquement des symptômes moteurs, avec des fonctions cognitives préservées, les autres présentent aussi des atteintes cognitives et comportementales plus ou moins sévères. On parle de *ALS-bi*, lorsque des troubles du comportement sont présents, de *ALS-eci*, lorsqu'il y a des troubles exécutifs et *ALS-neci* lorsqu'il n'y a pas de problème exécutif mais d'autres dysfonctionnements cognitifs. Ensuite 5 à 10 % des cas présentent à la fois des symptômes moteurs et cognitifs, c'est la catégorie *ALS-FTD*. A l'autre bout du spectre la DFT présente exclusivement des déficits cognitifs, des troubles du comportement ou de l'aphasie. *ALS*=*amyotrophic lateral sclerosis*. *FTD*=*frontotemporal dementia*. *ALS-eci*=*ALS with evidence of executive dysfunction*. *ALS-neci*=*ALS with no executive dysfunction but impairment in other cognitive domains*. *ALS-bi*=*ALS with behavioural changes*. *BV-FTD*=*behavioural variant of FTD*. *PPA*=*primary progressive aphasia*. *NFV*=*non-fluent variant*. *SV*=*semantic variant*. *LV*=*logopenic variant* **D'après** van Es MA et al. *Amyotrophic lateral sclerosis*. *Lancet*. 2017 Nov 4;390(10107):2084-2098. doi: 10.1016/S0140-6736(17)31287-4. Epub 2017 May 25. PMID: 28552366.

D'un point de vue génétique, de nombreux facteurs sont impliqués dans ces maladies. Les gènes concernant ces maladies peuvent se répartir en 3 catégories, ceux qui provoquent de manière générale un dysfonctionnement des protéines, ceux qui concernent la dynamique du cytosquelette, et enfin ceux qui sont en lien avec le métabolisme de l'ARN. Dans la première catégorie, on trouve des gènes parmi ceux déjà évoqués pour la SLA, comme le *C9ORF72* qui se retrouve aussi en cause dans la DFT, mais aussi *VCP*, *OPTN*, *SQSTM1* et bien d'autres[173]. Dans la seconde catégorie, il y a notamment *DCTN1*, *PFN1* et *TUBA4A* qui provoquent une altération du réseau de microtubules et donc influent sur la morphologie des neurones (avec des altérations diverses dans leur fonctionnement). Finalement, parmi les gènes en lien avec l'ARN, on trouve notamment deux gènes codants des RBP, *TARDBP* (TDP-43) et *FUS*[182]. Les deux protéines sont très fréquemment retrouvées sous forme d'inclusions cytoplasmiques dans les neurones[183, 184]. TDP-43 est le constituant majeur de ces agrégats, pour les trois MND évoquées. Deux tiers des cas de DFT sont associés à la présence d'agrégats ubiquitinylés (DFT-U), et 1/3 par la présence de la protéine Tau dans les inclusions. Parmi les différents syndromes de la DFT, 50 % des cas sont caractérisés par la présence de TDP-43[185]. La présence de FUS dans les agrégats cytoplasmiques est souvent détectée par immuno-réaction dans la plupart des cas de DFT-U, même si le gène en question ne présente pas forcément de mutation[183] (**voir figure 11**). Plus précisément, les corps d'inclusion de FUS sont détectés dans environ 9 % des cas de DFT (tout confondu). La présence de FUS peut-être détectée indépendamment de celle de TDP-43 dans une grande partie des cas[186]. Les mutations de FUS sont responsables de 4 % à 5 % des formes fSLA[185] mais moins de 1 % des cas de forme sSLA[187]. Pour TDP-43, 3 à 4 % des cas de fSLA sont concernés par la mutation de ce gène, et environ 1 % dans la forme sSLA[182]. En revanche, 95 % des patients, toutes formes confondues de SLA, présentent des agrégats cytoplasmiques de TDP-43[184].





**Figure 11- Les agrégats de TDP-43 et FUS sont une caractéristique commune de la SLA et la DFT.**

(A-C) inclusions positives à l'ubiquitine ( $\alpha$ -ubiquitin, vert) (A) et TDP-43 ( $\alpha$ -TDP-43, rouge) (B) dans un cas de sSLA. Inclusions positives à l'ubiquitine ( $\alpha$ -ubiquitin, vert) (D) et à FUS ( $\alpha$ -FUS, rouge) (E) dans le cas de la DFT-U. TDP-43 et FUS ne co-localisent pas forcément, exemple d'un cas de DFT-TDP où les inclusions positives à l'ubiquitine ( $\alpha$ -ubiquitin, vert) (G) sont négatives pour FUS ( $\alpha$ -FUS, rouge) (H). La fusion des images (merge) est présentée (C, F, I), avec le marquage des noyaux (Dapi, bleu). **D'après** Neumann M, et Al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. 2006 Oct 6;314(5796):130-3. doi: 10.1126/science.1134108. PMID: 17023659. et Neumann M, et al, A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain*. 2009;132(Pt 11):2922-2931. doi:10.1093/brain/awp214

Etant donné l'importance que revêtent ces deux RBP, il semble primordial d'étudier leur fonctionnement et leurs interactions, de manière à mieux comprendre les mécanismes qui peuvent conduire à leur perte de fonction. Nous développerons, ainsi dans le prochain chapitre, les différentes caractéristiques de ces deux protéines de façon à dresser leurs points communs, puis nous montrerons qu'elles interagissent.

## II. TDP-43 et FUS, deux protéines semblables qui interagissent

### 1. La protéine TDP-43

*Transactive response DNA-binding protein 43 kDa* (TDP-43) est une protéine de liaison à l'ARN, issue du gène *TARDBP*, hautement conservée chez les eucaryotes. C'est une protéine exprimée de manière ubiquitaire, essentiellement nucléaire mais qui a la capacité de faire la navette entre le noyau et le cytoplasme. Elle assure d'ailleurs des fonctions très variées, qui vont de la prise en charge de l'ARNm (transcription, transport, stabilisation) à la prise en charge des microARN (miARN) et des longs ARN non codants (lncARN)[188]. TDP-43, avant d'être associée à toutes ces fonctions, a été identifiée comme un facteur de transcription associé au gène VIH-1[189] puis à un régulateur de l'épissage des pré-ARNm[190]. La protéine TDP-43 est essentielle pour le développement du système neuronal[191], de plus un niveau anormal de cette protéine dans les cellules peut induire la neurodégénérescence, mais aussi avoir des conséquences létales, comme l'ont montrées des études sur des modèles de rongeurs[192].

Finalement, comme évoqué précédemment, elle est liée à différentes maladies neurodégénératives, notamment on la retrouve sous forme d'agrégats insolubles dans les neurones de patients atteints de dégénérescences très variées. L'ensemble des pathologies dans lesquelles TDP-43 est impliquée (SLA, DFT, LATE, certains cas de la maladie d'Alzheimer ou de la maladie de Parkinson...) sont regroupés sous le terme de protéinopathies TDP-43. D'une part, le lien a été établi que les mutations dans le gène *TARDBP* peuvent causer la SLA[193] ou la DFT[194]. D'autre part, il a été mis en évidence que les RBP sont un constituant majeur des agrégats positifs à l'ubiquitine dans les neurones des malades[184].

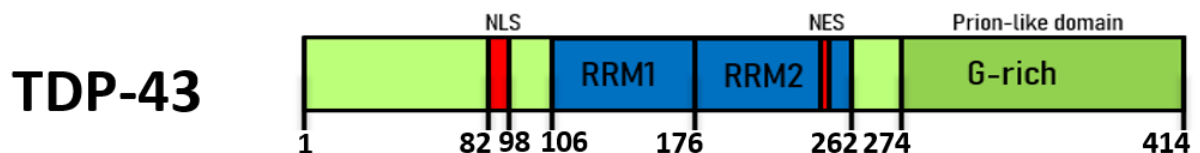
Dans cette partie nous allons étudier les caractéristiques de cette protéine, ces différents domaines, ainsi que les fonctions qu'elle exerce en condition physiologique.

#### a) Caractéristiques, domaines, et structure

La protéine TDP-43 est constituée de 414 acides aminés, elle est divisée en 4 domaines distincts (**voir figure 12**) :

- Un domaine Nter (N-terminal) principalement impliqué dans la dimérisation de la protéine (des AA 1 à 106).

- Deux domaines de liaison à l'ARN successifs, RRM1 et RRM2 (des AA 106 à 176 puis des AA 192 à 262) séparés par un segment de 14 résidus.
- Un domaine Cter (C-terminal) aussi appelé *Prion like-domain* (PrLD) ou domaine de faible complexité (*Low-Complexity Domain*, LCD) car il est riche en résidus glycines (des AA 274 à 414).
- Deux séquences distinctes le NLS (*nuclear localisation signal*) et le NES (*nuclear export signal*) lui permettent de faire la navette entre le noyau et le cytoplasme[156].



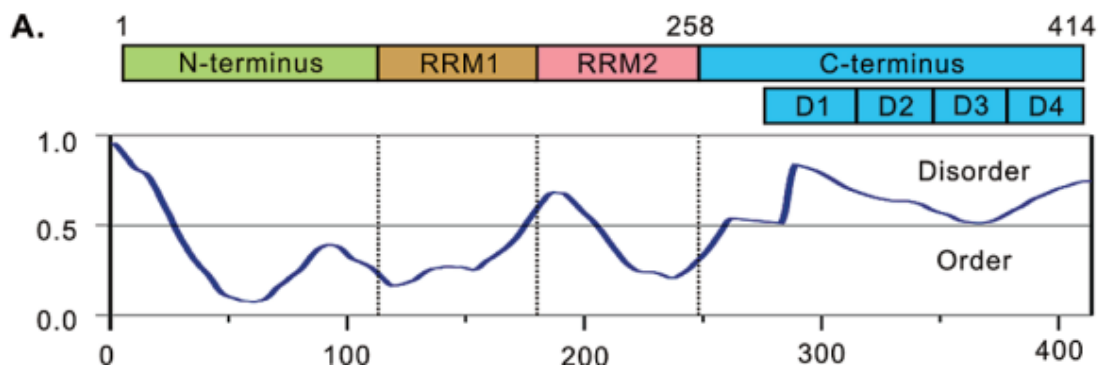
**Figure 12- Domaines de la protéine TDP-43.**

La protéine TDP-43 est pourvue d'un domaine structuré en position N-terminale, de deux domaines RRM (domaine de reconnaissance à l'ARN, *RNA recognition motif*) au centre, et d'un domaine de type prion (riche en glycine) en position C-terminale. Les domaines indiqués en rouge (NES, *nuclear exportation signal* et NLS, *nuclear exportation signal*) permettent à la protéine de se localiser et de faire la navette entre le noyau et le cytoplasme. Les chiffres indiquent la position des acides aminés aux intersections des différents domaines.

#### ▪ Le domaine de faible complexité ou PrLD

La protéine TDP-43 est pourvue d'un domaine de faible complexité en position Cter (Des AA 276 à 414), ce domaine de type prion est principalement constitué d'acides aminés polaires non chargés (glycines, asparagines et glutamines), ce qui en fait un domaine très désordonné (**voir chapitre 1,2,b**). Selon les modélisations informatiques PONDR (*Predictor of Natural Disordered Region*) 80 % du LCD de TDP-43 est désordonné[149] (**voir figure 13**). Cependant, une séquence riche en résidus aliphatiques comprise entre les acides aminés 321 et 330 abrite une hélice  $\alpha$ [106, 195]. Il a été montré que cette structure permet la dimérisation de la protéine par le biais de contacts intermoléculaires hélice-hélice, stabilisés par le biais d'interactions faibles. De ce fait, cette structure semble avoir un rôle important dans le processus de compartimentalisation par LLPS, puisqu'une mutation (A321G), et plus largement les mutations pathologiques à cet endroit, affectent la structure en hélice et réduisent la capacité à former des condensats[106, 196]. A l'inverse, la mutation G335D augmente la capacité de formation de l'hélice et dans le même temps la formation de condensats[106, 129]. De manière plus générale, le domaine PrLD est impliqué dans l'interaction protéine-protéine[101]. Ainsi, il a naturellement une forte tendance à l'agrégation et notamment, lorsqu'il est isolé, il est capable de former des fibres amyloïdes et des inclusions pathologiques[150]. Ainsi, il a été

montré que la maturation de condensats de TDP-43 pouvait conduire à l'apparition de fibrilles riches en brin  $\beta$  stabilisées par le segment d'acides aminés de 365-400[197].



**Figure 13-Le domaine PrLD de TDP-43 est principalement non structuré.**

L'analyse PONDR (VL3) sur la séquence de la protéine TDP-43 permet de mettre en évidence l'aspect désordonné du domaine de type prion C-terminal. **D'après** Chen AK, et al. *Induction of amyloid fibrils by the C-terminal fragments of TDP-43 in amyotrophic lateral sclerosis.* *J Am Chem Soc.* 2010 Feb 3;132(4):1186-7. doi: 10.1021/ja9066207. PMID: 20055380.

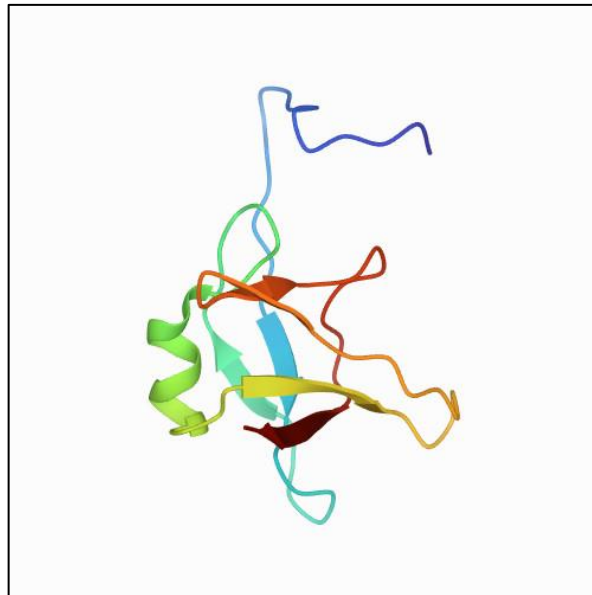
#### ▪ Le domaine Nter

Le domaine N-terminal n'entre pas dans une catégorie spécifique, c'est pourquoi nous garderons l'appellation générale Nter, mais il n'en reste pas moins important pour la fonction de la protéine. C'est un fragment de 10kDa qui regroupe les premiers acides aminés jusqu'au 100ème qui le connecte au RRM1. Une première étude proposant une simulation du Nter a d'abord proposé que c'est un domaine assez flexible, selon cette approche, il contient plusieurs motifs en feuillet  $\beta$  au niveau des 60 premiers AA, une hélice au niveau de résidus 25 à 32 permet la connexion de feuillet  $\beta$  croisés, puis il est connecté au RRM par une hélice au niveau de AA 89 à 100[198]. Plus tard, sa structure a été obtenue par RMN, ce qui a permis de confirmer la présence d'une hélice  $\alpha$  et de 6 petits brins  $\beta$ , qui s'organisent en 2 grands feuillets  $\beta$  (**voir figure 14**). Ce qui fait de ce segment un domaine assez rigide finalement. Il contient une certaine proportion d'acides aminés aliphatiques formant un cœur hydrophobe[199].

Il est notamment impliqué dans la dimérisation de la protéine. La délétion de ce segment (les 75 premiers AA) affecte la capacité de la protéine à s'autoassembler mais réduit aussi son activité dans l'épissage. En particulier, les 10 premiers AA semblent les plus critiques dans ces processus. Cela a été vérifié à la fois expérimentalement et par des modèles informatiques[198]. La Lysine 28 semble essentielle pour la bonne conformation du domaine, il a été montré qu'une mutation ponctuelle de ce résidu pouvait déstabiliser sa structure, altérer la fonction de la

protéine dans l'épissage, et induire une délocalisation de la protéine vers le cytoplasme[200].

Le domaine Nter contient aussi la séquence spécifique du NLS, au niveau des AA 78 à 100, responsable de translation de la protéine vers le noyau par le biais de l'importine  $\alpha$ . C'est une séquence importante pour la bonne localisation de TDP-43 comme le suggère les mutations qui peuvent avoir lieu à ce niveau, qui forcent TDP-43 à rester dans le cytoplasme et conduisent à l'agrégation de la protéine[201].



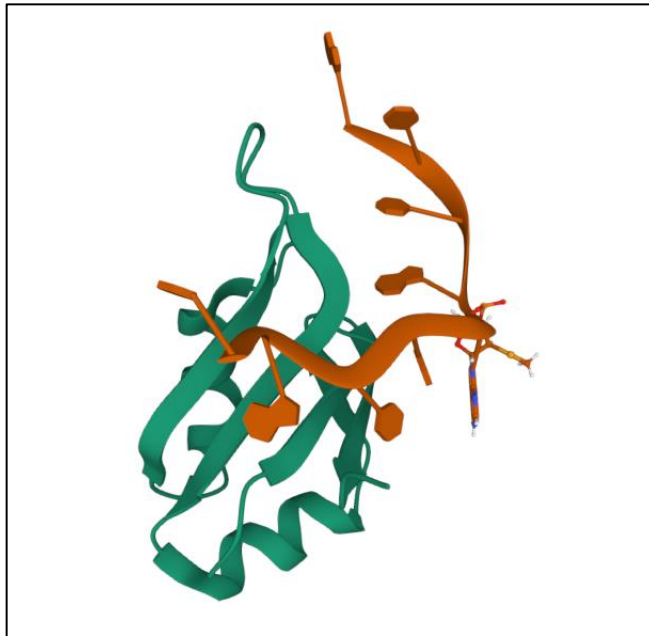
**Figure 14-Le domaine Nter de TDP-43 est structuré.**

Structure tridimensionnelle du domaine N-terminal de TDP-43 (protéine humaine exprimée chez *Escherichia coli*) obtenue par RMN. Ce domaine est structuré en 6 brins  $\beta$  et une hélice  $\alpha$ . **Source :** RCSB PDB modèle 2N4P, disponible ici <https://www.rcsb.org/structure/2N4P> **D'après** Mompeán M, et al. *The TDP-43 N-terminal domain structure at high resolution. FEBS J.* 2016 Apr;283(7):1242-60. doi: 10.1111/febs.13651. Epub 2016 Feb 11. PMID: 26756435.

#### ▪ Les domaines RRM

La protéine TDP-43 possède deux domaines RRM qui se succèdent en milieu de séquence, ce qui représente un fragment d'environ 20 kDa. Ce sont 2 RRM canoniques typiques de la famille des hnRNP (**voir chapitre I,1,a**), ils contiennent chacun environ 90 AA et suivent la topologie  $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 3'-\beta 3''-\beta 4$  avec deux brins particuliers et propres à TDP-43, les  $\beta 3'-\beta 3''$ , formant une épingle ( $\beta$  *Haipin*) qui sont essentiels pour la stabilité de la structure et la reconnaissance de l'ARN tout en étendant la surface de contact possible avec l'ARN[202]. La liaison des 2 RRM se fait par le biais d'un lien flexible constitué de 14 acides aminés (178 à 191) qui offre à chaque RRM des propriétés indépendantes, mais permet aussi une conformation étendue en établissant un pont avec 4 brins  $\beta$  (au lieu de 2 habituellement)[202]. Le RRM1 comme le RRM2 ont une spécificité de reconnaissance des motifs TG/GU chez

l'ADN/l'ARN, cependant le RRM1 a une affinité plus importante que le RRM2. Il semble alors que le RRM1 occupe le premier rôle dans la liaison à l'ARN, tandis que le RRM2 occupe une fonction de régulateur de structure de la spécificité de séquence dans la liaison de la protéine[203]. Partant de ces caractéristiques, nous allons étudier plus en détail dans le prochain chapitre les modes de liaison de TDP-43 à l'ARN et ses cibles spécifiques.



**Figure 15-Structure du domaine de liaison à l'ARN de TDP-43.**

Structure cristalline obtenue par diffraction des rayons X du RRM1 (vert) de la protéine TDP-43 humaine en complexe avec l'ARN (Orange). L'interaction se fait par contact direct entre la chaîne latérale des feuillets  $\beta$  avec les bases de l'ARN. **Source** RCSB PDB, modèle 4IUF, disponible ici : <https://www.rcsb.org/3d-view/4IUF/1> **D'après** Pan-Hsien Kuo, et al. *The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids*, *Nucleic Acids Research*, Volume 42, Issue 7, 1 April 2014, Pages 4712–4722, <https://doi.org/10.1093/nar/gkt1407>

A côté de sa capacité de liaison à l'ARN, le RRM est capable de mettre en place des auto-interactions et peut donc participer à l'agrégation de la protéine, il a d'ailleurs été montré qu'il pouvait former des structures fibrillaires *In vitro*[204].

### **b) Mode d'interaction et ARN pris en charge par TDP-43**

La protéine TDP-43 se lie à plus de 6000 ARNm ce qui représente environ 30 % du transcriptome[205]. Grâce à ses deux RRM, elle se lie avec une grande affinité (de l'ordre du nanomolaire) en particulier aux séquences répétées GU[206]. Des acquisitions en RMN ont, en effet, mis en évidence que le tandem RRM12 se lie à de multiples répétitions GU avec coopérativité[202]. Plus précisément, des interactions intermoléculaires entre une boucle contenue dans le RRM1 et une poche du RRM2 de différentes molécules se mettent en place pour permettre aux protéines TDP-43 de se

répartir coopérativement sur les ARN riches en motif GU[56]. D'un autre côté, le RRM1 et le RRM2 de TDP-43 peuvent se lier indépendamment l'un de l'autre au motif GU de l'ARN mais aussi aux répétitions TG de l'ADN. Les deux RRM sont nécessaires pour la liaison à haute affinité à l'ARN, le RRM2 a une affinité plus importante pour les petites séquences GU ( $K_d = 379$  nM), tandis que le RRM1 permet une plus grande affinité pour les séquences plus longues ( $K_d = 65,2$  nM)[202]. Des structures en cristallographie ont permis de montrer l'interaction du feuillet  $\beta$  du RRM canonique en interaction avec l'ARN (**voir figure 15**). La spécificité de l'interaction se fait grâce à un appariement des bases de type Watson-Crick par le biais d'un réseau de liaisons hydrogène[207]. De plus, la liaison à de longues séquences est facilitée par la dimérisation de la protéine à l'aide de son domaine Nter. La cartographie des ARN à l'échelle du génome a confirmé la préférence de TDP-43 pour les domaines GU dans une large gamme de transcrit lui conférant ainsi un rôle important dans la régulation de l'épissage et de l'expression des gènes[203]. Par exemple, cela a été démontré par des expériences de iCLIP (*individual-nucleotide resolution CLIP*) qui ont permis de mettre en évidence que TDP-43 ciblait préférentiellement les séquences d'intron (74 % des cibles de TDP-43 sont des séquences d'intron dans les modèles cellulaires), ce qui met en évidence son implication dans l'épissage alternatif[70]. La protéine est aussi recrutée pour la prise en charge des ARNm impliqués dans le développement neuronal (Fermeture du tube neural)[70]. Pour finir, TDP-43 se lie à son propre ARNm au niveau de la séquence 3' UTR pour réguler son niveau d'expression, sous forme de boucle de rétroaction négative[63].

## 2. La protéine FUS

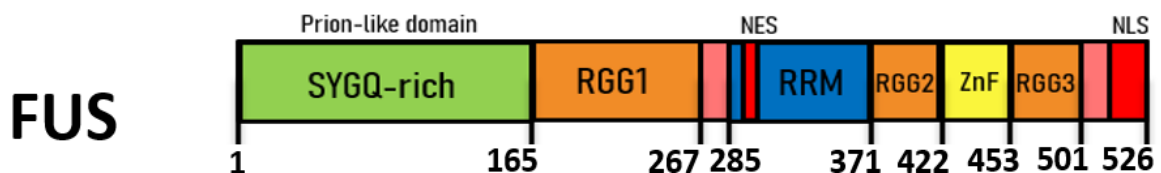
Tout comme TDP-43, FUS est une protéine de liaison à l'ARN impliquée dans certaines maladies neurodégénératives. La présence d'agrégats cytoplasmiques de cette protéine chez les patients a permis d'établir un lien avec la SLA en 2009[208]. FUS a été identifiée en premier lieu en tant qu'oncogène dans le liposarcoma, (*Fused in liposarcoma*). On retrouve aussi l'appellation FUS/TLS pour *Translocated in Liposarcoma*. C'est une protéine qui appartient à la famille FET (qui inclut aussi deux autres RBP, EWS et TAF15) et elle est aussi classée dans la catégorie des hnRNP sous le nom hnRNP P2. Tout comme TDP-43, FUS est une protéine principalement nucléaire, qui prend en charge une grande quantité d'ARN, et est impliquée dans son métabolisme à travers diverses fonctions. Dans cette partie, nous allons décrire plus en détail ses fonctions, sa structure, et son mode d'interaction avec l'ARN. FUS est très similaire à TDP-43 en matière de structure, de fonctions, et de localisation, elles partagent d'ailleurs certains ARN cibles.

### a) Caractéristiques, domaines et structure

FUS est une protéine d'environ 55 kDa constituée de 526 acides aminés rassemblés en plusieurs domaines caractéristiques (**voir figure 16**) :

- Un domaine de faible complexité de type PrLD en position Nter (AA de 1-165)
- Un RRM canonique au centre (AA 285 à 371)
- Trois domaines RGG (RRG1 de AA 176 à 267, RGG2 de AA 371 à 422, RGG3 de AA 453 à 501).
- Un domaine à doigt de zinc, ZnF entre le RGG2 et RGG3 (AA 422 à 453)
- Un domaine NES situé au niveau du RRM, et un domaine NLS à l'extrémité C-terminale

Le RRM, les RGG, ainsi que le ZnF, lui permettent d'interagir à la fois avec l'ARN et d'autres protéines. De plus, le comportement de FUS semble être régi par une compétition entre des interactions intra- et intermoléculaires régulées par des modifications post-traductionnelles[116]. Ce que nous verrons par la suite (**voir chapitre II.4.a**).



**Figure 16-Séquence de la protéine FUS.**

La protéine FUS est pourvue d'un domaine de type prion (riche en acides aminés SYGQ) en position N-terminale, de 3 domaines RGG (riches en arginines et glycines), d'un domaine unique RRM (domaine de reconnaissance à l'ARN) central, et d'un doigt de zinc (ZnF). Les domaines indiqués en rouge (NES, *nuclear exportation signal* et NLS, *nuclear exportation signal*) permettent à la protéine de se localiser. Les chiffres indiquent la position des acides aminés aux intersections des différents domaines.

#### ▪ Le domaine prion

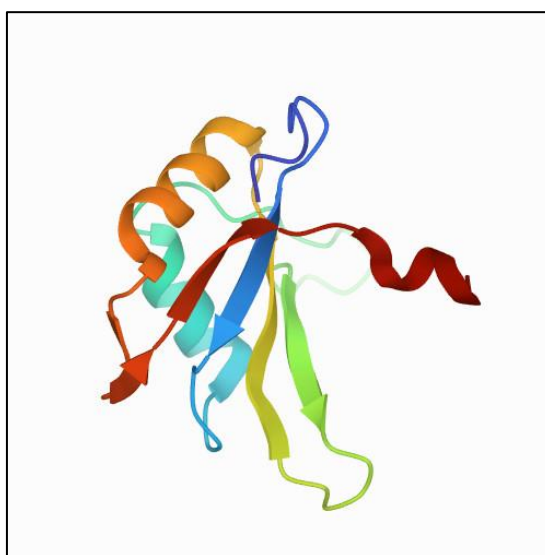
Placé du côté N-terminal de la protéine, le domaine PrLD de FUS est riche en glutamine, glycine, serine et tyrosine, il englobe les AA 1 à 165. Il permet des interactions hétérotypiques et homotypiques et donc l'oligomérisation de la protéine, comme la participation à de plus gros complexes. Plus précisément, il contient plusieurs segments LARKs (**voir chapitre I.2.b**) qui peuvent former des interactions homotypiques de type  $\beta$  croisés[116]. Plusieurs études ont démontré la capacité du domaine PrLD de FUS à former des structures fibrillaires[151, 209, 210]. Le cœur fibrillaire est situé principalement entre les AA 39 et 95 de la protéine[151]. Pour finir,



le domaine PrLD de FUS peut former des interactions intramoléculaires avec les RGG[211].

#### ▪ Le RRM

A la différence de TDP-43, FUS contient un unique RRM (AA de 285 à 371, **voir figure 16**) généralement connu pour reconnaître l'ARN, mais aussi l'ADN et les protéines[212]. C'est un RRM canonique similaire à celui d'autres RBP (de la forme  $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$  **voir chapitre I.1.a, voir figure 17**) dans sa forme mais dont la séquence en acides aminés diffère. Sa structure primaire contient 2 résidus lysines conservés, inclus dans une boucle qui connecte  $\alpha 1$  et  $\beta 2$ , cette boucle est nommée *KK-loop*. Cette dernière est impliquée dans l'interaction avec l'ARN et l'ADN double brin de manière dépendante de la charge et non par empilement de bases (*ring stacking*) avec les acides aminés aromatiques et les acides nucléiques comme cela s'observe de manière générale avec les RRM[213]. Les acides aminés phénylalanines réparties dans le RRM (F305L, F341L, F359L, et F368L) n'en restent pas moins importantes pour l'interaction, puisque leurs mutations abolissent la liaison à l'ARN[212]. De plus, la structure du RRM lui permet de se lier à une tige-boucle (*stem-loop*) de l'ARN en se liant du côté 3' de la boucle avec peu de spécificité de séquence, le feuillet  $\beta$  et l'hélice Cter rentrent en contact avec la tige de l'ARN par insertion d'une épingle  $\beta$  au sein du sillon majeur de l'ARN[211]. Par conséquent, le mode de liaison de FUS à l'ARN est non-spécifique. Enfin, tout comme le RRM de TDP-43, il a été montré que le RRM de FUS peut mettre en place des auto-assemblages et former des structures fibrillaires *In vitro*[204, 214].



**Figure 17-Structure du domaine de liaison à l'ARN de FUS.**

Structure tridimensionnelle obtenue par RMN du RRM de la protéine FUS humaine. RCSB PDB, modèle 6GBM disponible ici :<https://www.rcsb.org/structure/6GBM> **D'après** Loughlin FE, et al. *The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity.* *Mol Cell.* 2019 Feb 7;73(3):490-504.e6. doi: 10.1016/j.molcel.2018.11.012. Epub 2018 Dec 20. PMID: 30581145.

### ▪ Les RGG et Znf

A côté du RRM, les RGG et le Znf sont deux séquences prépondérantes pour la liaison à l'ARN. D'autant que l'interaction de FUS avec l'ARN semble être gouvernée davantage par les RGG que le RRM[215, 216].

Comme les autres protéines de la famille FET, FUS possède 3 RGG qui séparent les parties structurées de la protéine. Le RGG (**voir chapitre I.2.b**) est un domaine au rôle ambivalent qui permet à la fois la liaison à l'ARN mais aussi l'interaction protéine-protéine. Il reconnaît, par exemple, les structures en G-quadruplex (une organisation spatiale non canonique de l'ADN ou l'ARN, favorisé par les cations divalents[217]) de l'ARN[218]. Le domaine RGG contribue grandement à l'affinité de la protéine pour l'ARN, il a la capacité de remodeler la structure de l'ARN[211]. D'un autre côté, les RGG contribuent aux interactions protéine-protéine et plus largement aux mécanismes de LLPS. On peut définir les RGG comme des espaceurs autocollants (*Sticker and spacer*)[3], qui servent donc de lien entre les domaines avec la possibilité de mettre en place des contacts variés. On sait par exemple, que les domaines PrLD et RGG3 (Cter) de FUS interagissent par le biais de résidus arginine et tyrosine pour conduire la formation de condensats[107]. Le RGG est aussi important pour la localisation de FUS, notamment dans les granules de stress[78].

Le ZnF (**voir chapitre I.1.a**) est une structure particulière en épingle qui intervient aussi dans la liaison à l'ARN, il appartient à la famille des RanBP2[219]. Une étude a mis en évidence son mode d'interaction par RMN avec un ARN de motif UGGU. De manière simplifiée, la partie Nter du domaine (Q420-Q421-R422) se replie sur la base G<sub>2</sub>, reconnue par le biais de liaisons hydrogène. L'acide aminé F438 se superpose directement au G<sub>3</sub>, tandis que le U<sub>4</sub> est reconnu par la chaîne latérale du R441, et la chaîne principale du M436, F438. De cette façon, la reconnaissance est faite de manière spécifique. L'analyse approfondie de cette interaction conclut que ce domaine reconnaît spécifiquement les séquences NGGU où N peut être n'importe quelle base[211]. Enfin, le ZnF participe à l'autorégulation de FUS[211].

A la fin, de la séquence on retrouve le NLS qui interagit avec un important réseau de récepteur nucléaire tel que kapb2, TNPO1, ainsi que l'importine qui permet la régulation de la localisation de la protéine[220].

### **b) Mode d'interaction et ARN pris en charge par FUS**

FUS se lie à plus de 8000 ARN (au moins 5500 chez l'Homme), ce qui a été montré par des études génomiques sur la souris, des lignées cellulaires, ou directement chez l'humain[6]. Le motif de liaison de FUS le plus commun est GUGGU ou GUU[6, 221], cependant FUS prend en charge de très nombreux ARN, pas

seulement à l'aide de son RRM, mais aussi avec ses RGG et son ZnF. Ce qui fait que FUS est capable de reconnaître l'ARN, à la fois par sa structure et sa séquence, mais sans grande spécificité. Ainsi, les RGG à eux seuls peuvent se lier à l'ARN, mais ne permettent pas une grande affinité individuellement[222], en revanche les 3 RGG ensemble offrent une grande affinité pour l'ARN. Ainsi la partie RGG-ZnF-RGG peut se lier à l'ARN et à l'ADN avec la même affinité que la protéine pleine longueur[223]. D'un côté, le ZnF se lie à un motif GGU/A[219] par le biais notamment de 2 poches spécifiques à l'uracile qui se forment autour de la base par des interactions spécifiques, combinaison de liaisons hydrogène de type Watson-Crick et de type Hoogsteen sur les bords de la base[211]. D'un autre côté, grâce à ses RGG, FUS reconnaît les structures en G-quadruplex par complémentarité de forme[224], le RGG3 en particulier, permet la reconnaissance de ces structures, ce qui implique les résidus aromatiques Tyr et Phe, qui ne forment pas de structures très ordonnées mais plutôt des complexes dynamiques, comme cela a pu être suggéré par des études RMN[225]. Le domaine PrLD, de son côté, ne se lie pas à l'ARN, par contre les RRM et les RGG confèrent à la protéine une grande affinité pour l'ARN tandis que pris séparément le domaine RGG-Zn-RGG montre une coopérativité intermédiaire et le domaine PrLD-RRM a une grande affinité mais une coopérativité très réduite[53]. Enfin, Les RRM de FUS lui permettent de jouer le rôle de chaperonne, en remodelant la structure d'une boucle d'ARN pour que les bases soient mieux exposées pour la liaison à d'autres protéines partenaires[211]. Ce rôle a été décrit chez d'autres protéines comme les nucléolines[226]. De plus, la protéine reconnaît aussi les boucles d'ARN riches en motif AU[227]. La protéine a aussi une bonne affinité pour les structures secondaires de l'ARN, mais une affinité plus faible pour l'ARN en épingle et pour le simple brin excepté le poly-A et les motifs GGUG[222]

Concernant les cibles de FUS, la protéine est connue pour se lier aux régions 3'UTR, mais aussi au 5'UTR, aux exons, tout comme aux longs introns[6, 223]. Sur ces derniers, il a été mis en évidence que la protéine avait un mode de liaison en dent de scie[221]. Cette protéine a la particularité de prendre en charge les ARN naissants sur toute leur longueur très tôt après leur transcription[221]. Chez la souris, il n'a pas été observé de chevauchement particulier entre les sites de liaison à l'ARN entre FUS et TDP-43, cependant comme TDP-43, FUS est aussi impliquée dans la régulation de gènes en charge du développement neuronal[221]. Pour finir, tout comme TDP-43, FUS est capable de se lier à son propre ARNm pour autoréguler son expression[65].

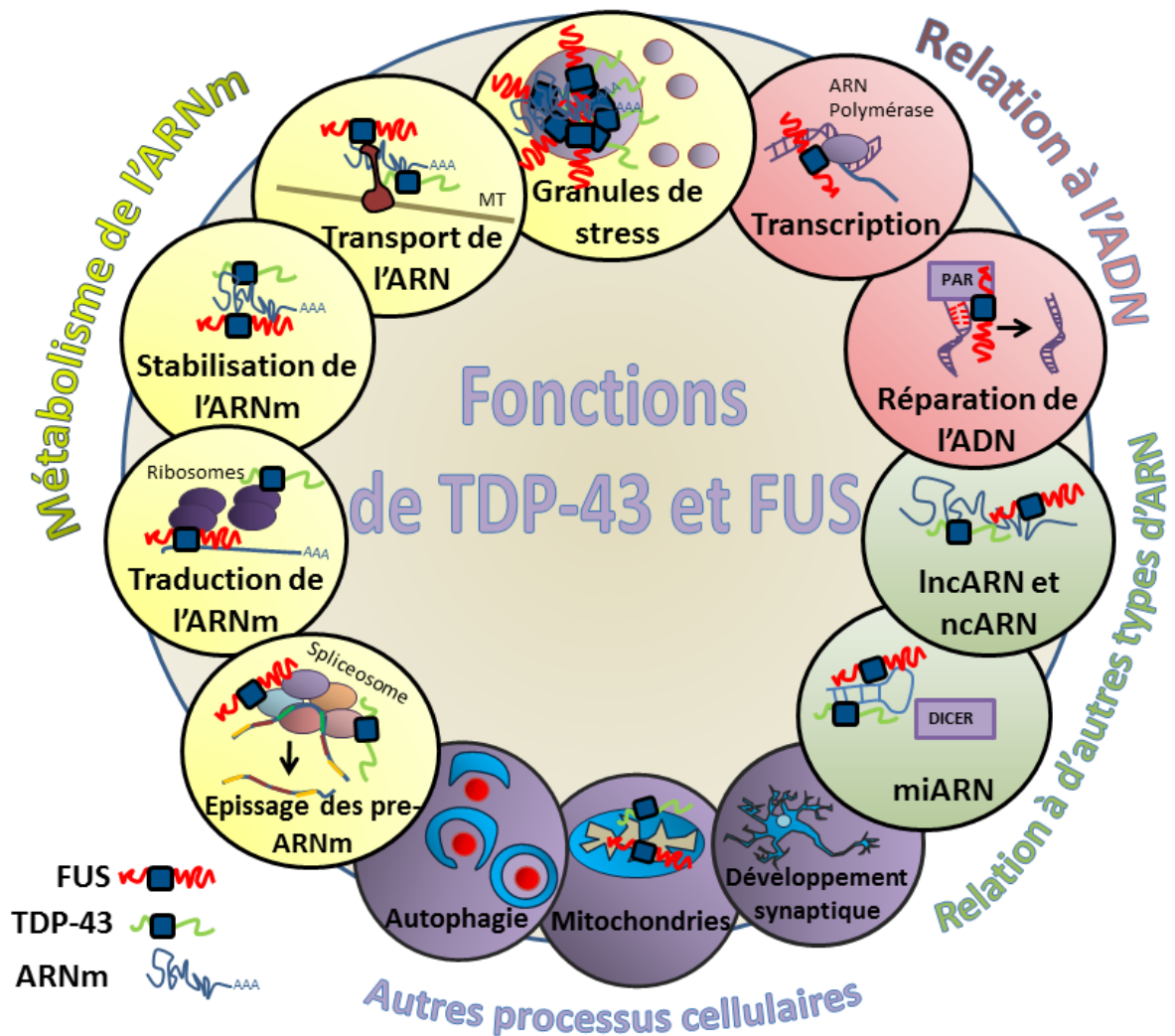
Pour résumer le mode de liaison de FUS, les RGG déstabilisent la structure de l'ARN et permettent son remodelage pour une meilleure prise en charge de l'ARN[219]. La complémentarité de forme permet la reconnaissance d'une boucle de

l'ARN par le RRM, par le biais du feuillet  $\beta$  et la partie Cter du RRM qui se replie sur l'ARN de manière à ce que 4 bases soient prises en charge. De plus, une épingle  $\beta$  vient au contact du sillon principal de l'ARN. Cependant, cette reconnaissance reste peu spécifique avec une affinité modérée[4]. Le RRM et le RGG agissent de manière complémentaire pour la reconnaissance de petites structures d'ARN. Pour conclure, FUS se lie de manière très coopérative et avec une grande affinité mais peu de spécificité pour une séquence particulière.

Nous avons vu dans cette partie les caractéristiques des deux protéines FUS et TDP-43 et constaté qu'elles ont des similitudes, nous avons vu leurs modes d'interaction avec l'ARN et leurs cibles, dans la continuité nous allons maintenant voir les fonctions auxquelles elles participent.

### 3. Fonctions de FUS et TDP-43

FUS et TDP-43 sont des protéines très étudiées, ce qui a permis d'identifier de nombreuses fonctions auxquelles elles sont associées. Elles ont notamment les capacités de se lier à l'ARN en tant que RBP ce qui leur permet d'effectuer une large palette de fonctions, qui leur est commune, dans la prise en charge et le métabolisme de l'ARN, comme son transport, son épissage et sa traduction. D'un autre côté, on trouve des fonctions plus spécifiques à FUS qui comprennent l'interaction avec l'ADN, c'est le cas de la transcription ou de la réparation de l'ADN (**voir figure 18**).



**Figure 18-TDP-43 et FUS sont associées à de nombreuses fonctions en lien avec l'ARN/ADN.**

Ensemble des fonctions pris en charge par TDP-43 et FUS, celles-ci incluent la prise en charge de l'ARNm, la relation à l'ADN, et à divers types d'ARN, mais aussi des fonctions plus globales comme l'autophagie, le fonctionnement des mitochondries et le développement synaptique. Le métabolisme de l'ARN qui inclut, sa stabilisation, son transport, sa traduction et son épissage, implique les deux protéines. Les fonctions en lien avec l'ADN comme la transcription ou la réparation de l'ADN endommagé sont plus spécifiques à FUS. La prise en charge des micro-ARN (miARN) et des longs ARN non codants (lncARN) est commune aux deux protéines. FUS comme TDP-43 se retrouvent dans les granules de stress en cas de stress cellulaire.

### a) Fonctions en lien avec l'ADN

#### ▪ Transcription et régulation des gènes

FUS est très impliquée dans la transcription, comme nous l'avons vu précédemment FUS se lie aux ARN naissants[221]. Le domaine PrLD de FUS lui permet de se lier à la RNAP2 (L'ARN polymérase II) pour réguler, son activité en orchestrant la phosphorylation de son domaine C-ter, ce qui a des implications sur l'initiation de la transcription au niveau de la chromatine[228]. D'un autre côté, FUS est connue pour réguler directement l'activité de l'ARN polymérase II en contrôlant sa

phosphorylation pendant la transcription[229]. De plus, FUS est aussi connue pour intervenir au cours de la transcription dans la régulation des signaux de la polyadénylation[230].

Le rôle de FUS dans la transcription est bien établi, cela est moins le cas pour TDP-43, qui semble, cependant, jouer un rôle indirect, notamment en participant aux processus de régulation de la transcription. En effet, la protéine a aussi la capacité de se lier aux ARN naissants[221, 231]. L'étude de la localisation de TDP-43 dans différentes structures nucléaires, a montré sa présence au niveau des fibres de perichromatine et des sites actifs de transcription[231]. Ces observations tendent à montrer que TDP-43 joue le rôle de facteur de transcription, et participe à régulation de la transcription de certains gènes[231]. Parmi ceux-là, on trouve en particulier SP-10 (en lien avec la spermatogénèse)[232], TNF- $\alpha$  (impliquée dans les maladies inflammatoires)[233], ou VPS4B (protéine impliquée dans le trafic endosomal)[234]. D'un autre côté, TDP-43 réprime l'expression des éléments transposables en s'y liant. C'est notamment le cas avec la transcription de l'HIV[235], ce sont des éléments source d'une grande instabilité génétique.

- **Prise en charge de l'ADN endommagé**

En plus, de se lier à l'ARN, FUS est capable d'interagir avec l'ADN simple et double brins[236], et est connue pour son implication dans la réparation de l'ADN. FUS participe à la réparation de l'ADN en association avec la protéine PARP (La polyADP-ribose polymérase, première protéine impliquée dans la réparation de l'ADN) par le biais de ses domaines RGG2 et PrLD [237, 238]. PARP1 arrive rapidement sur le site endommagé de l'ADN suivi de près par FUS qui dépend de PARP pour son recrutement comme pour son maintien sur le site[121]. L'étude de ce phénomène par AFM a permis de préciser le processus qui se met en place lors de la prise en charge de l'ADN endommagé : PARP-1 synthétise PAR au niveau des sites endommagés et recrute FUS, cela a pour effet de remodeler la chromatine en formant des compartiments transitoires au sein desquels l'ARN est réparé, enfin PARG induit la dissociation de ces condensats et la séparation de FUS[239]. De plus, FUS favorise la formation de boucle D sur l'ADN qui est un processus essentiel pour la réparation de cassure double brins de l'ADN[240]. Les défauts de prise en charge de l'ADN endommagé sont mis en cause dans le développement de pathologies neurodégénératives[212]. En particulier, les mutations de FUS dans le NLS peuvent induire un signal de réponse aux dommages de l'ADN inapproprié, corrélé à la formation d'agrégats cytoplasmiques de FUS[241].

Bien que le lien ne soit pas clairement avéré, TDP-43 aussi de son côté prend part à la stabilité de l'ADN et partage des partenaires d'interaction liés à la prise en

charge de l'ADN en commun avec FUS[242]. De plus, cette même étude a aussi montré que l'interactome de ces deux protéines était modifié suite à un dommage induit à l'ADN, de manière moins importante pour TDP-43 que FUS, ce qui laisse penser que FUS change son comportement suite à un dommage à l'ADN en renforçant ses interactions, tandis que TDP-43 participe indirectement aux mécanismes de réparation, par le biais d'interactions déjà existantes en l'absence de dommage[242]. Par ailleurs, d'autres travaux, ont aussi montré que TDP-43 était nécessaire à la réparation de l'ADN double brin en étant recruté sur les sites de cassure[243].

### **b) Fonctions en lien avec le métabolisme de l'ARNm**

#### **▪ Traduction**

Les RBP peuvent moduler la traduction des protéines de manière globale ou de manière spécifique. Dans le cas de TDP-43 et FUS, il a été montré qu'elles interagissent avec de nombreuses protéines impliquées dans la traduction, des sous-unités des ribosomes ou des facteurs de régulation de la traduction[244, 245]. Ainsi, la perte de cette interaction, à cause d'une mauvaise localisation de TDP-43 dans le cytoplasme, induit une réduction globale de la synthèse en protéine[246]. Les deux protéines peuvent réguler positivement ou négativement la traduction en fonction de leurs partenaires d'interaction, par exemple pour les ARNm *Camta1* ou *Mig12*, TDP-43 favorise leur traduction en promouvant l'association aux ribosomes. Par contre, elle réprime la traduction de *Rac1*[247]. Notons au passage que ces ARN, ont été associés au développement de symptômes moteurs ou cognitifs dans le cas de maladies neurodégénératives telles que la SLA, c'est ce que relève cette étude qui suggère que l'altération de leur traduction contribue au développement de ces pathologies[248].

FUS est impliquée dans la traduction locale des ARNm, par exemple, sa présence dans les complexes ACP-RNP est nécessaire pour une traduction efficace[249]. Ainsi, le recrutement de FUS dans des granules qui sont actives pour la traduction permet la production locale de protéines dans le cytoplasme. Cette étude précise aussi que les mutations en cause dans la SLA ne semblent pas interférer avec cette fonction. De plus, FUS se trouve associée avec le récepteur de signalisation NMDA ce qui suggère qu'elle participe à la régulation de la traduction des ARNm au niveau des sites synaptiques[250].

#### **▪ Epissage**

L'épissage alternatif est un processus effectué par un complexe de protéines nommé spliceosome qui permet la maturation de l'ARN pré-messager (pré-ARNm)

en ARNm en éliminant les introns et en conservant les séquences d'exon pour former la séquence codante qui sera traduite en protéine. TDP-43 en se liant aux domaines riches en motifs GU à proximité des séquences introniques permet la régulation de l'épissage et la conservation de différents exons. De plus, TDP-43 se lie à un grand nombre de protéines impliquées dans l'épissage (au moins 13 protéines distinctes, dont une majorité de la famille des hnRNP)[244]. Les cibles de TDP-43 dans l'épissage regroupent des pré-ARNm comme POLDIP3/SKAR, SORT1, STAG2, MADD, TNIK, et notamment un nombre important de protéines impliquées dans la neurodégénérescence comme hnRNPA1, Tau, et SMN. De plus, TDP-43 joue un rôle dans la répression de l'inclusion d'exons cryptiques[251]. Par ailleurs, TDP-43 est connue pour être impliquée dans la régulation de l'épissage alternatif de l'exon9 du gène CFTR (*cystic fibrosis transmembrane conductance regulator*) qui encode le régulateur transmembranaire de conductance de la mucoviscidose[190]. L'épissage par l'intermédiaire de TDP-43 permet de réguler la stabilité des ARNm. En particulier, la protéine se lie à son propre transcrite ce qui permet son autorégulation[64]. Il a été mis en évidence que l'altération de l'épissage est corrélée avec le développement de maladies neurodégénératives à la fois chez l'homme et les modèles animaux[252, 253] et cela pourrait être dû à la perte de fonction de TDP-43 causant un changement dans le métabolisme de l'ARN[254].

FUS aussi est impliquée dans l'épissage, ainsi les pré-ARNm forment la majeure partie de ses cibles[255]. Elle joue le rôle de facteur d'épissage en s'associant avec les composants du spliceosome. Cela a notamment été suggéré par une étude ancienne qui montre qu'elle régule le site 5' dans le pré-ARNm EA1[256]. Plus récemment, de larges études ont analysé le réseau d'exons à l'échelle du génome et ont établi que FUS intervient dans la prise en charge de plusieurs milliers d'exons[212]. Chez la souris, par exemple, une dérégulation de l'expression de FUS entraîne une altération de 3 202 exons, dont un certain nombre est associé à des gènes en lien avec la neurodégénérescence[257]. Par exemple, la dérégulation de la protéine FUS conduit à l'inclusion d'exons de MAPT qui code la protéine TAU, comme l'a montré cette étude sur des tissus neuronaux[258]. FUS stabilise préférentiellement les transcrits avec de très longues séquences introniques (>100kb) régulant ainsi l'épissage de gènes impliqués dans les mécanismes en lien avec la neurotransmission, l'exocytose, le transport de vésicules ou l'organisation des membranes[259]. Enfin, FUS se lie en particulier aux introns dans les gènes codants d'autres RBP, et régule donc leur épissage[255]. Par conséquent, FUS joue un rôle majeur dans l'épissage alternatif des pré-ARNm, comme l'indique sa déplétion qui induit une importante perturbation dans le métabolisme de l'ARN[6, 260].



### ▪ Stabilisation

Les protéines de liaison à l'ARN participent à la régulation de l'expression des gènes en favorisant ou en empêchant le recrutement de la machinerie de dégradation de l'ARN, c'est en cela que l'on parle de stabilisation de l'ARN. La capacité de TDP-43 à se fixer aux domaines 3' UTR participe à la stabilisation de l'ARNm[70]. A l'inverse un excès de TDP-43 conduit à l'altération de la stabilité de l'ARNm, c'est ce qui est suggéré par la surexpression de la protéine dans des modèles de cellules iPCS induisant un processus de phosphorylation oxydative[261]. FUS, comme TDP-43, se lie à la séquence 3'UTR de ses ARN cibles et participe ainsi à leur stabilisation. Par exemple, TDP-43 se lie au 3'UTR de *NEFL* (*Neurofilament light chain*) ce qui participe à sa stabilisation. Elle agit de même avec *Add2* et *HDAC6*[247]. De son côté, FUS est connue pour réguler la stabilité de plusieurs centaines de transcrits, elle stabilise au moins 330 ARNm et en déstabilise 44 au sein de cellules progénitrices neurales[262].

La déplétion de FUS et TDP-43 (mise sous silence de gènes) dans des cellules NSC-34 n'a pas montré d'effet particulier de stabilisation de certains ARN cibles (*Vegfa*, *Grn*, *Vps54*, *Nvl* et *Taf15*)[263]. FUS forme un complexe avec des facteurs contrôlant la polyadénylation et la prise en charge du 3' UTR (comme PAN2, PABPC1 et CPSF6), lorsque la concentration de FUS diminue, le complexe est affecté et donc on observe une décroissance de la stabilité de l'ARN (*GluA1*)[264], ce qui suggère un contrôle indirect de la stabilité de l'ARN.

### ▪ Transport de l'ARN

Nous l'avons évoqué, TDP-43 comme FUS, sont deux protéines principalement nucléaires mais qui font la navette entre le noyau et le cytoplasme, notamment parce qu'elles sont impliquées dans le transport de l'ARN. Si l'on considère le cas des neurones, ce sont des cellules complexes qui peuvent être très grandes avec une partie synaptique très éloignée du corps cellulaire. Cependant, la traduction locale au niveau de la synapse est nécessaire au bon fonctionnement de la cellule, il est donc nécessaire de transporter l'ARNm jusqu'au lieu de sa traduction. C'est donc le rôle des RBP (incluant les protéines Staufen), comme FUS et TDP-43 qui forment des RNP pour acheminer l'ARNm dans le cytoplasme[247].

TDP-43 est impliquée dans le transport axonal et dendritique et a été détecté à la fois au niveau pré- et post-synaptique dans des granules contenant de l'ARN. Son transport est bidirectionnel et fonctionne suivant le réseau de microtubules[265]. Il a aussi été mis en évidence que FUS participe au transport des granules d'ARN dans les dendrites de neurones et régule ainsi la traduction au niveau des synapses[266], ainsi elle se retrouve au niveau des terminaisons dendritiques[267]. De plus, elle interagit

avec le moteur microtubulaire KIF5, ce qui suggère que la protéine est transportée de manière active dans les neurones[268]. FUS comme TDP-43 sont transportés vers les neurites terminaux, en association avec les microtubules, grâce aux filaments d'actine, mais aussi par le biais de tubules du réticulum endoplasmique et de neurofilaments du cytosquelette[269].

### c) Prise en charge des LncARN et miARN

Les micro-ARN (miARN), des petits ARN de l'ordre d'une 20<sup>aine</sup> de bases, et les longs ARN non-codants (LncRNA, *Long-non coding RNA*), formés de séquences contenant plus de 200 nucléotides, sont des ARN non traduits en protéine qui sont impliqués dans la régulation de l'expression des gènes. Pour cela, ils s'apparient à l'ARN de manière quasi-complémentaire et spécifique. Leur dérégulation a des implications dans le développement de cancers et de diverses pathologies mais reste mal comprise[270].

FUS comme TDP-43, se retrouvent dans les « *paraspeckles* » des complexes nucléaires formés de protéines spécialisées et d'ARN, en particulier de lncRNA comme NEAT1\_2 (*Nuclear-Enriched Abundant Transcript 1\_2*) identifié dans les cerveaux atteints de la SLA. Il a été montré que TDP-43 et FUS pouvaient se lier directement à ce transcrit, notamment par des techniques de iCLIP. Ainsi, FUS contribue à la stabilité des *paraspeckles* en régulant le niveau d'expression de NEAT1\_2, et les mutations de FUS (en particulier au niveau du domaine Nter) affectent la formation de ces structures[271]. FUS est capable de se lier à au moins 30 % des lncRNA répertoriés, sur des sites distincts de TDP-43[272]. Enfin, la régulation croisée de NEAT1 et TDP-43 est essentielle pour leurs fonctions dans la promotion des différents états de différenciation des cellules souches[273].

Les miARN matures sont issus du clivage d'un ARN plus long (pri-miARN), un premier clivage se produit au sein du complexe de Drosha dans le noyau, un second dans le cytoplasme au niveau du Dicer. TDP-43 est impliqué dans les deux phases du processus. D'un côté, TDP-43 en interagissant avec les complexes nucléaires de Drosha et en se liant directement aux pri-miARN. D'un autre côté, dans le cytoplasme TDP-43 interagit avec le complexe de Dicer et favorise la prise en charge de certains miARN en s'y liant[274]. FUS participe à la régulation des gènes grâce au processus de mise sous silence des gènes par le biais des miARN. En effet, FUS se lie directement au miARN et interagit plus spécifiquement avec le cœur du complexe de miARN formé par la protéine AGO[275]. De plus, FUS se lie aux pri-miRNA naissants et favorise leur traitement en recrutant le complexe de Drosha au niveau de la chromatine. Cela est attesté par le fait que la déplétion de FUS dans les cellules altère

l'expression d'un nombre importants de miRNA[276]. Fait intéressant, les miARN mi-b2121 et miR-194 se lient au niveau du 3'UTR des transcrits de FUS et TDP-43 de manière à réguler leur niveau d'expression, il est alors possible que les deux protéines puissent s'autoréguler par le biais d'une boucle de rétroaction négative en régulant la biogénèse de ces miARN[277].

En plus de la prise en charge des lncARN et des miARN, TDP-43 prend aussi en charge les ARN non codant (ncRNA). En particulier, la protéinopathie TDP-43 chez la drosophile, a des conséquences sur les mécanismes responsables de la répression des retro-transposons. Cependant, l'importance de cette interaction dans le cas des MND est encore mal comprise[278].

#### d) Autres fonctions cellulaires

En plus de ces fonctions, directement en lien avec le métabolisme de l'ARN (et de l'ADN) TDP-43 et FUS peuvent être impliquées dans des processus plus larges comme le fonctionnement des mitochondries, l'autophagie et les fonctions neuronales et synaptiques (**voir figure 18**). Le dérèglement de ces dernières est, par ailleurs, suspecté d'être lié au développement des maladies neurodégénératives[247]. En effet, la dérégulation de ces fonctions de base peut avoir des conséquences plus large sur des processus cellulaires comme l'endocytose, la neuro-inflammation, l'auto-immunité, les mécanismes épigénétiques ce qui peut induire des perturbations globales au niveau du profil métabolique de l'organisme[251].

FUS et TDP-43 peuvent se trouver dans les **mitochondries** des neurones, la présence de ces RBP semble nécessaire pour leur bon fonctionnement, d'autant que FUS est capable de se lier à la sous unité  $\beta$  de l'ATP synthase[279]. De plus, ces protéines semblent s'accumuler au sein de ces organelles lors de mutations pathologiques. Un dysfonctionnement des mitochondries étant observé dans les maladies, peut-être est-il lié à un défaut de ces RBP ? Ce mécanisme reste à élucider[247].

L'**autophagie** est le mécanisme par lequel les composants intracellulaires sont dégradés et recyclés. Les protéines ATG (*autophagy-related*) forment un groupe de protéines impliquées dans ce processus. Il a été montré que TDP-43 régule l'autophagie notamment en se liant aux transcrits des ATG[247]. De plus, il a été mis en évidence que TDP-43, comme FUS, affecte directement l'expression de composants de la machinerie de l'autophagie, et des lysosomes lorsque ces protéines sont dérégulées ou mutées[280–282]. Ce mécanisme étant important pour le nettoyage des agrégats cytoplasmiques on peut suspecter qu'il est défaillant dans le cas de la SLA[247].

Pour finir, FUS et TDP-43 sont impliquées dans le **développement synaptique et neuronal**. Le dysfonctionnement de ces deux protéines laisse présager une altération de l'environnement synaptique. TDP-43 comme FUS peuvent être localisés au niveau des synapses. D'ailleurs, la baisse ou la réduction des niveaux de TDP-43 et FUS conduisent à une altération de la morphologie des neurites. Le lien semble être indirect et plus d'études seront nécessaires pour mieux comprendre ce mécanisme en lien avec les maladies neurodégénératives[247].

Nous avons donc vu que TDP-43 et FUS étaient impliquées dans de nombreuses fonctions qui pouvaient avoir des conséquences globales sur la régulation des processus cellulaires. Les protéines sont elles-mêmes régulées par le biais de modifications post-traductionnelles, ce qui a des conséquences sur leurs comportements et donc sur leurs fonctions. C'est ce que nous allons voir dans la prochaine partie.

## 4. Modifications et mutations de ces protéines

### a) Modifications post-traductionnelles

Les protéines, en général, sont sujettes à de nombreuses modifications post-traductionnelles (MPT) qui consistent, la plupart du temps, à attacher un groupement de manière covalent, à certains acides aminés par le biais d'une enzyme spécifique (**voir figure 19**). Elles peuvent modifier le comportement des protéines, en particulier elles permettent de réguler le processus de LLPS et les interactions protéine-protéine. Une altération dans le fonctionnement de ces MPT peut être liée avec la dégénérescence, puisque sous forme agrégées les protéines présentent de nombreuses MPT (hyper-phosphorylation, ubiquitinylation, mais aussi hypo-méthylation...). Ici nous allons décrire les principales MPT qui touchent les protéines FUS et TDP-43, et leurs effets sur le fonctionnement de ces protéines, puis nous verrons ce qu'il en est dans le cas pathologique.

#### ▪ Méthylation

La méthylation correspond à l'ajout d'un groupement méthyle sur un groupement amine (lysine) ou guanidinium (arginine) par le biais de la méthyltransférase. Pour, FUS il a été montré que la méthylation des arginines contenues dans le NLS et dans le RGG3 permettait de réguler la localisation de la protéine. En effet, FUS interagit avec TNPO1, qui est un récepteur d'importation nucléaire, et moins FUS est méthylé au niveau de la région c-terminale, plus cette interaction est importante[283]. De plus, la méthylation des arginines accroît la solubilité de la protéine tandis que la non-méthylation a tendance à favoriser la

transition vers un état de gel puis d'agrégation *in vitro*[284]. De la même manière, la présence d'une di-méthylation asymétrique au niveau des arginines des RGG, décroît la capacité de formation d'agrégats de la protéine, quand la sous-méthylation favorise son accumulation dans les granules de stress[285]. La méthylation semble donc être importante pour la solubilité de la protéine et la sous-méthylation pourrait être un facteur favorisant l'agrégation.

- **Acétylation**

L'addition d'un groupement acétyle sur les acides aminés lysines, catalysée par les acétyltransférases, se fait souvent en condition de stress cellulaire et peut influencer le comportement des protéines[286]. Chez FUS l'acétylation influence sa capacité à participer au processus de LLPS. Au niveau du domaine PrLD, cette MPT augmente la capacité à former des condensats et réduit l'agrégation[287]. L'acétylation de FUS permet aussi de réguler l'interaction avec l'ARN, l'acétylation des lysines K315 et K316 réduit à la fois l'interaction avec l'ARN et la formation d'inclusions cytoplasmiques[288].

Pour TDP-43, il a été mis en évidence que l'acétylation de lysines au niveau du RRM provoque une perte de la liaison à l'ARN et favorise la formation d'agrégats insolubles[289]. L'acétylation de la K84, au niveau du NLS de TDP-43, réduit son importation dans le noyau et induit sa délocalisation vers le cytoplasme[290]. Enfin, l'acétylation est liée à d'autres MPT puisqu'elle conduit à une augmentation des phosphorylations et des ubiquitinations des agrégats de TDP-43[286].

- **Phosphorylation**

La phosphorylation qui correspond à l'addition d'un groupement phosphate sur un acide aminé portant une fonction alcool (serine, thréonine) ou phénol (tyrosine), permet aussi de réguler le comportement des protéines. Elle a des conséquences très ambivalentes chez les protéines comme FUS et TDP-43 qui contiennent des domaines de faibles complexités, qui sont riches en acides aminés phosphorylables.

Chez FUS, la phosphorylation des serines et thréonines dans le LCD N-terminal peut conduire à la translocation de FUS du noyau vers le cytoplasme suite à un dommage à l'ADN comme décrit dans cette étude[291], néanmoins, une autre étude montre que la multi-phosphorylation du PrLD n'affecte pas sa localisation[292]. D'un autre côté, la phosphorylation du domaine Cter de FUS peut réduire l'interaction avec TNPO1 et l'importation de FUS au noyau[293], c'est aussi ce qui se passe quand FUS est phosphorylée au niveau de la tyrosine 526 contenue dans le PY-NLS, qui a pour conséquence d'altérer la liaison à Kap $\beta$ 2 et affecte donc la répartition noyau/cytoplasme[293]. Ces études révèlent le côté ambivalent de la phosphorylation

selon les sites touchés. Il semble, néanmoins, plus clairement établi que le rôle de la phosphorylation du PrLD permet de réguler le processus de LLPS, les contacts protéine-protéine, et plus largement la propension de la protéine à l'agrégation[113, 151, 294]. Par exemple, la phosphorylation par les DNA-PK dissout les gouttelettes de FUS. Plus précisément la phosphorylation des résidus Ser et Thr dans le cœur fibrillaire du PrLD de FUS interfère avec la formation de structures  $\beta$  croisés et donc d'agrégats fibrillaires[151].

Chez TDP-43, la phosphorylation du Nter de TDP-43 déstabilise la formation de condensats ce qui a des répercussions sur son implication dans l'épissage[295]. La phosphorylation du domaine PrLD accroît la solubilité de TDP-43 et prévient l'agrégation[286]. Cet effet protecteur est contraire aux observations des cas pathologiques, car la phosphorylation de multiples résidus dans le domaine PrLD est associée à l'agrégation de la protéine, c'est en effet une constante dans les agrégats neuronaux à la fois dans les formes familiales et sporadiques de la SLA (notamment au niveau des S409/410)[296]. Ainsi, il n'est pas clair si la phosphorylation de TDP-43 dans les neurones résulte d'un mécanisme pathologique ou au contraire d'une tentative des cellules pour dissoudre les agrégats de TDP-43. Il faudra plus de travail pour le déterminer, cependant on peut suspecter, que comme pour FUS, la phosphorylation de différents sites peut avoir des conséquences distinctes[116].

- **Ubiquitylation**

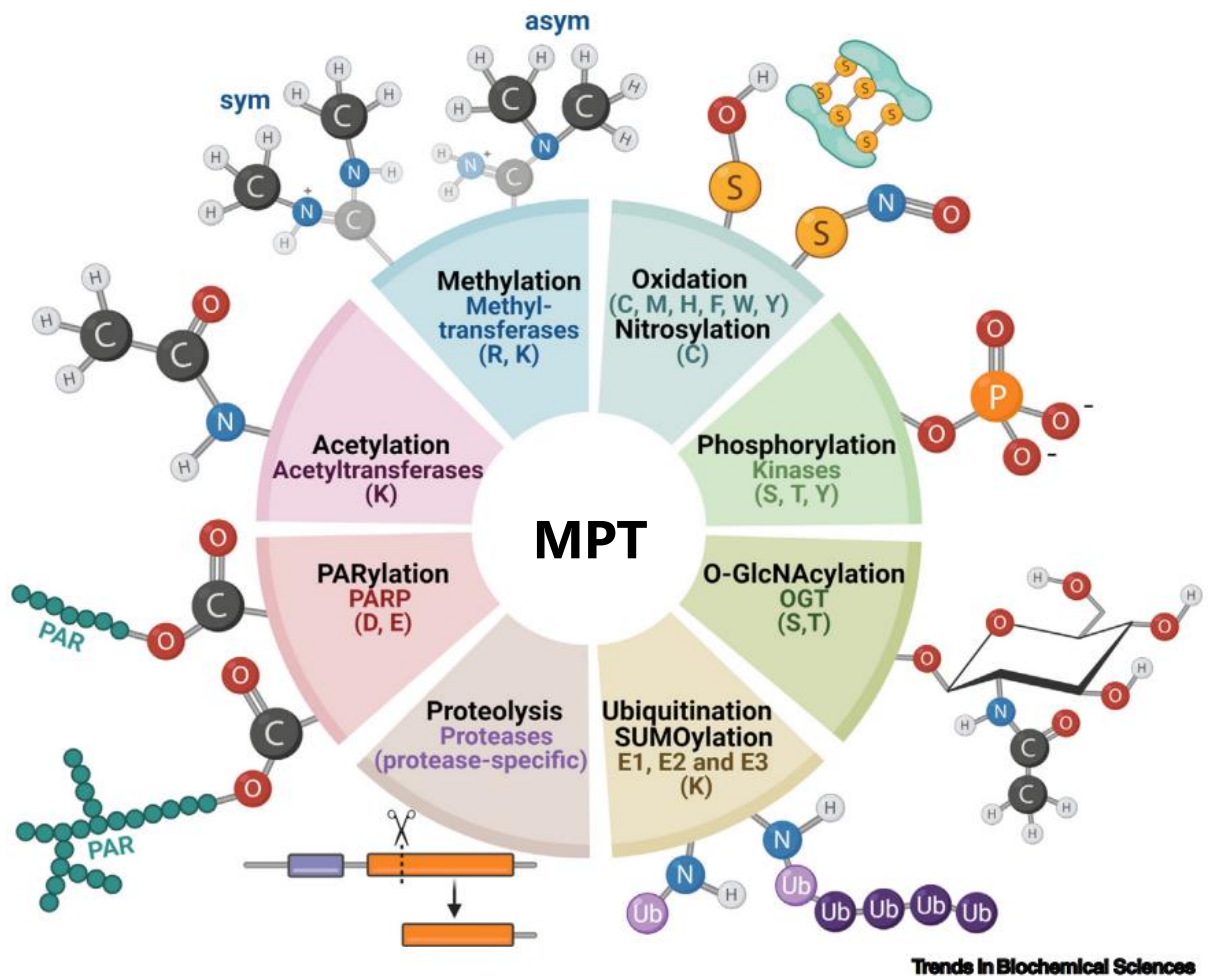
L'ubiquitylation est une MPT qui a une place importante dans la régulation et le recyclage des protéines. Elle consiste en l'ajout de marqueurs ubiquitine (une petite protéine) au niveau des lysines par le biais de ligases, dans le but de faciliter la dégradation des protéines. L'ubiquitylation est une caractéristique majeure des agrégats cytoplasmiques présents dans les neurones de patients, aussi bien pour FUS que pour TDP-43. Le facteur UBQLN en lien avec le protéasome peut se lier à la protéine TDP-43 ubiquitylée et stimuler son agrégation[297]. D'ailleurs, chez la drosophile, la surexpression de ce facteur accroît la sévérité du phénotype de neurodégénérescence[298]. C'est aussi le cas de UBQLN2 qui favorise l'accumulation de la protéine dans le cytoplasme et son agrégation[299]. Plusieurs études montrent que TDP-43 semble avoir un lien étroit avec le système du protéasome et de la machinerie à ubiquitine : l'ubiquitylation représente alors un facteur crucial pour la préservation du bon équilibre de la protéine TDP-43 au sein des cellules[300]. Chez FUS, l'ubiquitylation peut avoir lieu dans le RRM ou le ZnF ce qui permet sa dégradation[301]. De manière générale, l'ubiquitylation peut affecter la formation de condensats et l'agrégation[302] De plus, la forte concentration d'ubiquitine dans les inclusions de FUS conduit à la perturbation de l'homéostasie cellulaire, ce qui peut

conduire à une mauvaise régulation des voies de contrôle de la qualité des protéines en aval[303].

- **L'oxydation**

L'oxydation, notamment des résidus cystéines, peut jouer sur le comportement de la protéine, d'autant que le contexte de stress oxydatif au sein des cellules dans un contexte pathologique, représente un mécanisme important dans le développement de maladies[300]. TDP-43 est en effet capable de s'assembler par le biais d'oxydation des cystéines induisant la formation de ponts disulfures. Par exemple, l'oxydation des cystéines dans le RRM1 (Cys 173 et Cys 175) affecte la conformation de TDP-43 et conduit à un autoassemblage aberrant suspecté d'être à l'origine d'une agrégation irréversible[304]. Au contraire, dans le domaine Nter la mutation de deux cystéines (Cys39 et Cys50) a pour effet de réduire la capacité de TDP-43 à former des oligomères et donc nuit au bon fonctionnement de la protéine[305].

D'autres MPT peuvent avoir des implications sur le comportement de ces protéines comme la PARylation (ajout d'un ou plusieurs groupements ADP-ribose)[237], la SUMOylation (pour *Small Ubiquitin-Related Modifier*), l'o-GlcNAcylation[306] (ajout d'un monosaccharide), ou encore la nitrosylation[307] (ajout d'un groupement oxyde nitrique) (**voir figure 19**). Nous n'entrerons pas dans les détails pour celles-ci. Nous allons voir, à présent, la relation des MPT précédemment citées sur les maladies neurodégénératives, mais aussi les mutations, ainsi que les clivages retrouvés en conditions pathologiques.



**Figure 19-Le comportement des RBP est régulé par diverses modifications post-traductionnelles.** Présentation des différentes modifications post-traductionnelles que peuvent subir les protéines. Dans chaque cas les enzymes ainsi que les acides aminés concernés sont indiqués. Les groupements chimiques ajoutés lors des MPT sont mis en évidence sur le pourtour du cercle. **D'après** Sternburg EL, et al. *Post-translational modifications on RNA-binding proteins: accelerators, brakes, or passengers in neurodegeneration?* *Trends Biochem Sci.* 2022 Jan;47(1):6-22. doi: 10.1016/j.tibs.2021.07.004. Epub 2021 Aug 5. PMID: 34366183.

### b) MPT et mutations en lien avec les maladies

Les protéines agrégées sont généralement hyperphosphorylées[296], et ubiquitinylées[184] au sein des inclusions cytoplasmiques dans les neurones des malades. C'est un phénomène qui est bien documenté pour la protéine TAU dans le cas de la maladie d'Alzheimer, qui se retrouve hyper-phosphorylée et agrégée dans les neurones[308]. Il a ainsi été mis en évidence que la phosphorylation avait pour conséquence de favoriser l'agrégation de la protéine[309]. C'est aussi le cas de TDP-43 qui en plus d'être hyperphosphorylée, acétylée, ubiquitinylée[184], se retrouve sous forme clivée au sein des agrégats de patients[286]. Ainsi l'acétylation de la lysine K145 chez TDP-43 dans des conditions de stress oxydatif a été identifiée



comme une modification associée à la SLA[286]. De même que pour TDP-43, les inclusions de FUS sont positives à l'ubiquitine[183]. De plus, cette dernière est normalement di-méthylé au niveau des arginines dans le RGG (en particulier RGG3), ce qui est le cas en conditions normales, mais aussi dans les agrégats dans le cas de la SLA, en revanche, elle apparaît sous-méthylée dans le cas des agrégats dans la DFT[310]. De plus, chez les patients SLA, il a été mis en évidence un plus haut niveau d'acétylation sur les lysines du NLS (K510). L'acétylation de la K510 induit une perte de localisation nucléaire de FUS, et bien que le lien entre l'acétylation de FUS et le mécanisme de développement de la maladie ne soit pas clairement démontré, il a toutefois été proposé que des inhibiteurs de la désacétylase permettent de réduire la formation d'inclusion[288].

Maintenant que nous avons mieux compris l'implication des MPT sur le comportement des protéines et leur rôle dans les maladies, nous allons nous intéresser aux mutations et aux formes tronquées qui sont caractéristiques des pathologies.

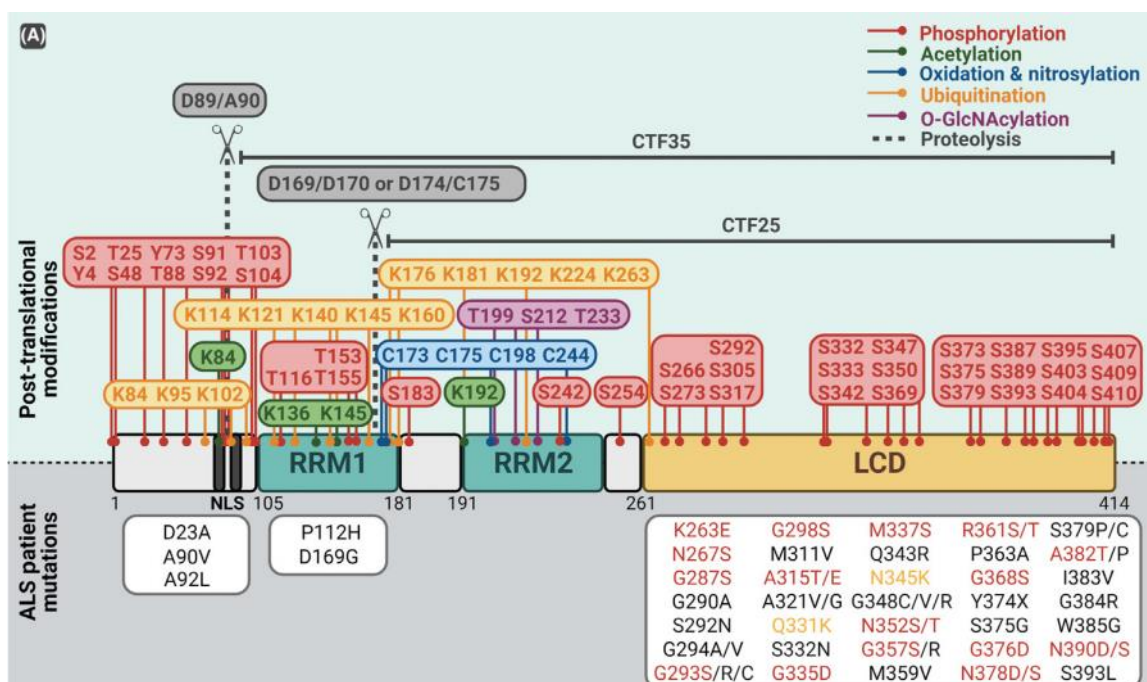
#### ▪ Mutations pathologiques

Il existe plus d'une cinquantaine de mutations pathologiques répertoriées pour la protéine TDP-43, la plupart sont situées dans le domaine PrLD (**voir figure 20**). Plus précisément, le site ALSod<sup>1</sup> qui collecte des informations sur les variants génétiques de la SLA recense, à ce jour, 69 mutations pour TDP-43 et 126 pour FUS. Un nombre important d'études montrent que les mutations permettent d'élever le niveau de TDP-43 dans la cellule en accroissant sa stabilité, et en altérant son autorégulation[311]. On observe des taux plus élevés de la protéine chez les patients de DFT sporadique, ce qui laisse penser que ces dysfonctionnements pourraient entraîner un mécanisme d'agrégation lié à l'âge, ou une neuro-inflammation[312]. De manière générale, les mutations dans le PrLD altèrent la séparation de phases de TDP-43, sa capacité à mettre en place des interactions protéine-protéine, sa demi-vie ainsi que sa répartition nucléo-cytoplasmique[313]. Des modèles de souris ont été développés, avec dans leurs génomes, des mutations ponctuelles dans le domaine PrLD de TDP-43. Ils montrent une perturbation de l'autorégulation de la protéine et plus généralement un gain de fonction qui induit une mauvaise modulation de l'épissage alternatif[314, 315]. Un autre exemple de conséquences des mutations pathologiques dans le domaine PrLD, avec la mutation A315T qui affecte le transport dendritique[316].

---

<sup>1</sup> Disponible ici : [http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene\\_id=TARDBP](http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=TARDBP) [consulté le 19-02-22].

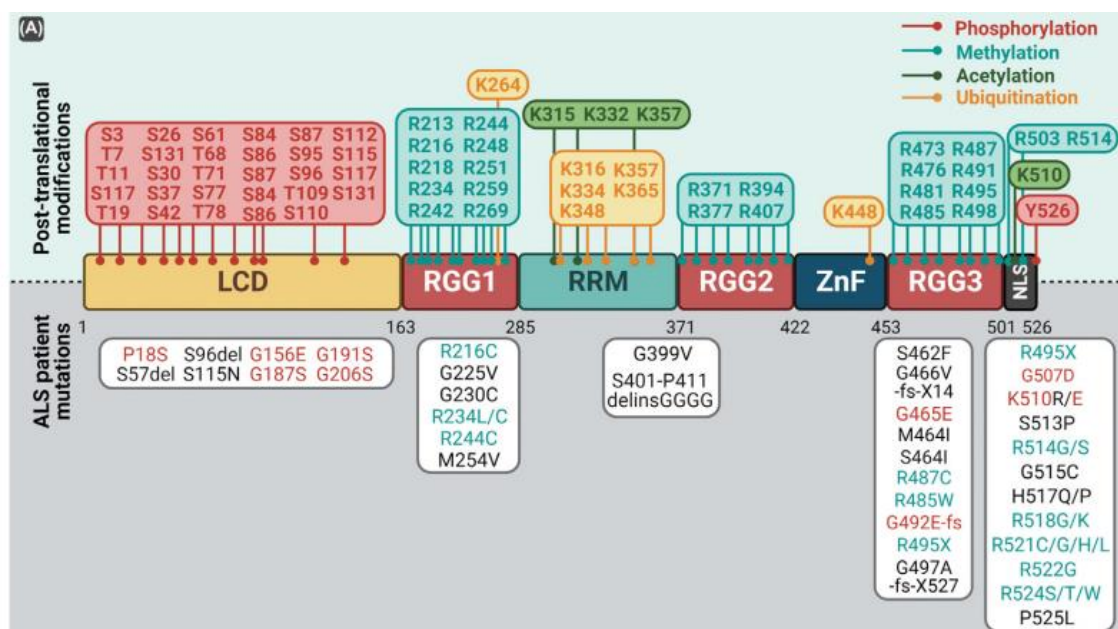
Quelques mutations ont aussi été mises en évidence dans le RRM, qui en général, ont un impact sur la liaison à l'ARN. Elles favorisent aussi l'agrégation ce qui suggère que la perte de liaison à l'ARN peut favoriser la toxicité de la protéine. Les conséquences de ces mutations, ont été étudiées sur des modèles cellulaires. La mutation K181E dans le RRM1 cause la perte de l'interaction avec l'ARN et accentue l'accumulation d'agrégats insolubles de la protéine dans le cytoplasme[317]. La mutation D169G favorise la formation d'agrégats insolubles, et surtout elle favorise la formation de formes tronquées (TDP25 et TDP35, voir paragraphe suivant)[318]. La double mutation F147L/F149L induit une perte de liaison à l'ARN et déstabilise la capacité de TDP-43 à autoréguler son niveau de transcription[63]. Les mutations N259S et K263E au niveau du RRM altèrent l'affinité de la protéine pour l'ARN et la conduisent aussi à s'agréger[319]. Enfin, la mutation située dans le NLS A90V induit une accumulation de la protéine dans le cytoplasme en la séquestrant sous forme d'agrégats insolubles[320].



**Figure 20-La protéine TDP-43 présente un nombre important de mutations et de sites de modifications post-traductionnelles.**

Représentation schématique de la séquence de la protéine TDP-43 avec l'ensemble des sites de modifications post-traductionnelles (haut), et des mutations associées aux maladies neurodégénératives (bas). On remarque un nombre important de mutations au niveau du domaine LCD en position C-terminale de la protéine. Les fragments issus de clivage protéolytique ainsi que le site du clivage sont indiqués par des ciseaux. Les chiffres indiquent la position des acides aminés concernés. Les mutations qui induisent un site de phosphorylation supplémentaire sont mises en évidence en rouge. Les mutations créant un site potentiel d'acétylation ou d'ubiquitinylation, sont mises en évidence en jaune. **D'après** : Sternburg EL, et al. Post-translational modifications on RNA-binding proteins: accelerators, brakes, or passengers in neurodegeneration? Trends Biochem Sci. 2022 Jan;47(1):6-22. doi: 10.1016/j.tibs.2021.07.004. Epub 2021 Aug 5. PMID: 34366183.

Contrairement à TDP-43 où les mutations sont principalement situées dans le PrLD, chez FUS un nombre important de mutations ont été identifiées au niveau du NLS (**voir figure 21**) ce qui affecte la localisation de la protéine en la redistribuant vers le cytoplasme[321]. Par exemple, un modèle de souris transgénique exprimant le mutant FUS-R521C présente des symptômes typiques de la SLA. Cela résulte d'une perturbation de l'épissage et d'un défaut dans la régulation de la croissance du réseau dendritique des neurones[322]. De manière générale, les mutations pathologiques sont connues pour amplifier les transitions de phase liquide vers l'état solide pour FUS, ce qui a été démontré *in vitro*[107]. La mutation R244C, au niveau du RGG1, induit une perte de l'interaction avec l'ARN (notamment HDAC1) et influe sur l'efficacité de la protéine dans les processus de réparation de l'ADN[321]. La mutation G156E dans le domaine PrLD augmente la tendance de la protéine à s'agréger, *in vitro* elle présente des caractéristiques différentes (gouttelettes plus grosses, changement de propriétés biophysiques) et favorise l'évolution des gouttelettes vers la formation de structures fibrillaires[121]. Enfin, les mutations de FUS peuvent affecter l'interaction de la protéine avec d'autres complexes, comme par exemple le U1-snRNP causant une altération dans les processus d'épissage[323].



**Figure 21-La protéine FUS présente un nombre important de mutations et de sites de modifications post-traductionnelles.**

Représentation schématique de la séquence de la protéine FUS avec l'ensemble des sites de modifications post-traductionnelles (haut), et des mutations associées aux maladies neurodégénératives (bas). Les numéros indiquent la position des acides aminés concernés. Les mutations qui induisent un site de phosphorylation supplémentaire sont mises en évidence en rouge, les mutations qui suppriment un site de méthylation (arginine) sont mises en évidence en bleu. **D'après :** Sternburg EL, et al. Post-translational modifications on RNA-binding proteins: accelerators, brakes, or passengers in neurodegeneration? *Trends Biochem Sci.* 2022 Jan;47(1):6-22. doi: 10.1016/j.tibs.2021.07.004. Epub 2021 Aug 5. PMID: 34366183.

### ▪ **Clivage ou protéolyse**

Des fragments de la protéine TDP-43 clivée ont été identifiés dans les agrégats de patients atteints de protéinopathies. Il s'agit de deux fragments côté C-terminal, en particulier de 25 kDa et 35 kDa qui sont issus du clivage protéolytique par caspase. D'un autre côté, jusqu'à récemment aucun fragment Nter n'avait été détecté[300], récemment l'existence de ces fragments a été mise en évidence, toutefois, ils resteraient dans le noyau et seraient rapidement dégradés, et donc seraient moins impliqués dans la toxicité[324]. La présence de fragments C-ter de TDP-43, tels que TDP25 ou TDP35, est très fréquente dans les inclusions cytoplasmiques de la protéine TDP-43[184, 325]. Par ailleurs, la distribution de ces fragments n'est pas homogène, ils se retrouvent en particulier dans les régions corticales et de l'hippocampe dans les cas de SLA, DFT et la maladie d'Alzheimer mais sont plus rares dans les échantillons de moelle épinière[326]. De plus, la présence de ces fragments Cter n'est pas une condition à la toxicité de TDP-43, puisqu'ils ne sont pas toujours présents, comme par exemple chez des modèles de souris qui montrent des signes de neurodégénérescence chez lesquelles leur présence n'a pas pu être détecté[300].

La première troncation TDP35 a lieu juste avant le RRM1 au niveau du NLS (AA A90V), la seconde TDP25 a lieu en plein milieu du RRM1 (au niveau de Asp174), ce qui altère sa structuration et donc sa capacité à se lier à l'ARN[327]. En conséquence, la forme TDP25 est particulièrement toxique pour les cellules[328]. De plus, des modèles de souris sur-exprimant le fragment TDP25, présentent des caractéristiques typiques de la SLA et de la DFT comme des déficits moteurs et cognitifs[329, 330]. La perte d'un RRM dans le fragment Cter de TDP-43 déstabilise la formation de granules de stress et la régulation de la transcription et conduit à la formation d'agrégats[331]. De son côté, la forme TDP35 conserverait sa capacité à se lier à l'ARN mais induirait la séquestration de la protéine pleine longueur sous forme d'agrégats cytoplasmiques[332]. Ainsi, les fragments TDP25 et TDP35 ont un effet négatif sur le métabolisme cellulaire et leur expression est associée à une augmentation de la toxicité, en particulier elles contribuent à augmenter la capacité de TDP-43 endogène à s'agréger[300]. Cet effet est vrai pour les deux formes tronquées, bien que TDP25 ait une toxicité plus importante que TDP35. D'un autre côté, un fragment de 35 kDa, semble être un isoforme naturel (AUG au niveau de l'AA 85) de la protéine, généré à partir d'un site de traduction alternatif. Il peut, par ailleurs, jouer un rôle physiologique dans la formation des granules de stress[333]. En plus de ces deux formes, de nombreuses troncations de 15 à 20 kDa (à la fin du RRM2) ont été détectées dans le cerveau humain. Elles sont issues de l'action d'endopeptidases

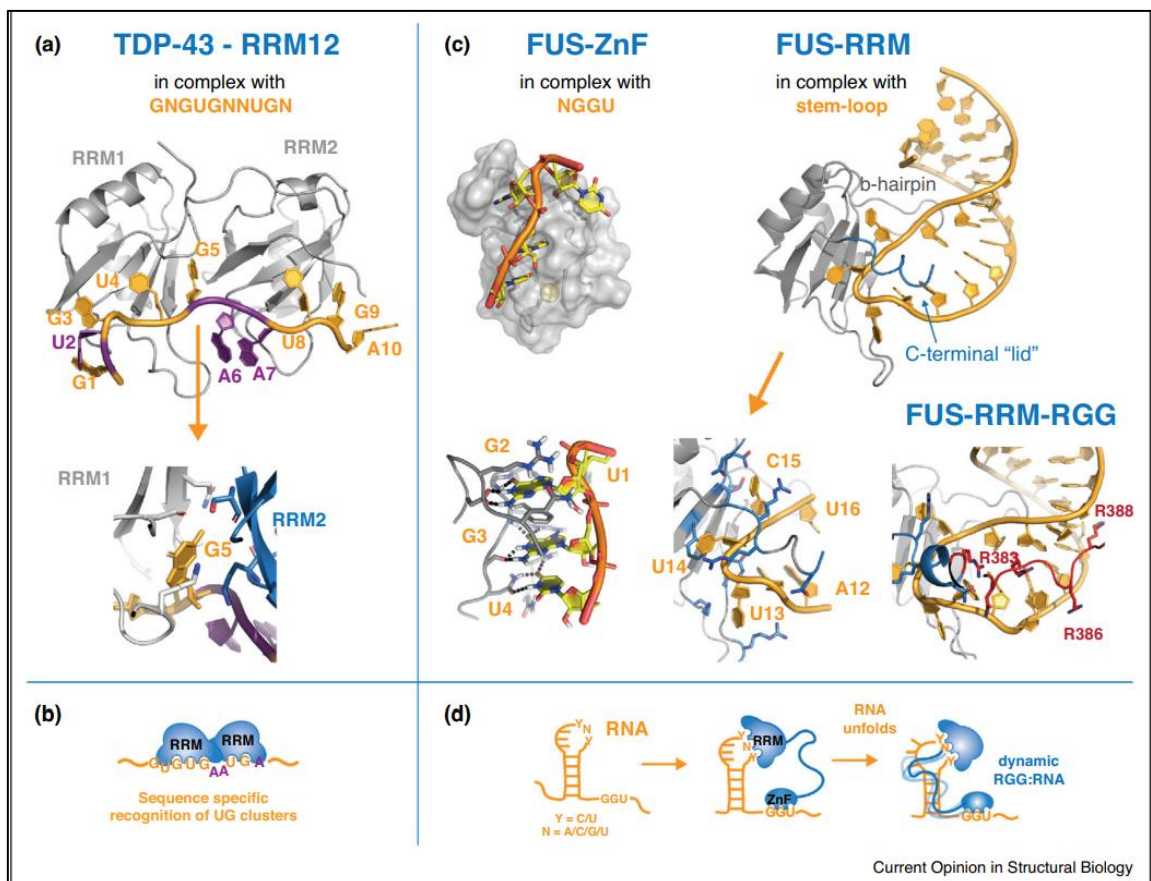
(AEP) ou d'enzymes Calpain. Elles sont suspectées de causer un stress cellulaire supplémentaire[300]. Parmi ces petits fragments TDP-16, à qui il manque les deux RRM, a une capacité particulière à former un cœur amyloïde rigide et bien ordonné, alors que TDP25 et TDP35 arborent une structure moléculaire polymorphe[334]. Par conséquent, ces fragments pourraient être responsables de la propagation d'agrégats de cellule en cellule sur le modèle prion (**voir chapitre I.3.d**)[324]. Pour finir, il est important de remarquer que l'on retrouve ces fragments présents dans les agrégats d'un large spectre de MND et pas seulement la SLA, ou la DFT, comme la maladie de Parkinson pour laquelle ils co-localisent avec l' $\alpha$ -synucléine ou certain cas de maladie d'Alzheimer où l'on retrouve TDP25. Cela suggère qu'ils pourraient simplement être des sous-produits de dégradation sans conséquence, néanmoins leur toxicité et leur capacité à induire l'agrégation et à séquestrer la protéine pleine longueur laissent penser qu'ils jouent un rôle important dans la rupture d'équilibre dans les processus cellulaires[324].

Quelques formes tronquées ont aussi été décrites pour FUS. En effet, certains patients possèdent des variants tronqués de la protéine, comme G478X et R495X qui sont associés à un développement précoce et agressif de la maladie[335]. La troncation R495X se fait au niveau du domaine Cter de FUS, il s'agit d'une coupure d'un petit fragment de 32 AA, au niveau du RGG3, incluant le NLS[336]. Cette dernière, a été étudiée dans un modèle cellulaire (HEK-293) et chez le poisson-zèbre, ce qui a montré que la protéine FUS tronquée avait tendance à s'accumuler dans le cytoplasme et à s'incorporer davantage dans les granules de stress. Par conséquent, cela induit une perte de fonction de la protéine. De plus, des mutations au niveau de l'exon 13 de FUS peuvent conduire à la production de formes tronquées de la protéine (p.G466VfsX14), lorsqu'elle est exprimée dans un modèle cellulaire, cette troncation est associée à une délocalisation de la protéine vers le cytoplasme et une réduction de son expression nucléaire normale[337].

Ces fragments des protéines TDP-43 et FUS perturbent le fonctionnement normal des protéines, et induisent un stress cellulaire. De plus, ils ne sont pas correctement dégradés, ce qui suggère qu'ils doivent jouer un rôle direct dans la progression des maladies neurodégénératives. On peut aussi supposer, et nous y reviendrons, que de tels tronctions altèrent la capacité d'interaction de ces protéines.

### c) Comparaison et mise en évidence des points communs entre les deux protéines

Comme nous venons de le voir TDP-43 et FUS sont des protéines très similaires, qui partagent différentes fonctions autour de la prise en charge de l'ARN et sont toutes deux essentielles à l'organisme. En effet, l'expression de ces protéines est très importante pour le bon développement neuronal et la survie des cellules, puisque la délétion des gènes codants ces protéines est létale chez la souris[338]. De plus, l'expression de FUS et TDP-43 est ubiquitaire, et est essentielle pour le développement, puis elle décroît avec l'âge dans tous les tissus, excepté les neurones moteurs. Elles partagent des similitudes en matière de structure, notamment, elles ont toutes deux un RRM et un LCD de type PrLD, cependant, TDP-43 a aussi un domaine Nter impliqué dans la dimérisation, alors que FUS possède un ZnF, ainsi que des RGG qui lui permettent aussi de se lier à l'ARN. Ces caractéristiques structurales

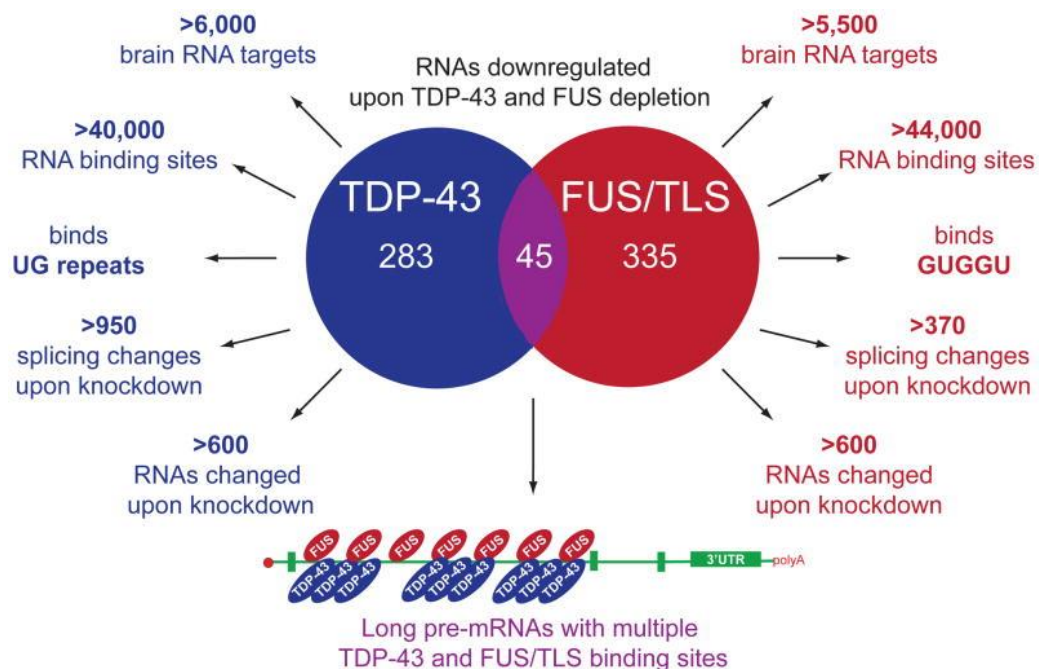


**Figure 22-Mise en évidence des différences d'interaction avec l'ARN entre TDP-43 et FUS.**

(a) Complexe des domaines RRM12 de TDP-43 avec l'ARN (GUGUAAUGAA), la reconnaissance est spécifique pour le motif GU (b). (c) Structure du domaine ZnF de FUS en interaction avec un fragment NGGU et du RRM de FUS en interaction avec une boucle d'ARN. Les RGG (bas) présentent une interaction dynamique avec le sillon mineur de l'ARN (modèle RMN). (d) modèle proposant un mode d'interaction pour FUS impliquant ses différents domaines de liaison à l'ARN, (RGG, RRM, ZnF), l'interaction se fait par spécificité de séquence et reconnaissance de forme. **D'après** : Loughlin FE, Wilce JA. TDP-43 and FUS-structural insights into RNA recognition and self-association. *Curr Opin Struct Biol.* 2019 Dec; 59:134-142. doi: 10.1016/j.sbi.2019.07.012. Epub 2019 Aug 31. PMID: 31479821

communes leur permettent de créer des contacts protéine-protéine, de se lier à l'ARN et de former des condensats. Cependant, elles présentent aussi des différences, notamment en termes de liaison à l'ARN, TDP-43 utilise ces deux RRM pour reconnaître l'ARN avec une grande affinité (de l'ordre du nM) et une grande spécificité, notamment pour les motifs GU, alors que FUS par le biais de son RRM, ses RGG et son ZnF se lie à tous types d'ARN avec peu de spécificité de séquence (**voir figure 22**).

Ainsi, nous avons vu que TDP-43 et FUS prenaient en charge un grand nombre d'ARN de façon à effectuer des fonctions variées, parfois distinctes et parfois qui se chevauchent. En dépit, d'un ensemble de cibles relativement distinctes (**voir figure 23**) [263], elles partagent des cibles communes, en particulier des gènes encodant des protéines essentielles au bon fonctionnement neuronal, la transmission synaptique et la plasticité neuronale[339, 340]. En conséquence, selon cette étude, qui a comparé le changement d'expression de gènes en fonction de la déplétion de ces deux protéines (parmi les 283 gènes qui sont impactés par la déplétion de TDP-43, ou les 335 gènes pour FUS, seul 45 gènes dépendraient à la fois de FUS et TDP-43), TDP-43 et FUS agiraient plutôt en parallèle et non en synergie[6].

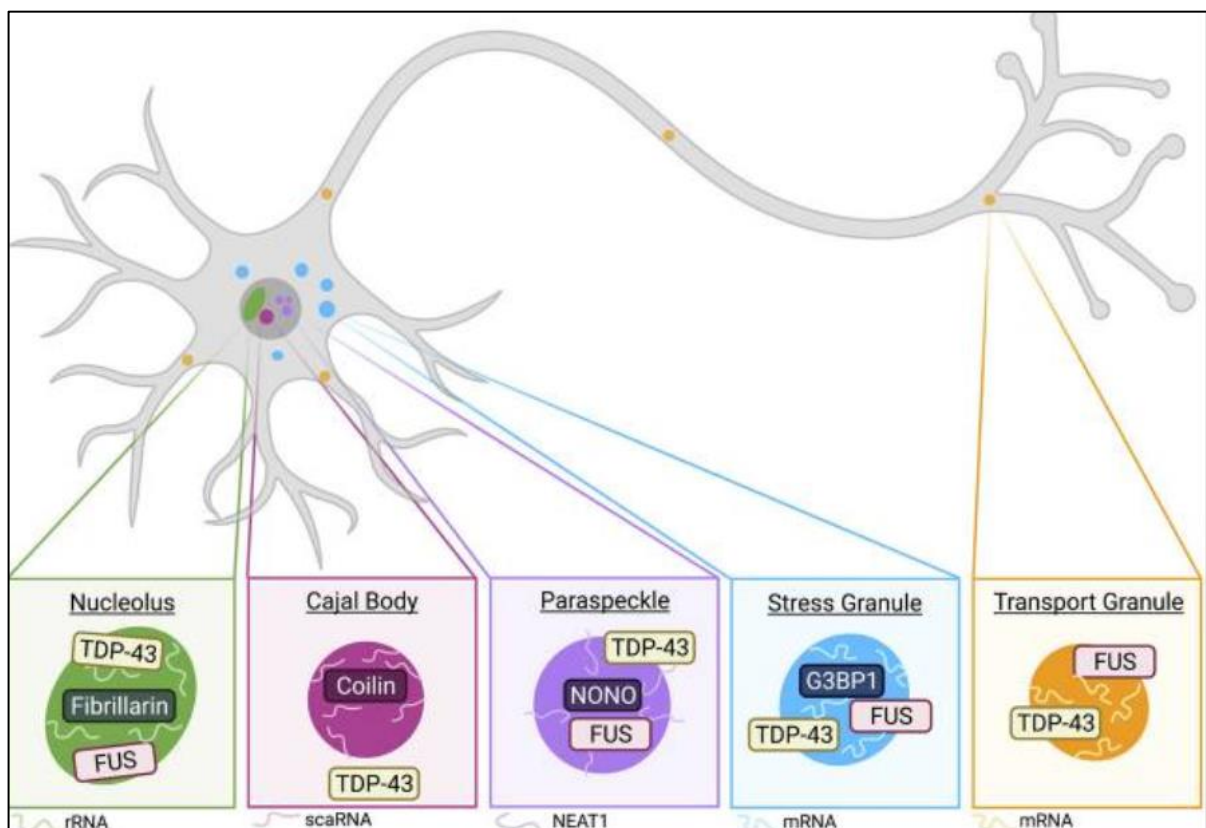


**Figure 23-Comparaison des cibles et des propriétés de liaison à l'ARN des protéines TDP-43 et FUS.**

Présentation des cibles de la protéine TDP-43 (en bleu) et FUS (en rouge). On remarque que leurs cibles sont largement distinctes, avec un recouvrement de 45 ARN dérégulés dans le cas de la déplétion de FUS et de TDP-43. FUS comme TDP-43 prennent en charge les long ARN (bas). **D'après :** Ling SC, et al. *Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis.* *Neuron.* 2013;79(3):416-438. doi:10.1016/j.neuron.2013.07.033

### ▪ Localisation

TDP-43 et FUS sont souvent localisées aux mêmes endroits, ce sont toutes deux des protéines principalement nucléaires, qui ont la capacité de faire la navette entre le noyau et le cytoplasme, notamment pour le transport de l'ARN[236, 341]. De plus, les deux protéines peuvent se trouver au niveau des synapses pour la traduction locale de l'ARNm[267, 342]. Elles peuvent aussi être localisées dans les complexes de Drosha[343] ou encore dans les mitochondries[247]. Nous avons vu précédemment (**voir chapitre 1.3.b**) qu'elles étaient toutes deux impliquées dans la formation de certains condensats en lien avec leurs activités comme les nucléoles, les *paraspeckles*, les granules de transport[3] ou encore les granules de stress[137, 344] (**voir figure 24**). Enfin, dans les cas pathologiques, comme on a pu le voir précédemment, leur localisation est perturbée, puisqu'on les retrouve toutes deux agrégées principalement dans le cytoplasme.



**Figure 24-FUS et TDP-43 sont localisées dans de nombreuses organelles en commun.**

Principalement nucléaire FUS et TDP-43 peuvent être exportées dans le cytoplasme. Elles participent à la formation de nombreux condensats comme les nucléoles (vert), les *paraspeckles* (violet), les granules de stress (bleu), ou les granules de transport (orange). Elles peuvent ainsi être transportées jusqu'au niveau des synapses. **D'après :** Carey JL, Guo L. Liquid-Liquid Phase Separation of TDP-43 and FUS in Physiology and Pathology of Neurodegenerative Diseases. *Front Mol Biosci.* 2022;9:826719. Published 2022 Feb 2. doi:10.3389/fmolb.2022.826719



- **Quantité relative**

Nous venons de voir que TDP-43 comme FUS avaient de nombreuses occasions de se rencontrer et d'interagir, on sait aussi qu'elles sont toutes deux exprimées de manière ubiquitaire. Une question qui émerge pour mieux comprendre leur relation est de savoir leur quantité relative au sein des cellules. L'une est-elle plus abondante que l'autre ?

Une étude qui a évalué la concentration de différentes RBP au sein de cellule a mesuré une concentration moyenne de FUS de 2,33  $\mu\text{M}$  (6,9  $\mu\text{M}$  dans le noyau) et de 0,98  $\mu\text{M}$  pour TDP-43 (2,65  $\mu\text{M}$  dans le noyau) soit l'équivalent d'environ un ratio 2,6/1[142]. Néanmoins, une étude de protéomique à grande échelle permet d'accéder aux quantités des protéines et de leur niveau d'expression, et relève que les deux protéines seraient présentes en quantité comparable[345]. Cette étude a analysé la composition de 29 tissus provenant de différents organes par RNAseq et spectrométrie de masse. Il apparaît donc que dans la majorité des tissus, y compris dans le cerveau, les quantités d'ARNm et de protéine exprimée pour chacune des protéines est similaire. Parfois même avec un léger excédent de TDP-43 en termes d'expression d'ARNm.

## 5. TDP-43 et FUS interagissent

### a) Mise en évidence de leur interaction

Comme nous l'avons vu dans les chapitres précédents les deux protéines FUS et TDP-43 partagent des caractéristiques et des fonctions similaires, on les trouve aussi aux mêmes endroits, nous nous posons donc la question de savoir si elles peuvent interagir l'une avec l'autre ? Dans cette partie, nous allons voir que plusieurs travaux, avec des techniques diverses, ont mis en évidence cette interaction.

L'outil de bio-informatique Pinot[346], qui recense les interactions protéine-protéine<sup>2</sup>, répertorie trois études pour le couple FUS/TDP-43. La première étude met en évidence la formation d'un complexe FUS/TDP-43 par le biais de manipulations d'immunoprécipitation couplées à la spectrométrie de masse et de *GST-pull down*[14]. En plus de mettre en évidence une interaction, cette étude suggère que l'interaction de FUS comme de TDP-43 est nécessaire pour réguler les mécanismes d'épissages de l'ARNm en particulier dans le cas de la HDAC6 (histone deacetylase 6) et participerait l'une et l'autre à son expression optimale, néanmoins les deux protéines seraient en compétition pour un site de liaison unique[14].

#### Détection des interactions protéine-protéine

Il existe de nombreuses techniques pour mettre en évidence une interaction protéine-protéine, parmi celles-ci, quelques-unes fréquemment rencontrées en biochimie sont la co-immunoprécipitation, le *pull-down* ou encore la PLA.

Chaque technique possède des avantages et des inconvénients, les critères majeurs de distinction sont la spécificité et la sensibilité. Pour plus de précisions, quelques techniques sont décrites et comparées dans la partie matériel et méthodes de ce rapport.

#### Boîte de texte 8-Techniques de détection des interactions protéine-protéine

Une deuxième étude confirme par immunoprécipitation l'interaction de FUS et TDP-43 dans les extraits nucléaires de fibroblastes, tout en pointant le dysfonctionnement des RBP (notamment le fait que leurs interactions soient perturbées par une mutation dans la ligase LRSAM1) dans la neurodégénérescence[347].

Enfin, une troisième étude dresse une carte des interactions de protéines impliquées dans les maladies neurodégénératives, parmi 5000 protéines, on trouve

<sup>2</sup> Disponible ici : [http://www.reading.ac.uk/bioinf/PINOT/PINOT\\_form.html](http://www.reading.ac.uk/bioinf/PINOT/PINOT_form.html) [Consulté le 11-04-22]

un lien entre TDP-43 et FUS. Cette carte a été établie par criblage en double-hybride chez la levure (Y2H) et confirmée par une analyse computationnelle[348].

A côté de ces trois études, d'autres études de protéomiques ont mis en évidence l'interaction de FUS et TDP-43. Les protéines partenaires de TDP-43 ont été isolées par immunoprécipitation, purifiées puis identifiées par spectrométrie de masse[244]. Cette étude a aussi montré que de manière générale TDP-43 se lie à des protéines en lien avec le métabolisme de l'ARN ou de l'ADN, notamment de la famille des hnRNP [55, 244]. De la même manière, une autre étude, grâce à la méthode TAP (*Tandem affinity purification*) associée à la spectrométrie de masse, a permis l'identification de TDP-43 avec un nombre important de protéines parmi lesquels les hnRNP, les protéines impliquées dans la prise en charge des miARN, mais aussi d'autres RBP comme FUS[16].

Enfin, certains travaux font état de cette interaction directement au sein de cellules grâce à des techniques complémentaires. Au sein des cellules (HeLa), par exemple, La PLA a permis de mettre en évidence l'interaction FUS/TDP43, ce qui est aussi confirmé une nouvelle fois par immunoprécipitation lors de ce travail, mais dans une faible proportion[16].

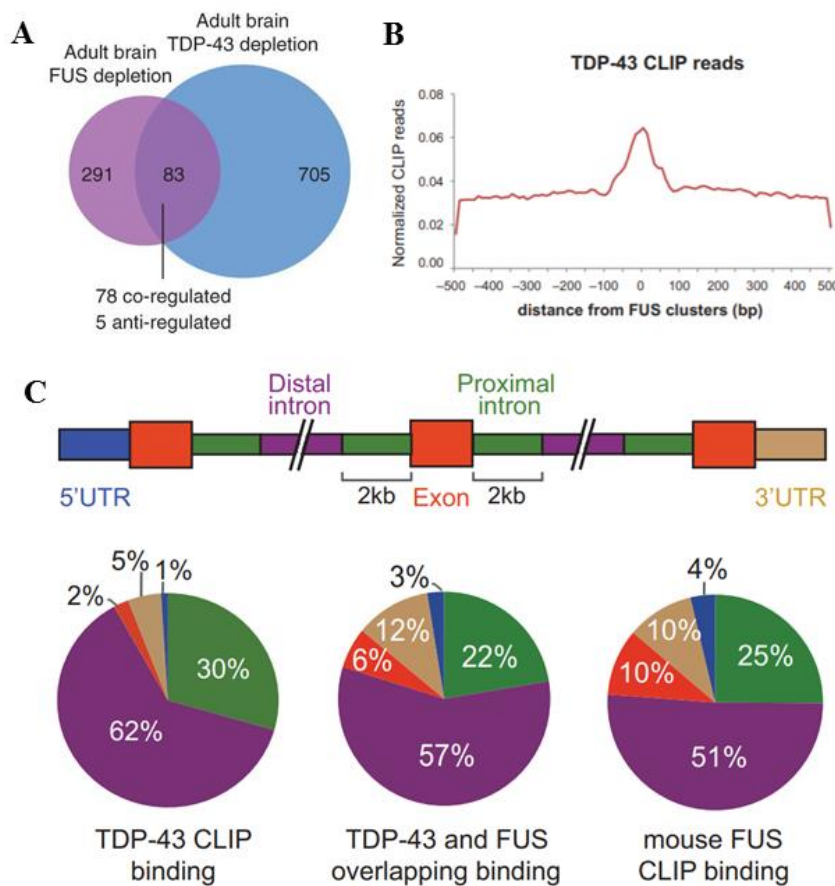
L'ensemble de ces travaux, par le biais de techniques distinctes, montre un lien entre ces protéines, ce qui suggère une participation à des mécanismes communs. Nous allons donc voir, plus en détail, quelle est la nature de cette interaction, en condition physiologique, et ce qu'elle advient en condition pathologique.

### **b) Conséquences sur leurs fonctions physiologiques ?**

On peut supposer de nombreuses fonctions communes à FUS et TDP au regard du fait qu'elles partagent un certain nombre d'ARN cibles[6, 262]. On a vu aussi que FUS comme TDP-43 semblaient jouer un rôle prépondérant dans les fonctions neuronales et qu'elles partageaient des localisations communes, nous allons donc tâcher de montrer maintenant que cet ensemble de faits converge vers une interaction nécessaire au bon fonctionnement physiologique des cellules.

Comme évoqué précédemment, chez la souris il y a 41 et 45 gènes (sur 13 627) qui dépendent à la fois de FUS et TDP-43, c'est aussi 78 à 86 évènements d'épissages qui sont régulés à la fois par FUS et TDP-43[6] (**voir figure 25, A**). Si l'étude semble indiquer que les deux protéines agissent en parallèle plutôt qu'en synergie, elle précise aussi que 6 312 gènes contiennent à la fois des sites de liaison pour TDP-43 et FUS avec un chevauchement de ces sites d'au moins 1 nucléotide dans au moins 2 727 gènes[6]. Ces sites sont parfois distincts, souvent très proches voir même superposés. Ainsi, il a été mis en évidence un enrichissement 2,2 fois plus important

de la liaison de TDP-43 autour des sites de liaison de FUS (**voir figure 25, B**). De plus, l'intersection des deux protéines se fait largement au niveau des introns, mais on observe un enrichissement de la liaison au niveau 3'UTR des deux protéines ensemble par rapport à leurs cibles individuelles (**voir figure 25, C**) [6]. L'étude de Lagier-tourenne *et al.* conclut donc que les deux protéines ont une fonction commune en relation avec les longs introns des gènes qui encodent des protéines impliquées dans les fonctions neuronales. De manière plus concrète, FUS comme TDP-43 sont capables de se lier directement au lncARN NEAT1\_2 au sein des domaine *paraspeckles*, on peut supposer que cette liaison commune est nécessaire à



**Figure 25-TDP-43 et FUS partagent des fonctions en commun par le biais de sites de liaisons qui se chevauchent.**

(A) Diagramme montrant un changement dans les événements d'épissages en fonction de la déplétion de FUS, ou de TDP-43. On remarque ainsi un certain nombre d'événements en communs. (B) Graphique illustrant la répartition des séquences de liaison de TDP-43 identifiées par iCLIP, par rapport à celle de FUS. La distribution de ces séquences montre une proximité des sites de liaison de TDP-43 avec ceux de FUS. (C) Schéma d'ARN divisé en différentes sections : 5'UTR (bleu), introns distants (violet), introns proximaux (vert), exons (rouge), 3'UTR (marron), illustrant les sites de liaison des protéines (haut). Les diagrammes (bas) montrent la répartition des séquences CLIP identifiée (en % selon les différentes sections de pre-ARNm considérées) individuellement pour les deux protéines (TDP-43 à gauche, FUS à droite), ainsi que la localisation des sites qui se chevauchent entre les deux protéines (au centre). On remarque un schéma de liaison commun qui diffère par rapport à celui des protéines seules, en particulier pour les introns distants qui représentent une majorité des sites de liaison des deux protéines, mais surtout pour le 3'UTR. **D'après** Lagier-Tourenne C, et al. *Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs.* *Nat Neurosci.* 2012;15(11):1488-1497. doi:10.1038/nn.3230

la bonne prise en charge de cet ARN, même si l'étude ne précise pas si leur interaction est simultanée et coopérative[123]. D'un autre côté, FUS et TDP-43 sont recrutées dans les granules de stress par le biais de leur domaine de liaison à l'ARN, mais aussi de leur PrLD. Si cette étude ne montre pas clairement d'interaction de FUS et TDP-43 au sein des GS, elle indique que le PrLD dans le cas de TDP-43 joue un rôle important dans le recrutement de la protéine ce qui laisse penser que cela se fait par l'intermédiaire d'un contact protéine-protéine, par le biais de diverses RBP associées au GS, comme TiA1 mais aussi éventuellement FUS[216]. De plus, la présence de marqueurs de granules de stress dans les inclusions cytoplasmiques suggère un mécanisme pathologique commun entre les protéines se trouvant dans ces complexes (ce que nous allons voir au prochain paragraphe). Enfin, comme évoqué plus haut, un mécanisme commun est suggéré entre TDP-43 et FUS, puisque les deux protéines sont nécessaires à la bonne prise en charge de l'ARNm HDAC6, toutefois elles partagent un site de liaison commun, ce qui indique plutôt une relation de compétition plutôt que de coopération[14].

D'un point de vue des travaux sur les modèles animaux, plusieurs études ont mis en évidence le rôle de l'interaction entre FUS et TDP-43. Chez les poissons-zèbres, par exemple, des travaux suggèrent un mécanisme génétique commun entre TDP-43 et FUS. De plus, FUS est capable de restaurer le phénotype moteur chez les poissons chez lesquels on a supprimé le gène *Tardbp*, cependant ce n'est pas réciproque[17]. Cela a aussi été démontré chez la drosophile, où des protéines homologues à FUS et TDP-43 (Respectivement CAZ et TBPH) ont montré des mécanismes génétiques communs, et interagissent l'une et l'autre[19]. Toujours chez la drosophile, et de manière similaire, un mécanisme commun entre les deux protéines est suggéré, notamment en lien avec la protéine SMN impliquée dans la SMA (amyotrophie spinale), même si l'interaction directe entre les deux protéines n'a pas été mise en évidence[349]. Cela converge avec l'expression des formes sauvages ou mutées de FUS et TDP-43 dans l'œil des drosophiles, où l'on peut observer des phénotypes de dégénération augmenter par la synergie entre les deux protéines[18]. Cette autre étude chez la drosophile suggère, quant à elle, que FUS et TDP-43 agissent indépendamment dans le cas de la prise en charge de séquences introniques d'une répétition toxique de pentanucléotides (UGGAA)[350]. Par ailleurs, cette même étude souligne que ces deux protéines jouent le rôle de chaperonne d'ARN. D'autres travaux suggèrent que TDP-43 et FUS (ainsi que TAF15 dans ce cas) participent à des mécanismes de régulation communs, en conditions physiologiques, comme le montre le chevauchement de gènes dérégulés en cas de suppression de ces protéines dans le striatum de la souris. Cette étude montre qu'un certain nombre de

gènes qui ne présentent pas de liaisons directes avec la protéine déplétée, sont malgré tout dérégulés, ce qui suggère une prise en charge indirecte par le biais d'autres partenaires d'interaction[262]. Enfin, chez la souris TDP-43 et FUS semblent coopérer dans les processus de maturation de l'ARNm puisqu'il a été mis en évidence un recouvrement des gènes altérés lors de mise sous silence de ces deux protéines dans des cultures de neurones primaires[351].

Ensemble ces travaux mettent en évidence un lien entre les deux protéines autour de fonctions communes, cependant certaines pointent une relation compétitive, quand d'autres indiquent que l'interaction serait plutôt coopérative et synergique. Nous garderons ce constat à l'esprit pour aborder la partie résultat de ce rapport, mais avant cela nous allons nous pencher sur les conséquences que peuvent avoir les conditions pathologiques sur le lien entre FUS et TDP-43, est-il renforcé ? Est-il perturbé ? Avant d'y répondre, nous poserons la question de savoir s'il existe un mécanisme pathologique commun entre ces deux protéines.

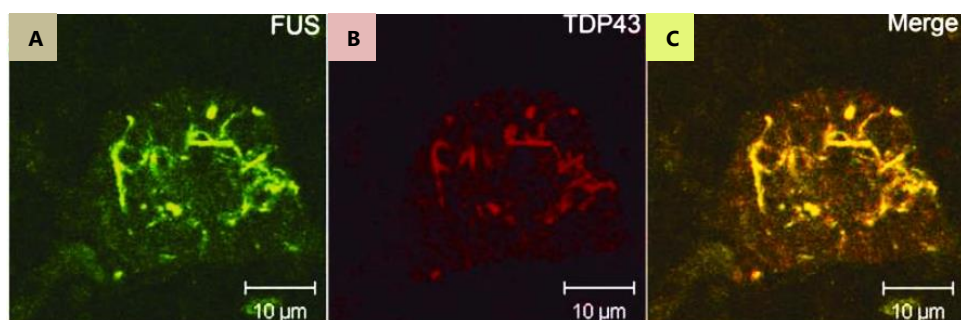
### c) En conditions pathologiques

#### ▪ Un mécanisme pathologique commun ?

Comme évoqué précédemment TDP-43 et FUS ont de nombreux points communs, en particulier, elles sont toutes deux des composantes des GS et se retrouvent sous forme d'agrégats cytoplasmiques ce qui laisse supposer un mécanisme commun dans le développement des MND. Mais qu'en est-il réellement ?

En dépit de la proportion importante de cas avec des agrégats de TDP-43 présents chez les patients DFT et SLA, il a été, dans un premier temps, considéré que les protéinopathies TDP-43 étaient indépendantes des protéinopathies FUS. En effet, les premières études sur les échantillons de patients laissaient penser que les agrégats de FUS et TDP-43 étaient généralement distincts. En effet, TDP43 est présent dans un nombre conséquent d'agrégats chez les patients SLA qui semblent contenir aussi d'autres protéines impliquées dans la réponse au stress, néanmoins, une très petite partie montre aussi des inclusions pour FUS[205, 352]. Cette étude qui montre sur des échantillons de patients que les mutations chez FUS sont en cause dans la forme familiale de la SLA, affirme que ces mêmes échantillons n'étaient pas positifs à la protéine TDP-43[353]. Ce qui est aussi suggéré par les premières études faisant le lien entre la protéine FUS et la DFT[183, 354]. Du côté de la DFT, même si à priori les inclusions neuronales cytoplasmiques semblent distinctes pour les deux protéines, elles partagent des caractéristiques similaires[355], les cas classés sous la terminologie DFT-FUS sont fréquemment associés à un marquage négatif de TDP-43 (DFT-U/TDP-43-négatif)[356].

Toutefois, l'ensemble de ces observations sont contrebalancées par d'autres travaux qui mettent en évidence une co-localisation de FUS et TDP-43 au sein des inclusions. Ainsi cette étude montre que TDP-43 et FUS sont corrélées par marquage d'immunofluorescence dans certains agrégats, retrouvés dans des échantillons de moelle épinière de patients atteints de SLA[20] (**voir figure 26**). Ce qui a été confirmé de la même manière par la suite[357], laissant penser qu'un mécanisme de dégénérescence pouvait être commun aux protéines de liaison à l'ARN. D'autres protéines sont aussi présentes dans ces agrégats, comme la protéine RGNEF, ces agrégats sont aussi marqués par des marqueurs de dégradation comme p62[357]. De manière complémentaire, un modèle de levure a montré des résultats similaires à ce qui a pu être observé dans les échantillons de patients. Dans celui-ci, TDP-43 et FUS sont surexprimées ce qui conduit à leur agrégation. Grâce à la fluorescence on peut observer une co-localisation des deux protéines au sein de ces agrégats[15]. Ces observations indiquent que la réponse à la question de la co-localisation de FUS et TDP-43 dans les inclusions n'est pas forcément évidente et reste débattue. Ainsi, plus récemment, une étude ayant analysé la composition d'agrégats de la protéine TDP-43 dont la formation a été induite au sein de cellules. Ces agrégats contiennent un nombre important de protéines impliquées dans le transport nucléocytoplasmique, et en dépit de plusieurs protéines de liaison à l'ARN détectées, on ne retrouve pas FUS[358].



**Figure 26-TDP-43 et FUS co-localisent au sein d'agrégats dans les neurones de patients atteints de la SLA.**

Echantillons de neurones de moelle épinière d'un patient atteint de SLA marqués par immunofluorescence et observés au microscope confocal. Les agrégats sont positifs à FUS en vert (A) et à TDP-43 en rouge (B), la superposition des deux canaux met en évidence la co-localisation des deux protéines en jaune (C). Ces agrégats sont aussi positifs à l'ubiquitine (non montré).

**D'après** Deng HX, et al. *FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis.* *Ann Neurol.* 2010 Jun;67(6):739-48. doi: 10.1002/ana.22051. PMID: 20517935; PMCID: PMC4376270.

Ces observations d'agrégats contenant à la fois FUS et TDP-43 sont soutenues par le fait qu'il a été mis en évidence que certaines protéines peuvent former des co-agrégats. Dans le cas, par exemple du fragment Cter de TDP-43 et de l' $\alpha$ -synucléine, il

a été observé une hybridation de fibres qui résulte d'une interaction synergique qui amplifie le processus d'agrégation[359]. Bien que la co-agrégation des deux protéines au sein de même structures reste discutée, ces éléments montrent qu'un mécanisme pathologique commun est envisageable.

- **Effet des mutations pathologiques et des MPT sur le binôme TDP-43 et FUS**

Comme décrit précédemment, il a été montré que TDP-43 et FUS pouvaient subir de nombreuses MPT, comme des mutations, ou des troncations qui pouvaient présenter un caractère pathologique. Dans cette partie, nous allons étudier leur impact sur l'interaction de FUS et TDP-43.

Pour commencer, il semble que les mutations peuvent accroître l'interaction entre les deux protéines, comme cela a été décrit par GST-pull down, dans le cas où TDP-43 porte des mutations pathologiques particulièrement agressives dans la SLA (Q331K et M337V). Dans ce cas, les mutations de TDP-43 confèrent une plus grande stabilité à la protéine et augmente l'interaction avec FUS, tout en perturbant son fonctionnement[16]. De même, chez la drosophile la mutation M337V chez TDP-43 avec FUS sauvage d'un côté, et la mutation R521H de FUS avec TDP-43 de l'autre, induisent un phénotype de dégénérescence amplifié par rapport à l'expression des formes sauvages des protéines. Cela suggère que les protéines interagissent de manière synergique pour amplifier le phénotype de dégénérescence[18]. On sait, par ailleurs, que les formes tronquées ou mutées peuvent séquestrer les formes pleine longueur[332, 360], par exemple la mutation P525L de FUS séquestre, au sein d'agrégats insolubles, différentes protéines de liaison à l'ARN comme hnRNP A1 ou SMN, on peut supposer qu'un mécanisme similaire est à l'œuvre avec TDP-43[360]. Dans la continuité, les mutations liées aux maladies neurodégénératives de FUS comme celles de TDP-43 ont un effet sur les granules de stress, en particulier leur taille, leur stabilité et leur longévité sont augmentées. Il est alors supposé que cela peut amplifier le recrutement d'un plus grand nombre de partenaires d'interaction tout en renforçant un mécanisme de dégénérescence commun[136, 361]. Ensemble ces résultats suggèrent que ces mutations peuvent induire une association aberrante des protéines entre elles, conduisant à des perturbations dans leurs fonctions respectives et/ou communes, ce qui suggère une convergence des mécanismes pathologiques de TDP-43 et FUS.

Cependant, ces observations peuvent être nuancées au regard d'autres résultats. Ainsi l'étude de Kim *et al.* n'observe pas de changement, ni d'altération de l'interaction entre les deux protéines (par CoIP), en présence des mutations pathologiques (D169G, A315T, M337V, et R361S). Cette contradiction notamment



dans le cas de M337V nécessite d'être clarifiée. Concernant les formes tronquées de TDP-43, l'étude de Kim *et al*, tend à montrer que le fragment de 35 kDa, issu du clivage de la protéine pleine longueur et retrouvé dans les agrégats de patients, TDP35 (**voir chapitre II.3.b**), tend à être exclu des complexes de la protéine TDP-43 pleine longueur mais conserve toujours la capacité de se lier à FUS[14]. De plus, dans les cellules reprogrammée (iPSC) issues de patients portant des mutations de la SLA, il a été montré que la mutation A382T chez TDP-43 n'affecte pas la localisation de FUS dans des cellules  $\gamma$  compris en condition de stress cellulaire, alors que la mutation de FUS P525L affecte grandement sa localisation, la protéine FUS se retrouve dans le cytoplasme en condition normale et son recrutement dans les granules de stress est amplifié en condition de stress. En revanche, cette étude ne se penche pas sur l'effet des mutations de FUS sur TDP-43[362]

Enfin, si plusieurs travaux, à l'aide de techniques et de modèles distincts, ont mis en évidence l'interaction entre FUS et TDP-43 dans les conditions normales et pathologiques, leur mode de liaison reste toutefois mal connu. Il a toutefois été mis en évidence que le domaine riche en glycine (PrLD) et le RRM2 de TDP-43 contribueraient à l'interaction avec FUS (le PrLD seul n'est pas suffisant), c'est ce qu'il ressort de manipulations de GST-pull down effectuées avec différentes troncations de TDP-43, ainsi *in vitro* les acides aminés compris entre 170 et 414 (soit après le RRM1) de la séquence de TDP-43 sont suffisants pour la liaison avec FUS[14].

En conclusion, nous avons mis en évidence un lien entre deux protéines similaires appartenant à la catégorie des RBP avec LCD. Elles sont toutes deux impliquées dans des fonctions autour de la prise en charge de l'ARN, et aussi dans les processus de séparation de phases. En dépit de la différence dans leurs modes de liaison à l'ARN, ces deux protéines peuvent être amenées à interagir ensemble sur des cibles partagées, et semblent s'organiser autour de fonctions communes dans la régulation des gènes, le transport de l'ARN et l'épissage. L'interaction entre ces deux protéines peut-être amplifiée par certaines mutations ce qui laisse envisager un mécanisme de neurodégénérescence commun, d'autant que ces deux protéines se retrouvent agrégées avec des marqueurs de stress dans certains échantillons de patients atteints de SLA ou DFT.

# Objectifs

---

Nous avons vu dans la première partie de ce rapport que TDP-43 et FUS sont deux protéines de liaison à l'ARN qui partagent des caractéristiques communes en matière de structure, de fonctions et de localisations. Nous avons détaillé, plus particulièrement, leur implication dans le métabolisme de l'ARN et le processus de séparation de phases. Nous avons aussi, constaté qu'elles sont décrites comme des marqueurs de certaines maladies comme la SLA, la LATE ou la DFT, puisqu'on les retrouve sous forme d'agrégats chez les patients. Enfin, nous avons mis en évidence plusieurs travaux qui répertorient une interaction entre ces deux protéines, toutefois, la nature de cette interaction reste mal comprise. **Le fait que ces protéines aux nombreux points communs partagent des localisations, et des fonctions communes nous amène à penser qu'elles peuvent avoir une influence l'une sur l'autre.**

La genèse de ce travail repose sur les questions suivantes : FUS et TDP-43 s'influencent-elles l'une et l'autre dans leur comportement ? Est-ce qu'elles ont tendance à se concurrencer ou bien à coopérer dans la prise en charge de l'ARN ? L'activité de l'une peut-elle conduire l'autre à s'agréger dans certaines conditions ? Ou bien au contraire, leur interaction est-elle nécessaire pour rester soluble ? Dans ce deuxième cas, quels facteurs peuvent perturber cette interaction au point de les conduire à s'agréger ?

Nous allons tâcher d'éclaircir certaines de ces questions dans ce rapport, bien que ce travail ne permette pas de répondre à toutes, il esquisse néanmoins quelques pistes. Ici, nous allons nous attacher dans un premier temps à étudier cette interaction, et la comparer à d'autres dans un contexte cellulaire en utilisant une nouvelle technique développée au laboratoire. Cette approche nous permettra d'analyser la capacité de ces protéines à générer des compartiments via le processus de LLPS excluant ou acceptant d'autres partenaires. Ces résultats seront mis en regard avec des données obtenues à différentes échelles sur la structure des condensats formés par ces protéines. Nous étudierons quels domaines sont particulièrement importants dans la capacité de ces protéines à interagir. Et finalement, nous nous intéresserons aux troncations pathologiques de TDP-43, pour voir l'influence qu'elles peuvent avoir sur les formes pleine longueur.

Au vu de leur implication dans les maladies neurodégénératives, et plus largement du fait que les protéines de liaison à l'ARN jouent un rôle important dans

l'émergence de ces maladies, il est capital de mieux comprendre le comportement de ces protéines et les mécanismes dans lesquels elles sont impliquées.

# Matériel et méthodes

---

Dans cette partie nous allons décrire les différentes techniques utilisées au cours de ce travail. Mais avant de rentrer dans les détails, la première partie vise à comparer les différentes méthodes pour mettre en évidence les interactions protéine-protéine, de façon à mieux introduire la méthode largement utilisée dans ce travail de thèse : le *Microtubule Bench*.

## I. Comparaison des méthodes pour l'identification de l'interaction protéine-protéine et ouverture sur le *microtubule Bench*

Dans ce rapport nous nous sommes focalisés sur l'interaction entre les protéines FUS et TDP-43, nous avons pu voir (**chapitre II.5**) que cette interaction avait été mise en évidence par des techniques diverses. Ici nous allons en décrire certaines.

Dans le but de mieux comprendre les mécanismes biochimiques, cellulaires et plus largement les pathologies liées à certaines protéines, il est important d'identifier les interactions d'une protéine avec son environnement. Identifier les partenaires d'une protéine permet ainsi de mieux comprendre les fonctions auxquelles les protéines participent et leurs implications au sein de la cellule. En effet, les contacts protéine-protéine sont au cœur des processus cellulaires, comme la transduction du signal, ou les fonctions génétiques et de réplication. Mieux connaître ces interactions permet de mieux comprendre le comportement ou la fonction d'une protéine, mais aussi contribue à mieux cerner les mécanismes en jeu dans le cas de pathologies comme le développement de comportements inadaptés comme la formation d'agrégats dans les maladies neurodégénératives. Ainsi, de très nombreux outils pour la détection des interactions protéine-protéine ont été développés ces dernières années.

Il existe des outils *In vitro*, à partir d'extraits cellulaire ou nécessitant des protéines purifiées ou et d'autres *In vivo* qui permettent d'étudier la protéine directement dans son milieu. Dans certains cas, les protéines peuvent être modifiées, en ajoutant des étiquettes fluorescentes par exemple. Chaque méthode possède des avantages et des inconvénients, les critères les plus importants sont la sensibilité, la spécificité. Il existe aussi un ensemble de techniques qui ont été développées pour prédire les interactions de protéine par simulation computationnelle, ou *In silico*,

cependant ici nous nous focaliserons sur les méthodes les plus couramment utilisées en biophysique et en biochimie, ainsi qu'une nouvelle méthode, le *Microtubule Bench* qui a été utilisée au cours de ce travail de thèse.

## 1. Méthodes conventionnelles

### a) Méthode *In vitro*

Ces méthodes nécessitent d'utiliser des extraits cellulaires ou de produire la ou les protéines d'intérêt.

#### ▪ **Co-immunoprécipitation**

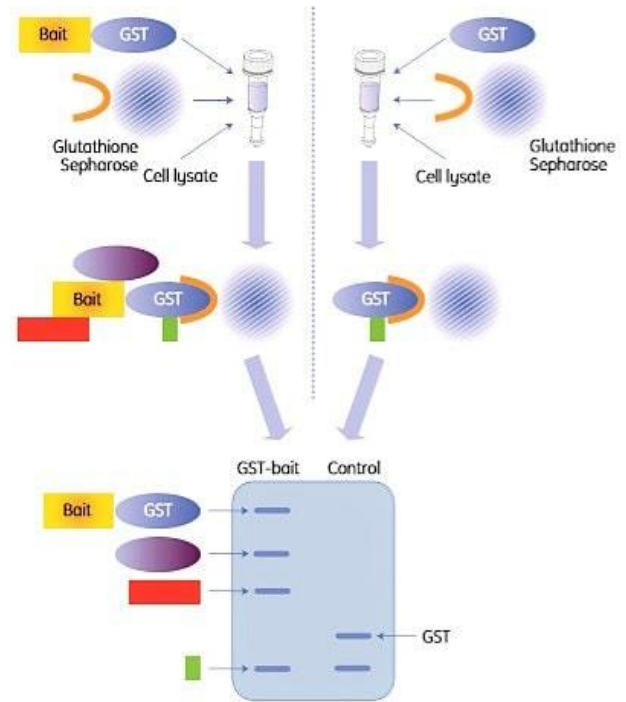
La co-immunoprécipitation est une technique très répandue qui permet d'identifier les partenaires d'interaction d'une protéine en particulier. Pour cela un anticorps est utilisé pour cibler la protéine d'intérêt (appât) ce qui permet ensuite d'extraire le complexe auquel elle appartient. Le complexe formé est capturé ou précipité sur un support sur lequel est immobilisée une protéine de liaison à l'anticorps. Par la suite, il est possible d'identifier les protéines (proies) faisant partie de ce complexe par *Western Blot*. Cette technique demande l'utilisation de différents anticorps et une répétition de la manipulation de manière à bien identifier les protéines cibles. Elle peut aussi révéler les interactions indirectes, par le biais de l'ARN par exemple.

L'avantage principale de cette technique est qu'elle permet d'identifier un grand nombre d'interactions, elle n'a pas besoin d'équipement complexe et de nombreux systèmes sont commercialisés, elle peut donc être assez facile à mettre en œuvre à condition que les anticorps soient disponibles, le principal inconvénient de cette technique est qu'elle requiert d'avoir accès à des anticorps spécifiques (qui sont assez coûteux) et que selon les anticorps les résultats peuvent varier[363].

Il existe des variantes de cette technique par le biais de billes super-paramagnétiques qui permettent de faciliter la mise en œuvre de la manipulation.

### ▪ Essai de *Pull Down*

La technique du *pull-down* est assez similaire à la co-immunoprécipitation et permet de détecter les interactions d'une protéine avec plusieurs partenaires en se basant sur l'affinité d'un ligand pour piéger les protéines partenaires, cependant, dans ce cas la protéine d'intérêt est étiquetée à la place d'utiliser des anticorps immobilisés. La protéine d'intérêt est immobilisée dans une colonne grâce à un ligand spécifique à l'étiquette qu'on lui a ajouté (c'est l'appât, par exemple une protéine conjuguée à la GST ou le His<sub>6</sub>), ensuite elle est mise en contact avec des extraits (lysate cellulaire, protéines purifiées...) contenant les potentielles protéines partenaires qui vont venir se complexer (proies), ensuite différentes étapes de lavages sont effectuées, avant de procéder à une élution pour analyse par électrophorèse et *western blot* [364].



**Figure 27-Principe de la technique du *pull-down*.**

Une protéine appât est conjuguée à la GST puis mise en contact avec un extrait cellulaire pour piéger ses protéines partenaires. L'élution du complexe dans une colonne où est fixé le Glutathione, puis l'analyse sur SDS-PAGE permet d'identifier les protéines partenaires. **D'après** <https://www.sigmaaldrich.com/FR/fr/technical-documents/protocol/protein-biology/enzyme->

Le *pull down* est une technique intéressante pour détecter un grand nombre de partenaires. Cette technique peut également détecter des protéines de faible abondance, mais peut aussi détecter des interactions indirectes et la présence d'étiquette peut influencer le résultat. De plus, il y a une compétition avec les complexes endogène. Cette technique peut aussi être couplée à la spectrométrie de masse pour une meilleure identification des partenaires[365].

### ▪ TAP : Tandem affinity purification ou purification d'affinité en tandem

Tout comme les techniques précédentes, la TAP permet d'identifier les partenaires d'interaction d'une protéine d'intérêt, en effectuant un criblage à haut débit. C'est un dérivé du pull-down, en deux étapes de purification pour améliorer la spécificité. La protéine doit être produite avec une étiquette (La protéine A associée au peptide de liaison à la calmoduline, CBP, entre lesquels on a introduit un groupement clivable TEV, *tobacco etch virus*), puis une purification du complexe est effectuée, grâce à deux chromatographies successives : la première joue sur

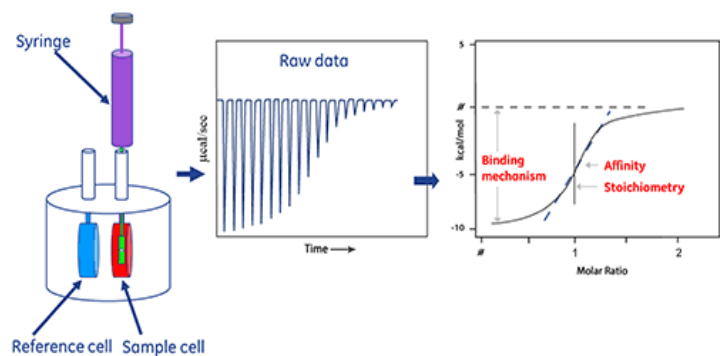
l'absorption de la protéine A sur les IgG immobilisés, tandis que la deuxième utilise l'affinité de la CBP avec la calmoduline après clivage. L'extrait est ensuite analysé par spectrométrie de masse[366].

Les inconvénients de cette technique sont que l'étiquette peut interférer dans l'interaction de la protéine avec certains partenaires, et elle manque aussi de spécificité et peut aussi conduire à de faux positifs mais de manière réduite par rapport au *pull down* classique[367].

- **ITC : *isothermal titration calorimetry* ou titration calorimétrique isotherme**

La titration calorimétrique isotherme est une méthode très quantitative qui

permet de mesurer les propriétés thermodynamiques, en particulier les constantes d'affinité ( $K_D$ ) des protéines ciblées. Elle est basée sur la mesure de la chaleur libérée ou absorbée lors de la liaison ou la dissociation des deux protéines. La mesure se fait dans un microcalorimètre qui détecte les différences de température par rapport à une référence lors de l'ajout du ligand dans une solution contenant la protéine à analyser.

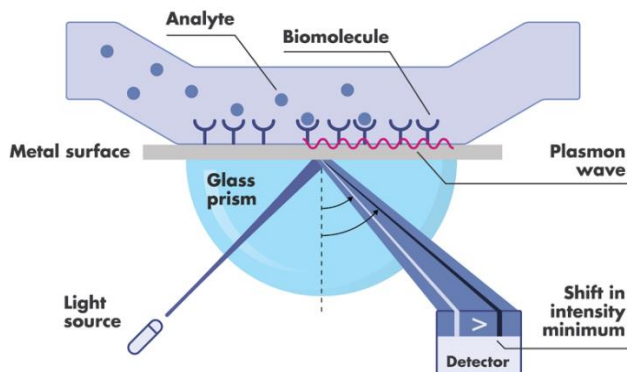


**Figure 28-Principe de l'ITC.**

Les échantillons sont chargés dans un microcalorimètre et l'échange de chaleur est mesuré à l'ajout d'un ligand de manière à mesurer une affinité de liaison. **D'après** <https://www.malvernpanalytical.com/fr/products/technology/microcalorimetry/isothermal-titration-calorimetry>

L'avantage de cette technique est qu'elle est non destructive et ne nécessite pas de marquage, ni d'étiquetage des protéines, elle est reproductible et très précise malheureusement, elle ne fonctionne pas systématiquement, notamment quand l'enthalpie d'interaction est très faible et pour les processus assez lents. Enfin, elle requiert la production d'une quantité importante de protéine[368].

- **SPR : Surface plasmon resonance ou résonance plasmonique de surface**



**Figure 29-Principe de la SPR.**

La protéine est immobilisée sur une surface métallique sur laquelle circule un plasmon. La mesure de l'indice de réfraction de la lumière permet de quantifier l'interaction de la protéine avec un partenaire. **D'après** <https://www.reactionbiology.com/services/biophysical-assays/surface-plasmon-resonance-spr>

solution contenant ses partenaires potentiels d'interaction[369, 370].

C'est une technique qui demande du matériel spécifique mais qui a l'avantage de donner des informations sur les constantes d'association d'un couple donné et permet des analyses cinétiques sans nécessiter de protéines étiquetées[371]. Néanmoins, l'immobilisation de la protéine peut affecter sa conformation et donc affecter l'interaction. Comparer à l'ITC, la SPR permet de mesurer des affinités de liaison plus élevées (mais de manière moins rigoureuse) tout en demandant une moins grande quantité de protéine[368].

- **Méthode structurale : La RMN et la cristallographie des rayons X**

La cristallographie aux rayons X permet d'accéder à la structure d'une protéine cristallisée avec une très haute résolution, au niveau atomique, grâce à la mesure de la diffraction des rayons X, ce qui permet de visualiser l'interaction d'une protéine avec des molécules et ses changements de conformation[372].

La résonance magnétique nucléaire, RMN, basé sur la mesure de la relaxation de noyaux atomique soumis à un champ magnétique, permet la détection des interactions protéine-protéine par le biais de différentes techniques comme, entre autre, l'étude de la perturbation du déplacement chimique (CSP), ou l'effet de relaxation dipôle-dipôle (NOE/NOESY), qui comparent l'évolution d'un spectre d'une protéine seule ou en mélange[373]. De plus, la RMN est un outil particulièrement intéressant pour étudier la structure des protéines, et peut-être utilisée pour l'étude

La SPR permet de déterminer les constantes d'association/dissociation d'une protéine et de son ligand en mesurant l'indice de réfraction de la lumière réfléchiée par une surface métallique dans laquelle des électrons circulent librement (plasmon). La variation de l'indice de réfraction étant proportionnel à l'ajout de matière sur la surface, on peut déterminer l'interaction d'une protéine immobilisée sur la surface, mise en contact avec une



des protéines contenant des IDR. Elle permet ainsi de définir les régions qui peuvent adopter des structures transitoires, et en particulier de mettre en évidence le processus de repliement suite à l'interaction (*folding upon binding*)[374].

Ces deux techniques ont l'avantage majeur de ne pas seulement détecter une interaction, mais surtout de donner des informations sur la structure, et ainsi de mieux comprendre l'interaction en jeu en identifiant les acides aminés impliqués, ainsi que les changements de conformation des protéines étudiées. La RMN a aussi l'avantage d'étudier les molécules en solution, ce qui permet de caractériser et d'obtenir plus d'information sur l'interaction avec le partenaire, comme le site de liaison, l'évaluation du  $K_d$ , ou encore l'aspect cinétique de l'interaction.

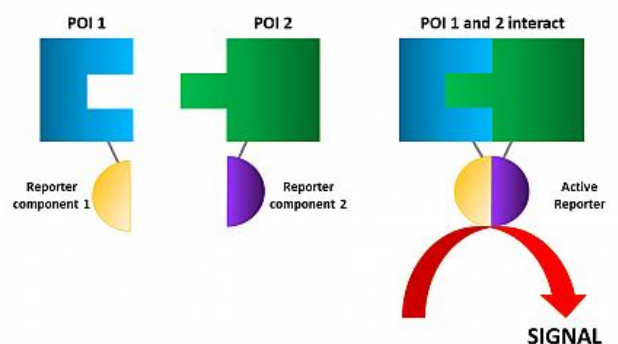
Cependant, elles sont souvent compliquées à mettre en place car elles demandent des échantillons de très haute qualité. Le processus de cristallisation est délicat, et demande des conditions particulières. En particulier, la cristallographie nécessite de travailler avec des protéines structurées, ce qui n'est pas toujours le cas, avec les protéines contenant des LCD. Pour la RMN, il est nécessaire d'effectuer un marquage isotopique de la protéine à l'azote  $^{15}\text{N}$  ou au carbone  $^{13}\text{C}$ .

### b) Méthode *In vivo* ou *In situ*

Ces méthodes permettent d'étudier la protéine dans son milieu comme au sein de cellules, mais elles n'excluent pas forcément de modifier la protéine (ajout d'une étiquette fluorescente par exemple), ce qui peut jouer sur le comportement de la protéine.

#### ▪ Essai de complémentation des fragments protéiques

Les protéines dont on veut étudier l'interaction sont conjuguées à une protéine rapportrice (comme la luciférase ou la GFP, par exemple) qui a été clivée, elles portent donc chacune une partie de cette protéine qui n'est pas fonctionnelle en tant que telle, mais lorsque les deux fragments sont suffisamment proches, dans le cas où les deux protéines d'intérêt interagissent, ils se rassemblent et transmettent un signal, souvent de fluorescence ou de la luminescence, qui peut être détecté.



**Figure 30-Principe de l'essai de complémentation des fragments protéiques**

Une protéine d'intérêt (POI1) est conjuguée à un fragment d'une protéine rapportrice. De même pour POI2. L'interaction des deux protéines partenaires induit la complémentation des fragments de la protéine rapportrice et donc la production d'un signal.

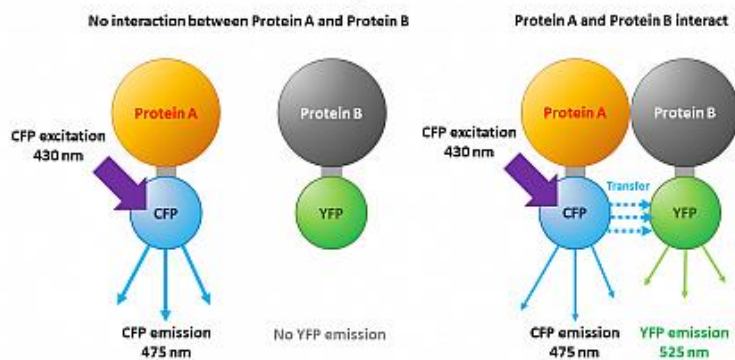
**D'après** <https://www.berthold.com/fr-fr/bio-analyses/applications/ipp/methodes-detude-des-interactions-protéine-protéine/>

Ce système est dynamique mais ne permet d'étudier l'interaction qu'entre deux partenaires d'un couple choisi. De plus, il nécessite de créer des protéines hybrides.

- **FRET : *Fluorescence Resonance Energy transfer* ou Transfert d'énergie entre molécules fluorescentes**

La technique FRET est une technique assez sensible qui permet d'analyser l'interaction entre deux protéines fusionnées à des groupements fluorescents. Lorsque les deux protéines sont très proches (distance inférieure à 10 nm) l'énergie absorbée lors de l'excitation du fluorophore porté par une protéine (le donneur, par exemple le CFP, *cyan fluorescent protéine*) est en partie transmise à l'autre (l'accepteur, la YFP, *yellow fluorescent protein*) ce qui induit une émission lumineuse à une longueur d'onde plus élevée qui peut être détectée au microscope de fluorescence[375].

La FRET est un système dynamique qui a l'avantage de pouvoir être utilisée dans des cellules vivantes, et qui permet ainsi d'avoir des informations sur la localisation, cependant elle offre un ratio signal sur bruit assez faible et peut-être délicate à utiliser notamment en association avec d'autres marquages.



**Figure 31-Principe de la technique FRET.**

La protéine d'intérêt A est conjuguée à la CFP, tandis que la protéine B est conjuguée à la YFP. Si les deux protéines sont assez proches, il y a transmission d'une partie de l'énergie de la CFP à la YFP qui provoque l'émission d'un signal de fluorescence à la longueur d'onde correspondante. **D'après**

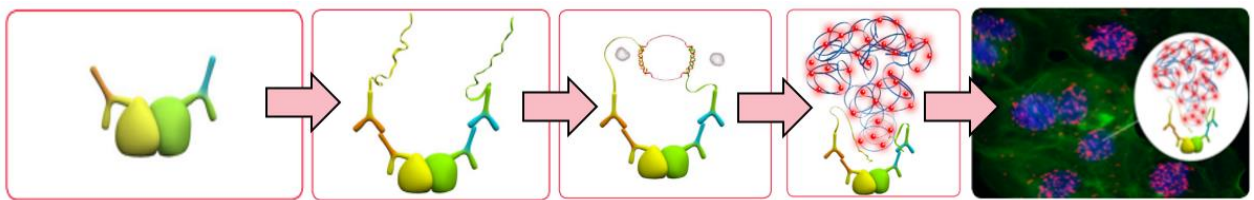
<https://www.berthold.com/en/bioanalytic/knowledge/glossary/fret/>

d'augmenter la résolution spatiotemporelle et offre une plus grande fiabilité mais demande un matériel plus sophistiqué. Cette technique peut aussi être utilisée en mesurant le photoblanchiment de l'accepteur plutôt que l'excitation[376].

Il existe des variantes à cette méthode comme la TR-FRET qui permet de réduire le bruit de fond et donc un gain de sensibilité en utilisant des fluorophores à plus longue durée de vie pour augmenter la résolution de fluorescence dans le temps, ou encore la FILM FRET qui mesure la durée de vie de fluorescence plutôt que l'intensité, permettant

- **PLA : *Proximity ligation assay*, essai de ligature de proximité**

La technique de PLA pour *Proximity ligation assay* est une technique *in situ* très sensible pour la détection d'interaction entre deux protéines. Elle est basée sur le principe de l'immunofluorescence en utilisant des anticorps primaires ciblant les protéines d'intérêt, puis des anticorps secondaires portant des fragments d'oligonucléotides, lorsque les fragments se trouvent à proximité (les deux protéines doivent se trouver à une distance inférieure à 40 nm[377]), ils peuvent être liés par le biais d'une enzyme, la ligase, puis amplifiés, par la polymérase pour augmenter le signal de fluorescence. La polymérase effectue plusieurs centaines de cycles pour étendre le fragment d'ADN (*Rolling circle amplification*) où viennent se conjuguer des oligonucléotides complémentaires portant des anticorps, permettant l'observation d'un signal ponctiforme au microscope optique à fluorescence[377].



**Figure 32-Principe de la PLA.**

Deux anticorps primaires spécifiques reconnaissent les protéines d'intérêts qui interagissent. Deux anticorps secondaires portant des oligonucléotides sont ensuite utilisés. Les séquences d'oligonucléotides sont ligaturées puis amplifiées ce qui permet d'obtenir un signal de fluorescence d'une grande sensibilité. **D'après** <https://www.sigmaldrich.com/FR/fr/products/protein-biology/duolink-proximity-ligation-assay>

La PLA est très sensible et peut détecter des concentrations de protéine de l'ordre du femtomole ( $10^{-15}$ ) et de manière très spécifique. C'est aussi une technique simple à mettre en œuvre à partir du moment où il existe des anticorps spécifiques aux protéines d'intérêt. Elle permet aussi d'avoir accès à la localisation des protéines d'intérêt dans les cellules ou plus largement un tissu. Néanmoins, la fixation des cellules supprime le côté dynamique du processus d'interaction[378].

- **Y2H : *Yeast two-hybrid*, Essai à deux hybrides chez la levure**

La technique de l'essai à deux hybrides est basée sur la modification génétique de levures. Les protéines d'intérêts sont conjuguées à des séquences d'ADN spécifiques, la protéine appât est fusionnée au domaine de fixation à l'ADN d'un facteur de transcription, tandis que la protéine proie est fusionnée au domaine d'activation de la machinerie basale de transcription. De cette façon, l'interaction des deux protéines d'intérêts permet l'activation des facteurs de transcription, et donc la croissance des levures ou l'émergence d'un phénotype particulier[379].

Cette technique a l'avantage de pouvoir être utilisée pour faire du criblage haut débit d'interaction. Elle est assez sensible mais peut conduire à de nombreux faux positifs ou négatifs, donc il est nécessaire de faire des contrôles et des validations avec des techniques complémentaires.

- **Autres techniques**

De nombreuses autres techniques existent comme l'utilisation de bactériophage (*phase display*), diverses méthodes de réticulation des complexes, mais aussi toutes les techniques de simulation *In silico* ou de dynamique moléculaire (on peut citer, par exemple, la base de donnée AlphaFold<sup>3</sup> basée sur une intelligence artificielle qui permet de prédire la structure tridimensionnelle d'une protéine à partir de sa séquence en acides aminés, et donc de mieux comprendre ses interactions) qui ont aussi leurs intérêts, avantages et inconvénients mais n'étant pas abordées dans ce rapport nous n'entrerons pas les détails.

---

<sup>3</sup> Disponible ici : <https://alphafold.com/>

## 2. Une nouvelle méthode : *le microtubule Bench*

### a) Description de la technique

Le *microtubule Bench* est une technique assez récente développée au laboratoire SABNP qui utilise les microtubules comme support pour analyser l'interaction protéine-protéine d'un couple de protéine d'intérêt. C'est donc une technique *in situ*, qui permet d'obtenir un bon signal de fluorescence en concentrant les protéines sur les microtubules. Le principe consiste à fusionner la protéine d'intérêt (Appât) avec un MBD (*Microtubule Binding Domain*) et une protéine rapportrice (GFP ou RFP), la protéine partenaire (proie) est, elle aussi, conjuguée à une protéine fluorescente. De cette façon, un signal de fluorescence correspondant à la proie sur les microtubules indique une interaction entre les deux protéines[380]. L'intérêt d'utiliser les microtubules est qu'ils offrent une grande surface de contact et donc la possibilité à un nombre important de protéine de s'y fixer, mais aussi que leur structure fibrillaire permet de distinguer clairement le signal par rapport au fond. Cette technique permet aussi de mettre en évidence les interactions indirectes par le biais de l'ARN par exemple.

Deux options sont possibles, soit la protéine proie peut être conjuguée, uniquement à une protéine fluorescente, soit elle peut être aussi conjuguée au MBD en plus de la protéine fluorescente. Dans le premier cas, cela permet de mettre en évidence l'interaction entre la proie et l'appât de par la présence de la proie sur les microtubules, dans le second cas, il y a une notion de miscibilité qui s'ajoute car les signaux de fluorescence peuvent être plus ou moins homogènes selon l'affinité des deux protéines. Dans les deux cas, le niveau interaction (ou la miscibilité) peut être quantifié en mesurant un coefficient de corrélation.

Le laboratoire SABNP, sous tutelle de l'INSERM, travaille en partenariat avec une start-up hébergée par l'université d'Evry, Synsight. Une plateforme de **criblage à haut débit automatisée**, basée sur la méthode du *microtubule Bench* a récemment été mise en place pour obtenir des résultats robustes et des images à très haute résolution. Cette approche permet aussi l'étude de molécules thérapeutiques.

La technique du *microtubule Bench* fait l'objet de plusieurs brevets.

#### Boîte de texte 9-Contexte pour la technique du *microtubule Bench*

<https://synsight.net/approach/>

Cette technique fait partie des principaux outils utilisés pour réaliser ce travail de thèse au cours duquel nous avons utilisé les deux variantes de la méthode.

### **b) Avantages et inconvénients**

Le *microtubule Bench* est une méthode sensible, spécifique et quantitative qui permet de détecter les interactions *in situ* en conservant un processus dynamique, car bien que liées aux microtubules, les protéines restent mobiles. Elle a aussi l'avantage d'utiliser un simple microscope optique à fluorescence, et peut-être automatisée pour du criblage à haut débit. Cependant, elle demande du travail de préparation de plasmides, mais est très efficace pour cribler des interactions une fois mise en place. La présence du MBD affecte la localisation de la protéine étudiée.

Cette méthode, qui présente l'intérêt de s'effectuer en contexte physiologique, est pertinente dans le cas de l'étude de protéines de liaison à l'ARN pour son côté dynamique où les interactions indirectes entrent en considération, d'autant plus que la propension à l'agrégation de ces protéines est un facteur limitant à leur étude *in vitro*. Elle permet aussi d'étudier la miscibilité de ces protéines entre elles comme nous le verrons par la suite.

### 3. Comparaison des méthodes

La compréhension d'une interaction entre deux protéines partenaires nécessite fréquemment l'utilisation de plusieurs techniques complémentaires. Certaines techniques sont pertinentes pour détecter une interaction, quand d'autres permettent d'obtenir plus de détails sur celle-ci. La combinaison de plusieurs techniques permet ainsi de valider l'interaction, et de mieux identifier ces implications. Ce tableau (**voir tableau 1**) résume et compare les différentes techniques évoquées précédemment.

Méthodes principales		Criblage/Auto-matisation	Sensibilité	Spécificité	Principaux avantages	Principaux inconvénients
<i>In vitro</i>	<b>CoIP</b>	Identification d'un nombre important de partenaires	Variable	Faible (risque de faux positif)	Simple et bien connue. Ne nécessite pas de clonage ou de production de protéine.	Besoin d'anticorps spécifiques. Détection d'interactions indirectes
	<b>Pull-down TAP</b>	Identification d'un nombre important de partenaires	Variable	Faible (Risque de faux positifs)	Permet de cibler des protéines peu abondantes. Simple et rapide	Nécessite la production et la purification d'étiquette. L'étiquette peut influencer le résultat
	<b>ITC</b>	Non	Elevée	Elevée (pour un couple donné)	Accès aux constantes d'association d'un couple sans immobilisation ou marquage. Très reproductible	Besoin d'une quantité importante de protéines. Difficulté pour les enthalpies d'interactions faibles. Difficile avec les IDR
	<b>RMN, DRX</b>	Non	Faible	Elevée	Identification de la structure et du site d'interaction	Equipement sophistiqué, cristaux de protéines difficiles à obtenir, peu adapté aux IDR
	<b>SPR</b>	Non	Très sensible mais variable selon le support et la longueur d'onde	Elevée (pour un couple donné)	Pas besoin de marquage, système flexible et analyse rapide en temps réel, accès à la cinétique	L'immobilisation de la protéine peut affecter sa conformation. Nécessite du matériel spécifique

Méthodes principales		Criblage/Automatisation	Sensibilité	Spécificité	Principaux avantages	Principaux inconvénients
<i>In vivo</i>	<b>PCA</b>	Haut débit	Bonne	Risque de recombinaison non-spécifique	Utilisable dans divers types cellulaires, permet de localiser l'interaction dans la cellule	Parfois peu sensible selon les protéines étudiées. La complémentation peut induire un biais. Protéines étiquetées.
	<b>Y2H</b>	Haut débit	Importante	Faible (Risque de faux positifs)	Automatisation pour criblage à haut débit	Risque de faux positifs et faux négatifs important. Noyau de la levure différent du contexte cellulaire. Protéines étiquetées
	<b>FRET</b>	Non	Plutôt haute (Variable selon la méthode)	Elevée (pour un couple donné)	Informations spatio-temporelles, notamment la localisation au sein de la cellule. Possibilité d'ajouter un marquage supplémentaire	Nécessite un équipement spécifique, distance/orientation du donneur et de l'accepteur peuvent influencer le résultat. Protéines étiquetées
	<b>PLA</b>	Non	Très sensible (signal amplifié)	Elevée (pour un couple donné)	Bon signal et facilité de mise en œuvre, identification de la localisation de l'interaction	Cellules fixées, besoin d'anticorps spécifiques et sondes onéreuses. Détection d'une co-localisation
	<b>MtB</b>	Haut débit, haut contenu (En développement)	Elevé (signal concentré sur les microtubules)	Elevée (pour un couple donné)	Bon signal de fluorescence, visualisation directe de l'interaction. Simple microscope optique à fluorescence. Possibilité de quantifier l'interaction	Détection d'une co-localisation. Utilisation des microtubules. Protéines étiquetées. Nécessite le clonage et la production de plasmide

**Tableau 1-Comparaison de méthodes de détection des interactions protéine-protéine**

Sensibilité : capacité à détecter une interaction faible, distinction d'un faible signal par rapport au bruit de fond. Spécificité : Les interactions détectées sont vraies, le risque de faux positif est faible. Criblage haut débit/haut contenu : test d'un nombre important de partenaires ou molécules/détection d'un nombre important de partenaires d'interaction ou d'effets [371, 378, 381–385].



## II. Méthodes utilisées au cours de ce travail

Dans cette partie nous allons décrire les techniques utilisées au cours de ce travail. Ce paragraphe se décompose en deux parties, l'une décrivant les expérimentations avec un modèle cellulaire, en particulier le *microtubule bench* dans ces différentes variantes, et l'autre décrivant les manipulations *In vitro*. Dans chaque cas, des plasmides spécifiques ont été produits, l'un pour l'expression cellulaire, l'autre pour l'expression bactérienne pour la production de protéines.

### 1. Expérimentations avec modèle cellulaire

#### **Préparation des plasmides pour le *microtubule bench* méthode proie/appât**

Dans cette version de *Microtubule Bench* une première protéine est amenée le long des microtubules, par exemple, TDP-RFP-MBD, l'autre protéine reste libre, FUS-GFP. On peut donc mettre en évidence une interaction dans un contexte cellulaire. En effet, si les deux protéines interagissent, directement ou indirectement la protéine libre co-localisera avec celle qui est fixée sur les microtubules, montrant un signal jaune sur les microtubules, dans le cas contraire, on aura des signaux distincts avec un réseau de microtubules bien marqué dans le rouge et un fond diffus pour le vert.

Des plasmides codants pour la séquence d'intérêt des protéines PL fusionnée à la RFP et au domaine de liaison aux microtubules de Tau (MBD) ont été produits comme décrit précédemment (Boca *et al.*, 2015[380]; Maucuer *et al.*, 2018[386]). Les séquences des protéines FUS et TDP-43 ont été amplifiées par le biais d'oligonucléotides contenant les sites de restriction Pac1 et Asc1. Après purification du fragment, la ligation est effectuée dans le plasmide intermédiaire RFP-MBD-pCR8/GW/TOPO au niveau des sites de restriction Pac1 et Asc1. Enfin, une recombinaison est effectuée (Gateway® LR Clonase® II Enzyme mix, INVITROGEN, cat n°11791020) pour obtenir la séquence souhaitée dans un plasmide d'expression de mammifère pEF-Dest51 (Invitrogen™). Les protéines proies fusionnées à la GFP ont été obtenues comme décrit précédemment (Boca *et al.*, 2015[380]), en insérant les séquences d'ADN codant les protéines d'intérêt dans un vecteur pEGFP-N1 (Clontech). Un contrôle par séquençage de l'ADN a été effectué pour chaque construction.

Désignation	Séquence	Etiquette	Oligonucléotides ou références
<b>FUS-RFP-MBD</b>	FL (1-526)	RFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>TDP43-RFP-MBD</b>	FL (1-414)	RFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>FUS-GFP</b>	FL (1-526)	GFP	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>TDP43-GFP</b>	FL (1-414)	GFP	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>YB1-GFP</b>	FL (1-324)	GFP	Boca, <i>et al.</i> Sci Rep. 2015 doi :10.1038/srep17304
<b>Lin28-GFP</b>	FL (1-209)	GFP	Boca, <i>et al.</i> Sci Rep. 2015 doi :10.1038/srep17304
<b>HuR-GFP</b>	FL (1-326)	GFP	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>G3BP1-GFP</b>	FL (1-466)	GFP	Boca, <i>et al.</i> Sci Rep. 2015 doi :10.1038/srep17304
<b>Larp6-GFP</b>	FL (1-491)	GFP	Boca, <i>et al.</i> Sci Rep. 2015 doi :10.1038/srep17304
<b>U2AF-GFP</b>	FL (1-475)	GFP	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>Sam68-GFP</b>	FL (1-443)	GFP	Pankivskyi, <i>et al.</i> Cell Mol Life Sci. 2021. Doi :10.1007/s00018-020-03610-y

Tableau 2-Plasmides produits et utilisés pour le *microtubule bench* méthode proie/appât

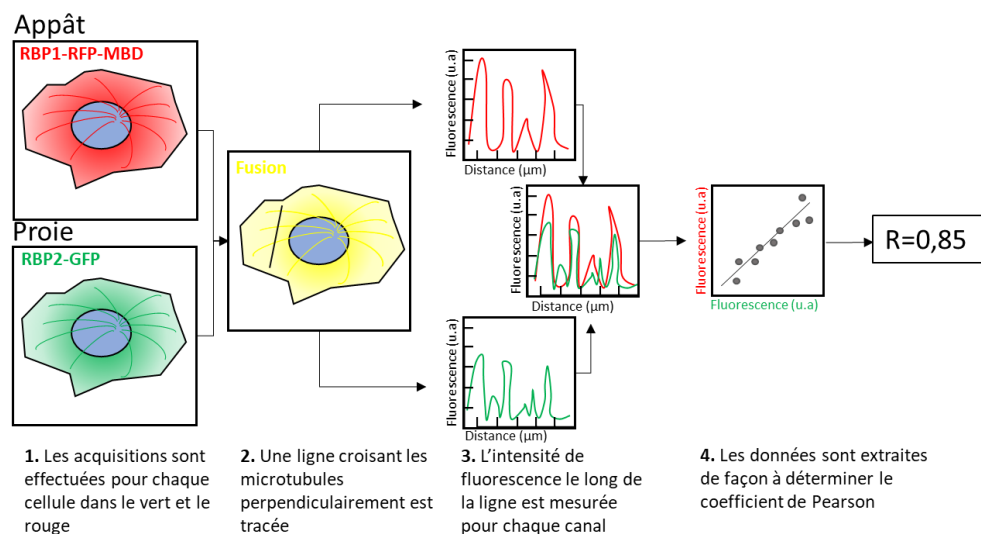
### Culture Cellulaire, transfection et fixation pour le *microtubule bench*

Des cellules U2OS, à la morphologie particulièrement adaptée pour la visualisation des microtubules, sont cultivées dans du milieu DMEM (*Dulbecco's Modified Eagle's Medium*, riche en glucose, Sigma) contenant du sérum bovin (FBS) à 10 % et des antibiotiques pénicilline et streptomycine 100 µg/mL (GIBCO Life Technologies). Une fois la confluence atteinte les cellules sont transférées dans des boîtes 4 puits pour transfection. Les cellules sont transfectées avec 500 ng de plasmides et un ratio 1 :1 pour les couples RBP1-RFP-MBD/RBP2-GFP (FUS-RFP-MBD ou TDP-RFP-MBD avec HuR-GFP, G3BP1-GFP, YB1-GFP, Lin28-GFP, LARP6-GFP, SAM68-GFP, ou U2AF-GFP) par ajout de *lipofectamine 2000™ reagent (Invitrogen)* dans le milieu de culture. Un changement de milieu est effectué après 4 h pour éliminer la lipofectamine puis les cellules sont incubées toute la nuit à 37 °C avec 5 % CO<sub>2</sub>. Les cellules sont ensuite lavées au PBS et fixées dans un premier temps avec une solution de méthanol à -20 °C, 20 min, puis une solution à 4 % de paraformaldéhyde

dans du PBS pendant 30 min à 37 °C. Cette méthode de fixation améliore grandement la visualisation de la structure des microtubules. Les cellules sont finalement montées sur des lames pour la visualisation en microscopie de fluorescence.

### **Analyse de corrélation, système proie/appât**

Les images des cellules transfectées ont été acquises avec un microscope inversé (Axiovert 220, Carl Zeiss 5 MicroImaging, Inc, Hamamatsu C10600, Axxio Vision software) équipé d'un objectif immergé dans l'huile 63x/1.4. Les temps d'exposition ont été ajustés dans les deux couleurs pour obtenir des niveaux d'intensité comparable. Les acquisitions sont ensuite traitées avec le logiciel ImageJ, elles sont rognées, et les niveaux d'intensité et de contraste sont ajustés dans chaque canal pour faire ressortir les microtubules. La méthode suivie pour la mesure de la corrélation des signaux a été décrite précédemment[387] et est résumée sur la **figure 33**. Dans un premier temps, il s'agit de tracer un trait (environ 20 µm) perpendiculairement aux microtubules et tangent au noyau (pour éviter les biais liés à la décroissance de la fluorescence quand on s'éloigne du noyau). Le profil de l'intensité de la fluorescence suivant ce trait est relevé dans chacun des canaux. Les données sont extraites puis de façon à comparer la répartition de la fluorescence le long du trait pour chaque couleur, un coefficient de Pearson est calculé avec le logiciel Excel (fonction COEFFICIENT.CORRELATION). Un minimum de 3 traits par cellule et de 10 cellules ont été analysées pour chaque condition.



#### **Calcul du coefficient de corrélation pour la méthode Proie/Appât**

**Figure 33-Etapes pour le calcul du coefficient de corrélation pour la détection d'interactions avec le microtubule Bench système proie/appât**

### Hybridation *In situ* pour marquage de l'ARN

Pour certaines manipulations, l'ARN (polyA) a été mis en évidence au sein des cellules par hybridation *in Situ*. Pour cela les cellules ont été incubées après fixation avec une sonde d'oligonucléotides poly-T couplée à une séquence dig (polyT-dig Sigma HA09131354-004) pendant 2 h à 37 °C dans un tampon d'hybridation (0,005 % BSA, 1mg/mL d'ARN de levure, 10 % sulfate de dextran, 25 % formamide dans du SSC 2x). Des lavages ont ensuite été effectués avec du tampon SSC 4X puis 2X (1,75 % NaCl, 0,88 % citrate de sodium, pH 7,0). Un anticorps primaire anti-dig *chicken* dilué dans le tampon SSC 2X contenant 0.1 % Triton x-100 a ensuite été utilisé suivi d'un anticorps secondaire  $\alpha$ -*chicken* 650 (Dye light 650 Goat pAB chk IgY ab96954 Abcam). Les anticorps primaires comme secondaires ont été incubés pendant 2 h à la dilution 1 :1000. La visualisation des noyaux se fait à l'aide du DAPI (Sigma-Aldrich) après une incubation des lamelles durant 10 s dans une solution à 0,66 mg/mL.

### Préparation des plasmides pour le *microtubule bench* méthode de compartimentation

Les plasmides HuR-RFP-MBD, Sam68-RFP-MBD, FUS-GFP-MBD, FUS-RFP-MBD et TDP-43-GFP-MBD ont été obtenus comme décrit précédemment pour la technique proie/appât (**voir tableau 3**). Concernant les formes tronquées, différents oligonucléotides (**voir tableau 4**), contenant les sites de restrictions Pac1 et ASC1 ont été utilisés pour isoler et amplifier la séquence tronquée de la protéine souhaitée. La séquence est ensuite insérée dans un plasmide RFP-MBD-pCR8/GW/TOPO ou GFP-MBD-pCR8/GW/TOPO suivi d'une étape de recombinaison (Gateway® LR Clonase® II Enzyme mix, INVITROGEN, cat n°11791020) dans le plasmide final d'expression cellulaire pEF-Dest51 (Invitrogen™). De manière à s'assurer de l'obtention de la bonne séquence, à chaque étape des contrôles sur gel d'agarose ont été effectués et le plasmide final a été séquencé.

Désignation	Séquence	Etiquette	Oligonucléotides ou référence
<b>FUS-GFP-MBD</b>	FL (1-526)	GFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>FUS-RFP-MBD</b>	FL (1-526)	RFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>TDP43-GFP-MBD</b>	FL (1-414)	GFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>TDP43-RFP-MBD</b>	FL (1-414)	RFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>Sam68-RFP-MBD</b>	FL (1-443)	RFP-MBD	Pankivskyi, <i>et al.</i> Cell Mol Life Sci. 2021. Doi :10.1007/s00018-020-03610-y
<b>HuR-RFP-MBD</b>	FL (1-326)	RFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692

**Tableau 3-Plasmides utilisés pour la technique du *microtubule bench*, méthode de compartimentation**

Désignation	Séquence	Etiquette	Oligonucléotides ou référence
<b>TDP RRM-GFP-MBD</b>	98-264	GFP-MBD	GCGTTAATTAAGCCACCATGGCAGTCCAGAAAACA TCCGAT/GCGGGCGCGCCTGGTGCTTAGGT TCGGCATTGGA
<b>TDP Nter-GFP-MBD</b>	1-102	GFP-MBD	GCGTTAATTAAGCCACCATGTCTGAATATATTCGGG TAACCG/GCGGGCGCGCCTGTGTTTTCTGGACTGCT CTTTTAC
<b>FUS RRM-GFP-MBD</b>	284-380	GFP-MBD	GCGTTAATTAATGAATTCAGACAACAACACCATC /GCGGGCGCGCCTTGCGAGTAGCAAATGAGACCTT G
<b>FUS PrLD-GFP-MBD</b>	1-164	GFP-MBD	GCGTTAATTAAGCCATGGCCTCAAACGATTATACCC /GCGGGCGCGCCatACTGCTGCTGTTGACTGGTT
<b>FUS dLCD</b>	277-525	RFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>TDP dPrLD-RFP-MBD</b>	1-277	RFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>TDP25-RFP-MBD</b>	208-414	RFP-MBD	GCGTTAATTAATGCGGGAGTTCTTCTCTCAGTACG GGG/ GCGGGCGCGCCATCATTCCCCAGCCAGAAGACTTA GAA
<b>TDP35-RFP-MBD</b>	102-414	RFP-MBD	GCGTTAATTAATGACGGAATATGAAACACAAGTG AAAG/GCGGGCGCGCCATCATTCCCCAGCCAGAAG ACTTAGAA
<b>TDP PrLD</b>	270-414	RFP-MBD	Maucuer, <i>et al.</i> J Cell Sci. 2018, doi :10.1242/jcs.214692
<b>mCherry-MBD</b>	FL	MBD	Boca M, <i>et al.</i> Sci Rep. 2015 doi:10.1038/srep17304

**Tableau 4-Plasmides produits pour les formes tronquées pour la technique du *microtubule bench*, méthode de compartimentation**

### Transfection des cellules et acquisition pour la méthode de compartimentation

Les cellules U2OS ont été préparées, transfectées et fixées comme décrit précédemment pour la méthode proie/appât. Pour l'étude de la miscibilité des protéines en fonction de leur concentration, différents ratios de plasmides ont été utilisés. Dans ce cas les cellules sont transfectées avec des quantités variables de plasmides, par exemple pour un ratio TDP-RFP-MBD/FUS-GFP-MBD 1/10 on mettra 500 ng de plasmide TDP-RFP-MBD pour 5 µg de plasmide FUS-GFP-MBD. Pour éviter tous biais liés aux différences d'intensité de fluorescence lors de l'analyse, les paramètres sont réajustés pour ramener les intensités à des niveaux comparables dans le vert et le rouge. Les ratios réels sont calculés en prenant en compte les paramètres ajustés (temps d'exposition [Expo], intensités mesurées [I] et quantité de plasmides transfectés [plasmide]) suivant la formule :

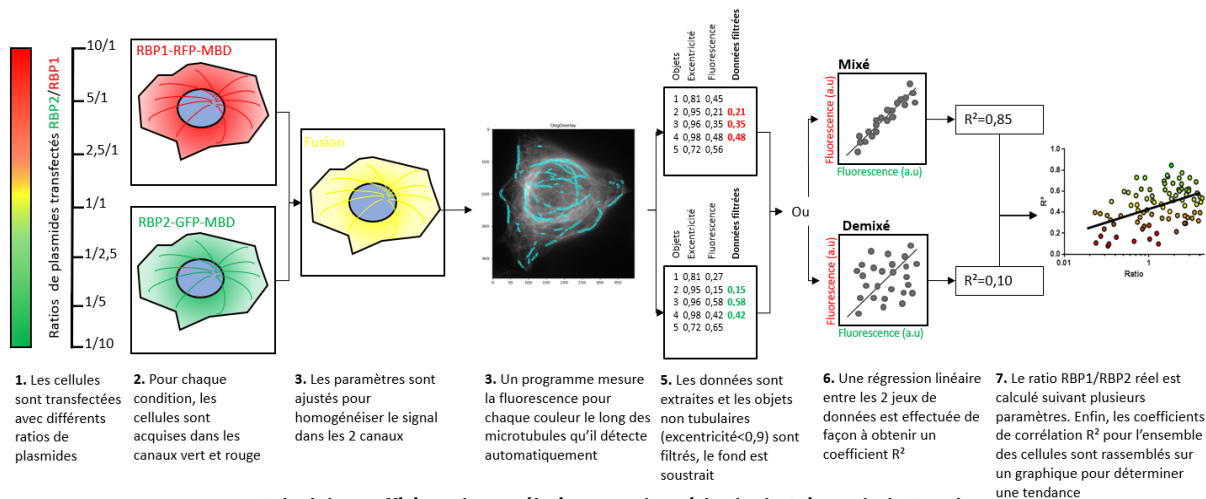
$$\text{Ratio corrigé} = \text{Ecart} \left( \frac{I_{\text{vert}}}{I_{\text{rouge}}} \right) \times \text{Ecart} \left( \frac{\text{Expo}_{\text{vert}}}{\text{Expo}_{\text{rouge}}} \right) \times \frac{\text{Plasmide}(\text{RBP-GFP-MBD})}{\text{Plasmide}(\text{RBP-RFP-MBD})},$$

$$\text{où Ecart} = \left( \frac{I_{\text{vert}}}{I_{\text{rouge}}} \right)_i / \text{moyenne} \left( \frac{I_{\text{vert}}}{I_{\text{rouge}}} \right).$$

Pour les formes tronquées les cellules ont été transfectées avec un ratio 1 :1.

### Analyse des images, détection de la compartimentation, détermination des coefficients de corrélation

Les images ont été acquises de la même manière que précédemment, ici c'est principalement la méthode d'analyse qui diffère. Les cellules sont isolées une à une grâce au logiciel ImageJ et les paramètres d'intensité sont réajustés. Les images ainsi obtenues sont passées dans un programme spécifiquement mis en place pour l'analyse des corrélations (développé avec le logiciel CellProfiler) résumé sur la **figure 34**. Brièvement, l'algorithme applique un filtre sur les images, il identifie les cellules par leur noyau, puis il repère les objets tubulaires (de 5 à 50 pixels) au-dessus d'un certain seuil au sein de ces cellules. Il mesure ensuite les intensités de fluorescence dans chaque canal au sein de ces objets ainsi que des paramètres caractéristiques (formes, longueurs...). Les données générées sont ensuite retraitées pour supprimer les objets non tubulaires (excentricité <0,9) et soustraire le fond. Pour chaque cellule, un coefficient  $R^2$  est mesuré à partir de régression linéaire (Microsoft Excel) de l'intensité verte en fonction de l'intensité rouge. Les valeurs de corrélation sont ensuite rassemblées sur un même graphique.



**Figure 34-Etapes pour le calcul du coefficient de corrélation pour la méthode du *microtubule Bench***

### **Essai de ligature de proximité (PLA)**

Le kit *Duolink PLA technology kit* (Sigma Aldrich) a été utilisé pour les expériences de PLA. Pour cela des cellules Hela sont préparées la veille et mises en plaque 96 puits à la concentration de 20 000 cellules par puits. Elles sont lavées au PBS puis incubées avec du PFA (Paraformaldéhyde) 20 min à 37 °C pour fixation. Elles sont ensuite rincées au PBS avant d'être incubées dans un tampon de blocage (3 % BSA, 0,1 % triton X-100 dans PBS) pendant 60 min à 37 °C. Les anticorps primaires  $\alpha$ -FUS *rabbit* et  $\alpha$ -TDP-43 *mouse* sont incubés pendant 90 min à température ambiante dans le tampon de blocage en chambre humide. En parallèle, les contrôles négatifs sont incubés sans anticorps ou avec un seul des deux anticorps dans la solution de blocage. Dans une expérience indépendante des anticorps  $\alpha$ -U2AF (U2AF<sup>35</sup> rabbit polyclonal Ab, A303-667A, Bethyl, et U2AF<sup>65</sup> mouse monoclonal Ab, clone MC3) ont aussi été utilisés en contrôle. Les sondes de PLA  $\alpha$ -*rabbit* et  $\alpha$ -*mouse* (*Minus et Plus*) sont diluées à 1 :5 dans le tampon approprié fournit dans le kit (Duolink PLA, Sigma Aldrich) puis incubés pendant 1 h à 37 °C avec les échantillons. Les cellules sont ensuite lavées plusieurs fois avec un tampon de lavage (10 mM tris, 150 mM NaCl, 0,05 % Tween). On procède ensuite à l'étape de ligature en incubant la ligase à 1 :40 dans le tampon de ligation 30 min à 37 °C. Les échantillons sont ensuite lavés avec le tampon de lavage avant de procéder à la dernière étape d'amplification. La polymérase est diluée à 1 :80 dans le tampon d'amplification avant d'être déposée sur les échantillons pour incubation 100 min à 37 °C. Les cellules sont lavées avec un second tampon de lavage 1X (200 mM Tris, 100 mM NaCl) puis une dernière fois avec un tampon 0,1X. Elles sont incubées 2 min avec le DAPI, puis stockées dans du PBS en plaque à 4 °C en attendant la visualisation. L'acquisition est faite au confocal avec

l'appareil Opera Phenix Plus avec un objectif 40x immergé dans l'eau, et le logiciel Harmony. Les données sont traitées à l'aide des logiciels ImageJ et CellProfiler.

## 2. Expérimentations *In vitro*

Pour la production *In vitro* de protéines recombinantes utilisées pour les manipulations de biochimie et la technique de visualisation des agrégats par immunofluorescence, différents plasmides ont été produits pour les protéines PL ou les formes tronquées (**voir tableau 5**). Les formes tronquées ont été produites à partir de la séquence pleine longueur des protéines grâce à des oligonucléotides spécifiques (**voir tableau 6**). Pour cela, un plasmide intermédiaire pMS2 a été utilisé pour ajouter le tag HA à la séquence (TDP RRM, TDP Nter, TDP25, FUS PrLD). Un premier clonage a été effectué pour isoler la séquence d'intérêt entre les sites Nhe1 et BamH1, puis après purification le fragment a été ligaturé dans le plasmide pMS2 contenant le tag HA en position N-terminale. Un deuxième clonage a été effectué pour récupérer la séquence d'intérêt associée au tag HA avant d'effectuer la ligature dans le plasmide d'expression bactérienne pET22b (Novagen) entre les sites de restriction HindIII et Xho1. Le plasmide pET22b contient notamment un gène de résistance à l'ampicilline, une séquence Lac pour l'induction de la production de la protéine d'intérêt à l'IPTG et l'étiquette 6 histidines (His<sub>6</sub>) pour la purification de protéine. Certains fragments (TDP35, TDP PrLD) ont été clonés avec la séquence correspondant au tag Myc directement incluse dans les oligonucléotides. Les étiquettes HA ou Myc ont été ajoutées à l'ensemble des formes tronquées de manière à pouvoir les marquer par des anticorps spécifiques et donc les visualiser par immunofluorescence.

Dénomination	Etiquette	Séquence	Poids moléculaire (kDa)
<b>TDP-43</b>	<b>His<sub>6</sub></b>	<b>1-414</b>	<b>44</b>
TDP RRM	His <sub>6</sub> /HA	98-264	20
TDP PrLD	His <sub>6</sub> /Myc	270-414	17
TDP Nter	His <sub>6</sub> /HA	1-102	14
TDP25	His <sub>6</sub> /HA	208-414	25
TDP35	His <sub>6</sub> /Myc	102-414	35
<b>FUS</b>	<b>His<sub>6</sub></b>	<b>1-526</b>	<b>55</b>
FUS PrLD	His <sub>6</sub> /Ha	1-164	20

Tableau 5-Plasmides produits pour la production de protéines *In vitro*



Dénominations	Sites de restriction	Oligonucléotides directes	Oligonucléotides inverses
<b>TDP RRM</b>	Nhe1/BamH1	GCGGCTAGCGGCAGTCCAGAAAACA TCCGA	GCGGGATCCGTGCTTAGGTTCCGGCATTG G
	HindIII/Xho1	GCGAAGCTTATGGGCTACCCCTACGA CGTG	GCGCTCGAGGTGCTTAGGTTCCGGCATTG G
<b>TDP PrLD</b>	HindIII/Xho1	GCGAAGCTTATGGAACAGAACTGA TTAGCGAAGAAGATCTGCAGTTAGA AAGAAGTGAAGATTTGGTGGTAA	GCGCTCGAGCATTCCCCAGCCAGAAGAC TTAGAATCCAT
<b>TDP Nter</b>	Nhe1/BamH1	GCGGCTAGCGATGTCTGAATATATTC GGGTAACC	GCGGGATCCTGTTTTCTGGACTGCTCTTT TCAC
	HindIII/Xho1	GCGAAGCTTATGGGCTACCCCTACGA CGTG	GCGCTCGAGTGTGTTTTCTGGACTGCTCTTT TCAC
<b>TDP25</b>	Nhe1/BamH1	GCGGCTAGCGCGGGAGTTCTTCTCTC AGTA	CGCGGATCCCATTCCCCAGCCAGAAGAC TTAGA
	HindIII/Xho1	GCGAAGCTTATGGGCTACCCCTACGA CGTG	CTCGAGCATTCCCCAGCCAGAAGACTTA GA
<b>TDP35</b>	HindIII/Xho1	GCGAAGCTTATGGAACAGAACTGA TTAGCGAAGAAGAAGATCTGACGGA ATATGAAACACAAGTAAAAG	GCGCTCGAGCATTCCCCAGCCAGAAGAC TTAGAA
<b>FUS PrLD</b>	Nhe1/BamH1	GCGGCTAGCGATGGCGAGCAACGAT TATAC	GCGGGATCCGCTGCTGCTGTATACTGGT
	HindIII/Xho1	GCGAAGCTTATGGGCTACCCCTACGA CGTG	GCGCTCGAGGCTGCTGCTGTATACTGGT

**Tableau 6-Listes des oligonucléotides générés pour le clonage des plasmides pour la production de formes tronquées *In vitro***

### Production et purification des protéines

Les protéines recombinantes (FUS, TDP-43 et formes tronquées) associées à l'étiquette His<sub>6</sub> sont exprimées dans la bactérie *E. Coli* BL21 (DE3). Après transformation par choc thermique, les bactéries sélectionnées sur un support d'agarose-ampicilline portant le plasmide sont mises en culture à 37 °C dans 100 mL de milieu 2YT contenant de l'ampicilline et incubées toute la nuit. Cette pré-culture est ensuite transférée dans un volume de 1 L et mise en incubation à 37 °C jusqu'à ce que la densité optique atteigne 0,6, l'IPTG est ensuite ajouté à la concentration finale de 1 mM pour induire l'expression des protéines. La culture est maintenue sous agitation à la température de 37 °C pendant 4 h avant d'être récoltée par centrifugation. Les culots de centrifugations sont ensuite re-suspendus dans un tampon (Tris 25 mM, Urée 8 M, 2 M NaCl, 5 mM Imidazole, 0,5 mM DTT, pH 7.4) dans lequel est ajouté du PMSF et des inhibiteurs de protéases (*EDTA-free protease inhibitor cocktail Roche*). Les protéines étudiées étant, en effet, très peu solubles avec une forte tendance à l'agrégation, il est important de toujours les stocker dans l'urée à forte concentration. Les cellules bactériennes sont ensuite lysées par sonication sur

glace (Bioblock Vibracell sonicar, modèle 72412). Le surnageant est ensuite récupéré par ultracentrifugation.

Pour la purification, le surnageant précédemment obtenue est chargé dans une colonne contenant une résine de Ni-NTA-agarose (Qiagen) équilibrée avec le même tampon que précédemment. Le tout est incubé 2 h à 4 °C pour permettre aux protéines d'interagir avec la colonne. La colonne est ensuite lavée avec un gradient d'imidazole de 10 mM à 100 mM avant de procéder à l'élution avec des solutions à 250 mM, 500 mM, et 1 M d'imidazole. L'ensemble des lavages et éluions sont ensuite chargés sur un gel SDS-PAGE à 12 % pour identifier les fractions contenant la protéine d'intérêt et contrôler leur pureté. Les fractions ainsi identifiées sont rassemblées et passées dans une colonne de dessalage (PD10 Sephadex G-25M, GE-Healthcare) équilibrée à l'aide d'un tampon (Urée 8 M, Tris 25 mM, NaCl 200 mM, DTT 0,5 mM, pH 7,5) pour enlever le surplus d'imidazole. La solution protéique est finalement concentrée à l'aide de filtre (SpinX<sup>R</sup> Concentrator 5KMCO ou 10KMWCO, Corning) par centrifugations successives, avant d'être congelée par congélation flash dans l'azote, pour être stockée à -80 °C, en attendant les futures expérimentations.

### **Production de l'ARN 2Luc par transcription *In vitro***

Pour les expériences nécessitant de l'ARN, l'ARN 2Luc (environ 3000 bases) a été produit par transcription *In vitro* à partir d'un plasmide linéarisé pSP72-2Luc. Ce plasmide contient les séquences d'ADN codants les luciférases *Renilla reinformis* et *Photinus pyralis* assemblées par un groupement lieur. Le kit *RiboMAX<sup>TM</sup> Large Scale RNA Production systems-SP6 et T7 (Promega)* a été utilisé pour effectuer la transcription *In vitro* en suivant les instructions du fabricant. L'ARN synthétisé a été purifié en utilisant l'extraction au phénol/chloroforme.

### **Gel retard (*gel shift assay*)**

L'ARNm 2Luc est chauffé à 80 °C pendant 1 min puis refroidi à température ambiante avant d'être préparé de façon à atteindre une concentration de 150 à 200 ng/puits en mélange avec les protéines à différents ratios, dans un tampon Hepes (Hepes 20 mM, KCl 25 mM, MgCl<sub>2</sub> 2 mM, DTT 1 mM, pH 7,4). Les mélanges sont incubés 10 à 15 min à température ambiante avant d'être chargés sur gel d'agarose 1 % (BET 5 %, dans TAE 1X). Les gels sont placés dans du tampon TAE 1X pendant 1h à 25 V pour migration des complexes ARN/protéines, ils sont ensuite révélés à l'UV (Bio-Rad Gel Doc<sup>TM</sup> XRT, image Lab Software). La mesure des intensités des bandes a été faite avec le logiciel Image Lab. Les valeurs d'intensité ont été normalisées de manière interne à la ligne de migration (intensité relative de la bande de migration de l'ARN par rapport à l'intensité globale de la ligne, du puits à la

bande), et par rapport à la bande de migration de l'ARN seul, selon la formule suivante :

$$\begin{aligned} & \textit{Intensité normalisée de la bande} \\ & = \frac{\textit{Intensité relative de la bande}(x) - \textit{fond}(x)}{\textit{Intensité relative de la bande de l'ARN seul}(1) - \textit{fond}(1)} \end{aligned}$$

De manière à démontrer la miscibilité de TDP-43 dans la phase de FUS, les échantillons de protéines de FUS, TDP-43 et du mélange des deux ont été incubés en absence d'ARN pendant 50 min à la concentration de 3  $\mu\text{M}$  (TDP-43) ou 9 ; 3 ; 1 : 0,3 ; 0,1  $\mu\text{M}$  (FUS) dans un tampon Hepes. Après avoir été incubé 1 min à 80 °C l'ARNm 2Luc est ajouté aux échantillons à la concentration de 115 ng/puits (correspondant à des ratios protéine/nucléotides de 1/10 pour TDP-43, ou 1 :3 ; 1 :9 ; 1 :30 ; 1 :100 ; 1 :300 dans le cas de FUS). Dans le cas du mélange les ratios correspondent à la somme des concentrations de TDP-43 (3  $\mu\text{M}$ ) avec différentes conditions de concentrations pour FUS (9 ; 3 ; 1 : 0,3 ; 0,1  $\mu\text{M}$  soit 1 :2,3 ; 1 :5 ; 1 :8 ; 1 :9 ; 1 :9,7 protéine : nucléotides). Après 10 min d'incubation les échantillons sont chargés sur un gel d'agarose (BET 5 %, TAE 1X). Trois gels indépendants ont été effectués. La quantification de la quantité d'ARN en interaction avec les RBP a été effectuée en mesurant l'intensité de la bande correspondant à l'ARN libre pour chaque condition et normalisée par rapport à la bande de l'ARN libre (puits 1) en suivant la formule :

$$\begin{aligned} & \textit{ARN pris en charge (normalisé)} \\ & = 1 - \frac{\textit{Intensité relative de la bande}(x) - \textit{fond}(x)}{\textit{Intensité relative de la bande de l'ARN seul}(1) - \textit{fond}(1)} \end{aligned}$$

Dans le cas de la somme (bandes oranges sur le graphique **figure 56, D**), elle a été calculée en additionnant la valeur correspondant à l'ARN pris en charge dans le cas de la condition 2 (TDP-43 seule) avec les valeurs des conditions 3 ; 4 ; 5 ; 6 et 7 (FUS seule), de manière à comparer ces valeurs avec les conditions du mélange.

### **SDS-PAGE et western Blot**

Les protéines sont diluées à la concentration de 10  $\mu\text{g}/\mu\text{L}$ , dans un tampon (Tris 20 mM, KCl 20 mM,  $\text{MgCl}_2$  2 mM, DTT 1 mM, pH 7,5) avant d'être chargées dans un gel de polyacrylamide 15 %. La migration est effectuée pendant 90 min à 140 V sur glace. Les protéines sont ensuite transférées sur une membrane de nitrocellulose à 4 °C à 25 V toute la nuit. La membrane est ensuite colorée au rouge ponceau puis décolorée dans de l'acide acétique 1 % avant d'être incubée dans un tampon de blocage TBS-Tween (Tris 20 mM, 150 mM NaCl, pH7.6, 1 % Tween20) lait 5 % pendant 1 h. Les anticorps primaires sont dilués au 1 :5000 dans une solution de TBS-

Tween lait 4 % avant d'être incubés 2 h au contact de la membrane. Enfin, la membrane est incubée 45 min à température ambiante avec les anticorps secondaires dilués au 1:4000 dans du TBS-T lait 2 %. La lecture se fait par fluorescence avec l'appareil Amersham Typhoon bioimager à 680 nm et à 800 nm (anticorps secondaires LI-COR IRDye).

### **Gel de co-sédimentation**

Avant la préparation des échantillons, les tubes ont été prétraités avec le tampon d'incubation (HEPES 20 mM, KCl 25 mM, MgCl<sub>2</sub> 4 mM, DTT 1 mM, pH 7,4) contenant 1 % de lait. Les protéines sont diluées à la concentration de 45 µM pour TDP-43 et 135 µM pour FUS dans un tampon HEPES avec ou sans ARN (quantité correspondant à un ratio d'une protéine pour 10 nucléotides). Quand il y a de l'ARN, il est introduit en premier. Le mélange est ensuite incubé suivant des durées variables entre 5 et 15 min à température ambiante avant d'être centrifugé 4 min à 10 000 g. Le surnageant est transféré dans un nouveau tube tandis que le culot est re-suspendu dans le tampon leammliie (33 % Tris 0,5 M, 33 % Glycérol, 33 % SDS, Bromophénol) puis chauffé à 95 °C pour dénaturation. Chaque échantillon est finalement déposé sur gel d'acrylamide 10 % pour migration à 140 V avant marquage au bleu de coomasie (45 % éthanol, 9,2 % acide acétique, 0,25 % Bleu R250).

Pour les gels de sédimentation TDP-43/ARN ou FUS/ARN le protocole est similaire avec une concentration de protéine fixée à 20 µM et un gradient d'ARN de 1:500 à 1:5. Ces échantillons ont été incubés 15 min avant centrifugation, dénaturation et migration.

### **Visualisation des agrégats par immunofluorescence**

Cette technique a été mise en place dans le but d'étudier le comportement, la morphologie et la mixité d'agrégats de protéines seules ou en mélange avec d'autres protéines ou l'ARN. L'idée étant de fixer les protéines après incubation dans différentes conditions sur des lames de verres pour ensuite les marquer à l'aide d'anticorps et les visualiser en microscopie optique. Cette technique a l'avantage, de ne pas nécessiter de modification des protéines, qui pourrait altérer leur comportement (comme l'ajout de la GFP par exemple) hormis pour les formes tronquées qui portent des tags d'une dizaine d'acide aminés.

Différentes conditions d'incubation ont pu être mises en œuvre. La plupart du temps les protéines ont été mélangées dans un tampon Tris 20 mM, KCl 20 mM, MgCl<sub>2</sub> 2 mM, DTT 1 mM, pH 7,5, et incubées 2 h à 37 °C avant fixation. Des lamelles de verre ont été prétraitées à l'aide d'une solution de poly-lysine à 0,2 mg/mL pendant 15 min, puis par une solution de glutaraldéhyde à 2 % dans l'eau pendant

40 min. Après le temps d'incubation les échantillons sont déposés (20  $\mu$ L) sur les lamelles rincées au PBS. L'échantillon est laissé 5 min sur la lamelle avant ajout d'une nouvelle solution de glutaraldéhyde à 2 %. Après une incubation de 15 min, puis 3 lavages successifs au PBS, une solution diluée de NaBH<sub>4</sub> est ajoutée sur la lamelle de manière à réduire l'auto-fluorescence du glutaraldéhyde. Une fois les échantillons fixés, on peut procéder à un marquage d'immunofluorescence classique. Une première étape de blocage est effectuée avec une solution de BSA 1 %, triton 0,25 % dans du PBS. Les anticorps primaires sont dilués dans le tampon de blocage à 1 :500, avant d'être déposés sur la lamelle pour 2 h d'incubation à 37 °C en chambre humide. Après lavage au PBS, les anticorps secondaires dilués au 1 :1000 dans le tampon de blocage sont déposés sur les lamelles. Après 2 h d'incubation les lamelles sont lavées au PBS, puis montées (DAKO<sup>R</sup> Fluorescent Mounting Medium S3023) sur des lames de verres avant d'être placées à 4 °C. L'observation de lames marquées est effectuée avec un microscope optique (Leica Microsystems DM4B, Hamamatsu C11440 Digital Camera) avec un objectif immergé dans l'huile à un grossissement de 1,6x63x. Les acquisitions sont effectuées avec le logiciel Las X (2.0.0.14332). Les images ont été retraitées avec le logiciel ImageJ. Certaines acquisitions ont été prises avec un microscope confocal (modèle TCS SP8, Leica), grâce à la plateforme d'imagerie ImCy du généthon. Le marquage ARN a été effectué avec une sonde d'hybridation polyT-dig (Sigma HA09131354-004) ajouté dans le tampon au cours de l'étape de blocage. Un anticorps primaire  $\alpha$ -Dig chk, puis un anticorps secondaire  $\alpha$ -chicken 650 ont été ajoutés avec les autres anticorps dans le tampon de blocage au cours des temps d'incubation pour révéler l'ARN.

Anticorps	Espèce	Epitope	Référence
<b><math>\alpha</math>-TDP43</b>	Mouse	AA 1-260 (Nter)	mAB ABnova H00023435-M01
<b><math>\alpha</math>-TDP43</b>	Rabbit	(Cter)	pAB Protein Teck 12892
<b><math>\alpha</math>-FUS</b>	Mouse	YGGQQQSYGQ (PrLD)	Novus CLO 190 mAB-03309
<b><math>\alpha</math>-FUS</b>	Rabbit	AA 1-50 (PrLD)	Novus Bio pAB NB 100-565
<b><math>\alpha</math>-HA</b>	Rabbit	Tag HA (YPYDVPDYA)	H6908 Sigma
<b><math>\alpha</math>-Myc</b>	Mouse	Tag Myc (EQKLISEEDL)	sAB1305535 Sigma (9E10)
<b><math>\alpha</math>-Dig</b>	Chicken	Tag Dig	pAb AB51946 Abcam

Tableau 7-Anticorps utilisés en immunofluorescence

### **Analyse de corrélation au sein des agrégats**

De façon à déterminer la miscibilité des protéines en mélange au sein des agrégats marqués par immunofluorescence, un programme a été développé à l'aide du logiciel CellProfiler. Il permet d'attribuer à chaque image un coefficient de corrélation. Brièvement, il consiste à détecter les agrégats sous forme d'objets distincts du fond par leur intensité dans le rouge et dans le vert d'une taille comprise entre 20 et 2000 pixels. Un calque est défini pour chaque objet de façon à identifier des clusters de 2 à 150 pixels pour chaque couleur à l'intérieur de ces agrégats. Ce qui a pour effet de quadriller les agrégats. De cette façon, pour chaque zone, l'intensité de fluorescence est mesurée dans le rouge comme dans le vert. Les données générées sont ensuite utilisées pour calculer un coefficient de corrélation (Pearson) traduisant la répartition de la fluorescence au sein des agrégats par images analysés. Le coefficient varie entre -1 et 1, reflétant des agrégats distincts pour les valeurs négatives, des agrégats hétérogènes et clustérisés pour les valeurs faiblement positives, et enfin un mélange homogène pour une valeur proche de 1. Entre 8 et 12 images de tailles identiques, mais regroupant un nombre d'agrégats variable selon la condition ont été analysées.

### **AFM : microscope à force atomique**

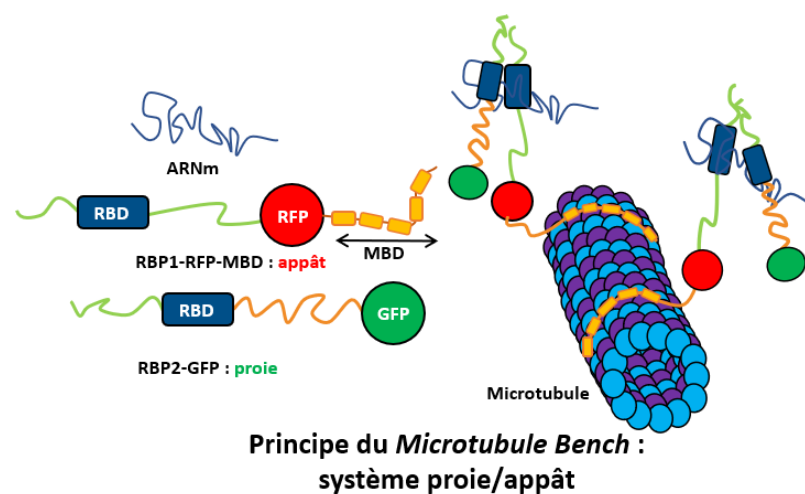
Les protéines pleines longueurs ou les formes tronquées sont incubées seules, avec l'ARN ou en mélange dans un tampon (Tris 10 mM, KCl 15 mM, MgCl<sub>2</sub> 2 mM, DTT 1 mM, Putrescine 10 mM, pH 6,8), selon des durées et des concentrations variables (en générale à 500 nM). Une goutte (10 µL) est ensuite déposée sur une lamelle de mica fraîchement clivée, puis trempée dans une solution diluée d'acétate d'uranyle. La lamelle est séchée aussitôt entre deux papiers filtres avant d'être montée sur un support. L'échantillon est ensuite scanné avec l'AFM en mode *Peakforce Tapping* à l'aide du logiciel Bruker Nanoscope V Multimode 8. Ce mode permet un enregistrement continu des courbes force-distance en utilisant une sonde *ScanAsist-Air*. Les images ont été acquises avec une résolution de 1536 points par lignes à une fréquence de balayage de 1,2 Hz. L'ensemble des images sont traitées avec le logiciel Nanoscope Analysis v1.8 pour mesurer l'aire, la hauteur et le nombre d'objets à l'aide de la fonction *particules analysis*.

# Résultats

## I. Etude du comportement des formes pleine longueur

### 1. Détection de l'interaction entre TDP-43 et FUS et comparaison avec d'autres protéines de liaison à l'ARN

La première phase de ce travail consiste à confirmer l'interaction de FUS et TDP-43, et la comparer avec d'autres RBP, avant d'aller plus loin dans la caractérisation de cette interaction. Pour cela, nous avons utilisé un outil original développé au laboratoire, le *microtubule bench (MtB)*. Brièvement, il s'agit d'une technique permettant de visualiser les interactions protéine-protéine au sein des cellules par fluorescence. Pour cela une protéine appât est conjuguée à un domaine de liaison aux microtubules (MBD) ainsi qu'à la protéine fluorescente RFP, tandis que la protéine proie est simplement conjuguée à la GFP. De ce fait, l'interaction entre les deux protéines d'intérêt peut être mise en évidence du fait de la co-localisation des protéines sur les microtubules. S'il n'y a pas d'interaction, la protéine proie ne vient pas sur les microtubules, tandis que les microtubules sont plus ou moins marqués dans le vert en fonction de l'interaction entre les deux protéines. Ce système est dynamique et permet de mettre en évidence des interactions directes comme indirectes, tout en concentrant les protéines sur les microtubules pour obtenir une bonne sensibilité (**voir figure 35**, pour plus de détails se référer à la partie matériel et méthodes).

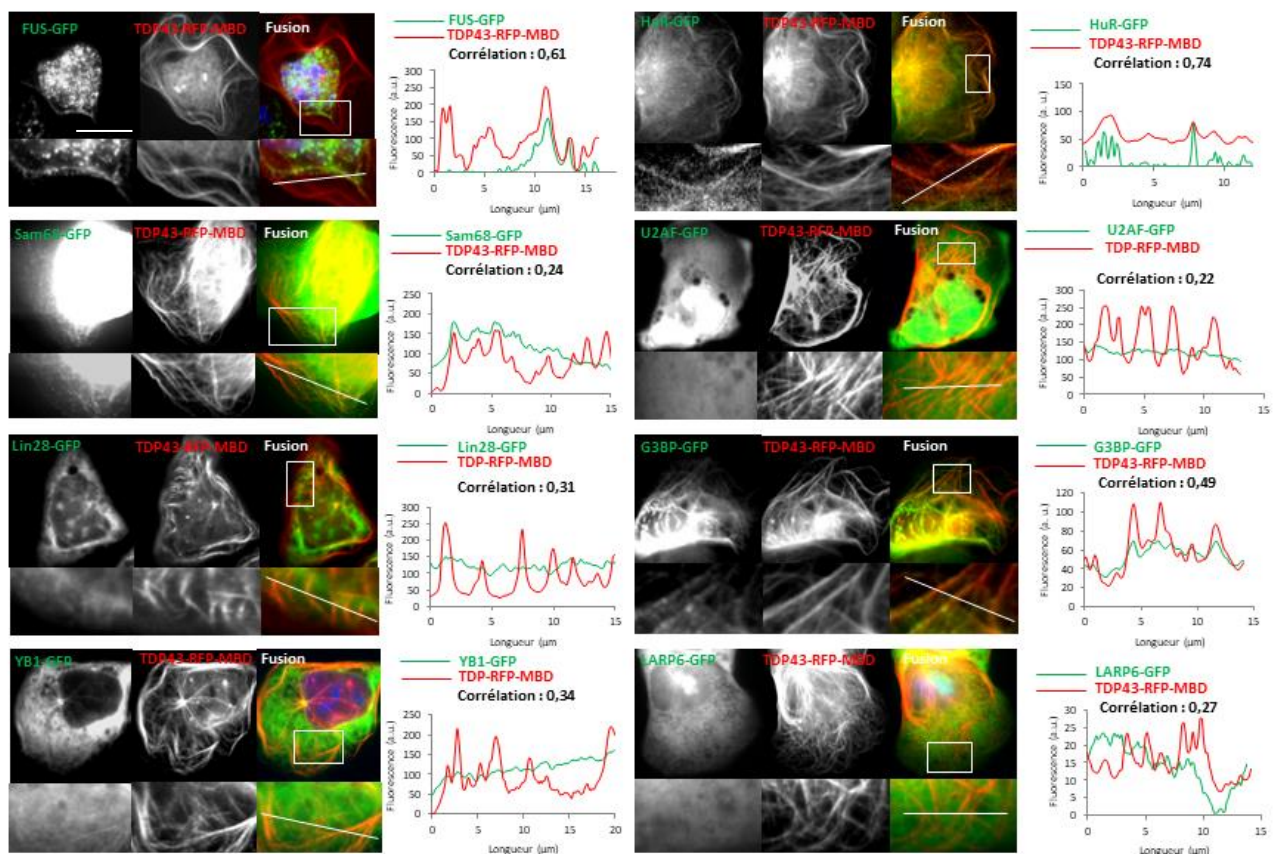


**Figure 35-Principe du *Microtubule Bench*.**

Dans le système proie/appât on souhaite détecter une interaction entre deux protéines, pour cela une protéine appât est conjuguée à un domaine de liaison aux microtubules (MBD) et à la protéine fluorescente rouge (RFP, rond rouge) ce qui la force à venir se lier aux microtubules, tandis qu'une protéine proie est uniquement fusionnée à la protéine fluorescente verte (GFP, rond vert). L'interaction est détectée selon le degré de co-localisation de la protéine proie sur les microtubules.

### a) L'interaction entre TDP-43 et FUS est plus importante qu'avec d'autres RBP

Nous avons donc utilisé ce système de manière à mettre en évidence l'interaction entre FUS et TDP-43 au sein des cellules, mais aussi comparer cette interaction avec diverses protéines de liaison à l'ARN. Différents plasmides, pour les formes pleine longueur (PL) des protéines, ont été produits pour cette expérience. Premièrement, nous avons utilisé TDP-43 en tant qu'appât, exprimé dans un plasmide TDP-RFP-MBD de manière à ce que la protéine se répartisse sur les microtubules. D'un autre côté, les protéines de liaison à l'ARN FUS, HuR, G3BP1, YB-1, Lin28, LARP6, SAM68 et U2AF ont été conjuguées à la GFP. Des cellules U2OS ont été co-transfectées avec un couple de plasmides (TDP-RFP-MBD/RBP2-GFP) pour chaque condition avant d'être fixées puis observées au microscope optique à fluorescence. Dans le canal rouge, on voit très clairement les microtubules (**voir figure 36**), tandis que pour les RBP testées, les microtubules sont plus ou moins distincts dans le vert.

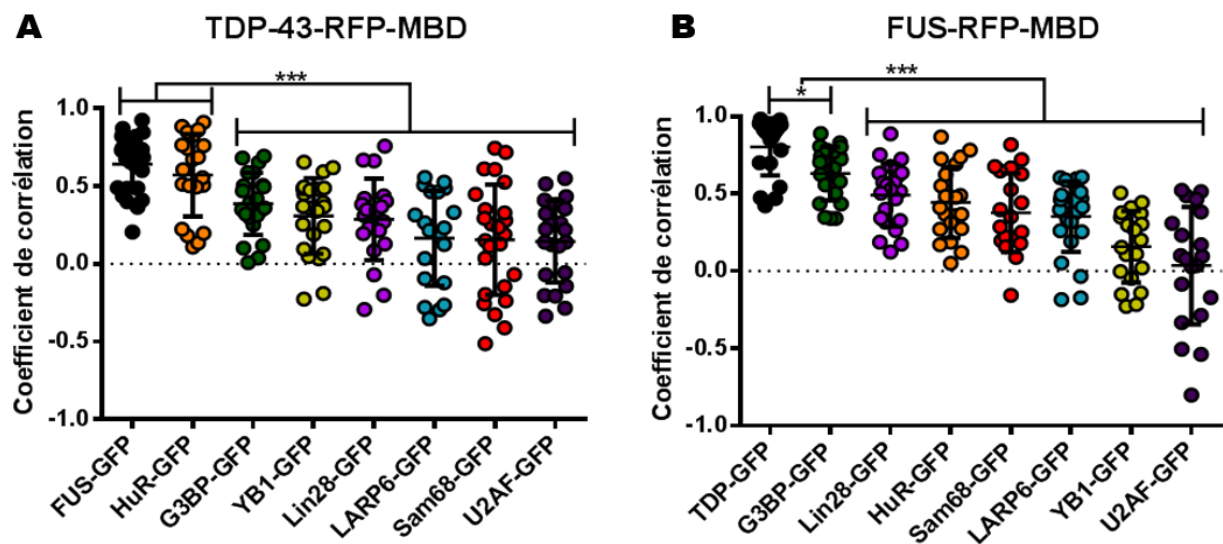


**Figure 36-Comparaison de l'interaction TDP-43-RFP-MBD avec diverses RBP grâce au système du microtubule bench.**

Des cellules U2OS ont été transfectées avec le plasmide TDP-43-RFP-MBD, qui marque les microtubules en rouge, et plusieurs autres RBP (G3BP1, Lin28, Sam68, YB-1, U2AF, HuR, LARP6 et FUS) dont le signal est visible en vert. Pour chaque condition une cellule représentative est présentée. La co-localisation sur les microtubules est plus ou moins importante selon le degré d'interaction entre les protéines, ce qui est mesuré par le biais d'un coefficient de corrélation. Les graphiques montrent la répartition de l'intensité de fluorescence le long d'une ligne coupant les microtubules (un pic correspond à un microtubule) de cette façon le coefficient de corrélation peut être mesuré dans chaque condition. Echelle : 20 μm.



Par exemple, LARP6, SAM68 ou encore U2AF ne sont pas ou très peu présentes sur les microtubules, on distingue principalement un fond continu. Notons au passage, qu'en majorité les RBP sont des protéines principalement nucléaires (en particulier SAM68 et U2AF), il est normal de les retrouver principalement dans le noyau, et donc que ce dernier soit saturé en fluorescence. Dans ce cas il n'y a pas d'interaction. En revanche, dans le cas de FUS, HuR et G3BP1, le réseau de microtubules est marqué, ce qui atteste de la présence de la protéine proie sur les microtubules et donc de l'interaction entre les deux protéines. FUS aussi est une protéine majoritairement nucléaire, pourtant on voit qu'elle interagit avec les microtubules, ce qui montre que le caractère nucléaire d'une protéine a peu d'impact sur la corrélation qui est mesurée. Enfin, les protéines YB-1 et Lin28 sont un peu entre deux et le réseau de microtubules est très faiblement marqué.

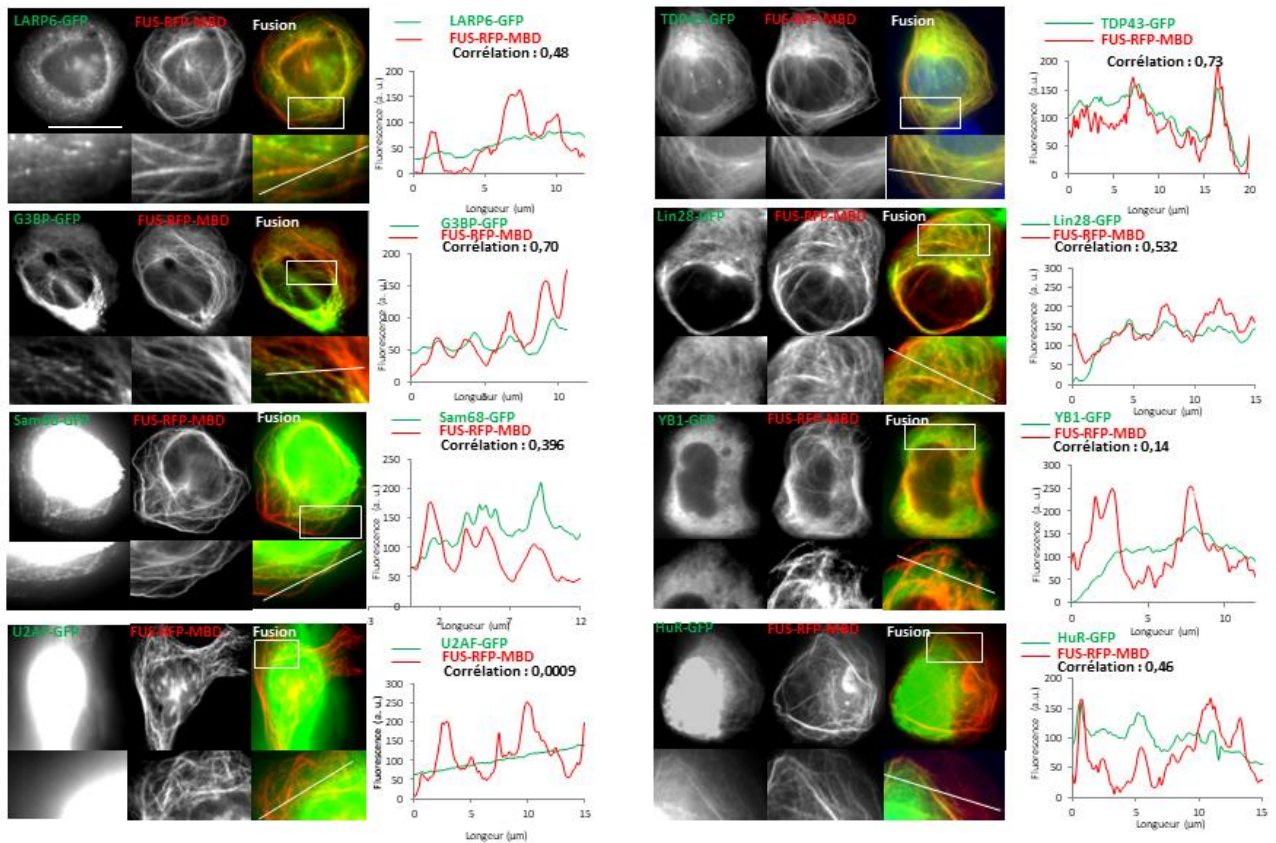


**Figure 37-L'interaction du couple TDP-43 et FUS est plus importante que pour d'autres protéines de liaison à l'ARN.**

Quantification des interactions observées avec la méthode du MtB pour différentes RBP avec TDP-43-RFP-MBD (A) et FUS-RFP-MBD (B). La mesure des profils d'intensité de fluorescence le long d'une ligne pour les deux couleurs (RFP et GFP) permet d'obtenir une valeur de coefficient de corrélation qui reflète l'affinité des protéines testées. Chaque protéine a été associée à une couleur arbitraire pour une meilleure comparaison des deux conditions. Analyse statistique par Test de T réalisée avec GraphPad Prism sur une moyenne de 20 mesures par condition. Moyenne avec écart-type. \* $p < 0,005$  ; \*\*\* $p < 0,0001$ .

De manière à mieux quantifier cette interaction, nous pouvons mesurer un coefficient de corrélation. Pour cela, un trait est tracé perpendiculairement au réseau de microtubules et l'intensité de fluorescence est mesurée le long de ce trait (**voir matériel et méthodes**, pour plus de précisions). Si le réseau de microtubules est bien défini dans les deux couleurs, du fait de la co-localisation des deux protéines, on

obtiendra des pics superposés (**voir figure 36, graphiques**) et donc un bon niveau de corrélation. Au contraire, si la protéine proie n'est pas sur les microtubules, son profil de fluorescence ne correspondra pas à celui de la protéine appât, il en résulte un coefficient de corrélation moindre. Une vingtaine de mesures ont été faites par condition. Lorsque l'on classe l'ensemble de ces mesures obtenues, pour chaque condition sur un même graphique, on peut comparer les niveaux d'interaction de TDP-43 avec les différentes RBP testées (**voir figure 37, A**). On observe donc que les valeurs de coefficients de corrélation sont plus importantes pour FUS et HuR (niveau comparable entre les deux) que pour les autres protéines pour lesquelles l'écart est significatif. Grâce à cette technique, nous avons pu détecter l'interaction entre FUS et TDP-43, la quantifier et vérifier qu'elle était plus importante en comparaison avec d'autres RBP. De manière à confirmer ces premières observations, nous avons réitéré l'expérience avec FUS en tant qu'appât et les même RBP en tant que proies. Ainsi, nous avons suivi la même procédure en effectuant les co-transfections avec les plasmides FUS-RFP-MBD/RBP2-GFP (ou RBP2 inclut l'ensemble des protéines testées précédemment y compris TDP-GFP). Dans ce deuxième cas, nous observons une bonne interaction entre FUS et TDP-43 avec un beau marquage des microtubules pour TDP-GFP et par conséquent un coefficient de corrélation qui s'élève en moyenne à 0,8. Pour les autres conditions, les coefficients sont plus faibles et s'échelonnent de 0,63 pour G3BP1 jusqu'à 0,035 pour U2AF (**voir figure 37, B**). Par conséquent, ici aussi nous avons mis en évidence que l'interaction TDP-43 et FUS était plus forte que celle de FUS avec d'autres protéines. Ainsi, au niveau des images, on voit que G3BP1, Lin28 et HuR marquent aussi les microtubules et dans une moindre mesure Sam68 alors que ce n'est pas le cas pour YB-1, LARP6 et U2AF (**voir figure 38**).



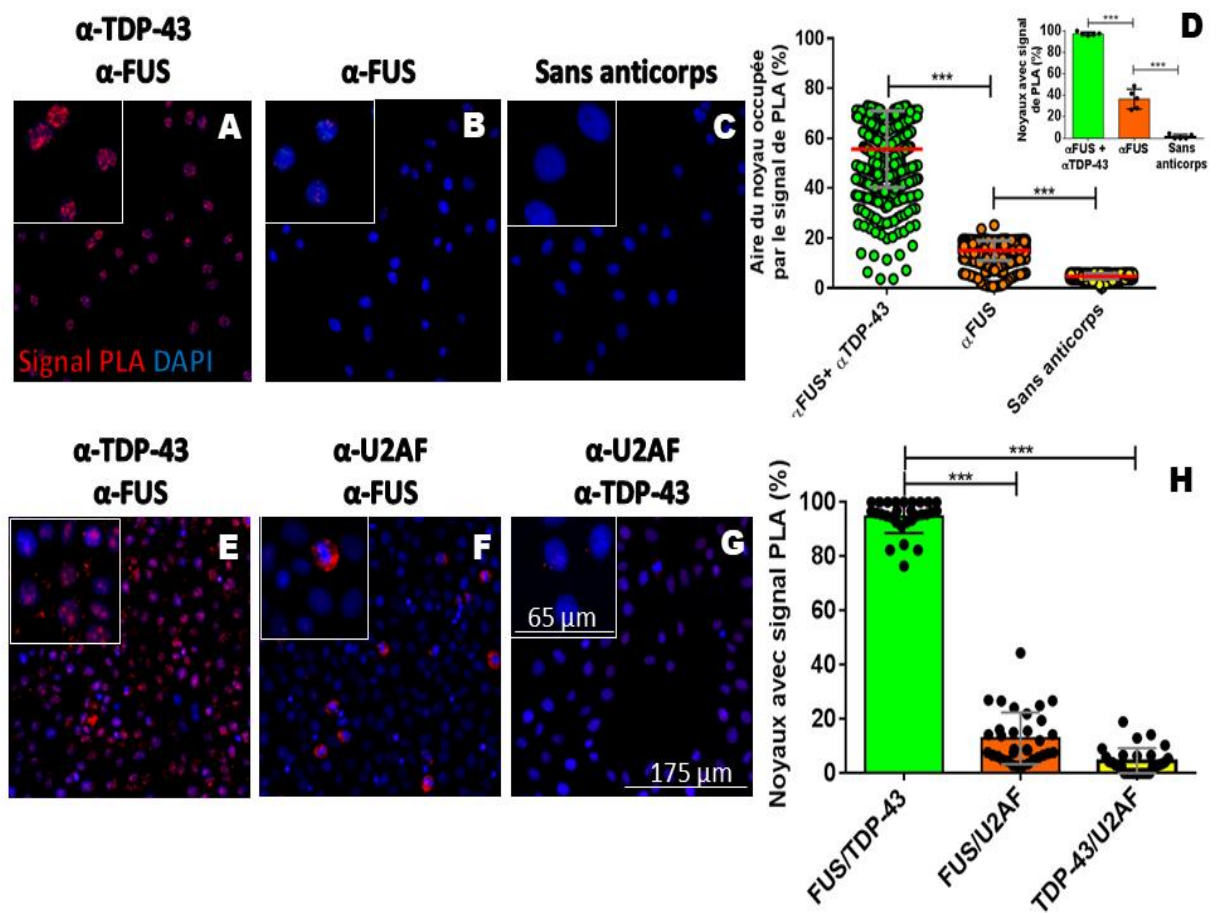
**Figure 38-Comparaison de l'interaction FUS-RFP-MBD avec diverses RBP grâce au système du *Microtubule Bench*.**

Les cellules U2OS ont été transfectées avec le plasmide FUS-RFP-MBD, qui marque les microtubules en rouge, et plusieurs autres RBP (G3BP, Lin28, Sam68, YB-1, U2AF, HuR, LARP6 et TDP-43) dont le signal est visible en vert. Pour chaque condition, une cellule représentative est présentée. La co-localisation sur les microtubules est plus ou moins importante selon le degré d'interaction entre les protéines, ce qui est mesuré par le biais d'un coefficient de corrélation. Les graphiques montrent la répartition de l'intensité de fluorescence le long d'une ligne coupant les microtubules (un pic correspond à un microtubule) de cette façon le coefficient de corrélation peut être mesuré dans chaque condition. Echelle : 20  $\mu\text{m}$ .

D'un point de vue de la littérature, l'interaction entre FUS et HuR, mais aussi G3BP1, a été détectée dans le cadre d'une d'étude de protéomique[388], en revanche ce travail ne liste pas TDP-43 dans les partenaires potentiels de FUS même si nous avons vu que l'interaction entre les deux protéines était confirmée par de nombreux travaux[5, 14, 348]. D'un autre côté, l'interaction de TDP-43 avec HuR (Code protéine : Q15717), G3BP (Q13283), YB-1 (P67809) et surtout FUS (P35637) avait déjà été mise en évidence[244]. De plus, cette même étude ne recense par d'interaction pour les protéines U2AF, Lin28, Sam68 ou encore LARP6. Cette première approche est donc cohérente avec ce qui est décrit dans la littérature.

### b) Confirmation de l'interaction à l'aide de la PLA

De manière à vérifier la véracité de cette interaction et la comparer avec une technique plus largement reconnue que le MtB, nous avons ensuite procédé à une manipulation de PLA (*Proximity ligation assay*). Des cellules HeLa ont été fixées avant d'être marquées par les anticorps  $\alpha$ -TDP-43 (souris) et  $\alpha$ -FUS (lapin), avant de poursuivre avec le protocole d'amplification du signal de PLA. Cette technique a bien confirmé l'interaction entre les deux protéines, en montrant un signal ponctiforme très intense au sein des noyaux (**voir figure 39, A**). D'ailleurs, aucun point n'a été détecté en dehors des noyaux, ce qui atteste de la localisation préférentiellement



**Figure 39-Confirmation de l'interaction FUS/TD-43 par la technique de PLA.**

Des cellules HeLa ont été fixées avant d'être marquées par les anticorps de TDP-43 et FUS (A et E), uniquement l'anticorps de FUS (B), sans anticorps primaire (C), avec le couple d'anticorps U2AF (ms)/FUS (rb) (F), ou U2AF (rb)/TDP-43 (ms) (G). Le protocole de PLA a ensuite été effectué de manière à révéler l'interaction du couple de protéine par fluorescence. Le graphique (D) présente la quantification du signal de PLA dans les cas A, B, C: en pourcentage de l'aire du noyau recouvert par le signal de PLA (graphique principal), et en pourcentage de noyaux arborant un signal de PLA (encart). Le graphique (H) représente le pourcentage de noyaux avec du signal dans le cas dans les conditions E, F et G. Acquisitions au microscope confocal. Echelle: 175  $\mu$ m, Zoom 65  $\mu$ m \*\*\* $p < 0,0001$ ;  $n(\text{FUS} + \text{TDP-43}) = 80$ ,  $n(\text{FUS}) = 1309$ ,  $n(\text{sans anticorps}) = 1228$

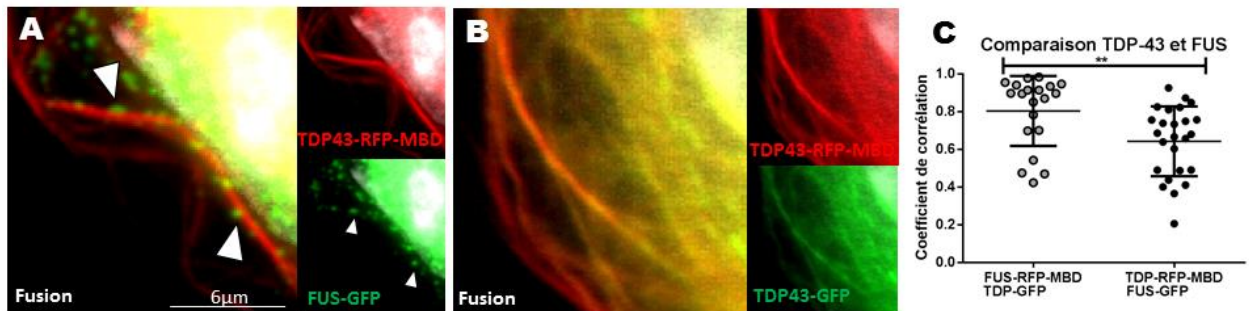
nucléaire de TDP-43 et FUS en conditions physiologiques. Deux contrôles ont été effectués, le premier avec uniquement l'anticorps de FUS, et le second sans anticorps primaire, de façon à vérifier que le signal ne venait pas d'un biais du système. Dans le cas où seul l'anticorps de FUS a été utilisé un léger signal plus épars et d'une faible intensité a été observé dans les noyaux résultant certainement d'une interaction non spécifique des sondes de PLA, tandis que sans anticorps aucun signal n'a été observé. La quantification du nombre de noyaux contenant du signal, et de la surface des noyaux recouverte par le signal de PLA, dans ces trois conditions confirment ces observations (**voir figure 39, D**). On note au passage que la quasi-totalité des noyaux étaient marqués avec en moyenne 60 % de leur surface recouverte.

Une expérience indépendante a été effectuée avec l'anticorps U2AF, qui est la protéine qui interagit le moins bien à la fois avec FUS et TDP-43 au regard des résultats de l'expérience précédente. Dans ce cas, on observe aussi une différence significative entre les marquages pour les couples FUS/TDP-43, avec un marquage très intense, le couple U2AF/FUS, quelques cellules marquées, ou le couple U2AF/TDP-43 pour lequel presque aucun signal n'a été observé (**voir figure 39, E, F, G et H**). L'utilisation de contrôle impliquant une protéine qui interagit peu avec TDP-43 ou FUS permet de confirmer les observations précédentes issues des résultats du MtB.

:

### c) FUS et TDP-43 présentent des comportements différents

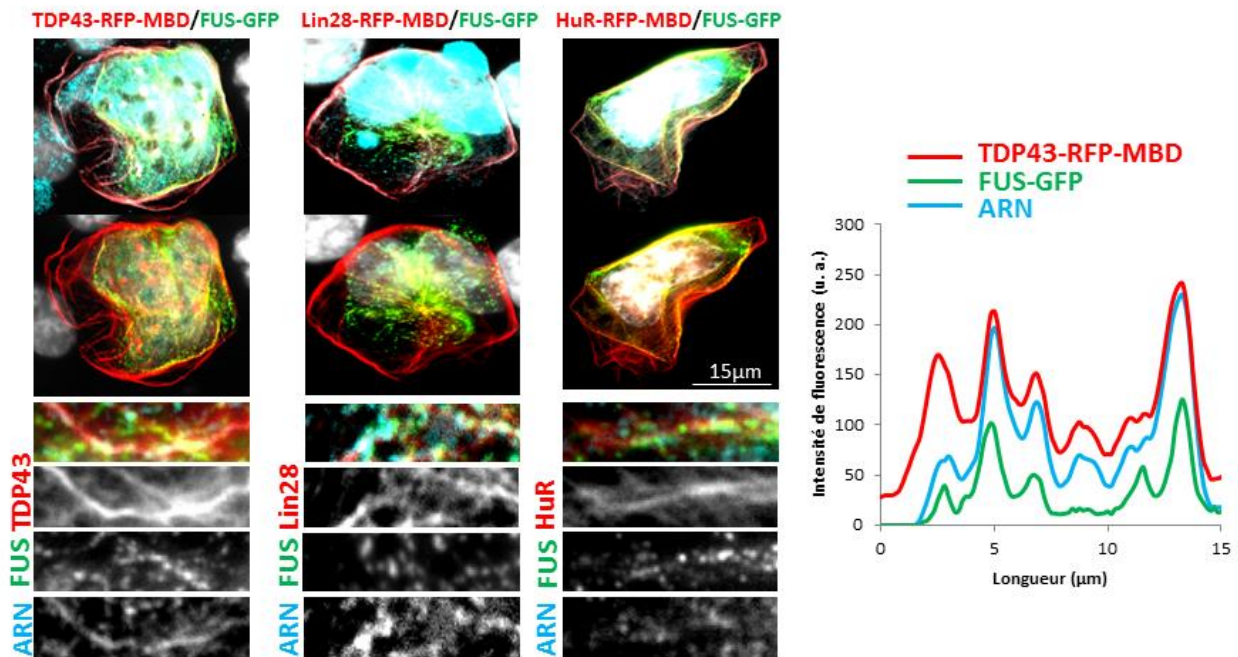
Sur les acquisitions de FUS-GFP/TDP-43-RFP-MBD de la manipulation de MtB précédemment présentée, on remarque que FUS a un comportement particulier, en tant que protéine nucléaire le marquage est très fort dans le noyau, mais dans le cytoplasme elle forme des petits granules (**voir figure 40, A et 41**). On remarque ainsi que ces granules viennent au contact des microtubules mais la protéine ne se disperse pas. Il semble alors qu'elles interagissent avec la protéine présente sur les microtubules sans totalement se mélanger. Cette différence de comportement peut se traduire par une différence dans les niveaux d'interaction des protéines dans les deux configurations. En effet, lorsque l'on compare les valeurs de coefficients de corrélation entre les conditions FUS-RFP-MBD/TDP-GFP et TDP-RFP-MBD/FUS-GFP, on remarque que le premier est plus faible que le second, ce qui laisse penser que le mélange des deux protéines ne se fait pas de manière réciproque (**voir figure 40, C**).



**Figure 40-L'interaction entre FUS et TDP-43 est asymétrique.**

Mise en évidence des comportements différents entre les deux protéines TDP-43 et FUS. FUS forme des granules dans le cytoplasme qui viennent en contact des microtubules (A), tandis que TDP-43 se disperse parfaitement le long des microtubules (B). Les coefficients de corrélation sont, par ailleurs, significativement différents avec une valeur plus élevée dans le cas de FUS-RFP-MBD/TDP-GFP que dans le cas TDP-RFP-MBD/FUS-GFP (C). Moyennes avec écart-types. Analyse de test de T \*\*:P<0,01. Echelle : 6  $\mu$ m.

De manière à vérifier le comportement de FUS, nous avons effectué des manipulations complémentaires, pour cela nous avons repris les protéines Lin28 et HuR et effectué des co-transfections avec les couples Lin28-RFP-MBD/FUS-GFP, HuR-RFP-MBD/FUS-GFP en ajoutant un nouveau contrôle pour le couple TDP-43-FUS. Ainsi, nous pouvons remarquer que FUS se comporte de la même manière quelles que soient les protéines partenaires. Elle forme des granules dans le cytoplasme (**voir figure 41**), qui viennent plus ou moins au contact des microtubules. De fait, si l'on se fie aux résultats obtenus précédemment, l'interaction de FUS avec Lin28 ou HuR est moins élevée qu'avec TDP-43 mais le coefficient de corrélation reste assez important en comparaison d'autres protéines, ce qui peut expliquer la présence de FUS sur les microtubules. A côté de cela, nous avons ajouté le marquage de l'ARN (sonde oligo-dT) ce qui a permis de mettre en évidence la présence d'ARN dans les granules mais aussi sur les microtubules. Cela relève le fait que l'interaction de ces diverses RBP peut se faire de manière indirecte par le biais de l'ARN. Enfin, la présence de FUS sous forme de granules peut être liée à une faible concentration cytoplasmique de la protéine qui est principalement nucléaire et peut-être moins bien exporté dans le cytoplasme.

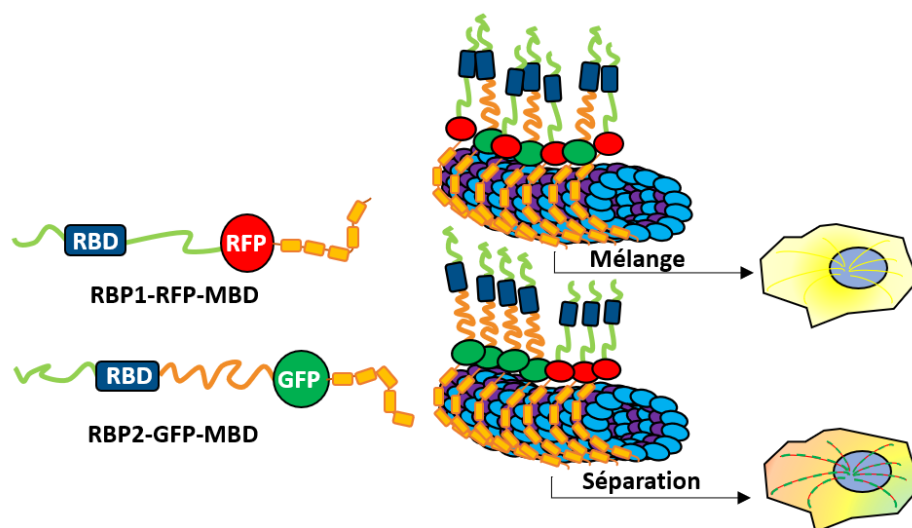


**Figure 41-FUS forme des granules dans lesquels co-localisent l'ARN quand elle est surexprimée**  
 Co-transfection de cellules U2OS avec le plasmide FUS-GFP et différentes protéines appâts TDP-43-RFP-MBD, Lin28-RFP-MBD, ou HuR-RFP-MBD. La protéine FUS forme des granules dans le cytoplasme qui co-localisent avec l'ARN et les microtubules, comme mis en évidence en nuance de gris. Le graphique montre une bonne co-localisation de l'ARN à la fois dans les granules mais aussi le long des microtubules (mesures effectuées dans le cas de TDP-43/FUS). ARN marqué par une sonde oligo polyT-dig et anticorps secondaire Alexa 647. Echelle : 15 μm.

En conclusion, ce résultat converge avec l'interaction entre FUS et TDP-43 décrite dans la littérature, en particulier les résultats de PLA qui avait déjà été mis en évidence[16]. Cependant la technique du MtB permet de mesurer un coefficient de corrélation en fonction de la quantité de protéines proies sur les microtubules, ce qui a l'avantage de permettre la quantification de cette interaction, et permet ainsi une comparaison avec d'autres RBP. Dans ces conditions, nous avons vu que l'interaction FUS/TDP-43 était plus importante qu'avec d'autres RBP. Nous avons relevé une différence d'interaction entre TDP-43 et FUS, et nous souhaitons par la suite caractériser davantage cette interaction et voir comment elle évolue avec différentes proportions de protéines.

## 2. La miscibilité entre TDP-43 et FUS dépend du ratio de ces protéines

Dans un deuxième temps, de manière à aller plus loin et à mieux caractériser l'interaction des deux protéines TDP-43 et FUS, nous avons voulu évaluer leur miscibilité. Pour cela, nous avons utilisé de nouveau, la technique du MtB avec une variante. Dans ce cas, les deux protéines d'intérêts sont conjuguées au MBD, ce qui les amène à se concentrer sur les microtubules, et donc à cohabiter sur ce réseau (**voir image 42**). Il est alors possible de mesurer leur affinité dans diverses conditions, en analysant la répartition de l'intensité de fluorescence le long des microtubules. Si l'affinité entre les deux protéines est bonne, le réseau de microtubules sera jaune et homogène, au contraire, il sera hétérogène avec des compartiments distincts rouges et verts dans le cas de la compartimentation des deux protéines, avec un ensemble d'états intermédiaires entre ces deux cas limites (qu'il est possible de quantifier). Dans cette expérience nous allons utiliser deux protéines en contrôle, rencontrées précédemment : Sam68 et HuR (**voir figure 43**).



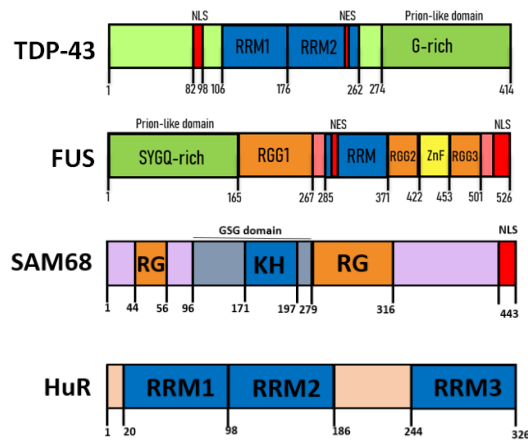
**Figure 42-Mesure de la miscibilité de protéines avec le MtB**

Deux protéines sont conjuguées au domaine de liaison aux microtubules (MBD), l'une est fusionnée à la RFP tandis que l'autre est fusionnée à la GFP. Dans ce cas les deux protéines vont se répartir sur les microtubules. En fonction de leur affinité, elles vont donc plus ou moins se mélanger le long de ce réseau donnant soit un signal homogène (haut) soit un signal discontinu avec des alternances de regroupements rouges ou verts (bas), reflétant la séparation des protéines en compartiments distincts.

Pour ce faire, les cellules ont été co-transfectées avec différents ratios de plasmides, la première condition étant de 10 fois plus de FUS que de TDP-43 soit 10/1, jusqu'à 10 fois plus de TDP-43 (1/10) en passant par des ratios intermédiaires 5/1 ; 2,5/1 ; 1/1 ; 1/2,5 ; et 1/5. De cette manière les protéines étudiées sont plus ou moins exprimées en fonction de la quantité de plasmides transfectés, mais la quantité totale de protéines reste globalement inchangée, et il en résulte une intensité de



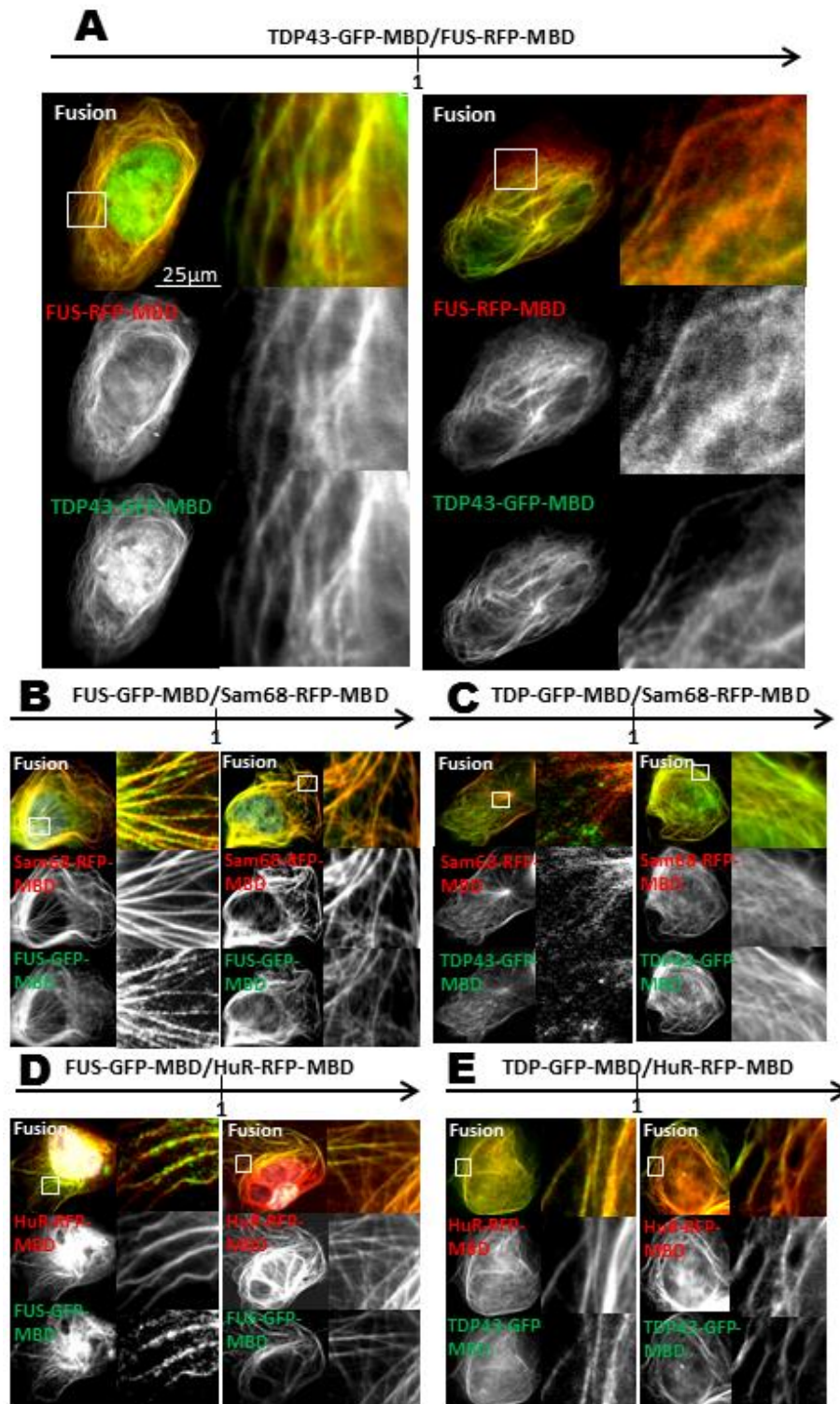
## Résultats



**Figure 43-Séquences des protéines étudiées.**

La miscibilité des protéines TDP-43 et FUS a été étudiée et comparée en mélange avec les protéines Sam68 et HuR. Les séquences des protéines FUS et TDP-43 ont été largement décrites précédemment, mais elles peuvent être comparées avec celles des protéines Sam68 et HuR qui sont deux autres protéines de liaison à l'ARN. La première ne possède pas de RRM mais un domaine KH ainsi que 2 domaines de faible complexité RG, la seconde, possède 3 RRM mais pas de LCD.

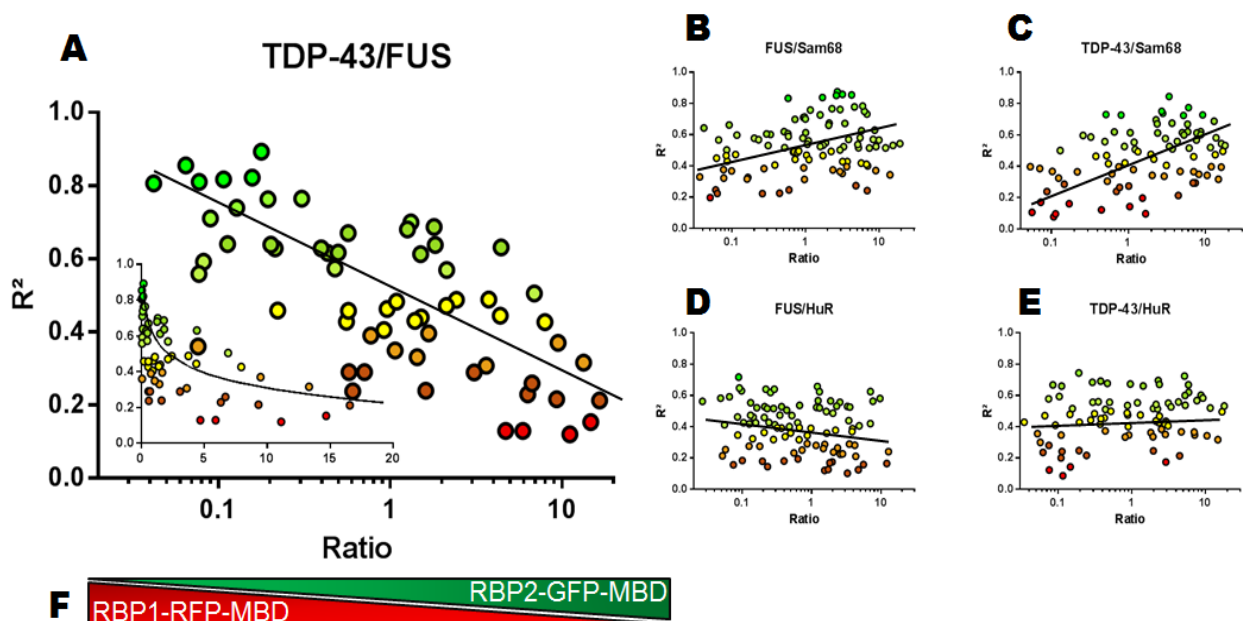
fluorescence plus faible ou plus intense. Pour pallier à un biais dans les analyses, les niveaux d'intensité ont été réajustés dans les deux canaux pour avoir des niveaux comparables. Cependant, le comportement des protéines sur les microtubules est différent selon le ratio étudié. Par exemple, dans le cas d'une forte expression de FUS avec peu de TDP-43 on observe une bonne miscibilité avec des microtubules homogènes plutôt jaunes (**voir images, figure 44, A**). Tandis qu'avec une proportion inverse le signal de FUS apparaît beaucoup plus discontinu, et forme des compartiments distincts du signal de TDP-43, le réseau de microtubules est donc moins homogène, on observe alors une dissociation entre les deux protéines.



**Figure 44-Cellules co-transfectées avec des ratios de plasmides différents pour plusieurs couples de RBP.**

Ces images présentent des cellules (U2OS) co-transfectées avec différents ratios de quantité de plasmides pour le couple de protéines TDP-43-GFP-MBD/FUS-RFP-MBD (A), ainsi que les couples FUS-GFP-MBD/Sam68-RFP-MBD (B), TDP-43-GFP-MBD/Sam68-RFP-MBD (C), FUS-GFP-MBD/HuR-RFP-MBD (D), TDP-43-GFP-MBD/HuR-RFP-MBD (E), grâce à la technique du MtB de manière à étudier leur miscibilité. Pour chaque couple deux cellules sont présentées pour les deux ratios limites RBP1-RFP-MBD/ RBP2-GFP-MBD, A gauche, excès de la protéine conjuguée à la RFP, à droite excès de la protéine conjuguée à la GFP. Dans chaque cas la cellule entière est montrée avec la superposition des deux couleurs, ainsi que les canaux séparés en nuances de gris, et un zoom est effectué sur les microtubules pour mettre en évidence leur aspect, afin de voir si les deux protéines se mélangent (jaune), ou si elles se séparent (alternance de vert et de rouge). Echelle : 25  $\mu\text{m}$  (Zooms, 12  $\mu\text{m}$ , identiques dans chaque condition)

Ce système permet de quantifier la miscibilité des protéines, par le biais de la mesure de coefficient de détermination ( $R^2$ ). Dans ce cas, un programme a été développé pour repérer automatiquement les microtubules au sein des cellules, il définit ensuite des zones le long de ces microtubules dans lesquelles les intensités de fluorescence sont mesurées. Un coefficient  $R^2$  est attribué pour chaque cellule analysée, en fonction de la répartition de l'intensité de fluorescence le long du réseau de microtubule pour chaque couleur. En analysant une centaine de cellules sur les différentes conditions, on observe alors une tendance assez marquée pour le couple TDP-43/FUS : TDP-43 est miscible dans la phase de FUS en faible quantité mais plus on augmente sa concentration, plus on observe un dé-mixage (**voir figure 45, A**). En effet, on observe une bonne corrélation des intensités de fluorescence le long des microtubules lorsqu'il y a peu de TDP-43 (les valeurs de coefficients montent jusqu'à 0,8 pour des faibles ratios et descendent en dessous de 0,2 pour des hauts ratios).



**Figure 45-La miscibilité de TDP-43 et FUS dépend de leur ratio.**

La technique du MtB a été utilisée pour différents ratios de quantité de plasmides pour le couple de protéines TDP-GFP-MBD/FUS-RFP-MBD de manière à étudier la miscibilité des deux protéines. (A) L'analyse de la répartition de la fluorescence, le long du réseau de microtubules dans les deux couleurs a permis d'obtenir un coefficient de détermination ( $R^2$ ) pour chaque cellule. Le ratio des deux protéines a été recalculé en fonction de l'intensité de fluorescence mesurée, la quantité de plasmide transfectée ainsi que de l'ajustement des paramètres d'acquisition. Sur le graphique, chaque point représente une cellule dont la couleur reflète sa valeur de coefficient de détermination, placé en fonction du ratio des deux protéines suivant une échelle logarithmique. Le gradient de ratio RBP1-RFP-MBD/ RBP2-GFP-MBD est toujours présenté de la protéine RFP à gauche vers un excès de protéine GFP vers la droite (F). Ce graphique montre une tendance pour le couple TDP-43 dont la miscibilité varie selon le ratio de ces deux protéines. Cette manipulation a été répétée avec d'autres RBP, en contrôle, pour lesquelles des tendances diverses ont été observées : Sam68-RFP-MBD/FUS-GFP-MBD (B), Sam68-RFP-MBD/TDP-43-GFP-MBD (C), HuR-RFP-MBD/FUS-GFP-MBD (D), HuR-RFP-MBD/TDP-43-GFP-MBD (E). Pour HuR la tendance est peu marquée dans les deux cas avec des valeurs de coefficients assez moyennes. Pour Sam68 l'effet est inverse par rapport au couple FUS/TDP-43, lorsqu'il y a peu de Sam68 on observe un bon mélange mais la dissociation est très forte en présence d'un excès de Sam68 qui n'accepte pas les autres protéines dans sa phase.

Cela met aussi en évidence un phénomène de transition d'une phase homogène vers une phase dé-mixée (**voir figure 45, A, encart**).

De manière à confirmer ce phénomène, et à vérifier que cela ne résultait pas d'un biais du système utilisé, nous avons procédé à la même manipulation avec d'autres couples de protéines. De nouvelles co-transfections ont été effectuées avec les plasmides HuR-RFP-MBD/TDP-43-GFP-MBD, Sam68-RFP-MBD/TDP-43-GFP-MBD, HuR-RFP-MBD/FUS-GFP-MBD, Sam68-RFP-MBD /FUS-GFP-MBD (**voir figure 45, B à E**). La même procédure a été suivie pour l'analyse. De cette manière, on observe des tendances très distinctes entre les divers couples de protéines : HuR montre des tendances assez peu marquées que ce soit avec FUS ou TDP-43 avec parfois des cellules homogènes et parfois hétérogènes avec une corrélation globalement assez moyenne (autour de 0,5). Ce que l'on peut aussi observer sur les images (**voir figure 44, D et E**) avec des cellules plus ou moins homogènes. En revanche, pour Sam68 un effet inverse au couple TDP-43/FUS est observé, avec une très forte séparation de phases dans le cas d'une forte proportion de Sam68, et une phase plus homogène en présence d'une grande quantité de TDP-43. Nous avons vu dans la première partie que Sam68 ne se mélange pas bien avec TDP-43. En effet, Sam68 a une forte capacité à s'auto-assembler, mais il est tout à fait possible que TDP-43 puisse l'accepter dans ses complexes en faible proportion. De la même manière avec FUS, la protéine Sam68 forme des compartiments lorsqu'elle est en forte proportion, tandis que pour une quantité importante de FUS les cellules sont plus homogènes (**voir figure 44, B et C**).

Sam68 et HuR sont deux protéines de liaison à l'ARN (**voir figure 43**), HuR possède 3 RRM et n'a pas de domaine de faible complexité, Sam68 possède un domaine KH pour la reconnaissance de l'ARN et des domaines de faible complexité RG, ces différences structurelles en comparaison de TDP-43 et FUS, que nous avons amplement décrites précédemment (**voir partie, II, 1, a, et II, 1, b**) peuvent conduire à l'émergence de ces comportements distincts.

Un contrôle impliquant un changement de couleur a aussi été effectué de façon à vérifier que l'effet observé ne provenait pas de la couleur choisie pour la condition TDP-43-RFP-MBD/FUS-GFP-MBD en mélange équimolaire pour laquelle on obtenait une valeur de coefficient de corrélation comparable dans les deux cas.

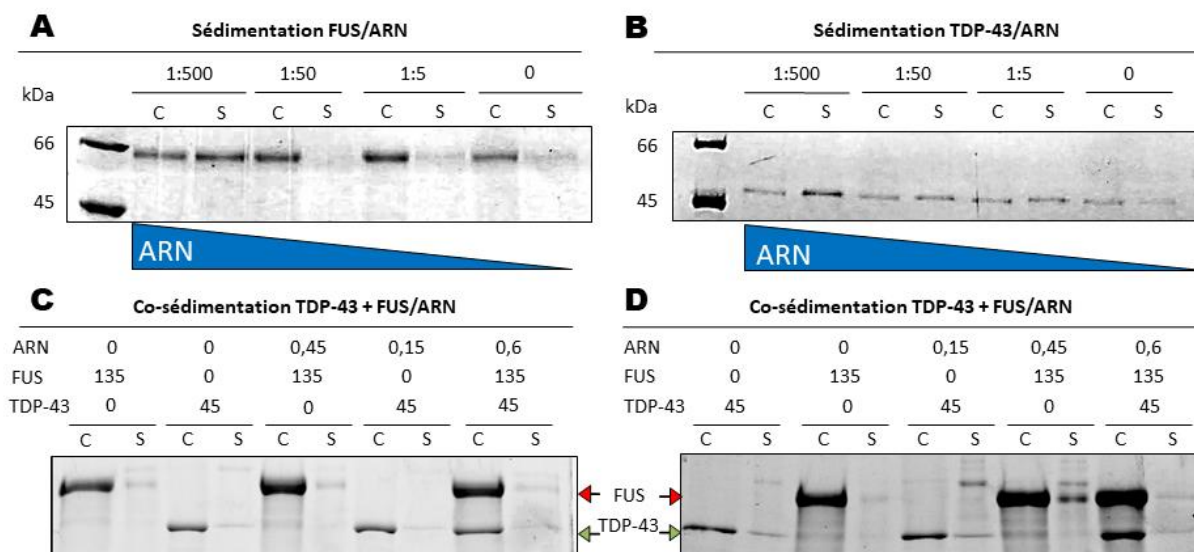
En résumé, nous avons mis en évidence l'interaction entre TDP-43 et FUS avec la technique du MtB et nous nous sommes aperçus que cette interaction était variable en fonction du ratio de ces deux protéines, avec un phénomène de miscibilité et de séparation de phases dépendant de la quantité de ces protéines. FUS acceptant plus

facilement TDP-43 dans sa phase, tandis que TDP-43 a tendance à se séparer de son côté. L'étape suivante est de mieux comprendre cette relation, et notamment leur rapport à l'ARN.

### 3. Etude du comportement des protéines *In vitro*

#### a) Première approche, les gels de sédimentation

Nous souhaitons maintenant savoir quel est le comportement de ces deux protéines, notamment vis-à-vis de leur liaison avec l'ARN. Font-elles preuve de coopération, ou bien de compétition dans leur interaction ? Pour cela nous avons donc testé une première approche qui est celle des gels de co-sédimentation. Dans cette expérience les échantillons sont incubés en mélange, dans différentes conditions, avant d'être centrifugés. Le culot ainsi que le surnageant sont déposés séparément sur un gel d'acrylamide (SDS-PAGE) de façon à identifier la présence des protéines dans la fraction soluble (surnageant) ou insoluble (culot). Les différentes conditions, sont alors comparées entre elles pour mettre en évidence un effet sur le comportement de/des protéine(s) étudiée(s). L'ARN utilisé pour cette manipulation est l'ARN 2Luc, et sera le même tout au long de ce travail. C'est un grand ARN d'environ 3000 bases qui rassemble deux segments codant la luciférase, il a donc une multitude de sites distincts sans spécificité pour une protéine ou une autre. Dans un premier temps, nous avons testé le comportement de FUS et TDP-43 (pleine longueur et non mutées) individuellement en présence d'ARN (**voir figure 46, A et B**). En présence d'une quantité décroissante d'ARN, avec des ratios qui vont d'une protéine pour 500 nucléotides, à une protéine pour 5 nucléotides. On remarque alors que l'ARN a pour effet de maintenir les protéines solubles et non agrégées. TDP-43 et FUS étant des protéines avec une très forte propension à l'agrégation, du fait de leur domaine de faible complexité, il est normal de les retrouver dans le culot en absence d'ARN. Par ailleurs, les deux protéines sont produites et conservées dans un tampon urée (8 M) pour prévenir leur agrégation, et diluées au dernier moment pour procéder aux manipulations. L'effet de solubilisation de ces RBP en présence d'une grande quantité d'ARN est cohérent avec ce qui est décrit dans la littérature[142].

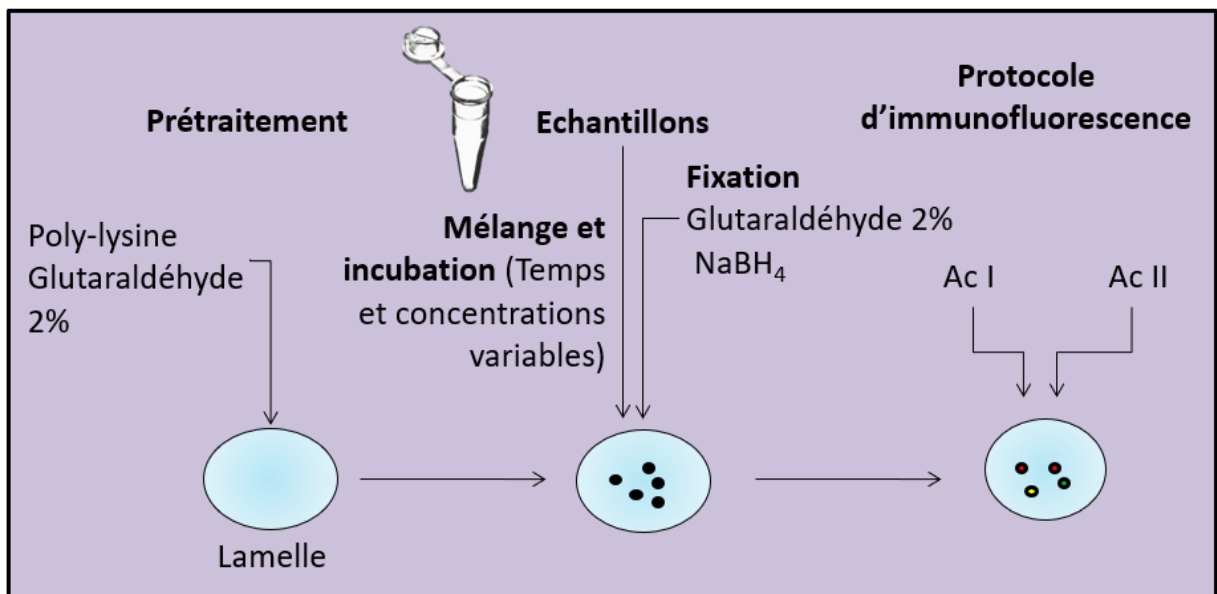


**Figure 46-Étude du comportement des protéines avec l'ARN *In vitro* par gels de sédimentation.** (A) Gel de sédimentation pour la protéine FUS avec un gradient d'ARN 2Luc. La protéine a été incubée à la concentration de 20  $\mu$ M avec différents ratios d'ARN (1:500, 1:50, 1:5, soit une protéine pour 500, 50 ou 5 bases) dans un tampon Hepes, 15 min avant d'être centrifugée. Le culot et le surnageant pour chaque condition ont été chargés séparément sur gel d'acrylamide (SDS-PAGE). L'ARN a tendance à garder la protéine soluble dans le surnageant, tandis que sans ARN la protéine est dans le culot. (B) Protocole similaire pour la protéine TDP-43 (10  $\mu$ M incubé 10 min) avec un effet semblable. (C, D) Gel de co-sédimentation pour le mélange TDP-43, FUS et ARN avec un ratio 1/3 entre TDP-43 et FUS, ratio d'ARN fixé à 1:10 par rapport à la quantité totale de protéine. Dans le premier cas (C) FUS a été introduit avant, dans le second cas (D) TDP-43 a été introduit avant. Quantité en  $\mu$ M. Incubation 4 min avant centrifugation. C=culot, S=surnageant.

Par la suite, nous avons effectué des expériences avec un mélange des deux protéines et d'ARN de façon à mieux comprendre le comportement du mélange en relation avec l'ARN. TDP-43 et FUS ont été mélangées avec un ratio 1/3 (en accord avec [142]) le ratio protéine/nucléotides est de 1/10 de façon à se placer à un niveau proche de la saturation tout en laissant suffisamment d'espace pour que les protéines s'y fixent. Plusieurs essais ont été effectués, deux sont présentés ici, avec pour unique différence l'ordre d'introduction des protéines (**voir figure 46, C et D**). Dans le cas du mélange nous avons plutôt tendance à observer un basculement des protéines dans le culot mais cet effet est peu marqué. Cependant cette technique a l'inconvénient de présenter une certaine variabilité et donc peu de répétabilité. Ainsi, ces résultats sont peu concluants, d'autant qu'ils ne permettent pas de répondre à un certain nombre de questions. Nous ne savons pas si les protéines sont mélangées ou dissociées, ni dans quelle phase se trouve l'ARN, nous n'avons même aucune information sur la morphologie des agrégats formés.

### b) Observation des agrégats et condensats par immunofluorescence *In vitro*

De manière à pallier ce problème une nouvelle méthode a été développée. Il s'agit d'une méthode d'immunofluorescence, permettant de visualiser les agrégats ou les condensats formés *in vitro*, en les fixant sur des lamelles, de façon à étudier la morphologie et la miscibilité des complexes formés par les protéines suivant différentes conditions. Cela permet aussi d'analyser l'influence de l'ARN sur la miscibilité des protéines. Pour cela, les échantillons sont incubés seuls ou en mélange dans différentes conditions de concentration et de durée. Ils sont ensuite fixés sur les lames de verres prétraitées. Enfin, un protocole d'immunomarquage à l'aide d'anticorps spécifiques (primaires et secondaires) est effectué de façon à révéler les complexes de protéines par microscopie de fluorescence (**voir figure 47, et matériel et méthodes**).

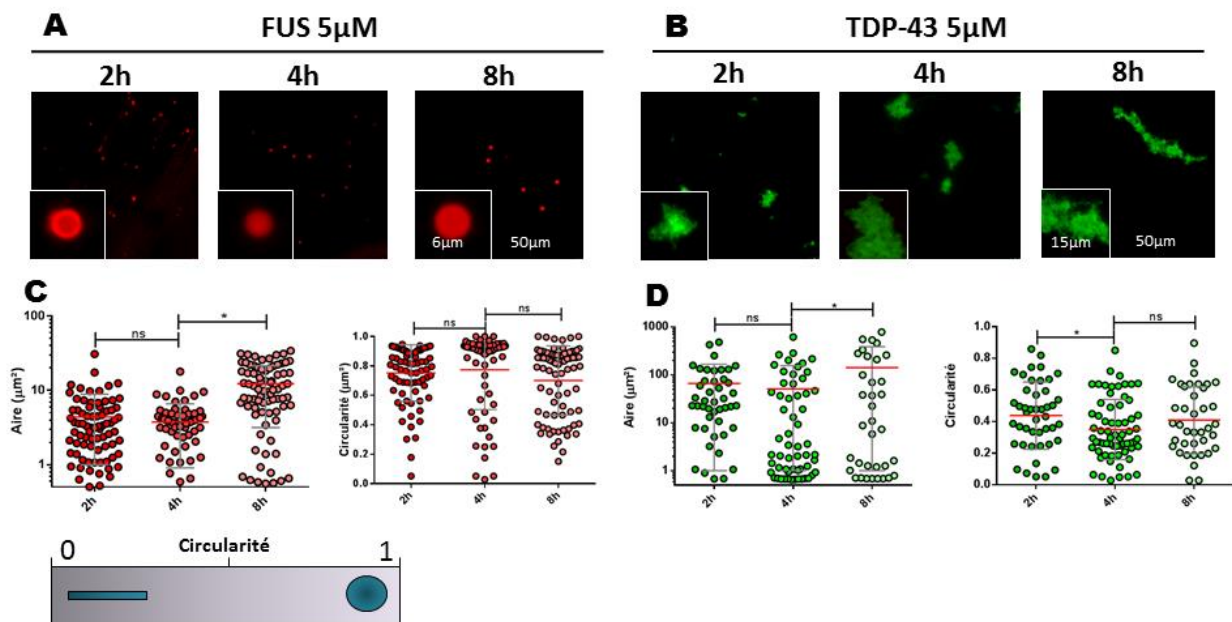


**Figure 47-Principe de la technique de visualisation des agrégats par immunofluorescence.**

De manière à étudier le comportement des protéines *In vitro* les échantillons de protéines (et d'ARN) sont incubés à des concentrations variables dans des tubes, avant d'être déposés sur des lamelles préalablement prétraitées (Poly-lysine puis glutaraldéhyde) puis fixés avec du glutaraldéhyde. Du  $\text{NaBH}_4$  est ajouté de façon à inhiber l'auto-fluorescence du glutaraldéhyde. Par la suite, un protocole d'immunofluorescence classique est effectué avec des anticorps primaires (Ac I) spécifiques aux protéines étudiées, puis des anticorps secondaires (Ac II) fluorescents pour marquer les échantillons et les visualiser en microscopie optique à fluorescence.

Après de nombreux tests pour optimiser le protocole et valider la technique nous avons pu tester les protéines TDP-43 et FUS individuellement et obtenir des informations sur leurs comportements. Ainsi nous remarquons d'emblée qu'en dépit de leur caractéristiques communes (comme décrit précédemment) les deux protéines ont un comportement très distinct (**voir figure 48, A et B**). FUS forme des

gouttelettes semblables aux condensats formés suite au processus de séparation de phases comme ce qui peut être observé par ailleurs[145, 151]. TDP-43, à l'inverse, forme des agrégats assez volumineux aux contours irréguliers et inhomogènes. Ces observations sont très répétables et se sont confirmées dans l'ensemble des manipulations suivantes. Ainsi, la morphologie de FUS reste très similaire que ce soit en fonction du temps d'incubation, ou de sa concentration (**voir figure 48 et 49, A**) avec des gouttelettes dont l'aire évolue entre 2 et 10  $\mu\text{m}^2$  en moyenne, et augmente légèrement avec le temps d'incubation. De plus, les gouttelettes pour FUS se forment à partir d'une certaine concentration, au-delà de 2,5  $\mu\text{M}$  (En dessous on voit seulement des points plus petits plus dispersés, dont la taille et la circularité sont inférieures aux gouttelettes, **voir figure 48, C**). Tandis que les agrégats de TDP-43 sont plus massifs, et s'assemblent en objets plus gros avec le temps, ou la concentration, jusqu'à dépasser les 100  $\mu\text{m}^2$  (**voir figure 48 et 49, B**).



**Figure 48-Evolution des objets formés *In vitro* par les protéines TDP-43 et FUS avec le temps**

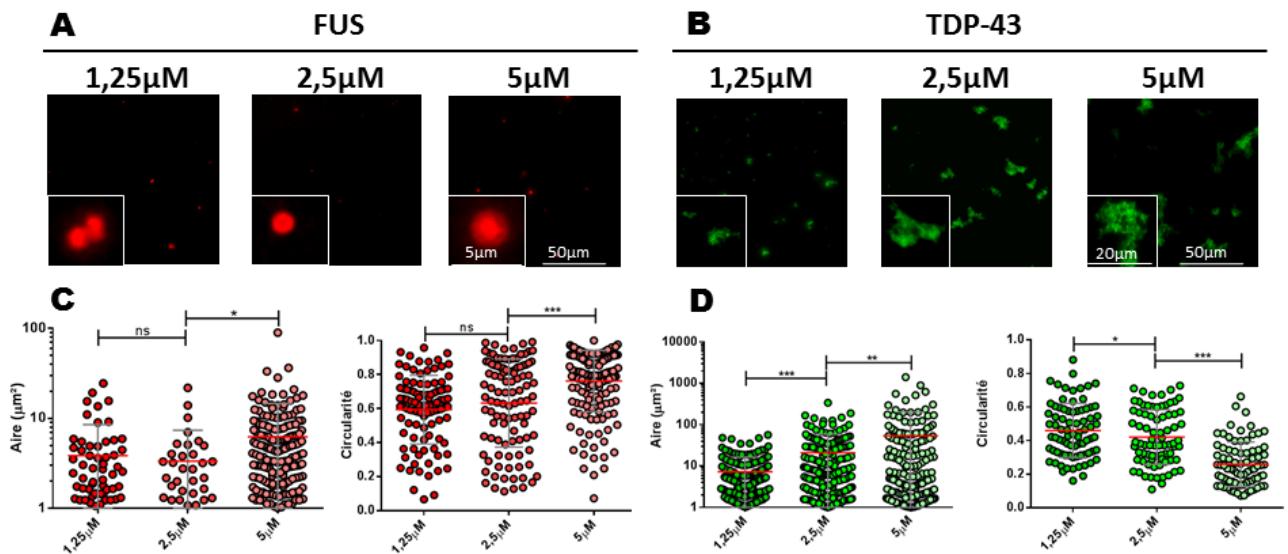
(A) Acquisitions pour la protéine FUS incubée à 5  $\mu\text{M}$  sur des échantillons fixés à plusieurs temps d'intervalles et marqués par immunofluorescence. La protéine forme des gouttelettes caractérisées par une bonne circularité qui varie légèrement en taille au cours du temps (C). Acquisitions pour la protéine TDP-43 pris dans de mêmes conditions. La protéine forme de gros agrégats caractérisés par une circularité plus faible que pour FUS (C). Le temps favorise l'apparition de gros objets. Echelle : 50  $\mu\text{m}$  (fenêtre large), 6  $\mu\text{m}$  (Zoom FUS) et 15  $\mu\text{m}$  (Zoom TDP-43). \*\*\*:  $P < 0,001$ ; \*:  $P < 0,05$

La circularité est un paramètre utilisé pour caractériser la forme des agrégats observés. Elle permet de décrire la forme d'un objet en lui attribuant une valeur, si l'objet est rond, la circularité est proche de 1, si l'objet est plutôt linéaire, la circularité s'approche de 0. Ainsi, les gouttelettes de FUS ont une circularité élevée quelle que



soit la condition, tandis que les agrégats de TDP-43 sont caractérisés par une circularité relativement faible qui a tendance à décroître avec la taille des agrégats.

Ces deux protéines étant très enclines à s'agréger de par leur domaine de faible complexité, il est normal de les retrouver agrégées *In vitro*, en s'auto-assemblant. Ces premières observations indiquent que les deux protéines adoptent des comportements distincts dans les mêmes conditions, ce qui nous amène à nous interroger sur leurs facultés à se mélanger.



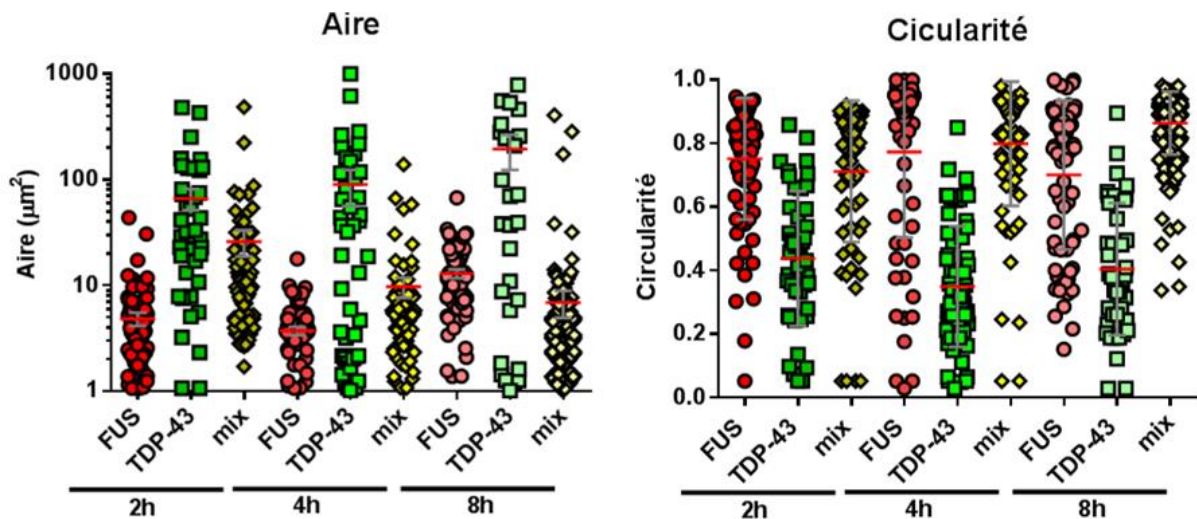
**Figure 49-Evolution des objets formés *In vitro* par les protéines TDP-43 et FUS en fonction de la concentration**

(A) Acquisitions de la protéine FUS incubée 2h à différentes concentrations 1,25 μM, 2,5 μM, 5 μM, par immunofluorescence sur des échantillons fixés. Les gouttelettes telles qu'observées précédemment ne sont visibles qu'au-delà de 2,5 μM, en dessous seuls des petits points assez épars sont observables (C). Acquisitions pour la protéine TDP-43 pris dans de mêmes conditions. La protéine forme des agrégats plus gros avec une concentration plus élevée et leur circularité décroît avec la taille (C). Echelle logarithmique dans le cas de la mesure des aires. Echelle : 50 μm (fenêtre large), 6 μm (Zoom FUS) et 20 μm (Zoom TDP-43). \*\*\*:P<0,0001; \*:P<0,05.

### c) Première approche pour le mélange des protéines *In vitro*

En première approche pour le mélange, nous avons choisi d'opter pour un ratio où TDP-43 est en moindre proportion par rapport à FUS (Ratio FUS/TDP-43 d'environ 1/3), cela se justifie à la fois par la mesure de la concentration des protéines au sein des cellules (qui a été publiée précédemment[142]) mais aussi par les manipulations précédentes où une quantité de FUS de 5 μM est nécessaire pour avoir des gouttelettes tout en se plaçant dans des conditions où la taille des objets de TDP-43 reste raisonnable. Ainsi, lorsque l'on effectue un mélange des deux protéines dans ces proportions, et que l'on observe leur comportement au cours du temps, comme précédemment, on remarque que les paramètres pour le mélange restent

relativement constants, en particulier en comparaison de TDP-43 seul. On mesure alors une aire des objets assez faible et une circularité moyenne très élevée, ce qui est comparable aux paramètres observés pour FUS (**voir figure 50**). En revanche, on pourrait s'attendre à avoir une population d'objets assez gros qui suit le comportement de TDP-43, qui a tendance à grossir avec le temps. Ainsi, l'aire moyenne de ces objets est limitée en dessous de  $10 \mu\text{m}^2$  et à une circularité supérieure à 0,8 après 8 h. Ce qui suggère une solubilisation de TDP-43 dans la phase de FUS, empêchant ainsi la protéine TDP-43 de s'agglomérer dans de gros agrégats au cours du temps, et cela sans l'intermédiaire de l'ARN. Ces observations sont le fruit d'une expérience préliminaire qui nous amène à nous interroger sur la miscibilité de FUS et TDP-43, nous avons donc souhaité, par la suite, approfondir ce phénomène avec différents ratios de protéines *In vitro*.



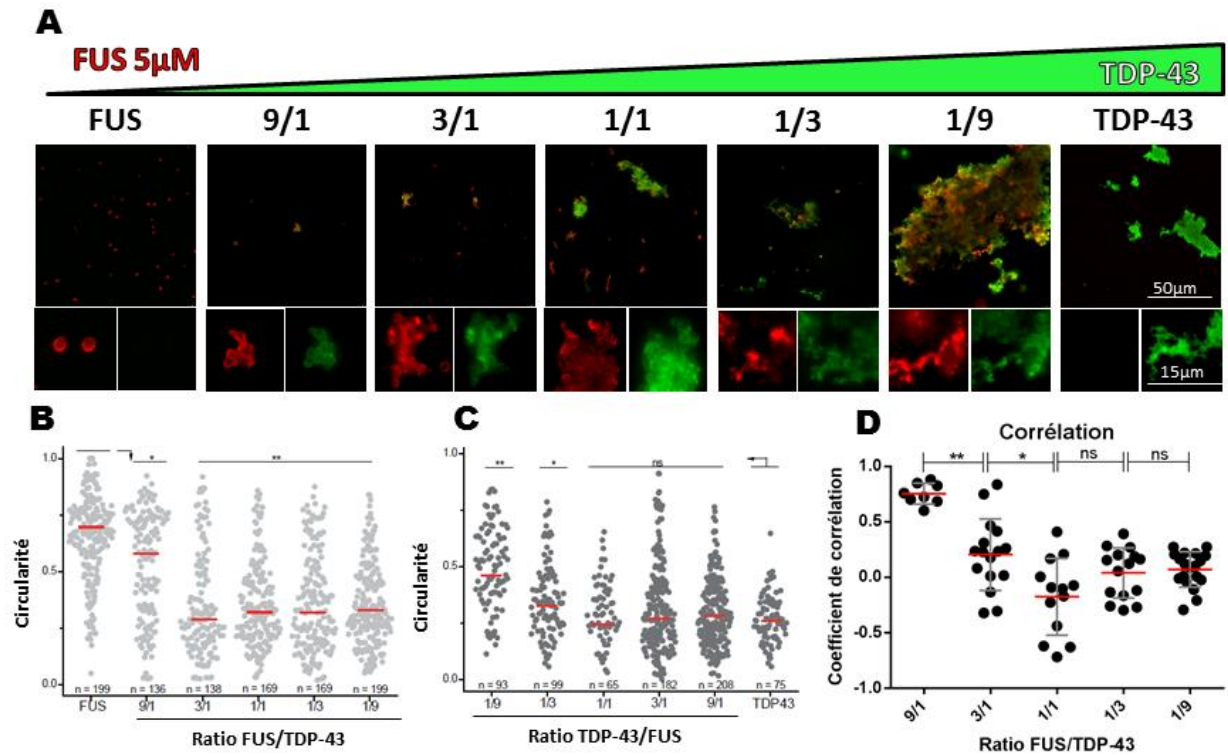
**Figure 50-Le mélange de TDP-43/FUS suit le comportement de la protéine FUS.**

Incubation 2 h, 4 h et 8 h, des échantillons TDP-43 (Vert) et FUS (rouge) seul et en mélange (jaune) dans des proportions où FUS est en excès. Lorsque l'on mélange TDP-43 et FUS dans ces proportions, on remarque que le mélange suit le comportement de la protéine FUS avec une aire des objets assez faible et une circularité élevée. Aire en  $\mu\text{m}^2$  (échelle logarithmique). Moyennes avec écart-types.

#### d) Etude du mélange *In vitro* avec différents ratios

De manière à mieux comprendre comment se comportent les deux protéines en mélange, nous avons effectué différentes solutions contenant un mélange des deux protéines à différents ratios. FUS a été fixée à  $5 \mu\text{M}$  et on augmente TDP-43 progressivement pour balayer des ratios FUS/TDP-43 de 9/1 à 1/9 sans ARN (**voir figure 51**). Dans ces conditions, on retrouve les gouttelettes pour FUS, en contrôle. Avec l'ajout de TDP-43, les objets évoluent vers des gouttelettes légèrement agglomérées et déformées que nous qualifierons de grappes. C'est ce que l'on

retrouve dans les conditions 9/1 et 3/1, ce qui est cohérent avec une miscibilité des protéines pour les proportions faibles en TDP-43.



**Figure 51-La miscibilité des protéines varie en fonction de leurs ratios.**

(A) Acquisitions pour les mélanges des protéines FUS et TDP-43 incubés 2h à différents ratios de protéine FUS/TDP-43, par immunofluorescence sur des échantillons fixés. FUS est fixée à 5 $\mu$ M et on augmente progressivement la quantité de TDP-43 pour augmenter le ratio. On observe des gouttelettes pour FUS seule qui évoluent vers des agrégats plus gros et plus dé-mixés au fur et à mesure de l'ajout de TDP-43, comme le montre la mesure de coefficient de corrélation (D). La circularité chute avec l'ajout de TDP-43 dans le rouge (B) comme dans le vert (C). Echelle : 50  $\mu$ m (fenêtre large), 15  $\mu$ m (Zoom). \*\*: P<0,01; \*: P<0,05.

De manière à évaluer la capacité des protéines à se mélanger, dans les différentes conditions, un programme a été développé qui permet de déterminer un coefficient de corrélation pour chaque image. Ce dernier découpe les agrégats en différentes zones en fonction des variations d'intensité. Il mesure ensuite l'intensité des différentes couleurs dans chacune de ces zones, puis ces valeurs sont comparées entre elles, pour attribuer un coefficient de corrélation. Il est alors possible de quantifier le mélange ou la séparation de ces protéines dans chaque condition en comparant ces coefficients de corrélation (**voir matériel et méthodes**). Ainsi, dans les structures s'apparentant à celles de FUS (caractérisées par une petite taille et une bonne circularité) on a un coefficient de corrélation élevé (9/1) qui diminue vers une valeur relativement moyenne au sein des grappes (1/3) (**voir figure 51, A et D**). En

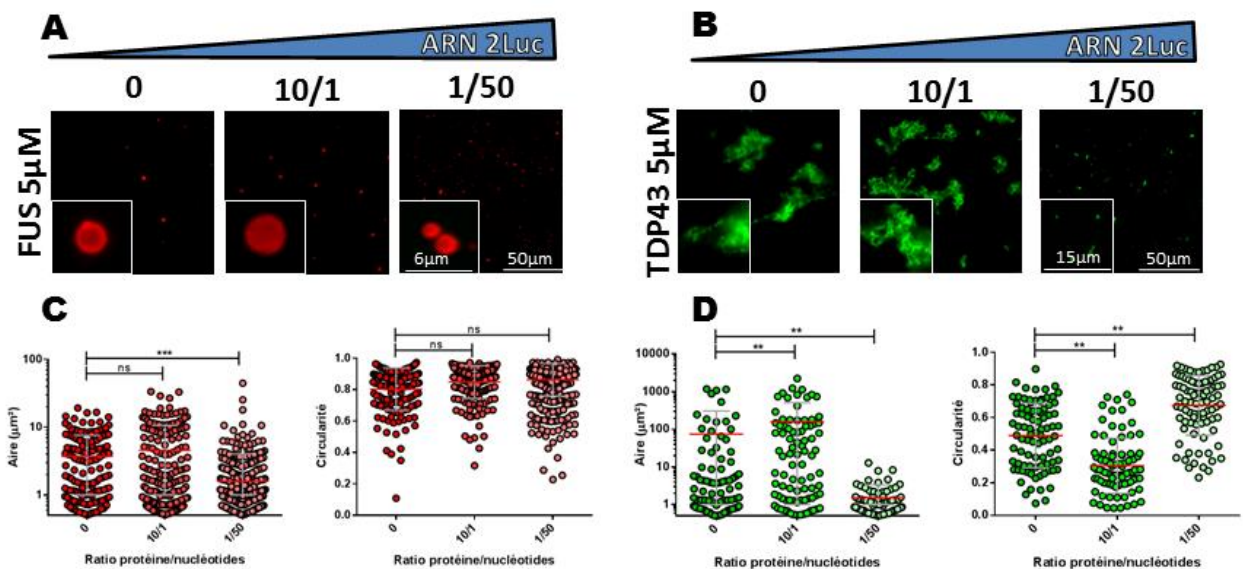
revanche, dès lors que l'on augmente davantage la concentration de TDP-43 (ratio 1/1), on remarque la formation de gros agrégats qui cohabitent avec de plus petits objets, avec un coefficient de corrélation plus faible. En dépit de cette faible corrélation, ces agrégats peuvent contenir la protéine FUS, donc on retrouve une certaine interaction entre les deux protéines, mais elles se mélangent de manière peu homogène, ainsi on remarque des gouttelettes de FUS piégées dans les agrégats de TDP-43. Lorsque l'on augmente encore la quantité de TDP-43, les agrégats sont plus gros et il y a de moins en moins de gouttelettes autour, elles se retrouvent piégées au sein des agrégats de TDP-43. Au-delà du ratio 1/1 la valeur des coefficients de corrélation mesurés reste assez faible. Ces observations sont supportées par les mesures de circularité qui décroissent avec l'ajout de TDP-43 (**voir figure 51, B et C**). Les tendances identiques que suivent les mesures de circularité dans le vert comme dans le rouge illustrent le mélange relatif des protéines au sein des mêmes structures. Ces résultats sont cohérents avec ce qui a pu être observé avec la manipulation de MtB. TDP-43 est soluble dans la phase de FUS quand sa proportion est inférieure à celle de FUS tandis que quand sa quantité est égale ou dépasse celle de FUS la protéine s'agglomère en agrégats dans lesquels FUS peut se retrouver mais de manière plutôt hétérogène tout en conservant sa capacité à former des gouttelettes. Ces premières observations suggèrent que la proportion des protéines a une influence sur leur comportement. Nous allons maintenant nous intéresser à l'influence de l'ARN sur le comportement des protéines seules ou en mélange.

#### 4. Analyse de la structure des condensats de TDP-43 et FUS en présence d'ARN

##### a) Effet de l'ARN sur les protéines seules

Maintenant que nous avons mieux cerné le comportement des protéines en mélange sans ARN, et avons eu un aperçu de leur comportement *In vitro*, nous allons étudier l'influence de l'ARN sur ces protéines. Nous allons dans un premier temps, tester deux cas limites pour chacune des protéines : 10/1 soit 10 protéines par nucléotide, donc une saturation des ARN en protéines, et 1/50, soit une protéine pour 50 nucléotides, donc un excès d'ARN. Les échantillons sont préparés de la même manière que précédemment, avec une concentration fixe de protéines et différentes quantités d'ARN 2Luc. Ainsi pour FUS, on retrouve les gouttelettes observées précédemment sans ARN, ainsi qu'avec les différentes concentrations d'ARN. Une nouvelle fois, on remarque que la morphologie de FUS est assez constante, caractérisée par une circularité toujours très élevée. Cependant, ces

gouttelettes grossissent légèrement quand il y a peu d'ARN, tandis qu'elles deviennent plus petites lorsqu'il y a un excès d'ARN (**voir figure 52, A**). Pour TDP-43, l'effet est plus flagrant, les agrégats sont plus gros avec peu d'ARN, tandis qu'ils sont dissouts avec un haut ratio d'ARN (**voir figure 52, B**). Dans le premier cas, nous sommes en condition de nucléation, la faible quantité d'ARN, favorise l'accumulation de protéines et la croissance des agrégats, tandis qu'en excès, l'ARN permet de dissocier les agrégats de protéines en répartissant les protéines le long de la chaîne d'acides nucléiques[9, 142, 145]. C'est aussi cohérent avec ce qui a pu être observé pour les protéines individuelles avec la technique du gel de sédimentation. En conséquence, il est évident pour des protéines de liaison à l'ARN, que la liaison à l'ARN permet une bonne régulation de leur comportement, et c'est d'autant plus vrai pour TDP-43, pour laquelle un excès d'ARN est nécessaire pour maintenir la protéine soluble. Maintenant, que nous avons vérifié le rôle de l'ARN sur le comportement des protéines individuellement, qu'en est-il du mélange ?

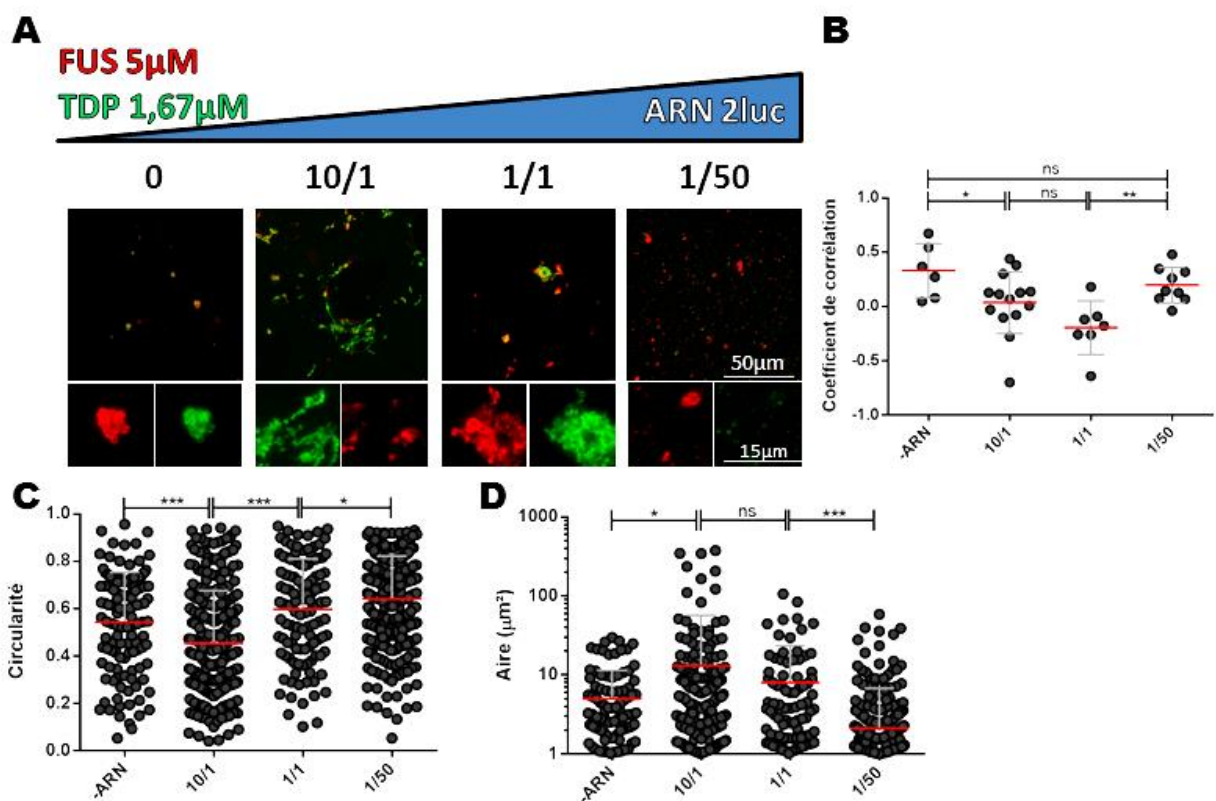


**Figure 52-L'ARN a un effet ambivalent sur les RBP.**

(A) Acquisitions de la protéine FUS (concentration 5 μM) en mélange à l'ARN dans 2 conditions de ratios protéines/nucléotides (10 protéines par nucléotide, et 1 protéine pour 50 nucléotides) incubées 2 h avant fixation. (B) Idem pour TDP-43 à la concentration de 12,5 μM. FUS forme des gouttelettes caractérisées par une forte circularité qui ne varie peu. En revanche, à forte concentration d'ARN les gouttelettes diminuent en taille (C), tandis qu'on observe une augmentation de la taille des objets de TDP-43 avec peu d'ARN et une forte diminution de leur taille en parallèle d'un regain de circularité à forte concentration d'ARN (D). Echelle : 50 μm (fenêtre large), 6 μm (Zoom FUS), 15 μm (Zoom TDP-43). Moyennes et écart-types. Test de T, \*\*\* : P<0,0001; \*\* : P<0,01

### b) Effet de l'ARN sur le mélange FUS et TDP-43

Après avoir observé l'effet de l'ARN sur les protéines seules, nous souhaitons donc savoir ce qu'il en est lors du mélange des deux protéines. La quantité d'ARN a-t-elle une influence sur leur mélange ou leur séparation ? De cette façon, nous avons repris différents ratios d'ARN avec le mélange TDP-43/FUS 1/3 car c'est un ratio qui permet la bonne cohabitation des deux protéines au sein de structures diverses : gouttelettes, et petits agrégats. Ainsi, sans ARN on retrouve des structures de type grappes avec un bon mélange des deux protéines (coefficient de corrélation moyen, voir figure 53, A et B). Lorsque l'on ajoute de l'ARN en ratio 10/1, soit 10 protéines par nucléotide, on induit un effet de nucléation qui favorise la formation de gros agrégats. Cet effet est couplé à la mise en concurrence des protéines, il en résulte une forte démixtion, avec un coefficient de corrélation avoisinant le 0. Avec un peu plus d'ARN (ratio 1/1), les agrégats diminuent en taille et reviennent à la forme de grappes, mais sont moins bien mélangées que sans ARN, puisque dans ces conditions les deux protéines sont toujours mises en concurrence du fait du manque d'ARN



**Figure 53- Comportement des protéines en mélange avec l'ARN par immunofluorescence**

(A) Acquisitions du mélange TDP-43/FUS au ratio 1/3 avec différents ratios d'ARN 2Luc (Ratio protéines/nucléotides : 10/1, 1/1, 1/50). On retrouve l'effet assez marqué de l'ARN avec la formation de gros agrégats quand il y a peu d'ARN, et un effet de solubilisation en excès d'ARN. (B) Coefficients de corrélation pour les différentes conditions. L'ARN favorise un effet de compétition entre les protéines avec une séparation lorsqu'il est en faible quantité, et un mélange lorsqu'il est en excès. Echelles : 50 µm et 15 µm (Zoom). Mesures de circularité (C) et d'aires (D) dans chaque condition. Moyennes et écart-types. \*:  $p < 0,05$  ; \*\*:  $p < 0,01$  ; \*\*\* :  $p < 0,0001$

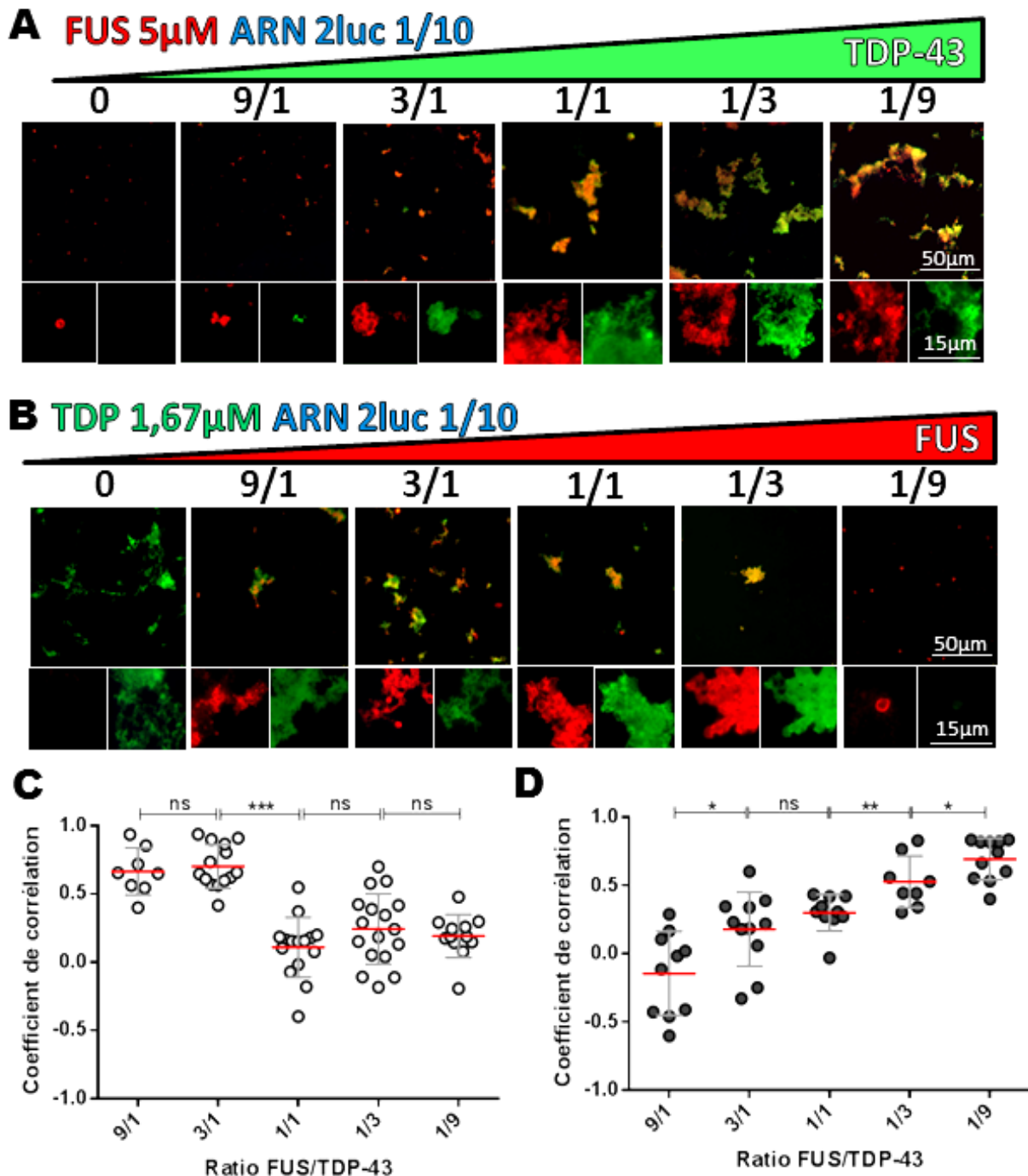
disponible. Enfin, à un ratio plus important, supérieur et à 1/10 et jusque 1/50, il y a suffisamment d'ARN pour que les protéines se répartissent librement sur ce dernier, ce qui a pour conséquence de former des petits objets, on retrouve ainsi l'effet de solubilisation des protéines. Nous observons alors une meilleure corrélation en comparaison de la condition précédente, et semblable au mélange sans ARN. On peut supposer que l'excès d'ARN permet aux protéines de se répartir comme elles le souhaitent sur leurs sites de prédilections mais peut aussi favoriser leur interaction.

Cette manipulation nous montre que le comportement du mélange vis-à-vis de l'ARN est identique à celui des protéines seules, on note une grande similitude avec le comportement de TDP-43 (qui est pourtant présent en quantité réduite). Cependant, une faible proportion d'ARN induit une compétition entre les deux protéines. Il faut donc avoir suffisamment d'ARN pour être dans des conditions favorables à leur interaction.

### c) L'ARN améliore le mélange des protéines

Pour continuer avec l'ARN, nous souhaitons maintenant voir son effet sur le comportement des protéines en mélange à différents ratios de protéines. Nous avons donc repris la gamme de ratios effectuée précédemment à laquelle nous avons ajouté de l'ARN dans les proportions moyennes 1/10 (1 protéine pour 10 nucléotides) par rapport à la quantité totale de protéines (FUS + TDP-43) de manière à favoriser la cohabitation des protéines sur ce dernier, sans induire une compétition trop importante, ni trop de site de liaison qui pourraient induire une dispersion des protéines ce qui ne serait pas pertinent pour étudier la relation de FUS et TDP-43 en présence d'ARN. Deux gammes ont été effectuées, l'une avec FUS fixée (à la concentration de 5  $\mu$ M) et un gradient de TDP-43, et l'inverse avec TDP-43 fixée (1,67  $\mu$ M) et un gradient de FUS de façon à vérifier la réciprocity du phénomène. L'ARN est ajusté dans chaque cas pour maintenir le ratio protéine/nucléotide constant. Dans le premier cas, on retrouve la morphologie des objets avec de petites gouttelettes de FUS. Lorsque l'on ajoute du TDP-43, on passe par la formation de grappes (ratios 1/9 et 3/1), puis en mélange équimolaire on commence à voir de gros agrégats avec la protéine TDP-43 qui impose sa structure (**voir figure 54, A**). D'un point de vue de la corrélation, on a un mélange assez homogène quand TDP-43 est en minorité, on retrouve la solubilisation de la protéine dans la phase de FUS. Tandis qu'en excès, on a un système très séparé avec des gouttes de FUS piégées dans les agrégats formés par TDP-43 avec une corrélation proche de 0 pour les ratios 1/1, 1/3 et 1/9 (**voir figure 54, C**). La structure des assemblages formés est très semblable à ce qui était observé sans ARN. De ce fait, l'ARN ne semble pas affecter particulièrement l'organisation des protéines dans ces conditions, en particulier en matière de

morphologie des condensats. Il apparaît, cependant, qu'il améliore leur mélange en décalant leur séparation vers de plus hauts ratios (3/1 bonne corrélation de 0,6, alors qu'elle était de 0,2 sans ARN (**voir figure 51, D**), sans pouvoir empêcher la séparation lorsqu'il y a un excès de TDP-43 (Ce ratio d'ARN ne permet pas de prévenir l'agrégation de la protéine TDP-43).



**Figure 54-L'ARN favorise la miscibilité des protéines TDP-43 et FUS**

Acquisitions de mélanges des protéines FUS et TDP-43 incubées 2 h avec de l'ARN 2Luc à la quantité 1/10 (1 protéine pour 10 nucléotides). (A) La quantité de FUS est fixée à 5  $\mu$ M, et on augmente progressivement la quantité de TDP-43 pour augmenter le ratio. (B) TDP-43 est fixée à 1,67  $\mu$ M avec un ratio protéine/nucléotide fixé à 1/10, tandis que l'on augmente la quantité de FUS. (C et D) Coefficients de corrélation mesurés dans les deux cas, FUS fixée (C), et TDP-43 fixée (D). On observe un phénomène réciproque avec une bonne miscibilité des deux protéines en faible proportion de TDP-43. Echelle : 50  $\mu$ m (fenêtre large), 15  $\mu$ m (Zoom). \*\*\*:P<0,0001 ; \*:P<0,05 ; ns=non significatif.

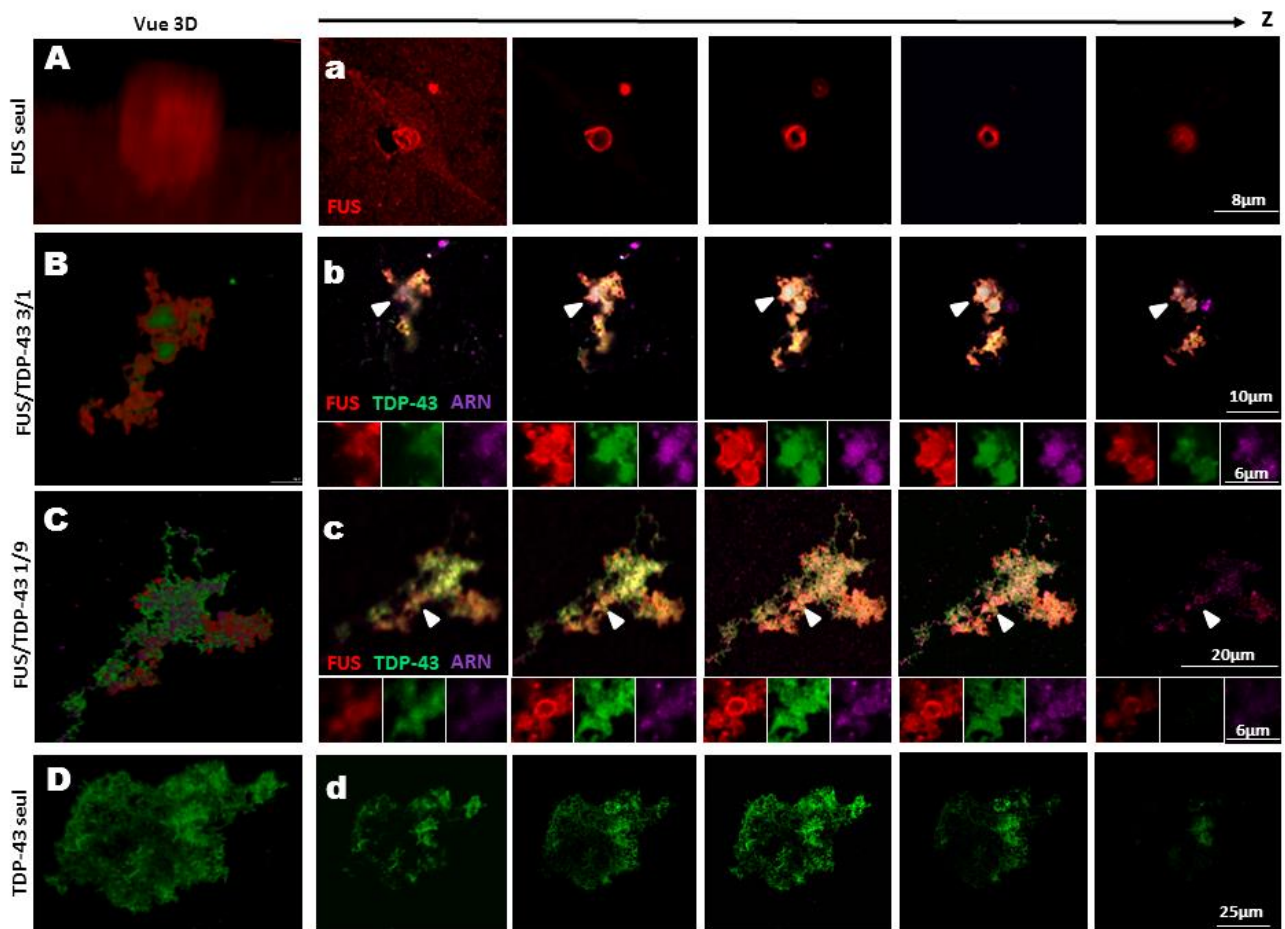


Dans le cas inverse, où TDP-43 est fixée, on retrouve le même effet de transition qui commence avec de gros agrégats de TDP-43 très hétérogènes avec des inclusions de FUS, puis le mélange évolue vers des grappes assez homogènes, et enfin on retrouve les gouttelettes lorsqu'il y a une faible proportion de TDP-43 (**voir figure 54, B**). Les coefficients de corrélation attestent de cette évolution progressive (**voir figure 54, D**). Ces observations étant convergentes, elles confirment le comportement des deux protéines et l'asymétrie de leur miscibilité qui se vérifie aussi en présence d'ARN. Ainsi, lorsqu'il y a peu de TDP-43, l'ARN favorise la miscibilité de ces deux protéines.

#### **d) Observations des structures formées par les protéines en 3 dimensions**

De manière à mieux caractériser les structures formées par les protéines TDP-43 et FUS nous avons effectué des acquisitions complémentaires au microscope confocal, qui permet d'avoir des informations en 3 dimensions, une meilleure résolution mais aussi plus de sensibilité dans la longueur d'onde. Nous avons donc observé les échantillons de FUS et TDP-43 seules, ainsi que les mélanges FUS, TDP-43, ARN dans les proportions 3/1 et 1/9 (avec l'ARN 1/10). Ces acquisitions ont permis de visualiser la forme des objets dans l'espace, chaque objet a été scanné sur 20 à 25 images en z, soit une dizaine de micromètres d'épaisseur, ce qui permet d'effectuer des reconstitutions en 3 dimensions des objets. En particulier, les gouttelettes formées par FUS sont bien sphériques (**voir figure 55, A**) et le marquage se fait sur le pourtour de l'objet. En revanche, nous ne pouvons expliquer si ce phénomène est dû à une organisation spécifique de la protéine, ou à un problème d'accessibilité de l'anticorps. Les agrégats de TDP-43 sont semblables à de grosses pelotes de filaments entremêlés, en revanche, il est intéressant de remarquer qu'en dépit de leur longueur importante (plusieurs dizaines de micromètres), ils sont assez plats (moins de 10  $\mu\text{m}$  d'épaisseur), mais cela peut être lié à la fixation des objets (**voir figure 55, D**). Les objets de type grappes, obtenus dans le cas d'un mélange 1/3 de TDP-43/FUS, sont des intermédiaires entre les gouttes et des agrégats, semblables à un assemblage de gouttelettes plus ou moins circulaires. De manière intéressante, le signal de TDP-43 se retrouve à l'intérieur des gouttelettes, un argument supplémentaire dans le sens de la solubilisation de TDP-43 dans la phase de FUS (**voir figure 55, B**). Pour la condition 1/9, on retrouve les agrégats de TDP-43 très hétérogènes avec des inclusions de FUS toujours semblables à des gouttes mais engluées dans ces agrégats (**voir figure 55, C**). Du côté de l'ARN, il se retrouve à la fois corrélé au marquage de FUS et de TDP-43. De plus, il semble plus concentré dans les zones denses, plutôt que dans les filaments. Le petit nombre d'acquisitions

effectuées au confocal et l'incertitude sur le marquage ARN en microscopie optique n'ont pas permis de quantifier et d'évaluer rigoureusement ce phénomène.



**Figure 55- Les acquisitions confocales donnent un aperçu de la structure 3D des objets.**

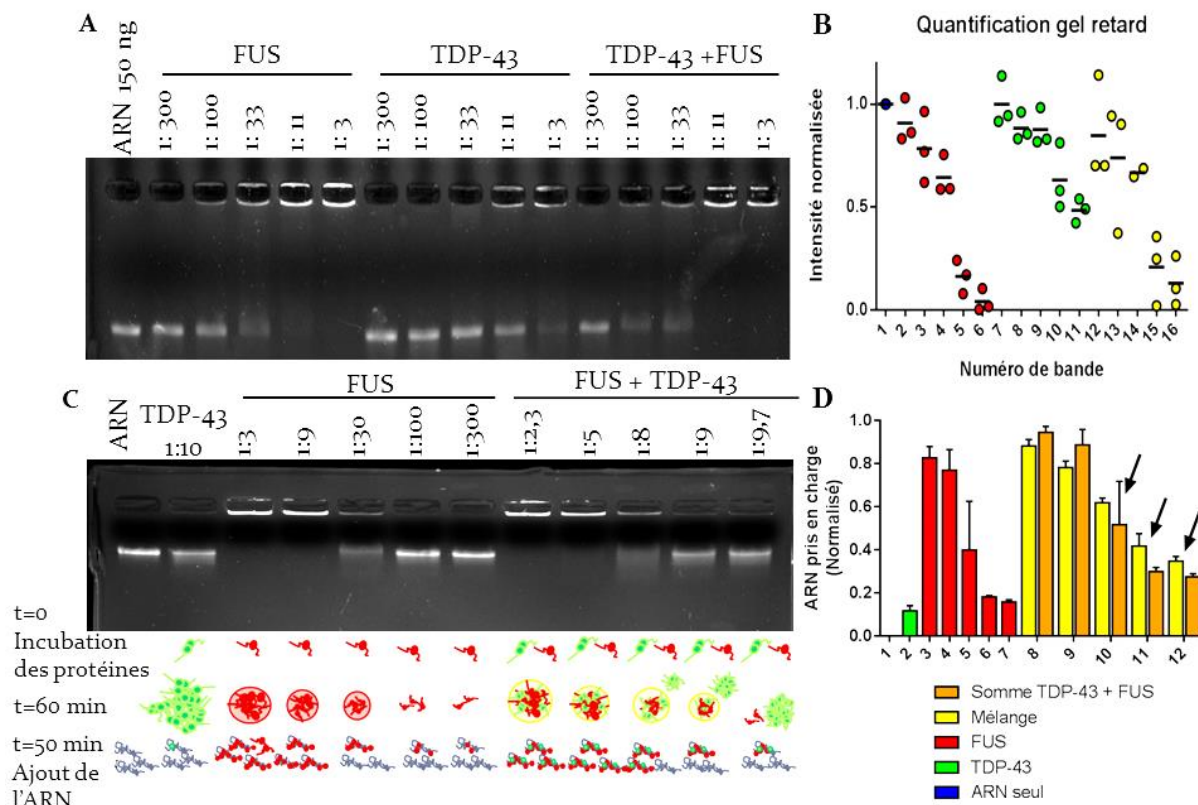
Aperçus des objets en 3 dimensions grâce à l'acquisition confocale (A-D) et vues de différentes coupes en z dans l'épaisseur des objets (a-d). (A) Gouttelette typique de FUS pur parfaitement sphérique, l'objet étant très petit le rendu 3D est délicat. (B) Objet de type grappe, pour un mélange 3/1 FUS/TDP-43 marqué pour TDP-43 en vert, FUS en rouge et l'ARN (sonde oligo dT) en magenta. La flèche met en évidence le signal de TDP-43 inclus dans une gouttelette de FUS. (C) Agrégats pour la condition 1/9 FUS/TDP-43. L'objet est assez large mais plutôt plat (inférieur à 10 µm, il est peu homogène avec des inclusions de FUS et contient de l'ARN. (D) Agrégat de TDP-43 seul, l'objet mesure moins de 10 µm d'épaisseur pour une longueur d'environ 75 µm. Echelles variables indiquées dans chaque condition.

## 5. TDP-43 et FUS coopèrent pour se lier à l'ARN

### a) Première approche par gel retard

De façon à étudier la liaison des protéines avec l'ARN, nous avons effectué un essai de gel retard, une manipulation assez simple qui permet de mettre en évidence l'interaction entre protéine et ARN. Dans cette manipulation la prise en charge de l'ARN par la protéine est mise en évidence lorsqu'un retard de migration de l'ARN est

constaté. Différentes solutions de protéines/ARN ont été préparées avec une quantité fixée d'ARN 2Luc à 150 ng par puits. Les quantités de protéines sont augmentées progressivement pour augmenter le ratio protéine/nucléotide de 1 :300 à 1 :3. Les échantillons, FUS/ARN, TDP/ARN et TDP+FUS/ARN sont incubés quelques minutes avant d'être déposés sur gel d'agarose. Nous pouvons alors observer la migration de l'ARN (marquage au bromure d'éthidium, BET). L'ARN 2Luc seul nous sert de référence. On remarque qu'en présence d'une quantité croissante de FUS, l'ARN est de plus en plus retenu dans le puits, signe d'une accumulation des protéines le long de l'ARN. Ainsi, ce décalage intervient entre les ratios 1 :100 et 1 :33. Pour TDP-43, le décalage de migration intervient plus tard entre 1 :33 et 1 :11, où les complexes commencent à rester dans le puits. Au ratio 1 :3, il reste néanmoins une proportion de l'ARN qui migre signe que TDP-43 se lie de manière coopérative en saturant certains ARN en laissant simultanément d'autres ARN libres. Lorsque l'on mélange les deux protéines en proportion équimolaire, on remarque alors que le décalage intervient dès le ratio 1 :100, ce qui est plus proche du comportement de FUS, et suggère ainsi un effet synergique entre les deux protéines dans leur liaison à l'ARN (**voir figure 56, A**). Cette expérience a été répétée 3 fois avec des résultats similaires, qui ont été quantifiés et présentés sur un même graphique (**voir figure 56, B**). Pour cela l'intensité de la bande de migration de l'ARN 2Luc a été mesurée et normalisée (**voir matériel et méthodes**) pour les différentes conditions de manière à les comparer. On voit, ainsi, qu'en moyenne, les profils de FUS et TDP-43 sont différents et que le mélange se rapproche de celui de FUS. Si les deux protéines se liaient indépendamment, le décalage interviendrait plus tard dans le cas du mélange (puisque la quantité de FUS est divisé par 2 de façon à conserver un ratio équivalent à la quantité totale de protéine). Ces résultats semblent montrer que l'interaction entre FUS et TDP-43 est coopérative (Le terme de coopération se réfère à la définition qui est donnée au **chapitre 1.1b, partie bibliographie**, c'est-à-dire que la présence d'une protéine sur une cible favorise l'addition d'autres protéine dans un but commun).



**Figure 56-Expériences de gel retard montrant la coopération entre TDP-43 et FUS dans leur liaison à l'ARN**

(A) L'ARN en mélange avec différentes concentrations de protéines, seules ou en mélange, est chargé sur un gel d'agarose de façon à étudier sa liaison avec les protéines qui induisent un retard de migration. Pour FUS le décalage intervient entre les ratios 1:100 et 1:33 où l'ARN commence à rester dans le puits signe qu'il est saturé par les protéines. Pour TDP-43, ce décalage intervient plus tard (entre 1:33 et 1:11), tandis qu'avec le mélange des deux protéines le retard intervient très tôt (dès 1:100) ce qui atteste d'un effet synergique dans leur prise en charge de l'ARN par les deux protéines. Mélange équimolaire de protéines. Ratio en protéines : nucléotides (1 protéine pour 300 nucléotides). (B) La quantification sur 3 expériences indépendantes permet de mettre en évidence l'effet de retard de migration de l'ARN en mesurant l'intensité de la bande de migration rapportée à l'intensité totale du puits (et normalisée par rapport à l'intensité de la bande de l'ARN nu). (C, haut) Gel de contrôle de la solubilisation de TDP-43 dans la phase de FUS, dans celui-ci l'ARN est ajouté après l'incubation de TDP-43 (3  $\mu$ M) ou d'une gamme de FUS (9 ; 3 ; 1 ; 0,3 ; 0,1  $\mu$ M), ou de l'addition des deux protéines (suivant les mêmes concentrations). (C, bas) Schéma illustrant l'évolution du comportement des protéines seules ou en mélange suivant les données *In vitro* et permettant d'expliquer que la prise en charge de l'ARN est supérieure dans le cas du mélange car une proportion de TDP-43 est solubilisée dans la phase de FUS. (D) Quantification associée au gel C. Les flèches indiquent les conditions où la prise en charge dans le cas du mélange est supérieure à la somme de l'ARN pris en charge dans le cas des protéines seules (addition des conditions 2 + 3, 4, 5, 6, 7). Données issues de 3 gels indépendants.

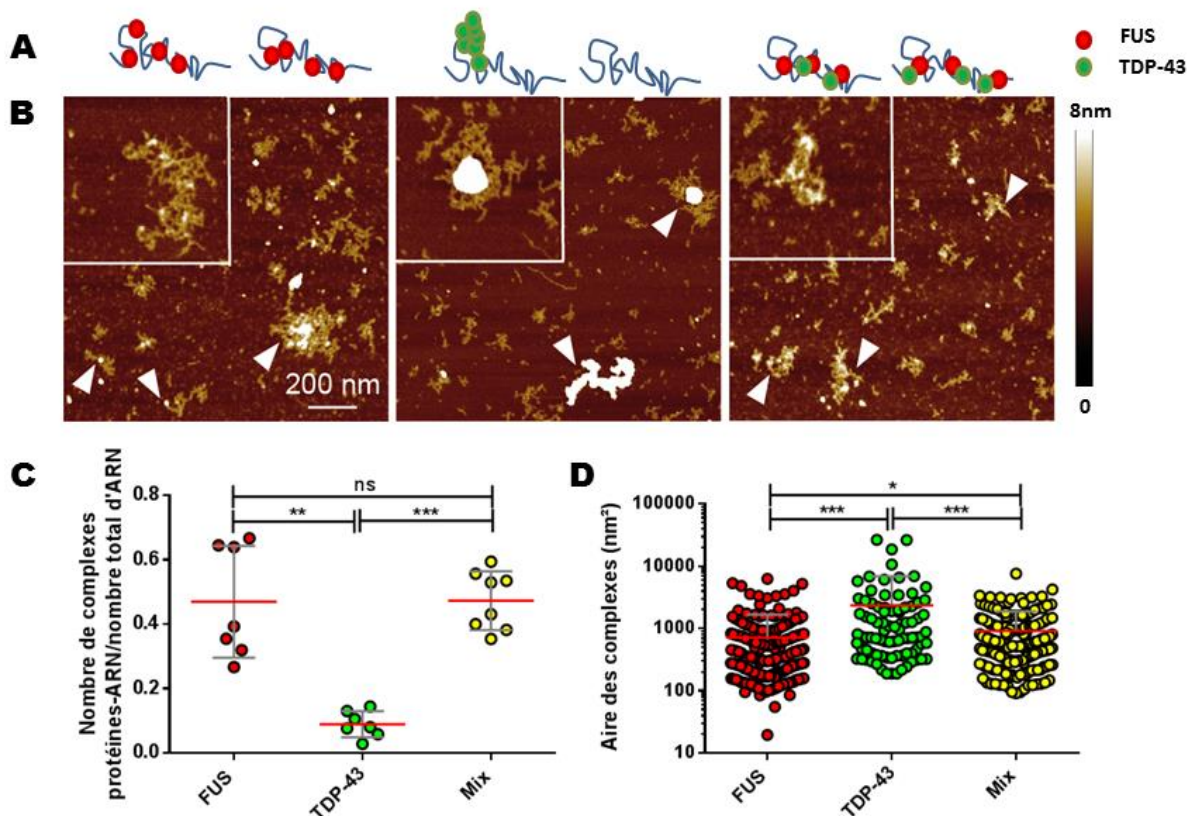
De manière à vérifier le phénomène de la solubilisation de TDP-43 dans la phase de FUS et à montrer que cela participe à une meilleure prise en charge de l'ARN, une manipulation de gel retard supplémentaire a été effectuée (**voir figure 56, C**). Dans cette dernière les protéines sont d'abord incubées seules puis l'ARN est ajouté dans

un deuxième temps (50 min après le début de l'incubation, contrairement à la manipulation précédente où il est ajouté dès le début). Au regard des expériences effectuées *In vitro*, on peut supposer que TDP-43 va s'agrèger en l'absence d'ARN mais que FUS va rester sous forme de condensat. Ainsi, la prise en charge de l'ARN dans le cas de TDP-43 est très faible puisque l'ajout d'ARN ne permet pas de solubiliser la protéine agrégée, tandis que dans le cas de FUS les protéines restent libres pour la prise en charge de l'ARN (**voir figure 56, D**). C'est ce que l'on observe sur le gel avec un gamme de concentration de FUS (ratio de FUS :nucléotide de 1 :3 à 1 :300) où la quantité d'ARN pris en charge dépend directement de la proportion de FUS. Ainsi, dans le cas du mélange (quantité de TDP-43 fixée additionnée à la gamme de FUS), on observe un retard semblable avec des bandes d'ARN moins intenses. En quantifiant ces observations sur 3 gels effectués indépendamment, on observe ainsi que la prise en charge de l'ARN est supérieure dans le cas du mélange que dans le cas de la somme de la prise en charge de FUS et de TDP-43 individuellement (**voir figure 56, E**). Cela montre que la solubilisation de TDP-43 dans la phase de FUS augmente la prise en charge de l'ARN.

#### **b) Influence de FUS sur l'interaction de TDP-43 avec l'ARN révélée par AFM**

L'AFM est une technique qui permet d'obtenir des informations à l'échelle de la molécule unique, grâce à une pointe qui scanne la surface d'un échantillon déposé sur une surface de mica. De cette manière, il est tout à fait possible d'observer l'ARN qui forme des filaments caractéristiques, ainsi que les protéines sous forme de petits points à l'échelle de quelques nanomètres. Différents échantillons ont été préparés pour les protéines PL (à la concentration de 500 nM) avec l'ARN pour un ratio 1/100 (1 protéine pour 100 nucléotides). Après 90 min d'incubation les échantillons ont été déposés sur le mica pour observation. On observe alors dans le cas de FUS, des complexes avec une bonne répartition de la protéine sur de nombreux ARN (**voir figure 57, B**). Ainsi, environ 50 % des ARN présents sur la surface sont pris en charge (**voir figure 57, C**). Dans le cas de TDP-43, la protéine s'agglomère par le biais d'auto-interactions sur un petit nombre d'ARN, ce qui peut être lié à sa capacité à reconnaître des séquences très spécifiques. On observe ainsi des complexes très denses dont la hauteur dépasse les 8 à 10 nm. Ce comportement est dû à la propriété de la protéine à s'auto-assembler, et non le fait de l'ARN, puisque nous sommes en condition d'un excès d'ARN (ratio 1/100). Moins de 10 % des ARN présents sur la surface sont liés à la protéine, mais ceux qui le sont forment des complexes avec une aire importante (**voir figure 57, D**). En mélange, on observe un nombre important d'ARN pris en charge avec une bonne répartition des protéines sur ces derniers et

surtout une absence d'objets très denses. Les complexes hauts et larges que l'on observe dans le cas de TDP-43/ARN ne semblent pas se former dans le cas du mélange (**voir tableau comparatif, figure 66**). C'est de nouveau 50 % des ARN qui sont pris en charge comme pour FUS, avec une aire moyenne plus réduite. Des essais complémentaires avec des temps d'incubation et de ratios ARN/protéine différents ont données des résultats similaires. Ces observations suggèrent que FUS permet à TDP-43 de se répartir sur l'ARN par le biais d'une interaction coopérative.



**Figure 57-TDP-43 se répartie sur l'ARN grâce à FUS.**

(A) Représentation schématique du comportement des protéines, FUS se répartie sur toute la longueur de l'ARN et prend en charge un nombre important d'ARN. TDP-43 se regroupe sur un petit nombre d'ARN en formant des objets compacts. En mélange, les deux protéines se répartissent sur une proportion importante d'ARN. (B) Acquisitions représentatives en AFM des protéines seules et en mélange à la concentration de 500 nM, avec un ratio en ARN de 1/100 (prot/nu), après 90 min d'incubation. Echelle : 200 nm de largeur pour 8 nm de hauteur. (C) Comptages du nombre de complexes protéines-ARN rapportés au nombre total d'ARN sur une surface identique pour chaque condition. (D) Aire des complexes mesurant plus de 4 nm de hauteur pour chaque condition. \*:  $p < 0,05$ ; \*\*:  $p < 0,01$ ; \*\*\* :  $p < 0,0001$

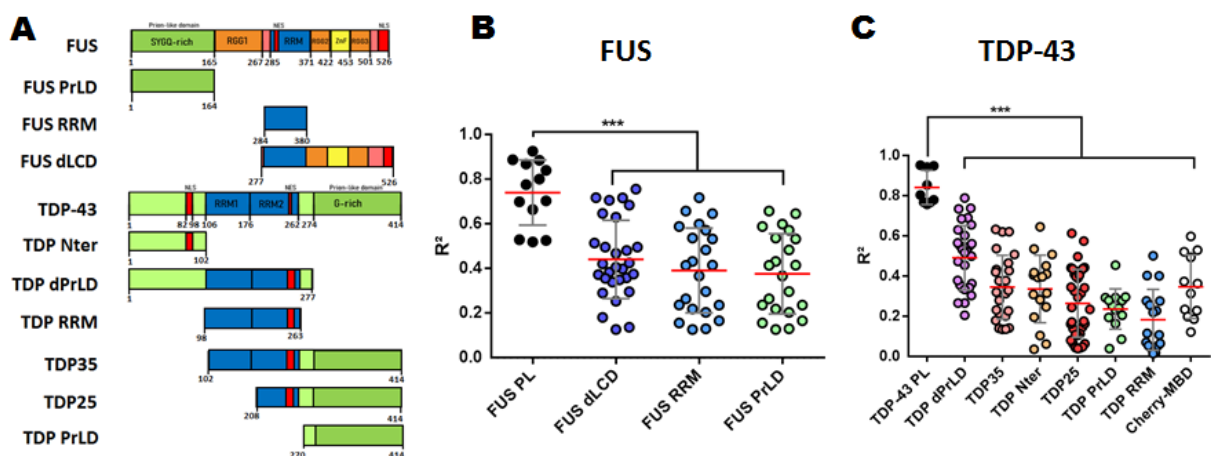
Maintenant, que nous avons montré que TDP-43 et FUS étaient capables d'interagir ensemble, et que cette interaction favorisait leur coopération dans la prise en charge de l'ARN, nous allons nous intéresser plus spécifiquement aux domaines impliqués dans cette interaction mais aussi aux formes tronquées qui peuvent perturber cette interaction.

## II. Etude des formes tronquées

### 1. Mesure de l'interaction des formes tronquées avec les formes PL grâce au MtB

#### a) Les formes tronquées induisent une séparation avec les formes PL

Des formes tronquées variées, ont été produites à la fois pour FUS et TDP-43 de façon à mieux étudier ces protéines, et notamment l'implication des différents domaines dans leur comportement. En première approche, nous avons utilisé le MtB pour mieux cerner la miscibilité des formes tronquées avec les protéines pleine longueur. Des plasmides codants ces formes tronquées ont donc été produits pour l'expression cellulaire. Ensuite, des plasmides pour l'expression bactérienne et la production de certaines de ces formes tronquées ont été produits pour les étudier *In vitro*. Concernant le MtB, les formes FUS PrLD, RRM et dLCD ont été étudiées pour FUS (**voir figure 58, A**). Tandis que pour TDP-43, les formes TDP Nter, dPrLD, RRM, PrLD, ainsi que deux formes clivées pathologiques TDP25 et TDP35 (**voir chapitre II.4.b, partie biblio**) ont été étudiées. Comme pour l'étude des formes PL, des cellules ont été transfectées avec des plasmide RBP-GFP-MBD et RBP-RFP-MBD. Ici, RBP correspond soit à une protéine PL soit à une forme tronquée. Nous avons d'abord étudié la miscibilité d'une protéine PL avec ses propres formes tronquées, avant de croiser les formes tronquées de TDP-43 avec FUS PL ou inversement, de façon à identifier les domaines impliqués dans l'interaction de TDP-43 et FUS.



**Figure 58-Mesure de la miscibilité des formes tronquées avec les formes PL grâce au microtubule Bench.**

(A) Différentes formes tronquées ont été produites pour FUS et TDP-43 de façon à étudier l'influence des différents domaines dans la miscibilité des protéines. (B) Graphique rassemblant les données des coefficients  $R^2$  mesurés pour FUS pleine longueur en mélange avec ses propres formes tronquées. (C) Graphique rassemblant les données des coefficients  $R^2$  mesurés pour TDP-43 pleine longueur en mélange avec ses propres formes tronquées. \*:  $p < 0,05$  ; \*\*:  $p < 0,01$  ; \*\*\* :  $p < 0,0001$ .

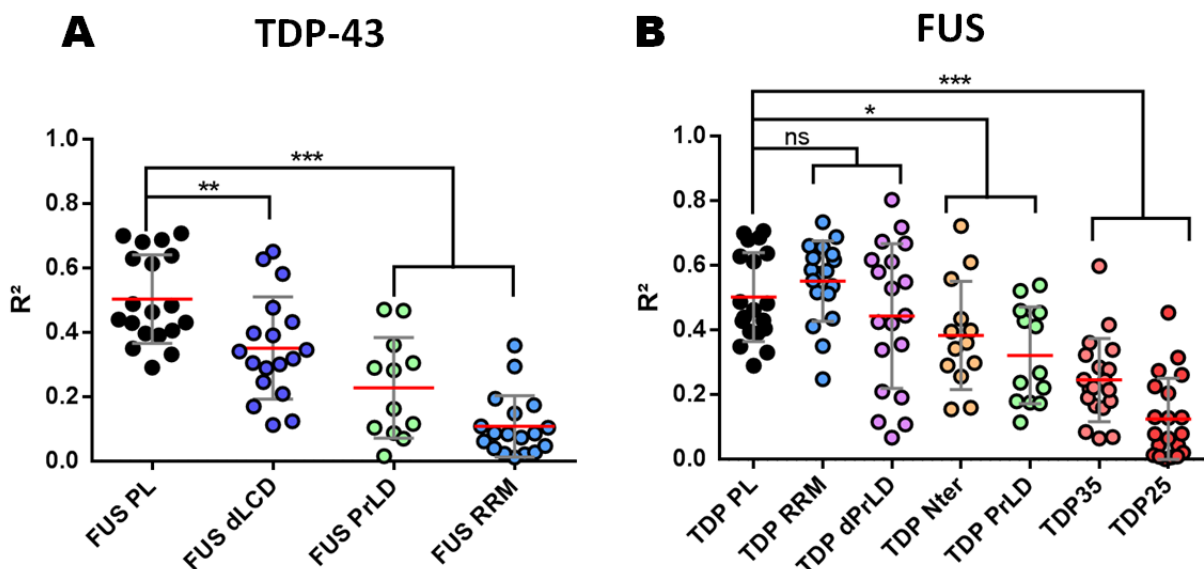
La corrélation d'une vingtaine de cellules par condition a été étudiée par le biais de l'analyse automatique, qui permet de générer un coefficient  $R^2$  par cellule. L'ensemble de ces coefficients ont été rassemblés sur un même graphique. Des transfections avec des couples de plasmides TDP-43-RFP-MBD/TDP-43-GFP-MBD d'un côté, et FUS-RFP-MBD/FUS-GFP-MBD d'un autre côté, ont été effectuées en contrôle pour lesquels on obtient un coefficient de référence reflétant le mélange homogène de mêmes protéines. Par ailleurs, la protéine Cherry qui n'a pas d'affinité particulière pour TDP-43 a été utilisée comme contrôle négatif. Nous avons donc un intervalle borné entre un mélange homotypique présentant une forte corrélation et une interaction aléatoire avec une interaction très faible, inférieur à 0,4. Ainsi, on remarque pour FUS (**voir figure 58, B**) une faible corrélation dès lors que la protéine est tronquée, elle a donc besoin à la fois du RRM et de son domaine de faible complexité pour former une phase homogène. Bien que le domaine de faible complexité soit impliqué dans les interactions protéine-protéine, à lui seul il ne suffit pas à former un bon mélange. En effet, pour FUS de nombreux domaines sont impliqués dans la liaison à l'ARN (Les RGG1/RGG2/RGG3, le RRM ou encore le ZnF), l'ARN étant au cœur de la miscibilité des RBP, en particulier dans le contexte du MtB, on peut donc supposer que dès lors que l'un de ces domaines est tronqué, la liaison à l'ARN est impactée, et donc le mélange est moins homogène. Concernant TDP-43 (**voir figure 58, C**), l'ensemble des formes tronquées testées induisent une diminution de la corrélation avec la forme PL avec des coefficients significativement moins élevés que la référence, en revanche, il y a des disparités dans les formes tronquées. Ainsi, la forme dPrLD chez qui il manque le domaine PrLD, mais qui conserve les domaines RRM et Nter, a une meilleure affinité que les autres formes tronquées. Il apparaît alors que le domaine PrLD, pourtant connu pour favoriser les contacts protéine-protéine, à lui seul ne peut promouvoir un bon mélange, il induit, par ailleurs, une forte séparation avec la forme pleine longueur. Il apparaît que ce domaine PrLD a plutôt tendance à s'auto-assembler de son côté, tout en étant incapable de se lier à l'ARN, ce qui provoque la séparation. Ensuite, la forme pathologique TDP35 est peu corrélée mais légèrement mieux que les autres en moyenne, avec un coefficient de corrélation au même niveau que le contrôle cherry-MBD, ce qui indique que cette interaction est non spécifique, pourtant TDP35 est connue pour interagir et séquestrer la forme pleine longueur dans le contexte pathologique[332]. Cet effet peut être dépendant du contexte, et la présence du MBD peut permettre à TDP35 de rester fonctionnelle le long des microtubules et non agrégée. Cette forme a perdu le Nter mais conserve la fonctionnalité de ses RRM. Nous verrons par la suite, qu'elle a tout de même un comportement assez proche de la protéine pleine longueur. La forme TDP25, quant à elle, est assez mal corrélée (en



dessous du contrôle GFP-Cherry). Ce qui suggère que son interaction avec la forme pleine longueur est affectée, d'autant que sa capacité à se lier à l'ARN est altérée du fait du clivage au niveau du RRM2. Étonnamment, le RRM est la forme tronquée avec le plus faible coefficient de corrélation, en dessous du contrôle Cherry-MBD. Pour résumer, TDP-43 s'assemble très bien avec elle-même mais la perte d'un de ces domaines affecte cette capacité, les domaines Nter et RRM apparaissent essentiels à cette interaction tandis que le domaine PrLD a tendance à favoriser la séparation, car il s'assemble avec lui-même et l'ARN ne peut pas jouer son rôle de médiateur. Ainsi, le RRM comme le Nter seuls, ne suffisent pas à induire une bonne miscibilité, puisque l'auto-interaction de la protéine PL mise en place par le domaine PrLD favorise sa séparation.

### b) L'interaction entre FUS et TDP-43 est favorisée par le RRM mais est rompue par les formes tronquées pathologiques de TDP-43

Par la suite, la miscibilité des formes tronquées a été testée de manière croisée avec les formes PL : TDP-43 avec les formes tronquées de FUS et inversement. Le mélange TDP-43 et FUS sert de référence avec un coefficient moyen autour de 0,5 pour un ratio 1/1. Ainsi, les formes FUS dLCD, PrLD et RRM sont toutes moins bien corrélées que la référence, indiquant que l'ensemble des domaines de FUS sont nécessaires pour la bonne interaction avec TDP-43 (**voir image 59, A**). La forme dLCD, qui possède à la fois le RRM et les RGG2 et RGG3 (mais ni le PrLD ni le RRG1),

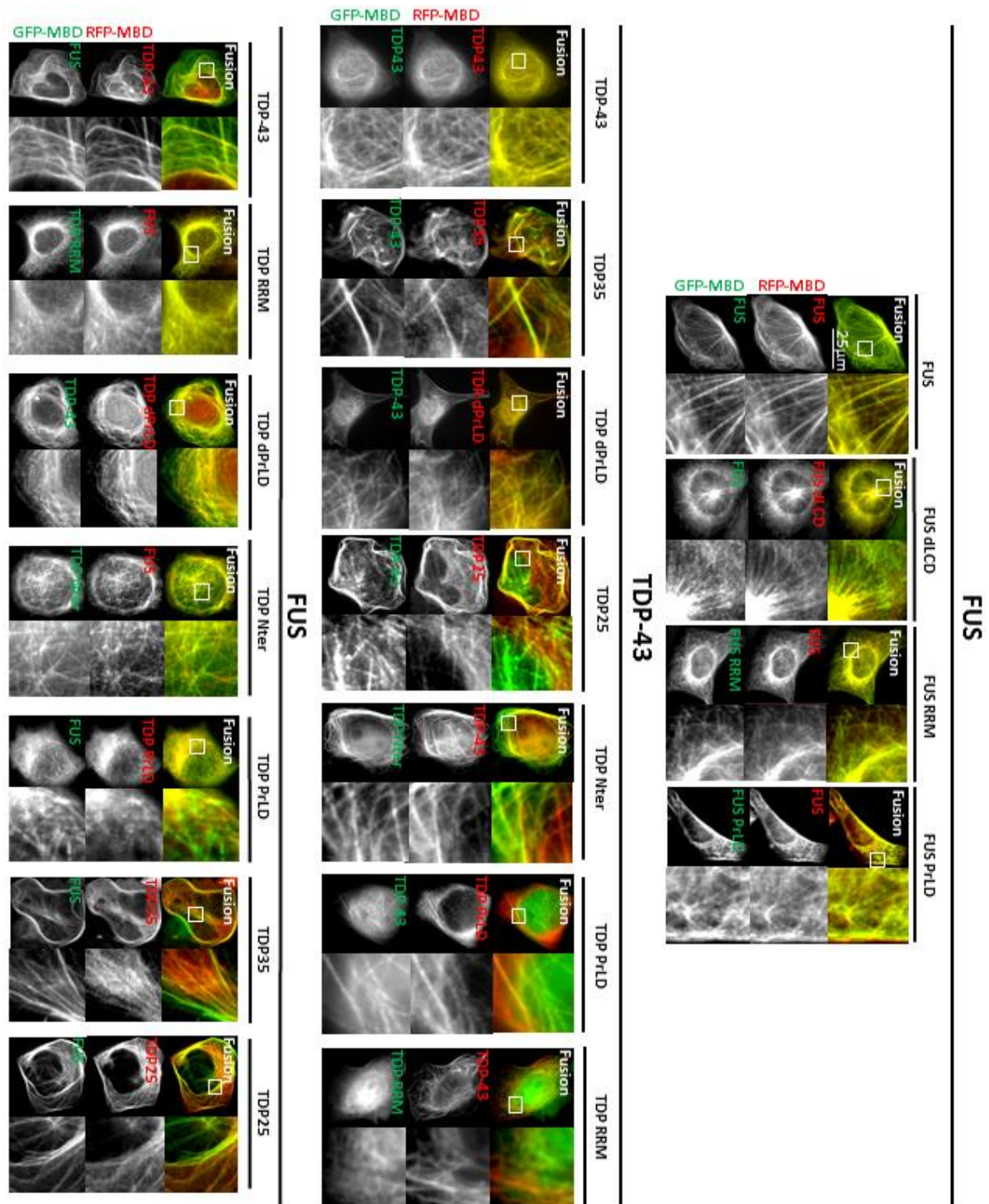


**Figure 59-Le RRM joue un rôle important dans l'interaction entre FUS et TDP-43, tandis qu'elle est affectée par les formes pathologiques de TDP-43.**

(A) Graphique rassemblant les données des coefficients  $R^2$  mesurés pour les formes tronquées de FUS avec TDP-43 pleine longueur, interaction moyenne TDP-43/FUS en contrôle pour un ratio 1/1. (B) Graphique rassemblant les données des coefficients  $R^2$  mesurés pour FUS pleine longueur en mélange avec diverses formes tronquées de TDP-43, dont deux formes pathologiques TDP25 et TDP35. (D). \*:  $p < 0,05$ ; \*\*:  $p < 0,01$ ; \*\*\*:  $p < 0,0001$ .

est la mieux corrélée de toutes, indiquant que le RRM et les domaines de faible complexité peuvent jouer un rôle, sauf que le RRM seul n'est pas suffisant pour induire un mélange homogène, il est même fortement décorrélié. D'un autre côté, nous savons que TDP-43 a tendance à former des phases homotypiques, on peut donc supposer qu'il accepte difficilement d'autres protéines d'autant plus quand ces protéines ne sont pas pleinement fonctionnelles (car il leur manque soit l'interaction avec l'ARN, soit les domaines de faible complexité, et que les deux sont nécessaires en même temps).

Du côté de FUS PL, deux formes tronquées sortent du lot avec une interaction semblable à la protéine TDP-43, il s'agit de TDP-RRM et TDP-dLCD (**voir image 59, B**). Le RRM présent dans les deux cas, semble donc jouer un rôle prépondérant dans cette interaction, quand le domaine prion, pourtant connu pour être impliqué dans les interactions protéine-protéine, n'est pas nécessaire. Le domaine PrID, avec le Nter, semblent toutefois capable de se mélanger avec la protéine FUS mais en moindre proportion. En revanche, les deux formes clivées pathologiques TDP35, et de manière plus extrême TDP25, sont fortement décorréliées avec FUS (**voir images mises en évidence, figure 60**). En dépit, de la présence des RRM, même si le RRM1 est altéré chez TDP25, ces deux formes pathologiques présentent une forte séparation. Certainement car elles préfèrent s'auto-assembler et ont une propension à l'agrégation plus importante que la protéine TDP-43 PL[318].



**Figure 60-Images représentatives pour les différentes formes tronquées en mélange avec les formes PL, analysées au MtB.**

On cherche à mesurer la miscibilité des formes tronquées avec les formes PL, en les exprimant dans des cellules U2OS, conjuguées au MBD de façon voir si elles se mélangent ou se séparent le long des microtubules. (A) Mélanges de FUS avec différentes formes tronquées correspondantes. (B) Mélanges de TDP-43 avec différentes formes tronquées correspondantes. (C) Mélanges croisés de FUS avec différentes formes tronquées de TDP-43. Echelle : 25 µm (Zooms : 12µm, identiques dans chaque condition)

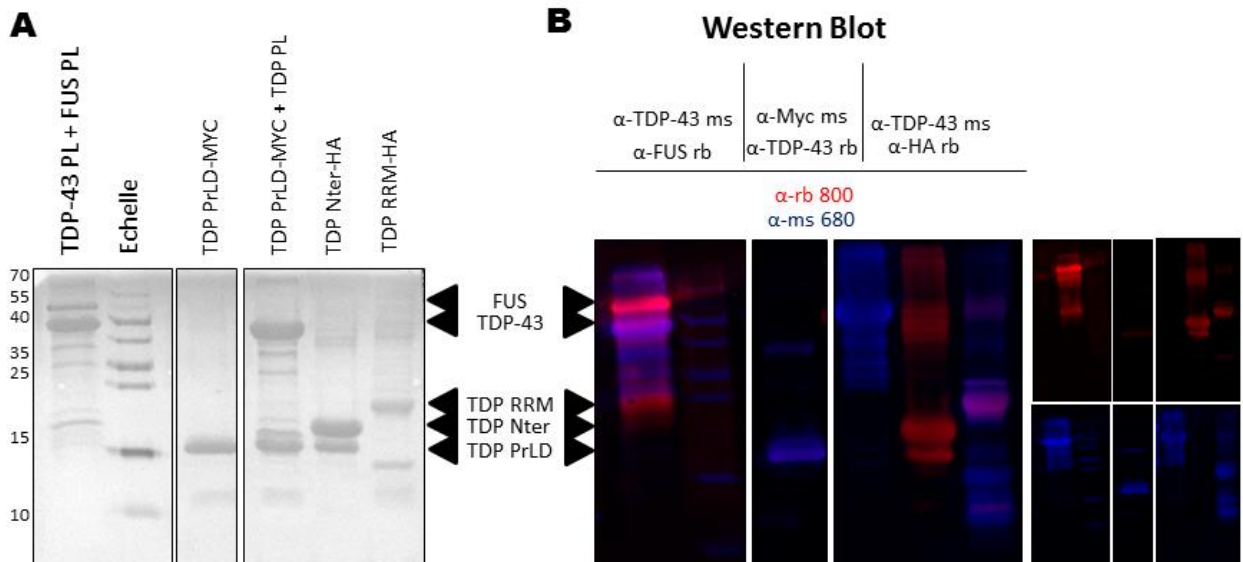
Ainsi, on peut supposer que l'interaction entre FUS et TDP-43 s'opère au contact de l'ARN en impliquant en particulier le RRM de TDP-43. Le domaine de faible complexité de TDP-43 semble plutôt induire une séparation (parce qu'il préfère interagir avec lui-même). Cependant, cette technique ne permet pas de savoir si ce contact se fait de manière directe ou par le biais de l'ARN. Ces observations sont cohérentes avec ce qui avait été décrit dans la littérature : le RRM2 et le PrLD de TDP-43 jouent un rôle dans l'interaction avec FUS, mais seul, le domaine PrLD n'est pas suffisant[14]. On peut néanmoins s'interroger sur les formes clivées pathologiques TDP25 et TDP35 qui conservent le RRM2 et le PrLD mais qui toutefois présentent une très forte décorrélation avec FUS. On peut supposer que leur capacité à auto-interagir est telle qu'elles ne se mélangent pas bien avec FUS PL. Nous allons les étudier plus en détail grâce aux observations *in vitro* par la technique d'immunofluorescence et l'AFM.

## 2. Etude des formes tronquées *In vitro*

### a) Contrôle des formes tronquées produites

Certaines formes tronquées ont été sélectionnées pour être produites et étudiées *In vitro*, c'est le cas notamment de TDP-PrLD, TDP-Nter, TDP-RRM, TDP25 et TDP35 et FUS-PrLD. Ces formes ont été produites avec des étiquettes (Myc ou Ha) de façon à les visualiser en fluorescence. Avant d'aller plus loin dans la caractérisation de ces formes tronquées, un test par *western blot* a été effectué pour vérifier le bon état de ces productions et la reconnaissance des étiquettes. Différents échantillons contenant les formes tronquées et les protéines PL ont été chargés sur gel d'acrylamide, puis des anticorps ont été utilisés pour marquer les protéines transférées sur une membrane. Un anticorps anti-Ha est utilisé pour marquer les formes tronquées TDP-Nter-Ha et TDP-RRM-Ha, et un anticorps anti-Myc pour la forme TDP PrLD-Myc (**Voir image 61**). Les formes TDP25 et TDP35 ne sont pas présentes sur ce contrôle mais TDP25 a été produite avec l'étiquette HA, tandis que TDP35 a été produite avec l'étiquette Myc. De ce contrôle, il ressort que les formes tronquées sont bien reconnues grâce à leurs étiquettes avec une bonne sensibilité. Il n'y a pas de croisement entre la forme TDP-PrLD et TDP-43 mais l'anticorps de TDP-43 (*ms*) utilisé reconnaît la forme TDP-RRM. Des tests similaires ont été effectués pour les formes TDP25 et TDP35 (données non montrées), et il ressort aussi un croisement de l'anticorps  $\alpha$ -TDP-43 pour ces formes tronquées (Epitope au niveau du RRM). Il

faudra garder cela en mémoire quand les formes tronquées seront étudiées par immunofluorescence par la suite.

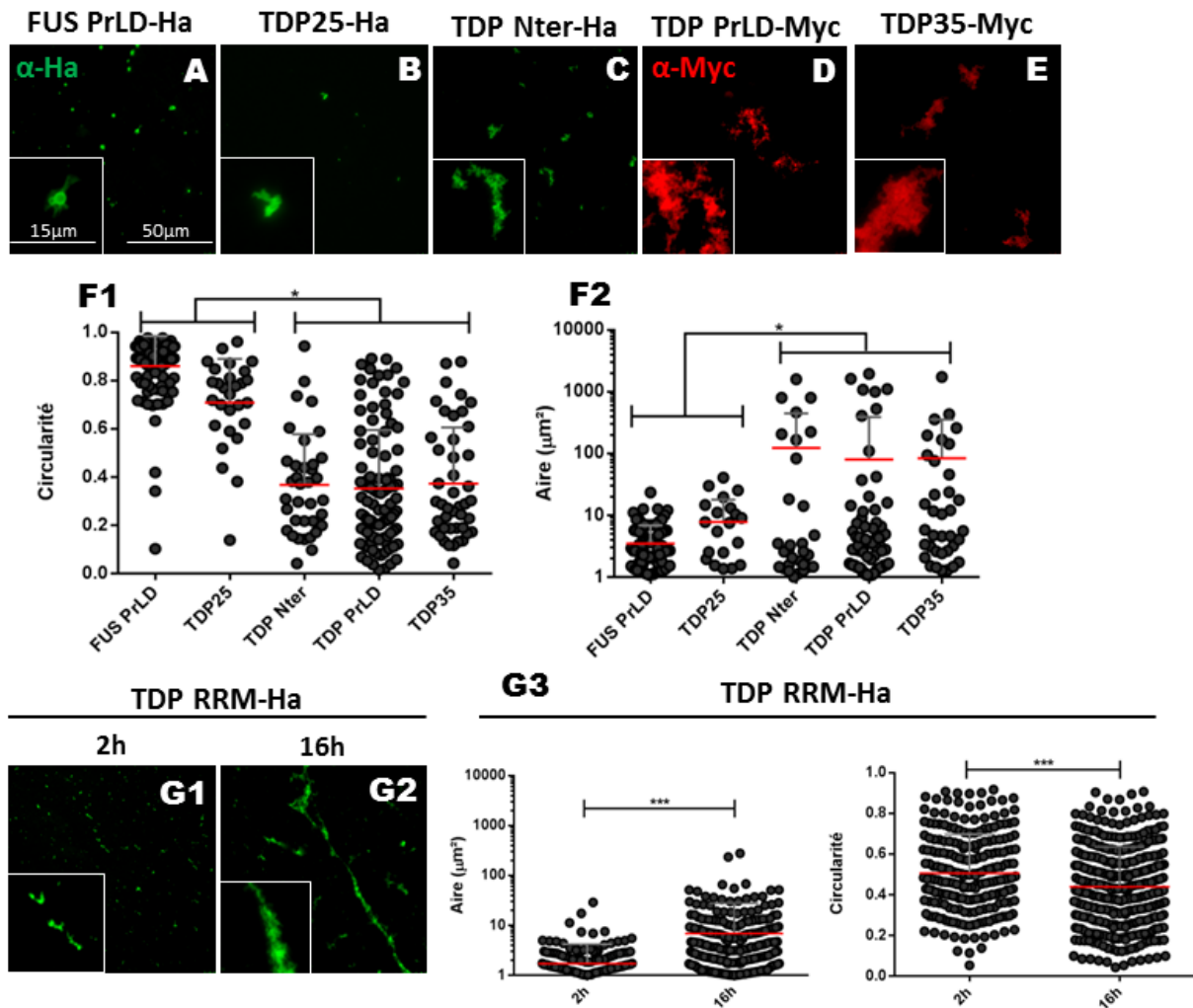


**Figure 61-Contrôle des formes tronquées des protéines par Western Blot.**

(A) Les formes tronquées produites ainsi que les protéines pleine longueur (PL) ont été chargées sur gel acrylamide 15% pour migration. La coloration au rouge ponceau permet de mettre en évidence les protéines totales. (B) Différents anticorps primaires ont été utilisés pour cibler les protéines et les formes tronquées portant des étiquettes HA ou Myc, puis des anticorps secondaires (IR dye) ont été utilisés pour marquer les protéines. Les anticorps primaires lapin (*rb*) sont révélés en rouge, tandis que les anticorps primaires souris (*ms*) sont révélés en bleu. Les formes TDP PrLD, TDP Nter, et TDP RRM sont bien reconnues grâce à leurs étiquettes Myc ou Ha, il n'y a pas de croisement entre TDP-43 et la forme TDP PrLD, en revanche l'anticorps TDP-43 *ms* reconnaît la forme TDP RRM.

### b) Aperçu du comportement des formes tronquées seules par immunofluorescence en microscopie optique

De la même manière que ce qui a été fait précédemment pour les formes PL, une première approche en immunofluorescence par microscopie optique a permis d'obtenir un aperçu du comportement des formes tronquées *In vitro*. Les formes tronquées FUS PrLD, TDP25, TDP Nter, TDP PrLD et TDP35 ont été incubées à la concentration de 5  $\mu$ M et déposées après 2 h d'incubation. Les trois premières ont été marquées par un anticorps anti-Ha, tandis que les deux dernières ont été marquées par un anticorps anti-Myc correspondant à l'étiquette avec laquelle elles ont été produites (**voir figure 62**). Pour la seule forme tronquée de FUS présentée ici, FUS PrLD on retrouve un comportement assez proche de la protéine PL avec la formation de petites gouttelettes mais légèrement déformées. Ces objets, tout comme ceux formés par FUS PL, sont caractérisés par une circularité importante et



**Figure 62-Aperçu du comportement des formes tronquées par immunofluorescence.**

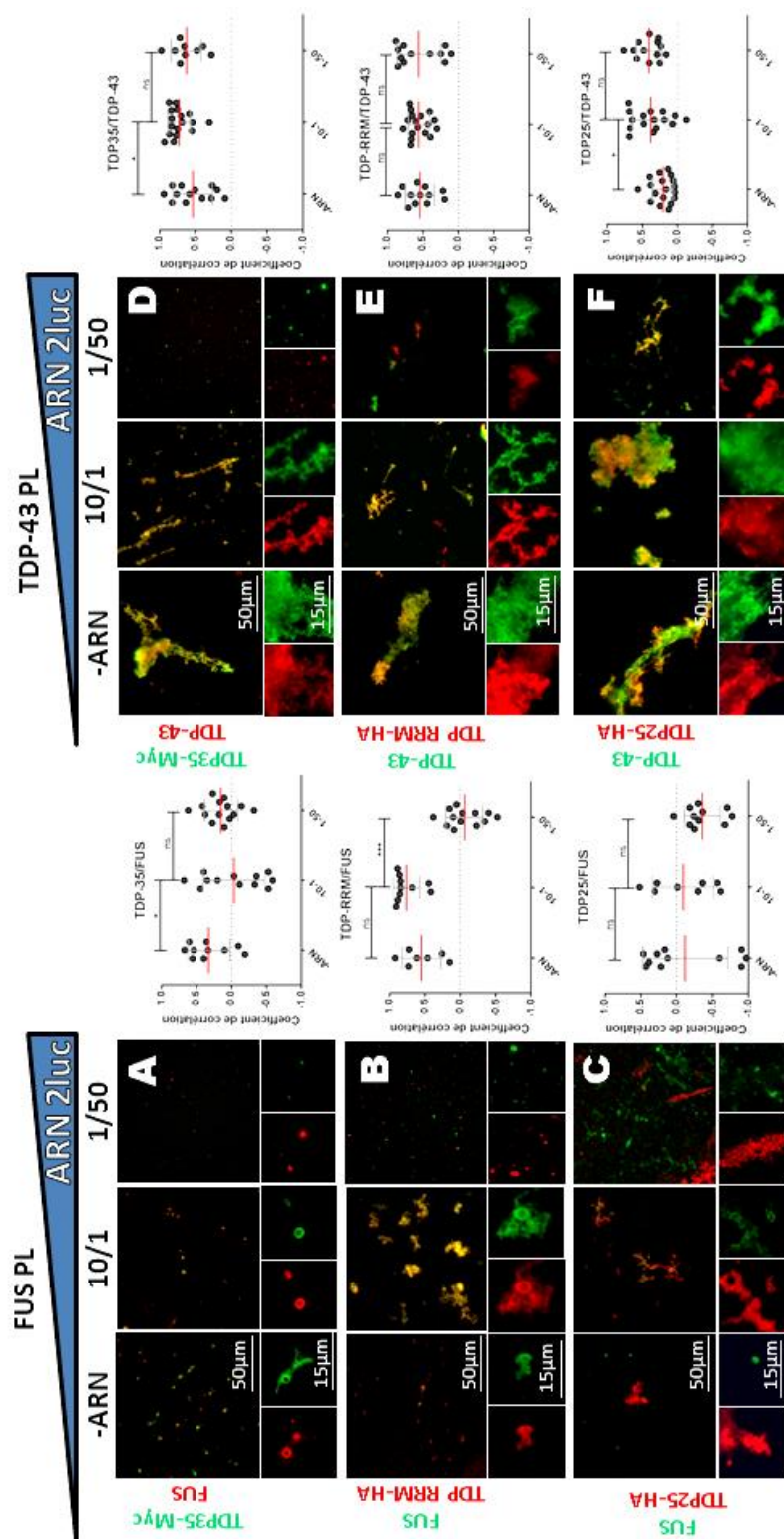
(A-F) Acquisitions des formes tronquées seules par immunofluorescence marquées par les anticorps anti-Ha (Vert) et anti-Myc (Rouge). L'ensemble des formes tronquées ont été incubées à la concentration de 5  $\mu\text{M}$ , 2 h avant dépôt : (A) FUS PrLD-HA, (B) TDP25-HA, (C) TDP Nter-Ha, (D) TDP PrLD-Myc, (E) TDP35-Myc. (G1, G2) La forme TDP RRM-Ha incubée à la concentration de 7,5  $\mu\text{M}$  est présentée pour deux temps d'incubation, pour montrer son évolution sous forme d'objets fibrillaires. (G3) Aire et circularité pour TDP RRM-Ha dans les 2 conditions présentées de 2 h et 16 h d'incubation avant fixation. (F1) Valeurs de circularité mesurées pour l'ensemble des formes tronquées. (F2) Aires mesurées pour l'ensemble des formes tronquées. Mesures effectuées pour 5 images dans chacune des conditions. Echelle : 50  $\mu\text{m}$  (fenêtre large), 15  $\mu\text{m}$  (Zoom). \*\*\*:  $P < 0,0001$  ; \*\*:  $P < 0,01$  ; \*:  $P < 0,05$

une aire réduite (**voir figure 62, A, F1 et F2**). Les agrégats des formes TDP Nter, TDP35, ou encore TDP PrLD sont assez semblables à ceux que peut former la protéine TDP-43 PL. Ce sont des gros agrégats amorphes caractérisés par une faible circularité (**voir figure 62, C, D et E**). D'un autre côté, TDP25 forme des petits agrégats, qui diffèrent un peu de ceux observés dans les autres cas, avec une circularité légèrement plus élevée (**voir figure 62, B**). Enfin, TDP RRM qui est plutôt soluble au premier abord, forme de tous petits objets après 2 h d'incubation (**voir figure 62, G1, G2 et**

**G3**), mais se révèle pouvoir former des gros agrégats (en l'absence d'ARN) d'apparence très filiformes, après 16 h d'incubation. Cette capacité du RRM à pouvoir former des fibres avait déjà été mise en évidence[204]. Cependant, ce qui nous intéresse ici, c'est la capacité de TDP-43 à interagir avec FUS par le biais des RRM, nous éviterons donc ces conditions par la suite d'autant qu'en présence d'ARN, le RRM reste soluble. Maintenant que nous avons un premier aperçu du comportement de ces formes tronquées *In vitro*, nous en avons sélectionné certaines, pour étudier l'influence de ces formes tronquées sur les protéines PL.

### c) Comportement des formes tronquées en mélange avec l'ARN et les formes PL

Maintenant, nous allons présenter les résultats les plus pertinents, en particulier concernant les deux formes pathologiques TDP25 et TDP35 mais aussi TDP-RRM. Selon le MtB les deux premières interagissent peu avec FUS tandis que la dernière est bien mélangée. Nous allons donc voir ce qu'il en est, et particulièrement les conséquences que cela peut avoir en présence d'ARN. Trois conditions ont donc été retenues dans chaque cas, sans ARN, avec une faible quantité d'ARN (10/1, 10 protéines par nucléotide) et un excès d'ARN (1/50 soit 1 protéine pour 50 nucléotides). Rappelons que l'ARN est particulièrement important dans la régulation des comportements des RBP, nous avons vu dans le cas des protéines saines (en particulier pour TDP-43) qu'à faible concentration l'ARN induisait un grossissement des agrégats, tandis qu'à haute concentration, il induisait une solubilisation des protéines (**voir figure 52 et 53**). Ainsi, nous souhaitons vérifier ici, comment cet effet peut être affecté dans le cas de formes tronquées. Des échantillons mélangeant protéines PL et formes tronquées ont été préparés à la concentration de 5  $\mu$ M (pour chaque protéine) avec des quantités variables d'ARN. Après 2 h d'incubation ces échantillons ont été déposés, puis fixés, avant de procéder au protocole de marquage d'immunofluorescence. Des anticorps primaires spécifiques aux formes PL ont été utilisés comme dans le cas des expériences précédentes, en revanche des anticorps  $\alpha$ -HA et  $\alpha$ -Myc ont été utilisés pour cibler les formes tronquées produites avec ces étiquettes.



**Figure 63-Le comportement des formes tronquées pathologiques est différent des formes PL par immunofluorescence *In vitro***

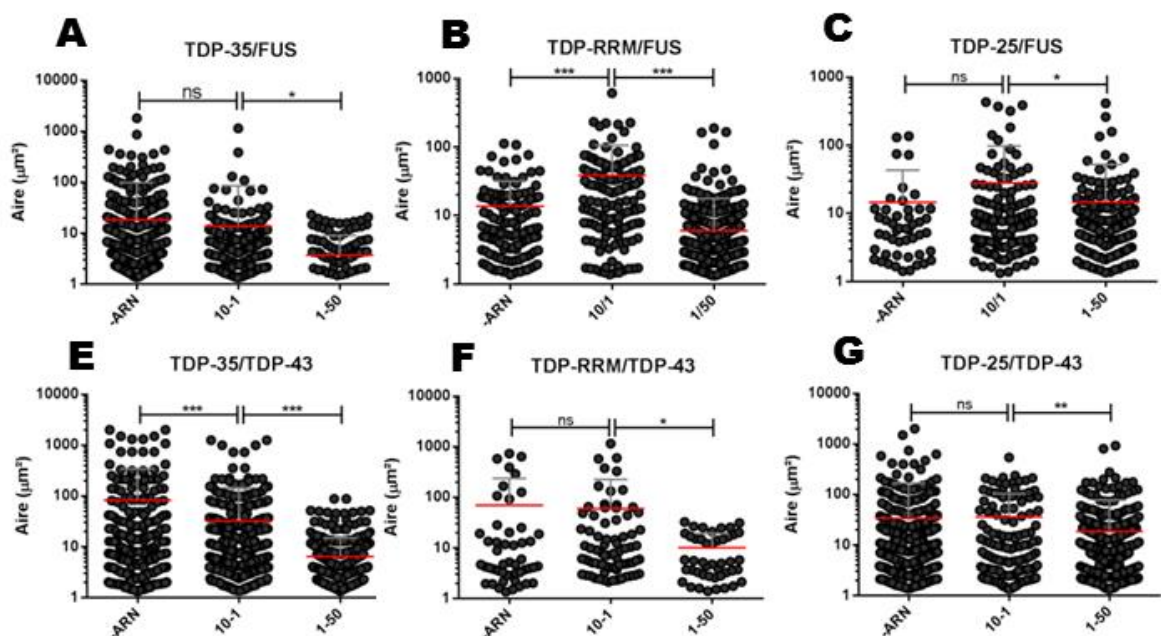
(A-C) Formes tronquées en mélange avec FUS pleine longueur (PL), échantillons incubés 2 h avant dépôt en mélange équimolaire à la concentration de 5 µM, sans ARN, peu d'ARN (10/1) et un excès d'ARN (1/50). (A) TDP35 (B) TDP-RRM et (C) TDP25. Les coefficients de corrélation mesurés dans chaque condition sont présentés sur les graphiques et indiquent la miscibilité des deux protéines au sein des objets. (D-F) Idem pour TDP-43 avec (D) TDP35 (E) TDP-RRM et (F) TDP25. Anticorps primaires spécifiques à la forme pleine longueur, et à l'étiquette (HA ou Myc) dans le cas des formes tronquées. Echelle : 50 µm (fenêtre large), 15 µm (Zoom). \*\*\*:  $P < 0,0001$  ; \*\*:  $P < 0,01$  ; \*:  $P < 0,05$



Pour la forme TDP35, on observe un comportement très proche de TDP-43, **(voir figure 63, A et D)**, ainsi en mélange avec la forme TDP-43 PL, on retrouve un comportement similaire vis-à-vis de l'ARN, avec lequel on observe un effet de solubilisation pour le ratio 1/50, les deux protéines étant relativement bien mélangées au sein des agrégats. Cependant, on ne retrouve pas l'effet de grossissement des agrégats pour le ratio protéine/nucléotide 10/1 **(voir figure 64, E)** lorsque l'on regarde l'évolution de la taille des agrégats. C'est aussi ce qui est marquant lorsque l'on regarde le cas du mélange avec FUS PL, on retrouve des gouttelettes mais il n'y a pas ou peu d'agrégats dans le cas d'un faible ratio d'ARN. La protéine TDP35 semble donc moins sensible à l'effet de grossissement des agrégats induits par une faible quantité d'ARN. D'un point de vue de la mesure des coefficients de corrélation, la forme TDP35 est cependant moins bien corrélée, ce qui montre qu'elle est moins miscible avec FUS que la forme PL de TDP-43. Cela est cohérent avec ce qui a été observé au MtB (dans le cas du mélange TDP-43/TDP35 le marquage de l'anticorps peut induire un biais dans la miscibilité). En résumé, la forme TDP35 a un comportement proche de la forme PL, avec toutefois une moins bonne miscibilité avec FUS.

Pour la forme TDP-RRM, on remarque un bon mélange, avec FUS PL, avec un coefficient de corrélation supérieur à 0,5 sans ARN, une faible quantité d'ARN semble même favoriser la miscibilité des protéines au sein d'agrégats plus importants que sans ARN. En revanche, avec un excès d'ARN **(voir figure 63, B et 64 B)** on sépare les protéines au sein de petits objets distincts. Le comportement d'agrégation/solubilisation est donc vérifié dans ce cas. Le bon mélange de FUS et TDP-RRM, confirme que ce domaine est impliqué dans l'interaction entre les deux protéines, et qu'en cas de faible quantité d'ARN, plutôt que d'induire une compétition, cette interaction favorise leur coopération. Avec TDP-43 PL **(voir figure 63, E)**, on remarque aussi des coefficients relativement moyens autour de 0,5 avec une légère séparation au sein des agrégats. Rappelons qu'au MtB la corrélation entre les deux était très faible. Au sein des cellules, différents paramètres peuvent influencer le comportement des protéines, en particulier un ensemble de partenaires d'interaction, comprenant aussi des protéines chaperonnes. La forme PL, par exemple, peut mettre en place des auto-interactions, par le biais de son PrLD, ce qui peut exclure le RRM et donc induire une séparation des deux espèces de TDP-43. Les chaperonnes peuvent aussi intervenir pour la forme pleine longueur mais peuvent être inefficaces dans le cas de la forme tronquée. D'un autre côté, nous avons vu aussi dans le paragraphe précédent que l'anticorps de TDP-43 ( $\alpha$ -TDP-43 *ms*) pouvait reconnaître la forme TDP-RRM, ce qui peut induire un biais dans ces mesures de

corrélation, en les augmentant artificiellement. On remarque, toutefois que les agrégats formés ont une morphologie assez proche de ceux de la forme PL. On remarque aussi que l'effet de nucléation est perdu en présence de cette forme TDP RRM (**voir figure 64, F**). Le plus surprenant reste que des petits agrégats persistent en présence d'une forte quantité d'ARN (même si la mesure de l'aire des objets montre quand même un effet de solubilisation). On peut aussi se poser la question de la bonne restructuration du RRM (concentré et stocké dans l'urée, comme les autres protéines produites) après dilution. Des tests par RMN ont été effectués et ont confirmé la bonne restructuration de la protéine.



**Figure 64-Evolution de l'aire des objets formés par les formes tronquées en mélange avec les formes PL en immunofluorescence *In vitro***

Graphiques présentant l'évolution de l'aire (en  $\mu\text{m}^2$ ) des objets pour les formes tronquées TDP35 (A), TDP RRM (B) ou TDP25 (C) avec différents ratio d'ARN, en mélange avec FUS PL, ou en mélange avec TDP-43 (E, F, G). Moyennes avec écart-types. Test de T. \*\*\*: P < 0,0001; \*\*: P < 0,01; \*: P < 0,05

Dans le cas de TDP25, on remarque un mélange assez moyen avec la protéine TDP-43 PL, au sein de gros agrégats assez hétérogènes (**voir figure 63, F**). Les agrégats étant déjà assez gros sans ARN, on ne remarque pas particulièrement d'effet de grossissement des agrégats avec peu d'ARN (**voir figure 64, G**). L'effet de dissolution est aussi affecté avec la persistance d'agrégats (moyennement corrélés), ce qui montre que la forme pathologique TDP25 a un effet sur le comportement de la protéine PL, en affectant son interaction avec l'ARN. Contrairement à ce que l'on a pu voir précédemment avec le MtB, TDP25 semble pouvoir cohabiter avec TDP-43 en formant des regroupements au sein des agrégats mais cela peut être simplement dû

à un biais dans le marquage comme évoqué précédemment. Comme pour TDP-RRM, cela peut s'expliquer par des comportements distincts au sein de la cellule et dans une étude simplifiée *In vitro*.

Plus intéressant encore, avec FUS FL, on observe une cohabitation d'objets rouges et verts distincts et donc une très mauvaise corrélation (**voir figure 63, C**). Ce qui est cohérent avec ce qui a été observé au MtB, et montre que cette forme tronquée de TDP-43 induit une vraie rupture avec son partenaire d'interaction FUS, en particulier car le RRM est tronqué. De cette manière la relation à l'ARN est totalement perturbée et l'effet de solubilisation est rompu avec la persistance d'agrégats pour TDP25.

### 3. Etude des formes tronquées par AFM

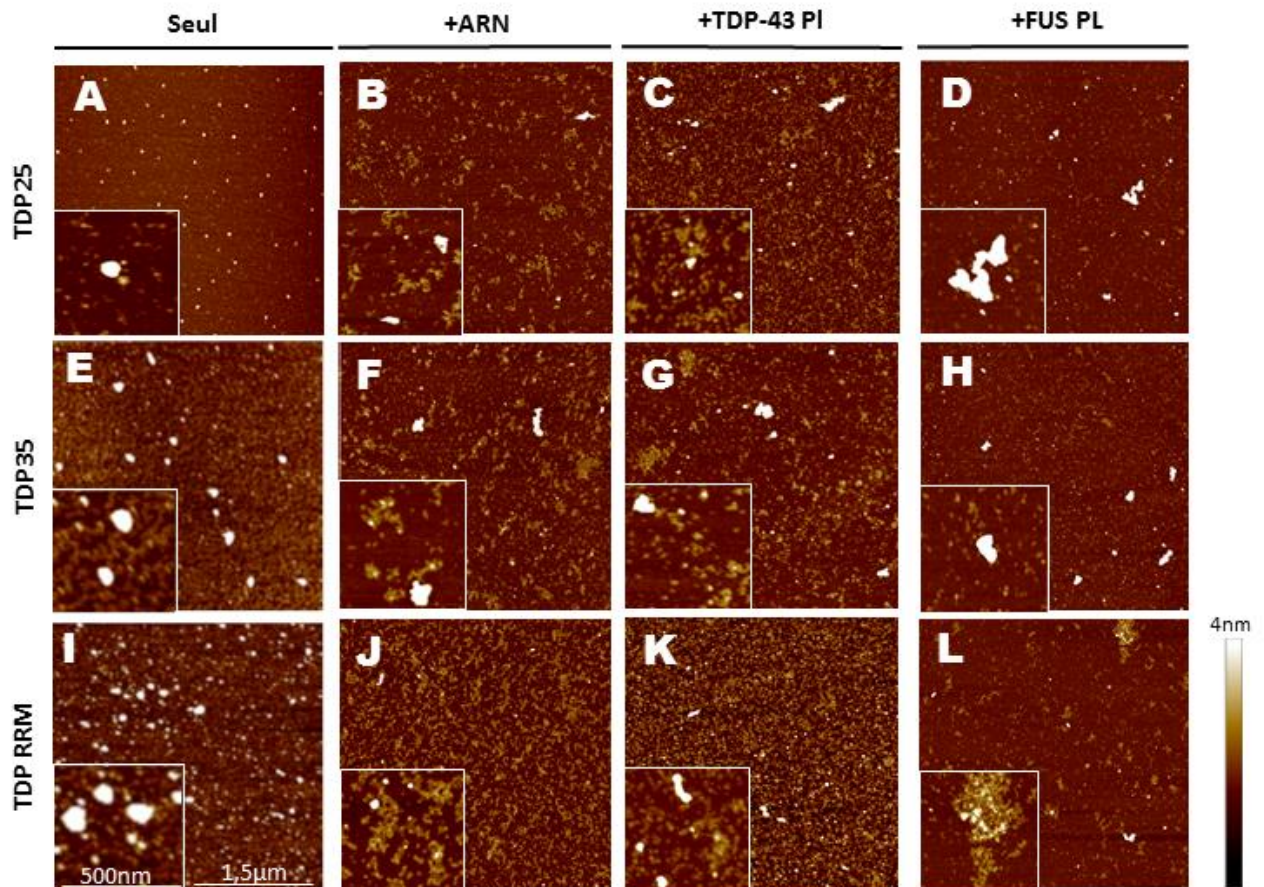
#### a) Le comportement des formes tronquées pathologiques diffère de celui des formes PL

De façon à compléter ces manipulations et à avoir une vision à l'échelle de la molécule unique du comportement des protéines et des formes tronquées avec l'ARN, nous avons utilisé l'AFM. Diverses formes tronquées ont été observées seules, en mélange avec les formes PL et avec ou sans ARN. Les échantillons ont été préparés de la même manière que précédemment, à différents temps d'incubation. En l'absence d'ARN, on voit des petits objets assez ronds pour la plupart des protéines (**voir figure 65, A, E, I**). Ces objets changent peu avec le temps, si la protéine est soluble, comme par exemple avec TDP-RRM, mais peuvent évoluer vers de plus gros objets comme dans le cas de TDP35. TDP25 a aussi la particularité de s'agréger très vite, il faut conserver des temps d'incubation courts, pour pouvoir voir des petits agrégats. Cette forme tronquée a notamment, la particularité de former des objets très fibrillaires à forte concentration, des sortes d'agrégats de fibrilles en plaques (**voir figure 68, B2**).

L'AFM permet de voir la liaison à l'ARN des protéines étudiées. D'après l'expérience précédente, on sait que FUS a tendance à bien se répartir sur l'ARN alors que TDP-43 se lie à l'ARN en formant des regroupements denses de protéines sur un nombre limité d'ARN. A présent, nous souhaitons voir comment l'interaction avec l'ARN est impactée dans le cas des formes tronquées, mais aussi étudier l'impact de ces formes tronquées en mélange avec les protéines PL. Pour cela les formes tronquées TDP25, TDP35 et TDP-RRM ont été incubées avec de l'ARN, seules (**voir figure 65, B, F, J**), mélangées avec TDP-43 PL (**C, G, K**), ou avec FUS PL (**D, H, L**) (mélange équimolaire). Chaque condition a été préparée avec un excès d'ARN par rapport à la quantité de protéine totale.

### ▪ Observations pour TDP25

Le clivage au niveau du RRM1 de TDP25 affecte sa capacité de liaison à l'ARN. Sur les images (**voir figure 65, B**), il apparaît alors qu'il y a un nombre important de petits agrégats qui ne semblent pas co-localiser avec l'ARN. On trouve néanmoins des points sur l'ARN qui peuvent résulter d'une interaction non spécifique. Le ratio ARN pris en charge/ARN total reste toutefois comparable à la forme PL (**voir figure 66, A**). Si en mélange avec la forme TDP-43 PL (**voir figure 65, C**), la forme TDP25 ne semble pas induire une grande différence de comportement, avec une tendance à la formation d'agrégats assez denses, cette forme présente un comportement distinct avec FUS. Ainsi, en mélange avec cette dernière on observe qu'il y a de gros agrégats,



**Figure 65-Les formes tronquées TDP25 et TDP35 induisent une perte de la coopération avec FUS**

(A-D) TDP25 incubée seule, 5  $\mu$ M, 20 min (A), à la concentration de 500 nM avec de l'ARN 4 nM pendant 40 min (B), avec TDP-43 1:1 (250 nM), ARN 5 nM, 30 min (C), avec FUS 1:1 (250 nM), ARN 5 nM, 30 min (D). (E-H) TDP35 incubée seule à la concentration de 1  $\mu$ M, 20 min (E), à la concentration de 50 nM, avec de l'ARN 4 nM, pendant 1 h (F), avec TDP-43 1:1 (25 nM), pendant 30 min (G), avec FUS 1:1 pendant 30 min (H). (I-L) TDP-RRM incubée seule à la concentration de 500 nM, 20 min (I), avec de l'ARN 5 nM, pendant 2 h (J), avec TDP-43 1:1 (125 nM), pendant 30 min (K), avec FUS 1:1 (125 nM) pendant 30 min (L). Echelle en xy de 1,5  $\mu$ m et zoom de 500 nm. Echelle en z de 4 nm

avec un peu d'ARN sur la surface, que l'on suppose être pris en charge par FUS (**voir figure 65, D**). Si le ratio d'ARN pris en charge du couple TDP25/FUS est comparable à celui du couple TDP-43/FUS PL (**voir figure 66, C**), il apparaît plutôt que cela reflète l'interaction de FUS avec l'ARN, tandis que TDP25 s'agrège de son côté. Ces données suggèrent ainsi que TDP25 a un comportement distinct de TDP-43 par rapport à l'ARN et que son interaction avec FUS est affectée. Ainsi, la protéine TDP-43, une fois clivée ne se répartit plus le long de l'ARN en présence de FUS, elle présente une forte tendance à l'agrégation que ni FUS et ni TDP-43 ne peut inhiber.

- **Observations pour TDP35**

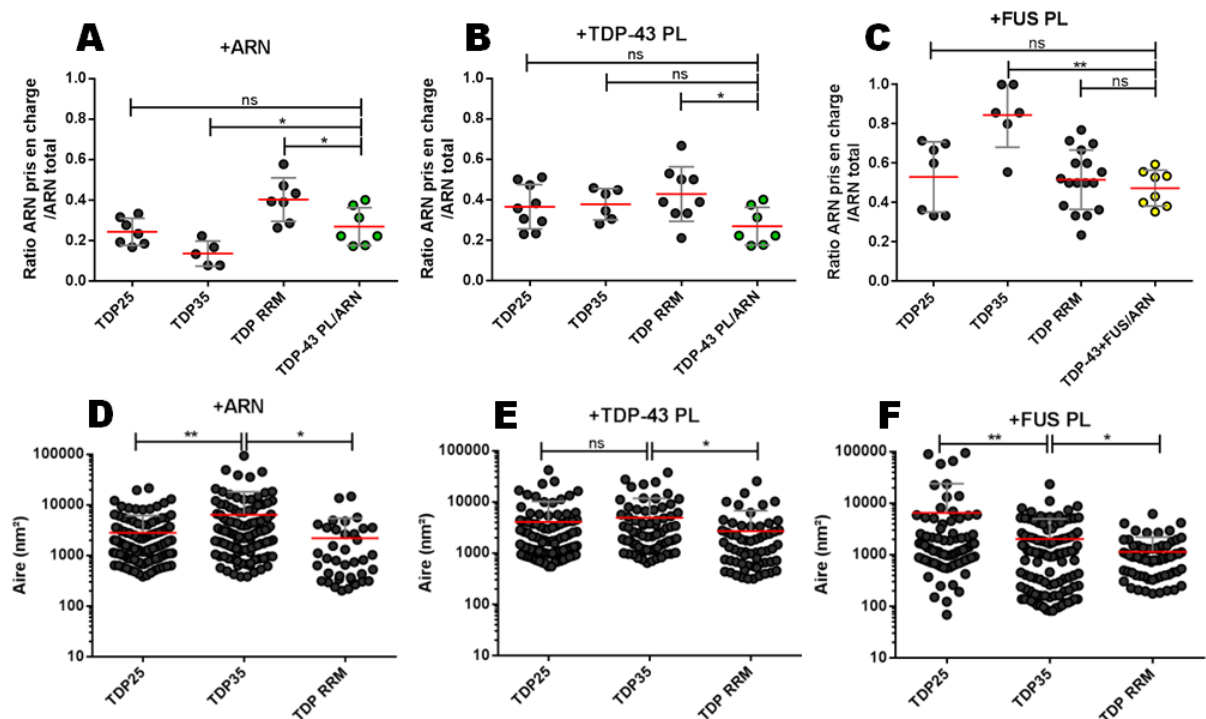
Pour TDP35, on remarque un comportement assez proche de TDP-43 dans sa liaison à l'ARN (**voir figures 65, F et 66, A et B**), avec la présence d'agrégats et des points sur l'ARN et un nombre réduit d'ARN pris en charge. Il apparaît toutefois que TDP35 est capable de compacter l'ARN (**comme illustré à la figure 68, A**), on peut alors supposer que l'ARN peut se retrouver piégé dans les agrégats. En présence de FUS, on observe un comportement similaire, avec la présence de quelques agrégats, et d'ARN sur la surface. Si le nombre d'ARN pris en charge est important en comparaison de la référence TDP-43/FUS c'est parce que les ARN sont pris en charge dans les agrégats (**voir figure 66, C et tableau 67**). On ne retrouve alors pas l'effet de répartition de la protéine le long de l'ARN comme avec TDP-43 PL, ce qui montre une nouvelle fois que le clivage de la protéine au niveau du domaine RRM1 affecte sa coopération avec FUS.

- **Observations pour TDP RRM**

Pour TDP-RRM, on a une très bonne liaison à l'ARN avec une meilleure répartition de la protéine et une moindre formation d'agrégats, ce qui est logique puisque cette forme tronquée est soluble (**voir figure 65, J**). De plus, cette protéine prend en charge un nombre plus important d'ARN que la forme PL (**voir figure 66, A**). En présence de la protéine TDP-43 PL (**voir figure 65, K**), on retrouve cependant des agrégats, avec des objets plus importants qui co-localisent avec l'ARN, caractéristiques de la forme PL, qui cohabitent avec des ARN pris en charge par quelques protéines. Il semble alors comme indiqué par le MtB que les deux formes interagissent très peu, présentent une liaison à l'ARN distincte et donc agissent chacune de leur côté. Le nombre d'ARN pris en charge (**voir figure 66, B**) est donc bien plus important que dans le cas de la référence TDP-43/ARN (avec peu de gros objets contrairement à TDP25, ou TDP35 (**voir figure 66, E**). En ce qui concerne le mélange avec FUS, on retrouve dans ce cas ce que l'on avait avec TDP-43 FL : des protéines bien réparties le long de l'ARN et très peu de gros objets. Ainsi, il y a autant d'ARN pris en charge que dans le cas du mélange des formes PL (**voir figure 66, C**).

Cela supporte l'hypothèse que les RRM de TDP-43 jouent un rôle dans sa coopération avec FUS pour la prise en charge de l'ARN.

### Quantification des objets



**Figure 66-Quantification des objets formés par les formes tronquées en mélange avec les formes PL en AFM.**

(A-C) Graphiques comparant les valeurs de ratios ARN pris en charge/ARN total pour chaque forme tronquée dans les conditions avec ARN uniquement (A), avec ARN+TDP-43 (B) et avec ARN+FUS (C). (D-F) Graphiques comparant les aires des objets dont la hauteur dépasse les 4 nm dans les conditions avec ARN uniquement (D), avec ARN+TDP-43 (E) et avec ARN+FUS (F). Echelle logarithmique. Test de T, moyennes et écart-types. \*\*\*:  $P < 0,0001$  ; \*\*:  $P < 0,01$  ; \*:  $P < 0,05$ .

La mesure de l'aire des objets dont la hauteur est supérieure à 4 nm permet de mieux distinguer la capacité des protéines à s'agglomérer ou au contraire à se répartir sur l'ARN (**voir figure 66, D, E, F et tableau 67**). Ce qui permet notamment de distinguer les cas où les ARN sont pris en charge parce que les protéines se répartissent correctement, des cas où l'on a un fort ratio ARN pris en charge/ARN total car l'ARN est piégé au sein d'agrégats. Ainsi, on remarque que TDP25, comme TDP35, forme des agrégats plus importants en aire que TDP RRM (**voir figure 66, D**) pour lequel il y a peu d'objets très gros. On retrouve la même différence entre TDP25/TDP35 et le RRM dans le cas du mélange avec TDP-43 (**voir figure 66, E**), et aussi dans le cas du mélange avec FUS. Dans ce dernier cas (**voir figure 66, F**), il apparaît cependant que TDP25 forme des agrégats bien plus importants que TDP35.

	FUS						TDP-43					
	TDP-43	FUS	TDP-25	TDP35	TDP RRM	TDP-43	TDP-25	TDP35	TDP RRM	TDP-25	TDP35	TDP RRM
Présence de gros objets	+	-	+	+	-	--	++	++	-	++	++	+
Densité d'ARN sur la surface	++	++	+	-	++	++	--	-	--	+	-	+

**Figure 67-Tableau récapitulatif pour les formes tronquées en AFM**

Tableau récapitulant les observations pour chaque forme étudiée en mélange avec de l'ARN selon deux critères : présence de gros objets et densité d'ARN sur la surface. Dégradé de couleur selon la fréquence des événements (fréquent : plus de 5 agrégats/ $\mu\text{m}^2$  ou plus de 40 ARN/ $\mu\text{m}^2$ , assez fréquent : plus de 3 agrégats/ $\mu\text{m}^2$  ou plus de 20 ARN/ $\mu\text{m}^2$ , rare : plus 1 agrégats/ $\mu\text{m}^2$  ou plus de 10 ARN/ $\mu\text{m}^2$ , très rare : Inférieur à 1 agrégats/ $\mu\text{m}^2$  ou inférieur à 10 ARN/ $\mu\text{m}^2$ ).

	Agrégats/ $\mu\text{m}^2$	ARN/ $\mu\text{m}^2$
++	Fréquent	5 40
+	Assez fréquent	3 20
-	Rare	1 10
--	Très rare	0 5

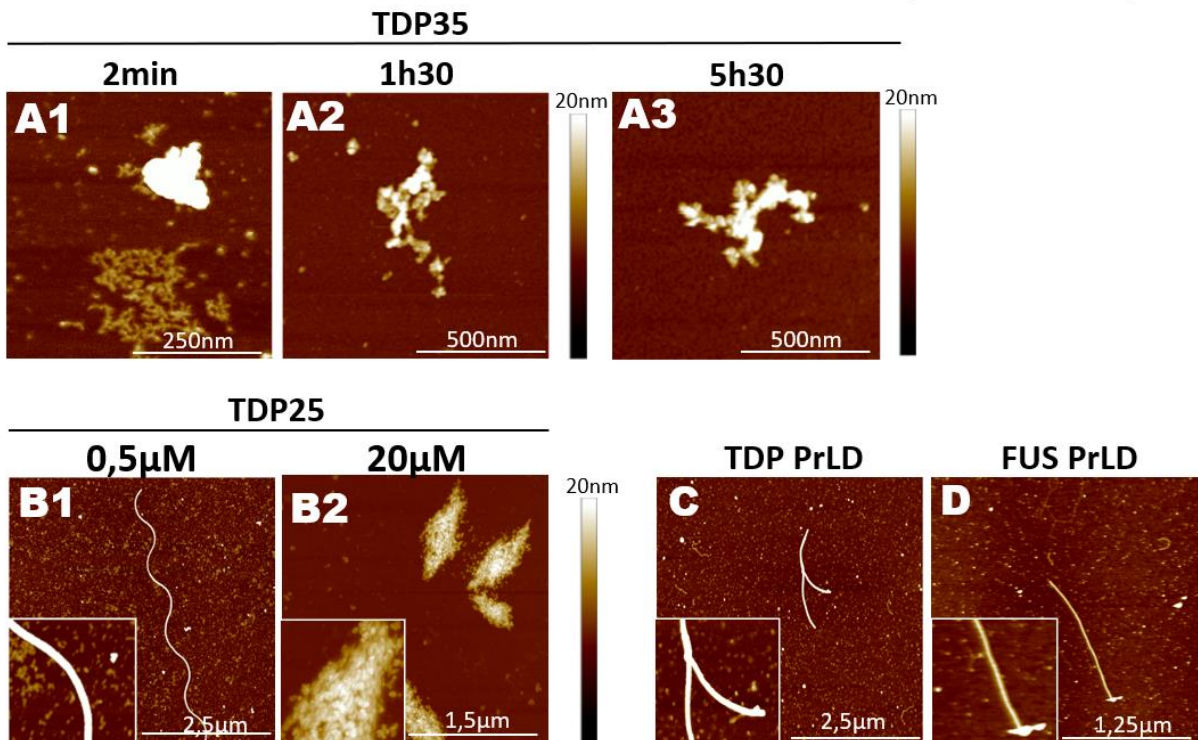
Ce tableau (**voir figure 67**) permet de récapituler et de mettre en avant des comportements distincts entre les formes tronquées. Certains points posent aussi questions, comme le fait pour TDP25, qu'il y a un certain nombre d'ARN sur lesquels on détecte des protéines, ou encore le fait que l'on observe de gros agrégats sur des temps plus longs avec l'absence d'ARN sur la surface (au cours du temps la surface se vide au profit de la formation de gros agrégats, cependant l'ARN ne devraient pas disparaître, s'il n'est pas pris en charge). Une interprétation possible est qu'avec le temps l'ARN peut être piégé de manière non spécifique dans les agrégats. Pour TDP-RRM, il y a une faible quantité d'ARN sur la surface en présence de FUS. Puisque les deux protéines sont solubles, on pourrait s'attendre à conserver l'ARN sur la surface mais dans nos conditions la surface est peu dense en ARN. Cela peut être dû à la variabilité des manipulations en AFM.

Ensemble, ces données d'AFM confirment que l'interaction entre FUS et TDP-43, qui leur permet de coopérer dans leur prise en charge de l'ARN, est favorisée en présence des RRM de TDP-43, mais que cette interaction est altérée dans le cas des formes tronquées de TDP-43. TDP25 et TDP35 ayant tendance à s'agréger de manière encore plus marquée que TDP-43, l'interaction avec FUS est rompue, et ne permet pas de les amener sur l'ARN.

### **b) Mise en évidence de certaines structures formées par les formes tronquées *in vitro***

Pour finir, cette dernière figure permet d'apporter un complément d'information à la figure précédente et permet d'illustrer un comportement remarquable des protéines étudiées qui peuvent former des objets à la morphologie particulière. Dans un premier temps (**voir figure 68, A**), les acquisitions effectuées pour TDP35 en mélange avec l'ARN à différents temps d'incubation mettent en évidence la formation d'objets très denses qui emprisonnent l'ARN, comme évoqué précédemment. Dès les

temps d'incubation très courts, on voit que TDP35 s'accumule sur l'ARN (**voir figure 68, A1**), comme peut le faire TDP-43. Mais rapidement on ne voit plus d'ARN sur la surface et on trouve uniquement des objets hauts et denses (**voir figure 68, A2**) dont la hauteur dépasse les 20 nm mais avec une morphologie très découpée. On retrouve aussi ces objets après 5,5 h d'incubation (**voir figure 68, A3**). Cela suggère que cette forme peut avoir la capacité de piéger les ARN au sein d'agrégats, ce phénomène pourrait donc induire un gain toxique de fonction.



**Figure 68-Les formes tronquées forment des structures particulières à l'échelle nanométrique.**

(A1-A3) TDP35 avec ARN incubée à différents temps et différentes concentrations (A1) 50 nM avec ARN 4 nM (1:250) 2 min (A2) 500 nM, ARN 66 nM (1:400), 1,5 h (A3) 250 nM, ARN 16,6 nM (1:200), 5,5 h. (B1) TDP25 incubée à la concentration de 500 nM, avec de l'ARN 5 nM (1:30) pendant 4 h ou (B2) à la concentration de 20 μM, 30 min sans ARN. (C) TDP PrLD incubée à la concentration de 5 μM, 1,5 h avant dépôt. (D) FUS PrLD incubée à la concentration de 1 μM, 5 h. Echelles variables indiquées dans chaque condition. Echelles en z de 4 nm, ou indiquées dans le cas contraire.

Pour TDP25, il a été observé de manière assez sporadique des objets fibrillaires assez remarquables, longs de plusieurs micromètres (**voir figure 68, B1**). Cette image a été obtenue à la concentration de 500 nM après 4 h d'incubation, mais la plupart du temps la protéine a tendance à former de petits agrégats comme mis en évidence à la figure précédente. En revanche, à haute concentration (20 μM après 30 min d'incubation) la formation d'objets ressemblant à des agglomérats de fibrilles est plus fréquente, ces objets sont très particuliers, et évoque les agrégats en forme de plaques retrouvés dans le cas des maladies neurodégénératives[395,396]. Les

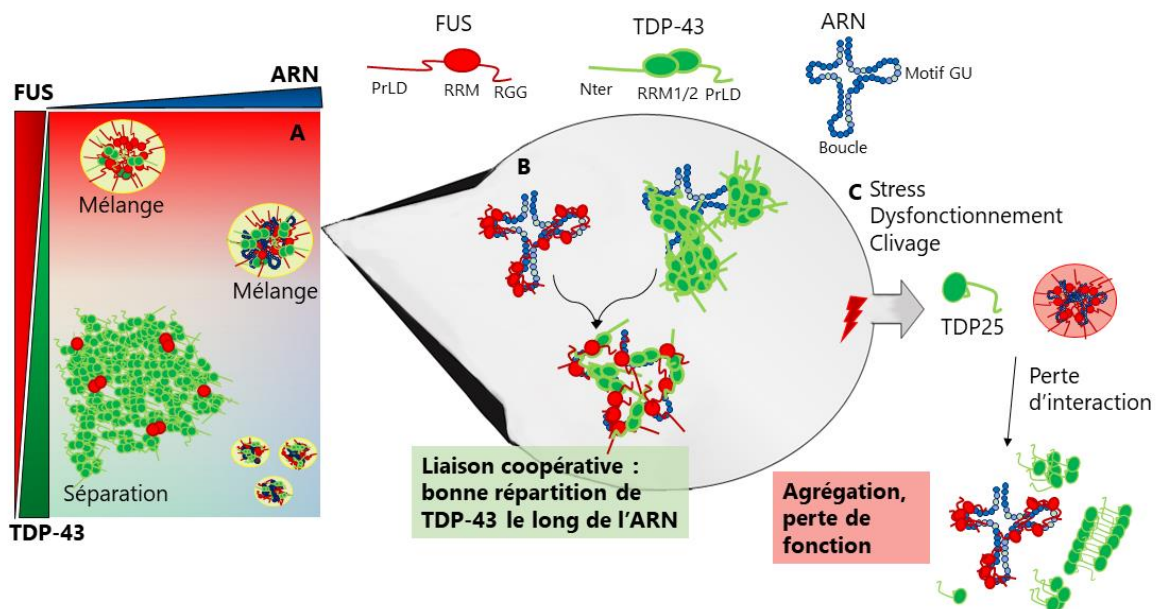


domaines PrLD sont connus pour leur capacité à former des fibres[151], c'est ce qui a été constaté ici, à la fois pour FUS PrLD et TDP PrLD. Ces objets ont pu être observés assez fréquemment pour ces deux constructions. Ces acquisitions montrent ainsi que la protéine TDP-43 clivée peut se structurer en fibres et confirment la tendance à s'organiser en fibres des domaines PrLD. Par ailleurs, ces objets sont observés dans les cas pathologiques, ce qui suggère un lien entre le clivage de la protéine TDP-43, la perte de fonction avec l'ARN due notamment à la perte d'interaction avec FUS, et la formation d'agrégats fibrillaires, souvent associés au développement de maladies neurodégénératives[154, 155, 391].

# Discussion

---

Au cours de ce travail, nous avons étudié deux protéines de liaison à l'ARN, FUS et TDP-43. Ces deux protéines sont assez similaires en matière de structure, elles sont pourvues de domaines de faible complexité de type prion favorisant les interactions multivalentes, et aussi de domaines RRM pour la reconnaissance de l'ARN. TDP-43 se lie à l'ARN avec une grande spécificité de séquence, tandis que FUS se répartit sur l'ARN avec une spécificité limitée. En dépit des fonctions communes qu'elles exercent, l'analyse de leurs interactomes n'indique pas que ce sont des partenaires préférentiels[323, 392], quand certaines études, ne mentionnent pas de lien entre les deux[242, 393]. Malgré tout, un certain nombre d'études ont mis en évidence une interaction entre ces deux protéines. Au cours de ce travail, nous avons donc confirmé l'interaction entre ces deux protéines, mais nous avons été plus loin dans sa caractérisation. Nous avons dans un premier temps montré que cette interaction était plus forte en comparaison avec d'autres RBP grâce à la technique du *MtB*. Nous avons aussi montré que cette interaction était asymétrique avec une miscibilité dépendante de la concentration. En effet, il apparaît que TDP-43 est solubilisée dans la phase de FUS, tandis que FUS est exclue de la phase de TDP-43. Nous avons par la suite montré que ces protéines adoptaient un mode de liaison coopératif sur l'ARN, et plus spécifiquement, les données d'AFM nous montrent que FUS participe à la répartition de TDP-43 le long des ARN en évitant de former des regroupements très denses de la protéine. Aussi l'ARN est important dans la régulation du comportement des protéines, en particulier pour TDP-43 qui s'agrège en son absence. Même si l'interaction de FUS et TDP-43 peut s'effectuer sans ARN, il améliore leur mélange. Toutefois, en faible quantité l'ARN induit une compétition entre les deux protéines. Enfin, le RRM de TDP-43 semble jouer un rôle important dans le cadre de cette interaction alors que des formes tronquées des protéines altèrent leur capacité d'interaction (**voir figure 69**).



**Figure 69-Modèle de l'interaction entre FUS et TDP-43 sur l'ARN**

(A) A large échelle, le mélange entre TDP-43 et FUS est dépendant de la quantité de TDP-43 : à faible concentration cette dernière est incluse dans la phase de FUS, alors qu'à haute concentration des structures compartimentées sont formées. L'ARN régule aussi le comportement des protéines en favorisant leur mélange au sein de petits objets. (B) A plus petite échelle on voit que FUS se répartit sur l'ARN (par spécificité de structure), tandis que TDP-43 s'accumule sur des séquences très spécifiques (Motifs GU). En mélange, l'interaction entre les deux protéines permet une bonne répartition de ces dernières sur toute la longueur de l'ARN. (C) Ce mélange est affecté en cas de stress, comme par exemple, avec le clivage de le TDP-43 en TDP25 qui n'a plus la capacité d'interagir avec FUS et peut conduire à l'émergence d'agrégats.

L'interaction mise en évidence ici, à la fois par les techniques du *MtB* et de PLA, rejoint ce qui avait été décrit précédemment par Ling *et al.*[16] ainsi que l'ensemble des travaux qui ont détecté l'interaction entre TDP-43 et FUS par des techniques de co-immunoprécipitation[347], de pull-down[14] mais aussi par Y2H[348]. Pour les résultats obtenus par PLA, on observe toutefois, une différence notable, les résultats présentés ici montrent une interaction assez importante entre les deux protéines avec un nombre de cellules possédant un signal supérieur à 90 %, tandis qu'il n'est que de 10 % dans l'étude de Ling *et al.* (Cela peut s'expliquer par une différence dans les anticorps utilisés). Grâce au *MtB*, nous allons plus loin en comparant cette interaction avec celles d'autres RBP. Ces résultats sont cohérents avec des études de protéomique recensant des interactions de TDP-43 avec HuR, et G3BP1 mais aussi YB1[244] et celles de FUS avec HuR et G3BP1[388]. Nous avons aussi étudié une notion de miscibilité entre les deux protéines qui n'avait pas été décrite précédemment. Le fait que TDP-43 soit soluble dans la phase de FUS, alors que FUS est exclue de la phase de TDP-43, peut avoir une importance dans leur fonction et notamment dans leur relation avec l'ARN. Ainsi, nous avons vu que FUS et TDP-43

faisaient preuve de coopération dans leur mode de liaison à l'ARN avec une meilleure répartition de TDP-43 sur l'ARN en présence de FUS mais que leur mélange dépend de la quantité d'ARN. L'étude de Kim *et al.*, suggère un rôle plutôt compétitif entre les deux protéines dans la prise en charge de l'ARNm HDAC6 avec un recouvrement de leurs sites de liaison[14]. Ici, nous avons utilisé un ARN assez long (3000 bases) avec de multiples sites de liaison, l'ARN 2Luc, de façon à mimer un comportement général avec l'ARN, et non une spécificité de reconnaissance avec un ARN particulier. Ainsi, cela suggère que TDP-43 et FUS auraient de manière générale un comportement coopératif, mais que la prise en charge de certaines cibles, en particulier HDAC6, résulte d'une compétition entre les deux protéines, simplement du fait de la redondance de leur liaison qui induit le même effet[14]. Par ailleurs, plusieurs travaux suggèrent que TDP-43 et FUS participent à des mécanismes communs, en particulier dans le cas du développement de mécanismes pathologiques[17–19, 349, 351]. Cette ambivalence peut s'expliquer par le fait que selon les fonctions étudiées et les ARN cibles, elles adoptent un mode d'interaction différent. Il apparaît aussi que FUS est plus polyvalente et qu'elle peut remplacer TDP-43 dans certaines de ces fonctions. L'étude de Kabashi *et al.* nous éclaire particulièrement sur la différence de comportement entre TDP-43 et FUS en montrant que FUS est capable de prendre en charge les fonctions de TDP-43 pour restaurer le phénotype moteur chez le poisson zèbre dont on a supprimé le gène *Tardbp*, alors que l'inverse ne se vérifie pas[17].

Par ailleurs, la proximité, voir la superposition de leurs sites d'interaction sur l'ARN a été mise en évidence[6], en particulier cette étude a montré un enrichissement 2,2 fois plus important de TDP-43 autour des sites de liaison de FUS. Ce qui est cohérent avec le mode d'interaction coopératif proposé ici, où FUS est nécessaire à une meilleure répartition de TDP-43 sur l'ARN. D'un autre côté, l'étude de Lagier-Tourenne *et al.* présente le rôle de FUS et TDP-43 comme distincts (avec peu de gènes dépendants à la fois de FUS et TDP-43 suivant leur déplétion), et conclut à une interaction parallèle, davantage que synergique, notamment dans le cas de la prise en charge des longs introns[6]. Toutefois, l'enrichissement des sites de liaison de TDP-43 près de ceux de FUS, tout comme le fait que FUS soit connue pour prendre en charge rapidement les ARN naissants[221], suggèrent que la première serait recrutée auprès de la seconde, peut-être pour l'assister dans ses fonctions, ou simplement parce que la présence de FUS facilite la liaison de TDP-43, ou encore que le chevauchement de leurs sites de liaison est nécessaire pour la bonne régulation des transcrits par le biais d'une relation compétitive. Les données présentées ici, ont tendance à appuyer l'hypothèse de la coopération.

Le processus de coopération décrit ici correspond à l'addition de protéines sur une même cible dans un but commun (**voir chapitre I,1,b**). Le mode d'interaction coopératif est courant chez les RBP. Dans le cas de TDP-43 la coopérativité semble essentielle à sa bonne prise en charge de l'ARN, la protéine est capable d'interagir avec elle-même pour une association optimale avec l'ARN. L'étude de Rengifo-Gonzales *et al.* montre que la liaison coopérative de TDP-43 à l'ARN passe par le RRM[56]. Plus précisément, la liaison coopérative de la protéine sur l'ARN passe par une interaction inter-protéines qui se met en place entre le RRM1 d'une protéine et le RRM2 d'une autre grâce à des résidus présents dans la 3<sup>ème</sup> boucle du RRM1 et ceux d'une poche du RRM2. Cette organisation permet à TDP-43 de se lier à l'ARN, en particulier de se concentrer au niveau des motifs GU, tout en limitant les interactions multivalentes entre LCD qui pourraient conduire à son agrégation massive. Le fait que FUS soit exclue des compartiments riches en TDP-43 peut permettre d'éviter la désorganisation de cette association spécifique TDP-43/ARN qui permet à TDP-43 de rester soluble et fonctionnelle. De son côté, FUS est aussi capable de se lier à l'ARN de façon coopérative en multimérisant sur l'ARN[53], mais ne présente pas de spécificité marquée pour des séquences particulières. Il est donc possible que FUS se répartisse le long des ARNs, bloquant potentiellement l'accès à d'autres RBP à leurs sites spécifiques respectifs. Cependant, FUS est capable d'accepter dans sa phase une certaine proportion de la protéine TDP-43 favorisant la formation d'un sous-compartiment. En contexte de forte compétition entre les RBP pour la prise en charge des ARN, ce système dynamique, pourrait ainsi permettre à TDP-43 d'avoir accès à son site spécifique GU. Cette coopération entre TDP-43 et FUS permettrait donc à la cellule de voir ses fonctions biologiques comme la transcription ou l'épissage assumées sans qu'il soit nécessaire d'apporter un surplus d'énergie pour répartir convenablement les acteurs le long de l'ARN.

La mise en avant de l'interaction entre deux protéines impliquées dans le développement de maladies a déjà été mise en évidence comme dans le cas de TDP-43 qui interagit avec l'Ataxin2 par le biais du RRM[394]. Cette interaction a par ailleurs une influence sur son comportement, car elle contribue à accroître la toxicité de la protéine TDP-43. Il a aussi été montré que FUS pouvait interagir de manière directe avec le complexe SMN, pointant aussi vers une convergence des mécanismes biologiques sous-jacents au développement de maladies des motoneurones[395]. L'étude de Ling *et al.* qui montre que les mutations de TDP-43 (M337V et Q331K) amplifient l'interaction avec FUS suggère un mécanisme pathologique commun entre les deux protéines[16]. Si plusieurs études pointent donc l'existence d'un mécanisme pathologique commun entre les deux protéines, dans notre cas, nous montrons

plutôt que l'interaction entre TDP-43 et FUS est nécessaire au maintien de la fonction normale de ces protéines. Leur miscibilité, et l'interaction au sein de phases hétérotypiques en présence d'ARN semblent essentielles au maintien de leur solubilité.

L'émergence du principe de LLPS apporte un autre regard sur le comportement de ces protéines. On sait désormais que les protéines possédant des domaines de faible complexité sont particulièrement impliquées dans la formation de condensats. Ces compartiments dynamiques sans membrane sont potentiellement nécessaires au comportement physiologique des protéines[107]. La technique de visualisation par immunofluorescence des échantillons fixés, développée ici, permet de mettre en évidence la morphologie des assemblages de protéines, avec en particulier une tendance pour FUS à former des gouttelettes. Nous avons aussi montré que la concentration relative des protéines, tout comme la quantité d'ARN étaient des paramètres importants dans la régulation du comportement des protéines. Ainsi, une quantité réduite de TDP-43 permet une meilleure miscibilité dans les condensats formés par FUS. Cela est pertinent lorsque certaines études tendent à montrer que la dérégulation de ce mécanisme de LLPS peut conduire à l'agrégation des protéines[106, 121].

Par ailleurs, nous sommes en présence de protéines qui ont une forte tendance à l'agrégation. Différents paramètres permettent donc de les maintenir solubles, comme certaines MPT, mais aussi des partenaires d'interaction et notamment des chaperonnes. Dans nos conditions simplifiées, il apparaît que l'interaction de TDP-43 avec FUS permet de maintenir la protéine en phase liquide et non de s'agréger. L'ARN est un paramètre indispensable puisqu'en grande quantité (comme au sein des cellules, en particulier dans le noyau), il offre un support aux protéines et permet de les réguler notamment en induisant un effet de solubilisation. Ce que l'on a vérifié *In vitro*. De plus, les différents ratios de protéines étudiés amènent à penser qu'une moindre quantité de TDP-43 par rapport à FUS (en dessous de 1/1 et jusqu'à 1/3) en présence d'une grande quantité d'ARN serait pertinente pour le bon fonctionnement des protéines et notamment pour prévenir l'agrégation de TDP-43. Ce qui rejoint les travaux de Maharana *et al*[142]. Toutefois au regard des études de protéomiques[345], la quantité des protéines TDP-43 et FUS semble être équivalente dans les neurones moteurs. Les protéines TAF15 et EWSR1 qui font partie de la famille des protéines FET avec FUS, sont tout à fait semblables dans leur structure et leur comportement[223]. Il est ainsi probable que le comportement observé ici entre FUS et TDP-43 soit valable pour les autres protéines de la famille FET, et qu'elles contribuent ainsi à la prévention de l'agrégation de TDP-43 en la solubilisant dans

leur phase. En outre, on remarque qu'avec les formes tronquées pathologiques de TDP-43, en particulier TDP25 dont la liaison à l'ARN est affectée, l'effet de solubilisation est perdu. On sait que FUS peut jouer le rôle de chaperonne avec l'ARN en remodelant sa structure par le biais de son RRM[211], de cette manière les bases sont mieux exposées pour la liaison à d'autres protéines partenaires. On peut facilement supposer que c'est le cas avec TDP-43. D'un autre côté, les condensats formés par LLPS de TDP-43 sont régulés par sa liaison à l'ARN, ou encore stabilisés par des interactions avec d'autres protéines comme la chaperonne HSP70[147]. De plus, la formation de condensats contenant deux protéines a notamment été mise en évidence entre TDP-43 et hnRNPA2 avec des gouttelettes dans lesquelles co-localisent les signaux des deux protéines[396], ce qui est semblable à ce que nous avons pu observer lors de ce travail.

Une étude a montré que FUS avait tendance à s'agrèger plus rapidement que TDP-43 *In vitro*[397], or ici nous observons (à des concentrations comparables) que TDP-43 a tendance à former des agrégats massifs lorsque FUS forme plutôt des condensats à l'apparence de gouttelettes caractérisées par une forte circularité. Cela montre qu'en dépit de leurs structures semblables, les deux protéines ont des comportements distincts. Ainsi cette observation, est appuyée par le fait que les molécules sont plus labiles dans les inclusions formées par FUS que celles de TDP-43, observées dans les cellules par la méthode FRAP[398]. Ce qui pourrait aussi expliquer que TDP-43 a plus de facilité à s'insérer dans les condensats de FUS, alors que l'inverse ne se vérifie pas. Par ailleurs, l'étude des gènes modificateurs de toxicité chez la levure (qui peuvent avoir des homologues chez l'humain)[397] suggère que le comportement de FUS et TDP-43 est différent, notamment au regard de leur rapport à l'agrégation. Malgré cela, le fait que ces deux protéines peuvent se retrouver au sein de mêmes agrégats dans les maladies neurodégénératives suggèrent un mécanisme pathologique commun[15, 18, 20, 357]. En particulier, l'étude sur le modèle de drosophile de Lanson *et al.* suggère un mécanisme synergique entre FUS et TDP-43 qui est amplifié par la présence de mutations sur l'une ou l'autre des protéines. Cette interaction synergique a pour effet d'amplifier le phénotype de neurodégénérescence. De ce fait, ils supposent que la présence de mutations chez l'une des protéines a un effet sur le fonctionnement physiologique de l'autre l'entraînant dans un processus pathologique commun. Ce même article propose que les deux protéines puissent s'influencer l'une et l'autre d'une manière dépendante à la dose, ou alors qu'elles participent à des fonctions parallèles en se croisant ponctuellement. La première option semble mieux converger avec ce qui est observé ici.

L'étude des différents domaines intervenant dans l'interaction entre TDP-43 et FUS montre que le domaine RRM de TDP-43 semble particulièrement impliqué. Cela est cohérent avec ce qui a été montré par GST-pull down[14], qui montre que le RRM2 et le PrLD de TDP-43 contribuent à la liaison à FUS. Cette même étude montre que le domaine PrLD seul n'est pas suffisant, et c'est aussi ce que l'on observe. En revanche, les formes tronquées semblent avoir un comportement bien particulier. Ici, nous montrons qu'elles manquent d'affinité avec la forme pleine longueur et que, plus largement, on observe une perte d'interaction avec FUS alors que ces formes contiennent à la fois le RRM (une partie du RRM2 pour TDP25) et le PrLD.

La protéine TDP-43 peut subir de nombreux clivages, les fragments TDP25 et TDP35 qui sont issus de clivages protéolytiques par caspase, ont été particulièrement étudiés du fait de leur présence dans les inclusions cytoplasmiques[300]. L'étude de Kim *et al.* met en évidence par CoIP que le fragment TDP35 peut se retrouver dans les complexes de TDP-43, de plus, ce fragment a aussi la capacité de se lier à FUS[14]. Les expériences réalisées ici, en particulier grâce à la technique du MtB, diffèrent de cette observation, où nous observons une mauvaise corrélation entre les deux protéines. Ces résultats ne signifient pas que TDP-43 est incapable de se lier à TDP35 mais qu'au sein des cellules ces deux formes peuvent co-exister en formant des compartiments distincts. Néanmoins, *In vitro* et en AFM, on remarque que TDP35 présente un comportement altéré, tout comme TDP25 vis-à-vis de FUS. D'autres travaux ont mis en évidence que TDP35 garde une bonne liaison à l'ARN, cette forme tronquée est aussi plus encline à former des inclusions, et piège TDP-43 PL par le biais de son interaction avec l'ARN au sein des agrégats[332, 399]. De plus, cette même étude indique que la fragmentation de TDP-43 en TDP35, influence sa fonction dans la prise en charge de l'ARN[399]. Ces observations sont cohérentes avec ce que l'on observe ici, notamment par AFM, qui suggère que les modes d'interaction à l'ARN sont différents pour les fragments en comparaison de la forme PL. De plus, on observe que les formes clivées pathologiques semblent condenser et piéger l'ARN au sein d'agrégats, même si cela peut résulter d'interactions non spécifiques en l'absence de tout concurrent. On peut supposer que ce phénomène peut entraîner d'autres protéines comme FUS. Le fait que l'on retrouve ces fragments dans les inclusions[184], suggère, en effet, qu'ils ont un effet sur les formes PL, qui peut conduire à des mécanismes pathologiques.

Pour finir, la structure des agrégats formés par les protéines au sein des échantillons de patients SLA ou DFT fait débat. Certaines études considèrent que ce sont des structures fibrillaires de type amyloïde[155, 391] comme cela peut-être le cas avec la protéine  $\beta$ -amyloïde dans le cas de la maladie d'Alzheimer ou l' $\alpha$ -synucléine



dans la maladie de Parkinson[400]. D'autres montrent que ces agrégats peuvent s'apparenter à des structures filamenteuses[154] sans répondre aux critères de distinctions des fibres de type amyloïdes (Marquage Thioflavin S et Congo Red) ou encore avec un caractère plutôt amorphe[153]. Par ailleurs, une étude récente montre que les agrégats fibrillaires retrouvés dans le cas de la DFT ne contiennent pas de TDP-43, mais la protéine TMEM106B, et que les agrégats de TDP-43 présents ne sont pas fibrillaires[152]. Ici, nous avons simplement montré que les formes tronquées peuvent former des structures filamenteuses sans les caractériser davantage (ce phénomène étant variable et peu répétable). Ce qui permet toutefois d'établir un lien entre les maladies et les structures qui ont pu être observées par ailleurs en AFM[154]. Ainsi, cela nous permet de conclure que les formes tronquées pathologiques de TDP-43 : TDP25 et TDP35, provoquent d'une part, un effet délétère en perdant leur fonction, notamment dans la coopération avec FUS dans la prise en charge de l'ARN, ce qui doit avoir de multiples conséquences sur leurs fonctions communes. D'autre part, le fait que l'on retrouve ces formes tronquées dans les agrégats chez les échantillons de patients, et qu'elles forment des structures caractéristiques, suggère qu'elles sont impliquées directement ou indirectement dans le développement des maladies.

# Conclusion et perspectives

---

Longtemps, une protéine a été associée à une unique fonction assurée par une structure déterminée. Aujourd'hui, on sait qu'une protéine peut assurer une grande variété de fonctions et qu'elle peut avoir de nombreux partenaires d'interaction, tout en possédant des domaines non-structurés. Ce changement de paradigme est particulièrement marqué dans le cas des protéines intrinsèquement désordonnées qui ont une grande adaptabilité et de multiples partenaires d'interaction. Au sein de la famille des protéines de liaison à l'ARN, on trouve une grande proportion de protéines aux domaines non structurés, ces protéines sont notamment impliquées dans le processus de séparation de phases. Ce phénomène de compartimentation et de formation de condensats intervient notamment dans la prise en charge de l'ARN. Face à la complexité de ces mécanismes cellulaires découverts relativement récemment, il reste beaucoup à faire pour mieux comprendre comment l'on peut passer d'une altération des processus cellulaires de base, jusqu'au développement d'un dysfonctionnement global de l'organisme avec l'émergence de pathologies. Pour cela, les maladies neurodégénératives sont un véritable défi car les mécanismes auxquels elles sont associées restent encore mal compris.

Dans ce rapport, nous avons abordé le cas de protéines de liaison à l'ARN pourvues d'un de domaine de faible complexité, à travers l'exemple de deux protéines qui sont des marqueurs caractéristiques des maladies neurodégénératives : TDP-43 et FUS. Ces dernières sont des protéines impliquées dans de nombreuses fonctions en lien avec la prise en charge de l'ARN. Elles ont des modes de liaison différents, et en dépit d'un ensemble de cibles relativement distinctes, elles possèdent des cibles communes, peuvent interagir, s'influencer et moduler leur comportement.

L'étude des interactions protéine-protéine permet d'identifier les partenaires d'interaction d'une protéine, de mieux comprendre son mode de fonctionnement, ainsi que les fonctions qu'elle exerce. Dans ce domaine, il existe de nombreuses techniques. Ici nous avons utilisé des techniques originales qui nous ont permis de caractériser l'interaction entre TDP-43 et FUS, notamment en la comparant à l'interaction avec d'autres RBP, puis en abordant une notion de miscibilité et de compartimentation. Enfin, grâce à l'AFM nous avons pu montrer que les deux protéines pouvaient coopérer pour se lier à l'ARN, et qu'en particulier FUS participait à une meilleure répartition de TDP-43 sur l'ARN.

Le processus de coopération décrit ici se réfère à l'addition de plusieurs protéines sur une même cible dans un but commun, les protéines peuvent alors interagir de manière directe ou indirecte. Le fait que les deux protéines se répartissent sur l'ARN en mélange, tel qu'observé en AFM, suggère une interaction directe entre les deux protéines. Toutefois, les données présentées ici ne permettent pas d'identifier clairement comment s'établit ce contact protéine-protéine. Plusieurs expérimentations amènent à penser que le RRM joue un rôle, mais il serait pertinent d'étudier plus précisément quels acides aminés sont impliqués dans l'interaction entre FUS et TDP-43 par le biais d'analyses à l'échelle du résidu comme la RMN, ou encore des modélisations en dynamique moléculaire. Ainsi, nous pourrions cibler ces résidus, et vérifier si les mutations ou les MPT qui les touchent pourraient être à l'origine d'un changement dans l'interaction entre les deux protéines.

Nous n'avons pu le vérifier au cours de ce travail, mais l'on peut supposer que cette interaction est nécessaire au bon fonctionnement de ces protéines, et à la bonne prise en charge de l'ARN dans un contexte physiologique. De plus, on peut supposer que la perturbation de cette interaction peut induire des conséquences délétères, notamment au vu des nombreuses fonctions auxquelles participent ces protéines. Un dysfonctionnement dans cette interaction pourrait entraîner une mauvaise prise en charge de l'ARN, et plus largement une perte de fonction des protéines. C'est pourquoi, il serait intéressant d'étudier plus en détail les phénomènes qui peuvent interférer comme les mutations, ou encore les MPT.

Etudier cette interaction au sein d'un modèle cellulaire de neurones serait particulièrement intéressant pour identifier l'implication de celle-ci en condition physiologique. Par exemple, nous pourrions regarder la localisation des protéines FUS et TDP-43 dans des neurones, et voir comment elles réagissent en condition de stress cellulaire, mais aussi, voir comment la déplétion d'une protéine affecte le comportement de l'autre. Si le niveau d'expression de FUS est diminué dans la cellule, est-ce que l'on trouve toujours TDP-43 dans les granules de stress ? Et inversement ? L'induction de granules de stress permettrait de mieux visualiser la co-localisation au sein de mêmes compartiments en condition physiologique, mais aussi d'étudier le comportement de ces protéines suivant des niveaux d'expression endogène ce qui n'est pas le cas avec le MtB.

Nous avons vu que l'ARN et notamment sa concentration jouait un rôle essentiel dans la régulation du comportement de ces protéines. Aussi une grande quantité d'ARN solubilise les protéines et favorise leurs interactions, tandis qu'une faible quantité les force à s'agréger et les mets en compétition. Ici, nous avons choisi d'utiliser un long ARN qui possède une diversité de séquence et ne présente pas de

spécificité pour l'une ou l'autre des protéines, de manière à obtenir des informations sur un comportement général. Etudier d'autres types d'ARN avec des motifs particuliers (riches ou pauvres en motifs GU) ou encore des cibles connues pour l'une ou l'autre de ces protéines comme des séquences de longs introns, pourrait être intéressant de façon à voir comment se comportent les deux protéines dans un cas plus spécifique, et ainsi de mieux comprendre les paramètres qui régissent le comportement de ces protéines.

Ici, nous avons vu, en particulier avec le modèle d'immunofluorescence *in vitro* et l'AFM, que TDP25 et TDP35 avaient des comportements distincts des formes PL. Pour approfondir le sujet, il aurait été intéressant de faire des mélanges avec à la fois la protéine FUS, TDP-43 et les formes clivées pathologiques pour voir si ces dernières pourraient induire des changements sur le comportement des formes PL et pourraient être capables de perturber l'interaction entre TDP-43 et FUS PL.

Pour aller plus loin, il serait aussi intéressant de voir quelles mutations pathologiques, en particulier au niveau du RRM, pourraient affecter l'interaction entre TDP-43 et FUS. Un travail préliminaire sur ce sujet a été effectué, avant d'être mis de côté au profit d'autres résultats, ce point demanderait ainsi à être approfondi. On peut supposer que les mutations pathologiques au niveau du RRM de TDP-43, telles que K181E, D169G, N259S, K263E, et F147L/F149L qui affectent le comportement de la protéine, en particulier modifient son affinité à l'ARN, pourraient avoir un effet sur l'interaction avec FUS. En plus de l'étude de ces mutations par les méthodes décrites dans ce rapport (MtB, immunofluorescence *In vitro*, AFM) il pourrait être particulièrement intéressant de regarder le comportement des protéines TDP-43 et FUS au sein de modèle de cellules iPCS portant ces mutations.

De même, il serait pertinent d'étudier l'effet des MPT qui touchent le domaine RRM de TDP-43. On sait que les protéines présentent de nombreuses MPT, qui jouent un rôle dans leur régulation, mais que ces MPT, en particulier la phosphorylation, l'acétylation et l'ubiquitination sont particulièrement abondantes chez les protéines présentes dans les agrégats cytoplasmiques. Ainsi, l'acétylation des lysines K136, K145, K192, ou encore la phosphorylation des sérines S183, S242 et S254 au niveau des RRM pourraient induire un changement de comportement entre TDP-43 et FUS.

Enfin, approfondir les structures formées par les formes tronquées, telles qu'observées en AFM, et mieux comprendre quelles sont les conditions qui peuvent favoriser leur formation, en particulier en présence des formes PL, serait pertinent. Un travail qui serait aussi intéressant à effectuer avec des formes mutées ou ayant subi des MPT.

Pour conclure, ce travail offre d'une part, un point de vue original sur l'interaction et le comportement de deux protéines de liaison à l'ARN, en particulier dans leur prise en charge de l'ARN. Il apporte ainsi quelques réponses sur le comportement qu'adoptent les protéines pour se répartir sur l'ARN, comme sur l'organisation des protéines en condensats de différents degrés d'homogénéité. Et d'autre part, il soulève de nombreuses questions sur les conséquences physiologiques ou pathologiques de cette interaction. Enfin, la formation de condensats dont l'homogénéité dépend du ratio en protéine, nous amène à nous interroger sur les paramètres qui régulent ses interactions au sein des cellules. Quelles sont les autres protéines qui peuvent adopter ce type de comportement ? Quelles sont les conséquences et implications de ce phénomène ?

# Table des illustrations

<i>Figure 1-Les protéines de liaisons à l'ARN contrôlent la vie de l'ARN.</i>	5
<i>Figure 2-Le RRM reconnaît spécifiquement la séquence de l'ARN.</i>	7
<i>Figure 3-Structure des domaines KH.</i>	9
<i>Figure 4-Les protéines de liaison à l'ARN possèdent des modes de liaison variés.</i>	11
<i>Figure 5-Différents modes d'interaction entre les protéines de liaison à l'ARN.</i>	15
<i>Figure 6-Un ensemble de protéines de liaison à l'ARN avec des caractéristiques de type prion</i>	20
<i>Figure 7-Les transitions de phases liquide-liquide</i>	26
<i>Figure 8-La transition de phases peut conduire à l'agrégation des protéines.</i>	30
<i>Figure 9-Etapes de formation de fibres : exemple de la maladie d'Alzheimer.</i>	34
<i>Figure 10-Le spectre des MND est continu.</i>	42
<i>Figure 11-Les agrégats de TDP-43 et FUS sont une caractéristique commune de la SLA et la DFT.</i>	44
<i>Figure 12- Domaines de la protéine TDP-43.</i>	46
<i>Figure 13-Le domaine PrLD de TDP-43 est principalement non structuré.</i>	47
<i>Figure 14-Le domaine Nter de TDP-43 est structuré.</i>	48
<i>Figure 15-Structure du domaine de liaison à l'ARN de TDP-43.</i>	49
<i>Figure 16-Séquence de la protéine FUS.</i>	51
<i>Figure 17-Structure du domaine de liaison à l'ARN de FUS.</i>	52
<i>Figure 18-TDP-43 et FUS sont associées à de nombreuses fonctions en lien avec l'ARN/ADN.</i>	56
<i>Figure 19-Le comportement des RBP est régulé par diverses modifications post-traductionnelles.</i>	67
<i>Figure 20-La protéine TDP-43 présente un nombre important de mutations et de sites de modifications post-traductionnelles.</i>	69
<i>Figure 21-La protéine FUS présente un nombre important de mutations et de sites de modifications post-traductionnelles.</i>	70
<i>Figure 22-Mise en évidence des différences d'interaction avec l'ARN entre TDP-43 et FUS.</i>	73
<i>Figure 23-Comparaison des cibles et des propriétés de liaison à l'ARN des protéines TDP-43 et FUS.</i>	74
<i>Figure 24-FUS et TDP-43 sont localisées dans de nombreuses organelles en commun.</i>	75
<i>Figure 25-TDP-43 et FUS partagent des fonctions en commun par le biais de sites de liaisons qui se chevauchent.</i>	79
<i>Figure 26-TDP-43 et FUS co-localisent au sein d'agrégats dans les neurones de patients atteints de la SLA.</i>	82
<i>Figure 27-Principe de la technique du pull-down.</i>	89
<i>Figure 28-Principe de l'ITC.</i>	90
<i>Figure 29-Principe de la SPR.</i>	91
<i>Figure 30-Principe de l'essai de complémentation des fragments protéiques</i>	92
<i>Figure 31-Principe de la technique FRET.</i>	93
<i>Figure 32-Principe de la PLA.</i>	94
<i>Figure 33-Etapes pour le calcul du coefficient de corrélation pour la détection d'interactions avec le microtubule Bench système proie/appât</i>	102
<i>Figure 34-Etapes pour le calcul du coefficient de corrélation pour la méthode du microtubule Bench</i>	106
<i>Figure 35-Principe du Microtubule Bench.</i>	114
<i>Figure 36-Comparaison de l'interaction TDP-43-RFP-MBD avec diverses RBP grâce au système du microtubule bench.</i>	115
<i>Figure 37-L'interaction du couple TDP-43 et FUS est plus importante que pour d'autres protéines de liaison à l'ARN.</i>	116
<i>Figure 38-Comparaison de l'interaction FUS-RFP-MBD avec diverses RBP grâce au système du Microtubule Bench.</i>	118
<i>Figure 39-Confirmation de l'interaction FUS/TD-43 par la technique de PLA.</i>	119

<i>Figure 40-L'interaction entre FUS et TDP-43 est asymétrique.....</i>	<i>121</i>
<i>Figure 41-FUS forme des granules dans lesquels co-localisent l'ARN quand elle est surexprimée.....</i>	<i>122</i>
<i>Figure 42-Mesure de la miscibilité de protéines avec le MtB.....</i>	<i>123</i>
<i>Figure 43-Séquences des protéines étudiées.....</i>	<i>124</i>
<i>Figure 44-Cellules co-transfectées avec des ratios de plasmides différents pour plusieurs couples de RBP.....</i>	<i>125</i>
<i>Figure 45-La miscibilité de TDP-43 et FUS dépend de leur ratio.....</i>	<i>126</i>
<i>Figure 46-Etude du comportement des protéines avec l'ARN In vitro par gels de sédimentation.....</i>	<i>129</i>
<i>Figure 47-Principe de la technique de visualisation des agrégats par immunofluorescence.....</i>	<i>130</i>
<i>Figure 48-Evolution des objets formés In vitro par les protéines TDP-43 et FUS avec le temps.....</i>	<i>131</i>
<i>Figure 49-Evolution des objets formés In vitro par les protéines TDP-43 et FUS en fonction de la concentration.....</i>	<i>132</i>
<i>Figure 50-Le mélange de TDP-43/FUS suit le comportement de la protéine FUS.....</i>	<i>133</i>
<i>Figure 51-La miscibilité des protéines varie en fonction de leurs ratios.....</i>	<i>134</i>
<i>Figure 52-L'ARN a un effet ambivalent sur les RBP.....</i>	<i>136</i>
<i>Figure 53- Comportement des protéines en mélange avec l'ARN par immunofluorescence.....</i>	<i>137</i>
<i>Figure 54-L'ARN favorise la miscibilité des protéines TDP-43 et FUS.....</i>	<i>139</i>
<i>Figure 55-Les acquisitions confocales donnent un aperçu de la structure 3D des objets.....</i>	<i>141</i>
<i>Figure 56-Expériences de gel retard montrant la coopération entre TDP-43 et FUS dans leur liaison à l'ARN.....</i>	<i>143</i>
<i>Figure 57-TDP-43 se répartie sur l'ARN grâce à FUS.....</i>	<i>145</i>
<i>Figure 58-Mesure de la miscibilité des formes tronquées avec les formes PL grâce au microtubule Bench.....</i>	<i>146</i>
<i>Figure 59-Le RRM joue un rôle important dans l'interaction entre FUS et TDP-43, tandis qu'elle est affectée par les formes pathologiques de TDP-43.....</i>	<i>148</i>
<i>Figure 60-Images représentatives pour les différentes formes tronquées en mélange avec les formes PL, analysées au MtB.....</i>	<i>150</i>
<i>Figure 61-Contrôle des formes tronquées des protéines par Western Blot.....</i>	<i>152</i>
<i>Figure 62-Aperçu du comportement des formes tronquées par immunofluorescence.....</i>	<i>153</i>
<i>Figure 63-Le comportement des formes tronquées pathologiques est différent des formes PL par immunofluorescence In vitro.....</i>	<i>155</i>
<i>Figure 64-Evolution de l'aire des objets formés par les formes tronquées en mélange avec les formes PL en immunofluorescence In vitro.....</i>	<i>157</i>
<i>Figure 65-Les formes tronquées TDP25 et TDP35 induisent une perte de la coopération avec FUS.....</i>	<i>159</i>
<i>Figure 66-Quantification des objets formés par les formes tronquées en mélange avec les formes PL en AFM.....</i>	<i>161</i>
<i>Figure 67-Tableau récapitulatif pour les formes tronquées en AFM.....</i>	<i>162</i>
<i>Figure 68-Les formes tronquées forment des structures particulières à l'échelle nanométrique.....</i>	<i>163</i>
<i>Figure 69-Modèle de l'interaction entre FUS et TDP-43 sur l'ARN.....</i>	<i>166</i>

# Liste des tableaux

---

<i>Tableau 1-Comparaison de méthodes de détection des interactions protéine-protéine.....</i>	<i>99</i>
<i>Tableau 2-Plasmides produits et utilisés pour le microtubule bench méthode proie/appât.....</i>	<i>101</i>
<i>Tableau 3-Plasmides utilisés pour la technique du microtubule bench, méthode de compartimentation.....</i>	<i>104</i>
<i>Tableau 4-Plasmides produits pour les formes tronquées pour la technique du microtubule bench, méthode de compartimentation.....</i>	<i>104</i>
<i>Tableau 5-Plasmides produits pour la production de protéines In vitro.....</i>	<i>107</i>
<i>Tableau 6-Listes des oligonucléotides générés pour le clonage des plasmides pour la production de formes tronquées In vitro.....</i>	<i>108</i>
<i>Tableau 7-Anticorps utilisés en immunofluorescence.....</i>	<i>112</i>

# Liste des boîtes de texte

---

<i>Boîte de texte 1-Prédire les domaines de faible complexité.....</i>	<i>19</i>
<i>Boîte de texte 2-Modes de liaison des LCD.....</i>	<i>22</i>
<i>Boîte de texte 3-Les protéines chaperonnes.....</i>	<i>27</i>
<i>Boîte de texte 4-Fibres amyloïdes.....</i>	<i>32</i>
<i>Boîte de texte 5-Propagation des fragments.....</i>	<i>33</i>
<i>Boîte de texte 6-Facteurs de risque pour la SLA.....</i>	<i>38</i>
<i>Boîte de texte 7-Traitements existants et perspectives pour la SLA.....</i>	<i>39</i>
<i>Boîte de texte 8-Techniques de détection des interactions protéine-protéine.....</i>	<i>77</i>
<i>Boîte de texte 9-Contexte pour la technique du microtubule Bench.....</i>	<i>96</i>



# Références

---

1. Dumurgier J, Sabia S. [Epidemiology of Alzheimer's disease: latest trends]. *Rev Prat*. 2020;70:149–51.
2. Uversky V. The roles of intrinsic disorder-based liquid-liquid phase transitions in the “Dr. Jekyll–Mr. Hyde” behavior of proteins involved in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Autophagy*. 2017. <https://doi.org/10.1080/15548627.2017.1384889>.
3. Carey JL, Guo L. Liquid-Liquid Phase Separation of TDP-43 and FUS in Physiology and Pathology of Neurodegenerative Diseases. *Front Mol Biosci*. 2022;9:826719.
4. Loughlin FE, Wilce JA. TDP-43 and FUS—structural insights into RNA recognition and self-association. *Curr Opin Struct Biol*. 2019;59:134–42.
5. Ling S-C, Polymenidou M, Cleveland DW. Converging mechanisms in ALS and FTD: Disrupted RNA and protein homeostasis. *Neuron*. 2013;79:416–38.
6. Lagier-Tourenne C, Polymenidou M, Hutt KR, Vu AQ, Baughn M, Huelga SC, et al. Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat Neurosci*. 2012;15:1488–97.
7. Shelkownikova T. Modelling FUSopathies: focus on protein aggregation. *Biochem Soc Trans*. 2013. <https://doi.org/10.1042/BST20130212>.
8. Buratti E, Baralle FE. TDP-43: gumming up neurons through protein–protein and protein–RNA interactions. *Trends Biochem Sci*. 2012;37:237–47.
9. Mann JR, Gleixner AM, Mauna JC, Gomes E, DeChellis-Marks MR, Needham PG, et al. RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. *Neuron*. 2019;102:321–338.e8.
10. Khong A, Matheny T, Jain S, Mitchell SF, Wheeler JR, Parker R. The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Mol Cell*. 2017;68:808–820.e5.
11. Zhang P, Fan B, Yang P, Temirov J, Messing J, Kim HJ, et al. Chronic optogenetic induction of stress granules is cytotoxic and reveals the evolution of ALS-FTD pathology. *eLife*. 2019;8:e39578.
12. Li YR, King OD, Shorter J, Gitler AD. Stress granules as crucibles of ALS pathogenesis. *J Cell Biol*. 2013;201:361–72.

13. Dassi E. Handshakes and Fights: The Regulatory Interplay of RNA-Binding Proteins. *Front Mol Biosci.* 2017;4:67.
14. Kim SH, Shanware NP, Bowler MJ, Tibbetts RS. Amyotrophic Lateral Sclerosis-associated Proteins TDP-43 and FUS/TLS Function in a Common Biochemical Complex to Co-regulate HDAC6 mRNA. *J Biol Chem.* 2010;285:34097–105.
15. Kryndushkin D, Wickner RB, Shewmaker F. FUS/TLS forms cytoplasmic aggregates, inhibits cell growth and interacts with TDP-43 in a yeast model of amyotrophic lateral sclerosis. *Protein Cell.* 2011;2:223–36.
16. Ling S-C, Albuquerque CP, Han JS, Lagier-Tourenne C, Tokunaga S, Zhou H, et al. ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *Proc Natl Acad Sci U S A.* 2010;107:13318–23.
17. Kabashi E, Bercier V, Lissouba A, Liao M, Brustein E, Rouleau GA, et al. FUS and TARDBP but not SOD1 interact in genetic models of amyotrophic lateral sclerosis. *PLoS Genet.* 2011;7:e1002214.
18. Lanson NA, Maltare A, King H, Smith R, Kim JH, Taylor JP, et al. A Drosophila model of FUS-related neurodegeneration reveals genetic interaction between FUS and TDP-43. *Hum Mol Genet.* 2011;20:2510–23.
19. Wang J-W, Brent JR, Tomlinson A, Shneider NA, McCabe BD. The ALS-associated proteins FUS and TDP-43 function together to affect Drosophila locomotion and life span. *J Clin Invest.* 2011;121:4118–26.
20. Deng H-X, Zhai H, Bigio EH, Yan J, Fecto F, Ajroud K, et al. FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis. *Ann Neurol.* 2010;67:739–48.
21. Hentze MW, Castello A, Schwarzl T, Preiss T. A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol.* 2018;19:327–41.
22. Harrison AF, Shorter J. RNA-binding proteins with prion-like domains in health and disease. *Biochem J.* 2017;474:1417–38.
23. Pereira B, Billaud M, Almeida R. RNA-Binding Proteins in Cancer: Old Players and New Actors. *Trends Cancer.* 2017;3:506–28.
24. Gebauer F, Schwarzl T, Valcárcel J, Hentze MW. RNA-binding proteins in human genetic disease. *Nat Rev Genet.* 2021;22:185–98.
25. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science.* 2001;291:1304–51.
26. Mattaj JW. RNA recognition: a family matter? *Cell.* 1993;73:837–40.

27. Stefl R, Skrisovska L, Allain FH-T. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep.* 2005;6:33–8.
28. Daubner GM, Cléry A, Allain FH-T. RRM-RNA recognition: NMR or crystallography...and new findings. *Curr Opin Struct Biol.* 2013;23:100–8.
29. Dominguez C, Fiset J-F, Chabot B, Allain FH-T. Structural basis of G-tract recognition and encaging by hnRNP F quasi-RRMs. *Nat Struct Mol Biol.* 2010;17:853–61.
30. Singh M, Choi CP, Feigon J. xRRM: a new class of RRM found in the telomerase La family protein p65. *RNA Biol.* 2013;10:353–9.
31. Lunde BM, Moore C, Varani G. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol.* 2007;8:479–90.
32. Linder P, Jankowsky E. From unwinding to clamping - the DEAD box RNA helicase family. *Nat Rev Mol Cell Biol.* 2011;12:505–16.
33. Bhaskaran H, Russell R. Kinetic redistribution of native and misfolded RNAs by a DEAD-box chaperone. *Nature.* 2007;449:1014–8.
34. Noble CG, Song H. MLN51 stimulates the RNA-helicase activity of eIF4AIII. *PLoS One.* 2007;2:e303.
35. Valverde R, Edwards L, Regan L. Structure and function of KH domains. *FEBS J.* 2008;275:2712–26.
36. Lindquist JA, Mertens PR. Cold shock proteins: from cellular mechanisms to pathophysiology and disease. *Cell Commun Signal CCS.* 2018;16:63.
37. Méndez-Vidal C, Wilhelm MT, Hellborg F, Qian W, Wiman KG. The p53-induced mouse zinc finger protein wig-1 binds double-stranded RNA with high affinity. *Nucleic Acids Res.* 2002;30:1991–6.
38. Theunissen O, Rudt F, Guddat U, Mentzel H, Pieler T. RNA and DNA binding zinc fingers in *Xenopus* TFIIIA. *Cell.* 1992;71:679–90.
39. Lu D, Searles MA, Klug A. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature.* 2003;426:96–100.
40. Hudson BP, Martinez-Yamout MA, Dyson HJ, Wright PE. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol.* 2004;11:257–64.

41. Ramos A, Grünert S, Adams J, Micklem DR, Proctor MR, Freund S, et al. RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J.* 2000;19:997–1009.
42. Wilson R, Doudna JA. Molecular mechanisms of RNA interference. *Annu Rev Biophys.* 2013;42:217–39.
43. Raven JF, Koromilas AE. PERK and PKR: Old kinases learn new tricks. *Cell Cycle.* 2008;7:1146–50.
44. Casu F, Duggan BM, Hennig M. The arginine-rich RNA-binding motif of HIV-1 Rev is intrinsically disordered and folds upon RRE binding. *Biophys J.* 2013;105:1004–17.
45. de Breyne S, Yu Y, Unbehaun A, Pestova TV, Hellen CUT. Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proc Natl Acad Sci U S A.* 2009;106:9197–202.
46. Castello A, Fischer B, Frese CK, Horos R, Alleaume A-M, Foehr S, et al. Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol Cell.* 2016;63:696–710.
47. Beckmann BM, Castello A, Medenbach J. The expanding universe of ribonucleoproteins: of novel RNA-binding proteins and unconventional interactions. *Pflugers Arch.* 2016;468:1029–40.
48. Gebhardt A, Habjan M, Benda C, Meiler A, Haas DA, Hein MY, et al. mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nat Commun.* 2015;6:8192.
49. Rahman MA, Masuda A, Ohe K, Ito M, Hutchinson DO, Mayeda A, et al. HnRNP L and hnRNP LL antagonistically modulate PTB-mediated splicing suppression of CHRNA1 pre-mRNA. *Sci Rep.* 2013;3:2931.
50. Okunola HL, Krainer AR. Cooperative-binding and splicing-repressive properties of hnRNP A1. *Mol Cell Biol.* 2009;29:5620–31.
51. Nakamura H, Kawagishi H, Watanabe A, Sugimoto K, Maruyama M, Sugimoto M. Cooperative role of the RNA-binding proteins Hzf and HuR in p53 activation. *Mol Cell Biol.* 2011;31:1997–2009.
52. Kretov DA, Curmi PA, Hamon L, Abrakhi S, Desforges B, Ovchinnikov LP, et al. mRNA and DNA selection via protein multimerization: YB-1 as a case study. *Nucleic Acids Res.* 2015;43:9457–73.
53. Schwartz JC, Wang X, Podell ER, Cech TR. RNA Seeds Higher-Order Assembly of FUS Protein. *Cell Rep.* 2013;5:918–25.

54. Wang X, Schwartz JC, Cech TR. Nucleic acid-binding specificity of human FUS protein. *Nucleic Acids Res.* 2015;43:7535–43.
55. Buratti E, Brindisi A, Giombi M, Tisminetzky S, Ayala YM, Baralle FE. TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. *J Biol Chem.* 2005;280:37572–84.
56. Rengifo-Gonzalez JC, El Hage K, Clément M-J, Steiner E, Joshi V, Craveur P, et al. The cooperative binding of TDP-43 to GU-rich RNA repeats antagonizes TDP-43 aggregation. *eLife.* 2021;10:e67605.
57. Zhu J, Mayeda A, Krainer AR. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell.* 2001;8:1351–61.
58. Lal A, Mazan-Mamczarz K, Kawai T, Yang X, Martindale JL, Gorospe M. Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs. *EMBO J.* 2004;23:3092–102.
59. Wigington CP, Jung J, Rye EA, Belauret SL, Philpot AM, Feng Y, et al. Post-transcriptional regulation of programmed cell death 4 (PDCD4) mRNA by the RNA-binding proteins human antigen R (HuR) and T-cell intracellular antigen 1 (TIA1). *J Biol Chem.* 2015;290:3468–87.
60. Abrakhi S, Kretov DA, Desforges B, Dobra I, Bouhss A, Pastré D, et al. Nanoscale Analysis Reveals the Maturation of Neurodegeneration-Associated Protein Aggregates: Grown in mRNA Granules then Released by Stress Granule Proteins. *ACS Nano.* 2017;11:7189–200.
61. Dassi E, Re A, Leo S, Tebaldi T, Pasini L, Peroni D, et al. AURA 2: Empowering discovery of post-transcriptional networks. *Transl Austin Tex.* 2014;2:e27738.
62. Zanzoni A, Marchese D, Agostini F, Bolognesi B, Cirillo D, Botta-Orfila M, et al. Principles of self-organization in biological pathways: a hypothesis on the autogenous association of alpha-synuclein. *Nucleic Acids Res.* 2013;41:9987–98.
63. Ayala YM, De Conti L, Avendaño-Vázquez SE, Dhir A, Romano M, D'Ambrogio A, et al. TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* 2011;30:277–88.
64. Avendaño-Vázquez SE, Dhir A, Bembich S, Buratti E, Proudfoot N, Baralle FE. Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes Dev.* 2012;26:1679.

65. Zhou Y, Liu S, Liu G, Oztürk A, Hicks GG. ALS-associated FUS mutations result in compromised FUS alternative splicing and autoregulation. *PLoS Genet.* 2013;9:e1003895.
66. Dassi E, Zuccotti P, Leo S, Provenzani A, Assfalg M, D'Onofrio M, et al. Hyper conserved elements in vertebrate mRNA 3'-UTRs reveal a translational network of RNA-binding proteins controlled by HuR. *Nucleic Acids Res.* 2013;41:3201–16.
67. Mansfield KD, Keene JD. Neuron-specific ELAV/Hu proteins suppress HuR mRNA during neuronal differentiation by alternative polyadenylation. *Nucleic Acids Res.* 2012;40:2734–46.
68. Lu L, Zheng L, Si Y, Luo W, Dujardin G, Kwan T, et al. Hu antigen R (HuR) is a positive regulator of the RNA-binding proteins TDP-43 and FUS/TLS: implications for amyotrophic lateral sclerosis. *J Biol Chem.* 2014;289:31792–804.
69. Jangi M, Boutz PL, Paul P, Sharp PA. Rbfox2 controls autoregulation in RNA-binding protein networks. *Genes Dev.* 2014;28:637–51.
70. Tollervey JR, Curk T, Rogelj B, Briese M, Cereda M, Kayikci M, et al. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci.* 2011;14:452–8.
71. Sephton CF, Cenik C, Kucukural A, Dammer EB, Cenik B, Han Y, et al. Identification of Neuronal RNA Targets of TDP-43-containing Ribonucleoprotein Complexes♦. *J Biol Chem.* 2011;286:1204–15.
72. Dosztányi Z. Prediction of protein disorder based on IUPred. *Protein Sci Publ Protein Soc.* 2018;27:331–40.
73. Dunker AK, Silman I, Uversky VN, Sussman JL. Function and structure of inherently disordered proteins. *Curr Opin Struct Biol.* 2008;18:756–64.
74. Kumari B, Kumar R, Kumar M. Low complexity and disordered regions of proteins have different structural and amino acid preferences. *Mol Biosyst.* 2015;11:585–94.
75. Mier P, Paladin L, Tamana S, Petrosian S, Hajdu-Soltész B, Urbanek A, et al. Disentangling the complexity of low complexity proteins. *Brief Bioinform.* 2020;21:458–72.
76. King OD, Gitler AD, Shorter J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res.* 2012;1462:61–80.
77. Boisvert F-M, Chénard CA, Richard S. Protein interfaces in signaling regulated by arginine methylation. *Sci STKE Signal Transduct Knowl Environ.* 2005;2005:re2.

78. Thandapani P, O'Connor TR, Bailey TL, Richard S. Defining the RGG/RG Motif. *Mol Cell*. 2013;50:613–23.
79. Hughes MP, Sawaya MR, Boyer DR, Goldschmidt L, Rodriguez JA, Cascio D, et al. Atomic structures of low-complexity protein segments reveal kinked  $\beta$  sheets that assemble networks. *Science*. 2018;359:698–701.
80. Hughes MP, Goldschmidt L, Eisenberg DS. Prevalence and species distribution of the low-complexity, amyloid-like, reversible, kinked segment structural motif in amyloid-like fibrils. *J Biol Chem*. 2021;297:101194.
81. Michelitsch MD, Weissman JS. A census of glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel prions. *Proc Natl Acad Sci*. 2000;97:11910–5.
82. March ZM, King OD, Shorter J. Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain Res*. 2016;1647:9–18.
83. Couthouis J, Hart MP, Shorter J, DeJesus-Hernandez M, Erion R, Oristano R, et al. A yeast functional screen predicts new candidate ALS disease genes. *Proc Natl Acad Sci*. 2011;108:20881–90.
84. Dobra I, Pankivskiy S, Samsonova A, Pastre D, Hamon L. Relation Between Stress Granules and Cytoplasmic Protein Aggregates Linked to Neurodegenerative Diseases. *Curr Neurol Neurosci Rep*. 2018;18:107.
85. Banfi S, Servadio A, Chung M yi, Kwiatkowski TJ, McCall AE, DuVick LA, et al. Identification and characterization of the gene causing type 1 spinocerebellar ataxia. *Nat Genet*. 1994;7:513–20.
86. Lorenzetti D, Bohlega S, Zoghbi HY. The expansion of the CAG repeat in ataxin-2 is a frequent cause of autosomal dominant spinocerebellar ataxia. *Neurology*. 1997;49:1009–13.
87. Neumann M, Bentmann E, Dormann D, Jawaid A, DeJesus-Hernandez M, Ansorge O, et al. FET proteins TAF15 and EWS are selective markers that distinguish FTLD with FUS pathology from amyotrophic lateral sclerosis with FUS mutations. *Brain*. 2011;134:2595–609.
88. Kim HJ, Kim NC, Wang Y-D, Scarborough EA, Moore J, Diaz Z, et al. Prion-like domain mutations in hnRNPs cause multisystem proteinopathy and ALS. *Nature*. 2013;495:467–73.

89. Tada M, Doi H, Koyano S, Kubota S, Fukai R, Hashiguchi S, et al. Matrin 3 Is a Component of Neuronal Cytoplasmic Inclusions of Motor Neurons in Sporadic Amyotrophic Lateral Sclerosis. *Am J Pathol.* 2018;188:507–14.
90. Courey AJ, Holtzman DA, Jackson SP, Tjian R. Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. *Cell.* 1989;59:827–36.
91. Triezenberg SJ, Kingsbury RC, McKnight SL. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev.* 1988;2:718–29.
92. Frey S, Richter RP, Görlich D. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science.* 2006;314:815–7.
93. Pajkos M, Dosztányi Z. Functions of intrinsically disordered proteins through evolutionary lenses. *Prog Mol Biol Transl Sci.* 2021;183:45–74.
94. van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, et al. Classification of Intrinsically Disordered Regions and Proteins. *Chem Rev.* 2014;114:6589–631.
95. Linke WA, Kulke M, Li H, Fujita-Becker S, Neagoe C, Manstein DJ, et al. PEVK domain of titin: an entropic spring with actin-binding properties. *J Struct Biol.* 2002;137:194–205.
96. Bardwell JCA, Jakob U. Conditional disorder in chaperone action. *Trends Biochem Sci.* 2012;37:517–25.
97. Boquet I, Boujemaa R, Carlier MF, Préat T. Ciboulot regulates actin assembly during *Drosophila* brain metamorphosis. *Cell.* 2000;102:797–808.
98. Tompa P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.* 2005;579:3346–54.
99. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradović Z. Intrinsic disorder and protein function. *Biochemistry.* 2002;41:6573–82.
100. Wright PE, Dyson HJ. Intrinsically Disordered Proteins in Cellular Signaling and Regulation. *Nat Rev Mol Cell Biol.* 2015;16:18–29.
101. Ntountoumi C, Vlastaridis P, Mossialos D, Stathopoulos C, Iliopoulos I, Promponas V, et al. Low complexity regions in the proteins of prokaryotes perform important functional roles and are highly conserved. *Nucleic Acids Res.* 2019;47:9998–10009.



102. Ekman D, Light S, Björklund AK, Elofsson A. What properties characterize the hub proteins of the protein-protein interaction network of *Saccharomyces cerevisiae*? *Genome Biol.* 2006;7:R45.
103. Perutz MF, Johnson T, Suzuki M, Finch JT. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc Natl Acad Sci.* 1994;91:5355–8.
104. Vernon RM, Chong PA, Tsang B, Kim TH, Bah A, Farber P, et al. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife.* 2018;7:e31486.
105. Mittag T, Parker R. Multiple Modes of Protein–Protein Interactions Promote RNP Granule Assembly. *J Mol Biol.* 2018;430:4636–49.
106. Conicella AE, Zerze GH, Mittal J, Fawzi NL. ALS Mutations Disrupt Phase Separation Mediated by  $\alpha$ -Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Struct Lond Engl* 1993. 2016;24:1537–49.
107. Wang J, Choi J-M, Holehouse AS, Lee HO, Zhang X, Jahnel M, et al. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell.* 2018;174:688–699.e16.
108. Bedford MT, Clarke SG. Protein arginine methylation in mammals: who, what, and why. *Mol Cell.* 2009;33:1–13.
109. Mt B, A F, Mb Y, S C, P L, S R. Arginine methylation inhibits the binding of proline-rich ligands to Src homology 3, but not WW, domains. *J Biol Chem.* 2000;275.
110. Tripsianes K, Madl T, Machyna M, Fessas D, Englbrecht C, Fischer U, et al. Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat Struct Mol Biol.* 2011;18:1414–20.
111. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, et al. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 2004;32:1037–49.
112. Ojala PM, Yamamoto K, Castaños-Vélez E, Biberfeld P, Korsmeyer SJ, Mäkelä TP. The apoptotic v-cyclin-CDK6 complex phosphorylates and inactivates Bcl-2. *Nat Cell Biol.* 2000;2:819–25.
113. Monahan Z, Ryan VH, Janke AM, Burke KA, Rhoads SN, Zerze GH, et al. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* 2017;36:2951–67.
114. Courchaine EM, Lu A, Neugebauer KM. Droplet organelles? *EMBO J.* 2016;35:1603–12.

115. Poudyal RR, Pir Cakmak F, Keating CD, Bevilacqua PC. Physical Principles and Extant Biology Reveal Roles for RNA-Containing Membraneless Compartments in Origins of Life Chemistry. *Biochemistry*. 2018;57:2509–19.
116. Portz B, Lee BL, Shorter J. FUS and TDP-43 Phases in Health and Disease. *Trends Biochem Sci*. 2021;46:550–63.
117. Sephton CF, Yu G. The function of RNA-binding proteins at the synapse: implications for neurodegeneration. *Cell Mol Life Sci CMLS*. 2015;72:3621–35.
118. Decker CJ, Teixeira D, Parker R. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J Cell Biol*. 2007;179:437–49.
119. Protter DSW, Rao B, Treeck BV, Lin Y, Mizoue L, Rosen M, et al. Intrinsically disordered regions contribute promiscuous interactions to RNP granule assembly. 2017. <https://doi.org/10.1101/147561>.
120. St George-Hyslop P, Lin JQ, Miyashita A, Phillips EC, Qamar S, Randle SJ, et al. The physiological and pathological biophysics of phase separation and gelation of RNA binding proteins in amyotrophic lateral sclerosis and fronto-temporal lobar degeneration. *Brain Res*. 2018;1693:11–23.
121. Patel A, Lee HO, Jawerth L, Maharana S, Jahnelt M, Hein MY, et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*. 2015;162:1066–77.
122. Lin Y, Zhou X, Kato M, Liu D, Ghaemmghami S, Tu BP, et al. Redox-mediated regulation of an evolutionarily conserved cross- $\beta$  structure formed by the TDP43 low complexity domain. *Proc Natl Acad Sci U S A*. 2020;117:28727–34.
123. Nishimoto Y, Nakagawa S, Hirose T, Okano HJ, Takao M, Shibata S, et al. The long non-coding RNA nuclear-enriched abundant transcript 1\_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. *Mol Brain*. 2013;6:31.
124. Udan-Johns M, Bengoechea R, Bell S, Shao J, Diamond MI, True HL, et al. Prion-like nuclear aggregation of TDP-43 during heat shock is regulated by HSP40/70 chaperones. *Hum Mol Genet*. 2014;23:157–70.
125. Martinez-Macias MI, Moore DA, Green RL, Gomez-Herreros F, Naumann M, Hermann A, et al. FUS (fused in sarcoma) is a component of the cellular response to topoisomerase I-induced DNA breakage and transcriptional stress. *Life Sci Alliance*. 2019;2:e201800222.

126. Izumikawa K, Nobe Y, Ishikawa H, Yamauchi Y, Taoka M, Sato K, et al. TDP-43 regulates site-specific 2'-O-methylation of U1 and U2 snRNAs via controlling the Cajal body localization of a subset of C/D scaRNAs. *Nucleic Acids Res.* 2019;47:2487–505.
127. Vogler TO, Wheeler JR, Nguyen ED, Hughes MP, Britson KA, Lester E, et al. TDP-43 and RNA form amyloid-like myo-granules in regenerating muscle. *Nature.* 2018;563:508–13.
128. West JA, Mito M, Kurosaka S, Takumi T, Tanegashima C, Chujo T, et al. Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J Cell Biol.* 2016;214:817–30.
129. Conicella AE, Dignon GL, Zerze GH, Schmidt HB, D'Ordine AM, Kim YC, et al. TDP-43  $\alpha$ -helical structure tunes liquid-liquid phase separation and function. *Proc Natl Acad Sci U S A.* 2020;117:5883–94.
130. Gao J, Wang L, Ren X, Dunn JR, Peters A, Miyagi M, et al. Translational regulation in the brain by TDP-43 phase separation. *J Cell Biol.* 2021;220:e202101019.
131. Altmeyer M, Neelsen KJ, Teloni F, Pozdnyakova I, Pellegrino S, Grøfte M, et al. Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). *Nat Commun.* 2015;6:8088.
132. Gilks N, Kedersha N, Ayodele M, Shen L, Stoecklin G, Dember LM, et al. Stress Granule Assembly Is Mediated by Prion-like Aggregation of TIA-1. *Mol Biol Cell.* 2004;15:5383–98.
133. Tourrière H, Chebli K, Zekri L, Courselaud B, Blanchard JM, Bertrand E, et al. The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J Cell Biol.* 2003;160:823–31.
134. Baradaran-Heravi Y, Van Broeckhoven C, van der Zee J. Stress granule mediated protein aggregation and underlying gene defects in the FTD-ALS spectrum. *Neurobiol Dis.* 2020;134:104639.
135. Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj AP, Kim HJ, et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell.* 2015;163:123–33.
136. Ding Q, Chaplin J, Morris MJ, Hilliard MA, Wolvetang E, Ng DCH, et al. TDP-43 Mutation Affects Stress Granule Dynamics in Differentiated NSC-34 Motoneuron-Like Cells. *Front Cell Dev Biol.* 2021;9:611601.

137. Khalfallah Y, Kuta R, Grasmuck C, Prat A, Durham HD, Vande Velde C. TDP-43 regulation of stress granule dynamics in neurodegenerative disease-relevant cell types. *Sci Rep.* 2018;8:7551.
138. Liu-Yesucevitz L, Bilgutay A, Zhang Y-J, Vanderweyde T, Vanderwyde T, Citro A, et al. Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One.* 2010;5:e13250.
139. Kato M, Han T, Xie S, Shi K, Du X, Wu L, et al. Cell-Free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers Within Hydrogels. *Cell.* 2012;149:753–67.
140. Babinchak WM, Haider R, Dumm BK, Sarkar P, Surewicz K, Choi J-K, et al. The role of liquid-liquid phase separation in aggregation of the TDP-43 low-complexity domain. *J Biol Chem.* 2019;294:6306–17.
141. Tziortzouda P, Van Den Bosch L, Hirth F. Triad of TDP43 control in neurodegeneration: autoregulation, localization and aggregation. *Nat Rev Neurosci.* 2021;22:197–208.
142. Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovskiy A, Poser I, et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science.* 2018;360:918–21.
143. Aarum J, Cabrera C, Jones T, Rajendran S, Adiutori R, Giovannoni G, et al. Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates. *EMBO Rep.* 2020. <https://doi.org/10.15252/embr.201949585>.
144. Navarro MG-J, Kashida S, Chouaib R, Souquere S, Pierron G, Weil D, et al. RNA is a critical element for the sizing and the composition of phase-separated RNA-protein condensates. 2018. <https://doi.org/10.1101/457986>.
145. Burke KA, Janke AM, Rhine CL, Fawzi NL. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol Cell.* 2015;60:231–41.
146. Mann JR, Donnelly CJ. RNA modulates physiological and neuropathological protein phase transitions. *Neuron.* 2021;109:2663–81.
147. Yu H, Lu S, Gasior K, Singh D, Vazquez-Sanchez S, Tapia O, et al. HSP70 chaperones RNA-free TDP-43 into anisotropic intranuclear liquid spherical shells. *Science.* 2021;371:eabb4309.
148. Van Treeck B, Parker R. Emerging Roles for Intermolecular RNA-RNA Interactions in RNP Assemblies. *Cell.* 2018;174:791–802.

149. Chen AK-H, Lin RY-Y, Hsieh EZ-J, Tu P-H, Chen RP-Y, Liao T-Y, et al. Induction of amyloid fibrils by the C-terminal fragments of TDP-43 in amyotrophic lateral sclerosis. *J Am Chem Soc.* 2010;132:1186–7.
150. Berning BA, Walker AK. The Pathobiology of TDP-43 C-Terminal Fragments in ALS and FTL. *Front Neurosci.* 2019;13:335.
151. Murray DT, Kato M, Lin Y, Thurber KR, Hung I, McKnight SL, et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell.* 2017;171:615–627.e16.
152. Jiang YX, Cao Q, Sawaya MR, Abskharon R, Ge P, DeTure M, et al. Amyloid fibrils in disease FTL-DTP are composed of TMEM106B not TDP-43. *Nature.* 2022. <https://doi.org/10.1038/s41586-022-04670-9>.
153. Capitini C, Conti S, Perni M, Guidi F, Cascella R, De Poli A, et al. TDP-43 inclusion bodies formed in bacteria are structurally amorphous, non-amyloid and inherently toxic to neuroblastoma cells. *PLoS One.* 2014;9:e86720.
154. Capitini C, Fani G, Vivoli Vega M, Penco A, Canale C, Cabrita LD, et al. Full-length TDP-43 and its C-terminal domain form filaments in vitro having non-amyloid properties. *Amyloid Int J Exp Clin Investig Off J Int Soc Amyloidosis.* 2021;28:56–65.
155. Robinson JL, Geser F, Stieber A, Umoh ME, Kwong L, Deerlin V, et al. TDP-43 skeins show properties of amyloid in a subset of ALS cases. *Acta Neuropathol (Berl).* 2012. <https://doi.org/10.1007/s00401-012-1055-8>.
156. Chien H-M, Lee C-C, Huang JJ-T. The Different Faces of the TDP-43 Low-Complexity Domain: The Formation of Liquid Droplets and Amyloid Fibrils. *Int J Mol Sci.* 2021;22:8213.
157. Hampel H, Hardy J, Blennow K, Chen C, Perry G, Kim SH, et al. The Amyloid- $\beta$  Pathway in Alzheimer's Disease. *Mol Psychiatry.* 2021;26:5481–503.
158. Kane MD, Lipinski WJ, Callahan MJ, Bian F, Durham RA, Schwarz RD, et al. Evidence for Seeding of  $\beta$ -Amyloid by Intracerebral Infusion of Alzheimer Brain Extracts in  $\beta$ -Amyloid Precursor Protein-Transgenic Mice. *J Neurosci.* 2000;20:3606–11.
159. Li J-Y, Englund E, Holton JL, Soulet D, Hagell P, Lees AJ, et al. Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation. *Nat Med.* 2008;14:501–3.
160. Desplats P, Lee H-J, Bae E-J, Patrick C, Rockenstein E, Crews L, et al. Inclusion formation and neuronal cell death through neuron-to-neuron transmission of  $\alpha$ -synuclein. *Proc Natl Acad Sci U S A.* 2009;106:13010–5.

161. Vaquer-Alicea J, Diamond MI. Propagation of Protein Aggregation in Neurodegenerative Diseases. *Annu Rev Biochem.* 2019;88:785–810.
162. Nonaka T, Masuda-Suzukake M, Arai T, Hasegawa Y, Akatsu H, Obi T, et al. Prion-like Properties of Pathological TDP-43 Aggregates from Diseased Brains. *Cell Rep.* 2013;4:124–34.
163. Shimonaka S, Nonaka T, Suzuki G, Hisanaga S-I, Hasegawa M. Templated Aggregation of TAR DNA-binding Protein of 43 kDa (TDP-43) by Seeding with TDP-43 Peptide Fibrils. *J Biol Chem.* 2016;291:8896–907.
164. Porta S, Xu Y, Restrepo CR, Kwong LK, Zhang B, Brown HJ, et al. Patient-derived frontotemporal lobar degeneration brain extracts induce formation and spreading of TDP-43 pathology in vivo. *Nat Commun.* 2018;9:4220.
165. Arseni D, Hasegawa M, Murzin AG, Kametani F, Arai M, Yoshida M, et al. Structure of pathological TDP-43 filaments from ALS with FTL. *Nature.* 2022;601:139–43.
166. Lee EB, Lee VM-Y, Trojanowski JQ. Gains or losses: molecular mechanisms of TDP43-mediated neurodegeneration. *Nat Rev Neurosci.* 13:38.
167. Sclérose latérale amyotrophique (SLA) / maladie de Charcot · Inserm, La science pour la santé. Inserm. <https://www.inserm.fr/dossier/sclerose-laterale-amyotrophique-sla-maladie-charcot/>. Accessed 12 Nov 2021.
168. Maladies neurodégénératives. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-neurodegeneratives>. Accessed 11 Nov 2021.
169. Swinnen B, Robberecht W. The phenotypic variability of amyotrophic lateral sclerosis. *Nat Rev Neurol.* 2014;10:661–70.
170. Majounie E, Renton AE, Mok K, Dopper EGP, Waite A, Rollinson S, et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* 2012;11:323–30.
171. Smeyers J, Banchi E-G, Latouche M. C9ORF72: What It Is, What It Does, and Why It Matters. *Front Cell Neurosci.* 2021;15:661447.
172. Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature.* 1993;362:59–62.
173. Renton AE, Chiò A, Traynor BJ. State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci.* 2014;17:17–23.

174. Taylor JP, Brown RH, Cleveland DW. Decoding ALS: from genes to mechanism. *Nature*. 2016;539:197–206.
175. Chiot A, Lobsiger CS, Boillée S. New insights on the disease contribution of neuroinflammation in amyotrophic lateral sclerosis. *Curr Opin Neurol*. 2019;32:764–770.
176. Fourier A. Vers un marqueur biochimique des dégénérescences lobaires fronto-temporales: variations quantitatives et profils protéiques de la protéine TDP43 dans différentes matrices biologiques. Theses. Université de Lyon; 2018.
177. Onyike CU, Diehl-Schmid J. The epidemiology of frontotemporal dementia. *Int Rev Psychiatry Abingdon Engl*. 2013;25:130–7.
178. Johnson JK, Diehl J, Mendez MF, Neuhaus J, Shapira JS, Forman M, et al. Frontotemporal lobar degeneration: demographic characteristics of 353 patients. *Arch Neurol*. 2005;62:925–30.
179. Young JJ, Lavakumar M, Tampi D, Balachandran S, Tampi RR. Frontotemporal dementia: latest evidence and clinical implications. *Ther Adv Psychopharmacol*. 2018;8:33–48.
180. Nelson PT, Dickson DW, Trojanowski JQ, Jack CR, Boyle PA, Arfanakis K, et al. Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. *Brain*. 2019;142:1503–27.
181. van Es MA, Hardiman O, Chio A, Al-Chalabi A, Pasterkamp RJ, Veldink JH, et al. Amyotrophic lateral sclerosis. *Lancet Lond Engl*. 2017;390:2084–98.
182. Picchiarelli G, Dupuis L. Role of RNA Binding Proteins with prion-like domains in muscle and neuromuscular diseases. *Cell Stress*. 4:76–91.
183. Neumann M, Rademakers R, Roeber S, Baker M, Kretzschmar HA, Mackenzie IRA. A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain*. 2009;132:2922–31.
184. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. 2006;314:130–3.
185. Ferrari R, Kapogiannis D, Huey ED, Momeni P. FTD and ALS: a tale of two diseases. *Curr Alzheimer Res*. 2011;8:273–94.
186. Seelaar H, Klijnsma KY, de Koning I, van der Lugt A, Chiu WZ, Azmani A, et al. Frequency of ubiquitin and FUS-positive, TDP-43-negative frontotemporal lobar degeneration. *J Neurol*. 2010;257:747–53.

187. Lai S-L, Abramzon Y, Schymick JC, Stephan DA, Dunckley T, Dillman A, et al. FUS mutations in sporadic amyotrophic lateral sclerosis. *Neurobiol Aging*. 2011;32:550.e1-4.
188. Hanson KA, Kim SH, Tibbetts RS. RNA-binding proteins in neurodegenerative disease: TDP-43 and beyond. *WIREs RNA*. 2012;3:265–85.
189. Ou SH, Wu F, Harrich D, García-Martínez LF, Gaynor RB. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol*. 1995;69:3584–96.
190. Buratti E, Dörk T, Zuccato E, Pagani F, Romano M, Baralle FE. Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J*. 2001;20:1774–84.
191. Sephton CF, Cenik B, Cenik BK, Herz J, Yu G. TDP-43 in central nervous system development and function: clues to TDP-43-associated neurodegeneration. *Biol Chem*. 2012;393:589–94.
192. Philips T, Rothstein JD. Rodent Models of Amyotrophic Lateral Sclerosis. *Curr Protoc Pharmacol*. 2015;69:5.67.1-5.67.21.
193. Banks GT, Kuta A, Isaacs AM, Fisher EMC. TDP-43 is a culprit in human neurodegeneration, and not just an innocent bystander. *Mamm Genome*. 2008;19:299–305.
194. Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, et al. TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science*. 2008;319:1668–72.
195. Lim L, Wei Y, Lu Y, Song J. ALS-Causing Mutations Significantly Perturb the Self-Assembly and Interaction with Nucleic Acid of the Intrinsically Disordered Prion-Like Domain of TDP-43. *PLoS Biol*. 2016;14:e1002338.
196. Franzmann TM, Alberti S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J Biol Chem*. 2019;294:7128–36.
197. Fonda BD, Jami KM, Boulos NR, Murray DT. Identification of the Rigid Core for Aged Liquid Droplets of an RNA-Binding Protein Low Complexity Domain. *J Am Chem Soc*. 2021;143:6657–68.
198. Zhang Y-J, Caulfield T, Xu Y-F, Gendron TF, Hubbard J, Stetler C, et al. The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum Mol Genet*. 2013;22:3112–22.



199. Mompeán M, Romano V, Pantoja-Uceda D, Stuani C, Baralle FE, Buratti E, et al. The TDP-43 N-terminal domain structure at high resolution. *FEBS J.* 2016;283:1242–60.
200. Mompeán M, Romano V, Pantoja-Uceda D, Stuani C, Baralle FE, Buratti E, et al. Point mutations in the N-terminal domain of transactive response DNA-binding protein 43 kDa (TDP-43) compromise its stability, dimerization, and functions. *J Biol Chem.* 2017;292:11992–2006.
201. Nishimura AL, Župunski V, Troakes C, Kathe C, Fratta P, Howell M, et al. Nuclear import impairment causes cytoplasmic trans-activation response DNA-binding protein accumulation and is associated with frontotemporal lobar degeneration. *Brain.* 2010;133:1763–71.
202. Lukavsky PJ, Daujotyte D, Tollervey JR, Ule J, Stuani C, Buratti E, et al. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat Struct Mol Biol.* 2013;20:1443–9.
203. Kuo P-H, Chiang C-H, Wang Y-T, Doudeva LG, Yuan HS. The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic Acids Res.* 2014;42:4712–22.
204. Agrawal S, Kuo P-H, Chu L-Y, Golzarroshan B, Jain M, Yuan HS. RNA recognition motifs of disease-linked RNA-binding proteins contribute to amyloid formation. *Sci Rep.* 2019;9:6171.
205. Pham J, Keon M, Brennan S, Saksena N. Connecting RNA-Modifying Similarities of TDP-43, FUS, and SOD1 with MicroRNA Dysregulation Amidst A Renewed Network Perspective of Amyotrophic Lateral Sclerosis Proteinopathy. *Int J Mol Sci.* 2020;21:3464.
206. Buratti E, Baralle FE. Characterization and Functional Implications of the RNA Binding Properties of Nuclear Factor TDP-43, a Novel Splicing Regulator of CFTR Exon 9 \*. *J Biol Chem.* 2001;276:36337–43.
207. Kuo P-H, Doudeva LG, Wang Y-T, Shen C-KJ, Yuan HS. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. *Nucleic Acids Res.* 2009;37:1799–808.
208. Kwiatkowski TJ, Bosco DA, Leclerc AL, Tamrazian E, Vanderburg CR, Russ C, et al. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science.* 2009;323:1205–8.
209. Luo F, Gui X, Zhou H, Gu J, Li Y, Liu X, et al. Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. *Nat Struct Mol Biol.* 2018;25:341–6.

210. Murakami T, Qamar S, Lin JQ, Schierle GSK, Rees E, Miyashita A, et al. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron*. 2015;88:678–90.
211. Loughlin FE, Lukavsky PJ, Kazeeva T, Reber S, Hock E-M, Colombo M, et al. The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity. *Mol Cell*. 2019;73:490-504.e6.
212. Sama RRK, Ward CL, Bosco DA. Functions of FUS/TLS From DNA Repair to Stress Response: Implications for ALS. *ASN NEURO*. 2014;6:1759091414544472.
213. Liu X, Niu C, Ren J, Zhang J, Xie X, Zhu H, et al. The RRM domain of human fused in sarcoma protein reveals a non-canonical nucleic acid binding site. *Biochim Biophys Acta*. 2013;1832:375–85.
214. Lu Y, Lim L, Song J. RRM domain of ALS/FTD-causing FUS characteristic of irreversible unfolding spontaneously self-assembles into amyloid fibrils. *Sci Rep*. 2017;7:1043.
215. Iko Y, Kodama TS, Kasai N, Oyama T, Morita EH, Muto T, et al. Domain architectures and characterization of an RNA-binding protein, TLS. *J Biol Chem*. 2004;279:44834–40.
216. Bentmann E, Neumann M, Tahirovic S, Rodde R, Dormann D, Haass C. Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43). *J Biol Chem*. 2012;287:23079–94.
217. Ishiguro A, Katayama A, Ishihama A. Different recognition modes of G-quadruplex RNA between two ALS/FTLD-linked proteins TDP-43 and FUS. *FEBS Lett*. 2021;595:310–23.
218. Takahama K, Miyawaki A, Shitara T, Mitsuya K, Morikawa M, Hagihara M, et al. G-Quadruplex DNA- and RNA-Specific-Binding Proteins Engineered from the RGG Domain of TLS/FUS. *ACS Chem Biol*. 2015;10:2564–9.
219. Nguyen CD, Mansfield RE, Leung W, Vaz PM, Loughlin FE, Grant RP, et al. Characterization of a Family of RanBP2-Type Zinc Fingers that Can Recognize Single-Stranded RNA. *J Mol Biol*. 2011;407:273–83.
220. Baade I, Hutten S, Sternburg EL, Pörschke M, Hofweber M, Dormann D, et al. The RNA-binding protein FUS is chaperoned and imported into the nucleus by a network of import receptors. *J Biol Chem*. 2021;296:100659.
221. Rogelj B, Easton LE, Bogu GK, Stanton LW, Rot G, Curk T, et al. Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Sci Rep*. 2012;2:603.

222. Ozdilek BA, Thompson VF, Ahmed NS, White CI, Batey RT, Schwartz JC. Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding. *Nucleic Acids Res.* 2017;45:7984–96.
223. Schwartz JC, Cech TR, Parker RR. Biochemical Properties and Biological Functions of FET Proteins. *Annu Rev Biochem.* 2015;84:355–79.
224. Imperatore JA, McAninch DS, Valdez-Sinon AN, Bassell GJ, Mihailescu MR. FUS Recognizes G Quadruplex Structures Within Neuronal mRNAs. *Front Mol Biosci.* 2020;7:6.
225. Vasilyev N, Polonskaia A, Darnell JC, Darnell RB, Patel DJ, Serganov A. Crystal structure reveals specific recognition of a G-quadruplex RNA by a  $\beta$ -turn in the RGG motif of FMRP. *Proc Natl Acad Sci U S A.* 2015;112:E5391–5400.
226. Ghisolfi L, Joseph G, Amalric F, Erard M. The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. *J Biol Chem.* 1992;267:2955–9.
227. Hoell JI, Larsson E, Runge S, Nusbaum JD, Duggimpudi S, Farazi TA, et al. RNA targets of wild-type and mutant FET family proteins. *Nat Struct Mol Biol.* 2011;18:1428–31.
228. Yang L, Gal J, Chen J, Zhu H. Self-assembled FUS binds active chromatin and regulates gene transcription. *Proc Natl Acad Sci U S A.* 2014;111:17809–14.
229. Schwartz JC, Ebmeier CC, Podell ER, Heimiller J, Taatjes DJ, Cech TR. FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. *Genes Dev.* 2012;26:2690–5.
230. Masuda A, Takeda J, Okuno T, Okamoto T, Ohkawara B, Ito M, et al. Position-specific binding of FUS to nascent RNA regulates mRNA length. *Genes Dev.* 2015;29:1045–57.
231. Casafont I, Bengoechea R, Tapia O, Berciano MT, Lafarga M. TDP-43 localizes in mRNA transcription and processing sites in mammalian neurons. *J Struct Biol.* 2009;167:235–41.
232. Abhyankar MM, Urekar C, Reddi PP. A novel CpG-free vertebrate insulator silences the testis-specific SP-10 gene in somatic tissues: role for TDP-43 in insulator function. *J Biol Chem.* 2007;282:36143–54.
233. Murata H, Hattori T, Maeda H, Takashiba S, Takigawa M, Kido J, et al. Identification of transactivation-responsive DNA-binding protein 43 (TARDBP43; TDP-43) as a novel factor for TNF- $\alpha$  expression upon lipopolysaccharide stimulation in human monocytes. *J Periodontol Res.* 2015;50:452–60.

234. Schwenk BM, Hartmann H, Serdaroğlu A, Schludi MH, Hornburg D, Meissner F, et al. TDP-43 loss of function inhibits endosomal trafficking and alters trophic signaling in neurons. *EMBO J*. 2016. <https://doi.org/10.15252/embj.201694221>.
235. Lalmansingh AS, Urekar C, Reddi P. TDP-43 Is a Transcriptional Repressor. *J Biol Chem*. 2011. <https://doi.org/10.1074/jbc.M110.166587>.
236. Zinszner H, Sok J, Immanuel D, Yin Y, Ron D. TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *J Cell Sci*. 1997;110 ( Pt 15):1741–50.
237. Jungmichel S, Rosenthal F, Altmeyer M, Lukas J, Hottiger MO, Nielsen ML. Proteome-wide identification of poly(ADP-Ribosyl)ation targets in different genotoxic stress responses. *Mol Cell*. 2013;52:272–85.
238. Mastrocola AS, Kim SH, Trinh AT, Rodenkirch LA, Tibbetts RS. The RNA-binding Protein Fused in Sarcoma (FUS) Functions Downstream of Poly(ADP-ribose) Polymerase (PARP) in Response to DNA Damage. *J Biol Chem*. 2013;288:24731–41.
239. Singatulina AS, Hamon L, Sukhanova MV, Desforges B, Joshi V, Bouhss A, et al. PARP-1 Activation Directs FUS to DNA Damage Sites to Form PARG-Reversible Compartments Enriched in Damaged DNA. *Cell Rep*. 2019;27:1809-1821.e5.
240. Baechtold H, Kuroda M, Sok J, Ron D, Lopez BS, Akhmedov AT. Human 75-kDa DNA-pairing Protein Is Identical to the Pro-oncoprotein TLS/FUS and Is Able to Promote D-loop Formation \*. *J Biol Chem*. 1999;274:34337–42.
241. Naumann M, Pal A, Goswami A, Lojewski X, Japtok J, Vehlow A, et al. Impaired DNA damage response signaling by FUS-NLS mutations leads to neurodegeneration and FUS aggregate formation. *Nat Commun*. 2018;9:335.
242. Kawaguchi T, Rollins MG, Moinpour M, Morera AA, Ebmeier CC, Old WM, et al. Changes to the TDP-43 and FUS Interactomes Induced by DNA Damage. *J Proteome Res*. 2020;19:360–70.
243. Mitra J, Guerrero EN, Hegde PM, Liachko NF, Wang H, Vasquez V, et al. Motor neuron disease-associated loss of nuclear TDP-43 is linked to DNA double-strand break repair defects. *Proc Natl Acad Sci U S A*. 2019;116:4696–705.
244. Freibaum BD, Chitta R, High AA, Taylor JP. Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. *J Proteome Res*. 2010;9:1104–20.
245. Kamelgarn M, Chen J, Kuang L, Jin H, Kasarskis EJ, Zhu H. ALS mutations of FUS suppress protein translation and disrupt the regulation of nonsense-mediated decay. *Proc Natl Acad Sci*. 2018;115:E11904–13.

246. Russo A, Scardigli R, La Regina F, Murray ME, Romano N, Dickson DW, et al. Increased cytoplasmic TDP-43 reduces global protein synthesis by interacting with RACK1 on polyribosomes. *Hum Mol Genet.* 2017;26:1407–18.
247. Birsa N, Bentham MP, Fratta P. Cytoplasmic functions of TDP-43 and FUS and their role in ALS. *Semin Cell Dev Biol.* 2020;99:193–201.
248. Neelagandan N, Gonnella G, Dang S, Janiesch PC, Miller KK, Kuchler K, et al. TDP-43 enhances translation of specific mRNAs linked to neurodegenerative disease. *Nucleic Acids Res.* 2019;47:341–61.
249. Yasuda K, Zhang H, Loisel D, Haystead T, Macara IG, Mili S. The RNA-binding protein Fus directs translation of localized mRNAs in APC-RNP granules. *J Cell Biol.* 2013;203:737–46.
250. Lagier-Tourenne C, Polymenidou M, Cleveland DW. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum Mol Genet.* 2010;19:R46–64.
251. Buratti E. Trends in Understanding the Pathological Roles of TDP-43 and FUS Proteins. *Adv Exp Med Biol.* 2021;1281:243–67.
252. Polymenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, et al. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci.* 2011;14:459–68.
253. Xiao S, Sanelli T, Dib S, Sheps D, Findlater J, Bilbao J, et al. RNA targets of TDP-43 identified by UV-CLIP are deregulated in ALS. *Mol Cell Neurosci.* 2011;47:167–80.
254. Ratti A, Buratti E. Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. *J Neurochem.* 2016;138 Suppl 1:95–111.
255. Nakaya T, Alexiou P, Maragkakis M, Chang A, Mourelatos Z. FUS regulates genes coding for RNA-binding proteins in neurons by binding to their highly conserved introns. *RNA N Y N.* 2013;19:498–509.
256. Lerga A, Hallier M, Delva L, Orvain C, Gallais I, Marie J, et al. Identification of an RNA binding specificity for the potential splicing factor TLS. *J Biol Chem.* 2001;276:6807–16.
257. Ishigaki S, Masuda A, Fujioka Y, Iguchi Y, Katsuno M, Shibata A, et al. Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. *Sci Rep.* 2012;2:529.

258. Orozco D, Tahirovic S, Rentzsch K, Schwenk BM, Haass C, Edbauer D. Loss of fused in sarcoma (FUS) promotes pathological Tau splicing. *EMBO Rep.* 2012;13:759–64.
259. Prashad S, Gopal PP. RNA-binding proteins in neurological development and disease. *RNA Biol.* 2021;18:972–87.
260. Masuda A, Takeda J-I, Ohno K. FUS-mediated regulation of alternative RNA processing in neurons: insights from global transcriptome analysis. *Wiley Interdiscip Rev RNA.* 2016;7:330–40.
261. Tank EM, Figueroa-Romero C, Hinder LM, Bedi K, Archbold HC, Li X, et al. Abnormal RNA stability in amyotrophic lateral sclerosis. *Nat Commun.* 2018;9:2845.
262. Kapeli K, Pratt GA, Vu AQ, Hutt KR, Martinez FJ, Sundararaman B, et al. Distinct and shared functions of ALS-associated proteins TDP-43, FUS and TAF15 revealed by multisystem analyses. *Nat Commun.* 2016;7:12143.
263. Colombrita C, Onesto E, Megiorni F, Pizzuti A, Baralle FE, Buratti E, et al. TDP-43 and FUS RNA-binding proteins bind distinct sets of cytoplasmic messenger RNAs and differently regulate their post-transcriptional fate in motoneuron-like cells. *J Biol Chem.* 2012;287:15635–47.
264. Udagawa T, Fujioka Y, Tanaka M, Honda D, Yokoi S, Riku Y, et al. FUS regulates AMPA receptor function and FTLD/ALS-associated behaviour via GluA1 mRNA stabilization. *Nat Commun.* 2015;6:7098.
265. Alami NH, Smith RB, Carrasco MA, Williams LA, Winborn CS, Han SSW, et al. Axonal Transport of TDP-43 mRNA Granules Is Impaired by ALS-Causing Mutations. *Neuron.* 2014;81:536–43.
266. Belly A, Moreau-Gachelin F, Sadoul R, Goldberg Y. Delocalization of the multifunctional RNA splicing factor TLS/FUS in hippocampal neurones: exclusion from the nucleus and accumulation in dendritic granules and spine heads. *Neurosci Lett.* 2005;379:152–7.
267. Fujii R, Okabe S, Urushido T, Inoue K, Yoshimura A, Tachibana T, et al. The RNA Binding Protein TLS Is Translocated to Dendritic Spines by mGluR5 Activation and Regulates Spine Morphology. *Curr Biol.* 2005;15:587–93.
268. Kanai Y, Dohmae N, Hirokawa N. Kinesin Transports RNA: Isolation and Characterization of an RNA-Transporting Granule. *Neuron.* 2004;43:513–25.
269. Muresan V, Ladescu Muresan Z. Shared Molecular Mechanisms in Alzheimer's Disease and Amyotrophic Lateral Sclerosis: Neurofilament-Dependent Transport of

sAPP, FUS, TDP-43 and SOD1, with Endoplasmic Reticulum-Like Tubules. *Neurodegener Dis.* 2016;16:55–61.

270. Hill M, Tran N. miRNA interplay: mechanisms and consequences in cancer. *Dis Model Mech.* 2021;14:dmm047662.

271. Gagliardi S, Pandini C, Garofalo M, Bordoni M, Pansarasa O, Cereda C. Long non coding RNAs and ALS: Still much to do. *Non-Coding RNA Res.* 2018;3:226–31.

272. Lourenco GF, Janitz M, Huang Y, Halliday GM. Long noncoding RNAs in TDP-43 and FUS/TLS-related frontotemporal lobar degeneration (FTLD). *Neurobiol Dis.* 2015;82:445–54.

273. Modic M, Grosch M, Rot G, Schirge S, Lepko T, Yamazaki T, et al. Cross-Regulation between TDP-43 and Paraspeckles Promotes Pluripotency-Differentiation Transition. *Mol Cell.* 2019;74:951-965.e13.

274. Kawahara Y, Mieda-Sato A. TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc Natl Acad Sci.* 2012;109:3347–52.

275. Zhang T, Wu Y-C, Mullane P, Ji YJ, Liu H, He L, et al. FUS Regulates Activity of MicroRNA-mediated Gene Silencing. *Mol Cell.* 2018;69:787-801.e8.

276. Morlando M, Dini Modigliani S, Torrelli G, Rosa A, Di Carlo V, Caffarelli E, et al. FUS stimulates microRNA biogenesis by facilitating co-transcriptional Drosha recruitment. *EMBO J.* 2012;31:4502–10.

277. Hawley ZCE, Campos-Melo D, Strong MJ. Novel miR-b2122 regulates several ALS-related RNA-binding proteins. *Mol Brain.* 2017;10:46.

278. Cirillo D, Agostini F, Klus P, Marchese D, Rodriguez S, Bolognesi B, et al. Neurodegenerative diseases: Quantitative predictions of protein–RNA interactions. *RNA.* 2013;19:129–40.

279. Deng J, Wang P, Chen X, Cheng H, Liu J, Fushimi K, et al. FUS interacts with ATP synthase beta subunit and induces mitochondrial unfolded protein response in cellular and animal models. *Proc Natl Acad Sci.* 2018;115:E9678–86.

280. Soo KY, Sultana J, King AE, Atkinson R, Warraich ST, Sundaramoorthy V, et al. ALS-associated mutant FUS inhibits macroautophagy which is restored by overexpression of Rab1. *Cell Death Discov.* 2015;1:15030.

281. Ling S-C, Dastidar SG, Tokunaga S, Ho WY, Lim K, Ilieva H, et al. Overriding FUS autoregulation in mice triggers gain-of-toxic dysfunctions in RNA metabolism and autophagy-lysosome axis. *eLife.* 2019;8:e40811.

282. Xia Q, Wang H, Hao Z, Fu C, Hu Q, Gao F, et al. TDP-43 loss of function increases TFE3 activity and blocks autophagosome-lysosome fusion. *EMBO J.* 2016;35:121–42.
283. Dormann D, Madl T, Valori CF, Bentmann E, Tahirovic S, Abou-Ajram C, et al. Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. *EMBO J.* 2012;31:4258–75.
284. Qamar S, Wang G, Randle SJ, Ruggeri FS, Varela JA, Lin JQ, et al. FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation- $\pi$  Interactions. *Cell.* 2018;173:720-734.e15.
285. Hofweber M, Hutten S, Bourgeois B, Spreitzer E, Niedner-Boblenz A, Schifferer M, et al. Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell.* 2018;173:706-719.e13.
286. Sternburg EL, Gruijs da Silva LA, Dormann D. Post-translational modifications on RNA-binding proteins: accelerators, brakes, or passengers in neurodegeneration? *Trends Biochem Sci.* 2021;;S0968-0004(21)00160-2.
287. Bock AS, Murthy AC, Tang WS, Jovic N, Shewmaker F, Mittal J, et al. N-terminal acetylation modestly enhances phase separation and reduces aggregation of the low-complexity domain of RNA-binding protein fused in sarcoma. *Protein Sci.* 2021;30:1337–49.
288. Arenas A, Chen J, Kuang L, Barnett KR, Kasarskis EJ, Gal J, et al. Lysine acetylation regulates the RNA binding, subcellular localization and inclusion formation of FUS. *Hum Mol Genet.* 2020;29:2684–97.
289. Cohen TJ, Hwang AW, Restrepo CR, Yuan C-X, Trojanowski JQ, Lee VMY. An acetylation switch controls TDP-43 function and aggregation propensity. *Nat Commun.* 2015;6:5845.
290. Morato JG, Hans F, Zweyendorf F von, Feederle R, Elsässer SJ, Skodras AA, et al. Sirtuin-1 Sensitive Lysine-136 Acetylation Drives Phase Separation and Pathological Aggregation of TDP-43. *bioRxiv*; 2021.
291. Deng Q, Holler CJ, Taylor G, Hudson KF, Watkins W, Gearing M, et al. FUS is phosphorylated by DNA-PK and accumulates in the cytoplasm after DNA damage. *J Neurosci Off J Soc Neurosci.* 2014;34:7802–13.
292. Rhoads SN, Monahan ZT, Yee DS, Leung AY, Newcombe CG, O’Meally RN, et al. The prionlike domain of FUS is multiphosphorylated following DNA damage without altering nuclear localization. *Mol Biol Cell.* 2018;29:1786–97.



293. Darovic S, Prpar Mihevc S, Župunski V, Gunčar G, Štalekar M, Lee Y-B, et al. Phosphorylation of C-terminal tyrosine residue 526 in FUS impairs its nuclear import. *J Cell Sci.* 2015;128:4151–9.
294. Owen I, Rhoads S, Yee D, Wyne H, Gery K, Hannula I, et al. The prion-like domain of Fused in Sarcoma is phosphorylated by multiple kinases affecting liquid- and solid-phase transitions. *Mol Biol Cell.* 2020;31:2522–36.
295. Wang A, Conicella AE, Schmidt HB, Martin EW, Rhoads SN, Reeb AN, et al. A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J.* 2018;37:e97452.
296. Neumann M, Kwong LK, Lee EB, Kremmer E, Flatley A, Xu Y, et al. Phosphorylation of S409/410 of TDP-43 is a consistent feature in all sporadic and familial forms of TDP-43 proteinopathies. *Acta Neuropathol (Berl).* 2009;117:137–49.
297. Kim SH, Shi Y, Hanson KA, Williams LM, Sakasai R, Bowler MJ, et al. Potentiation of amyotrophic lateral sclerosis (ALS)-associated TDP-43 aggregation by the proteasome-targeting factor, ubiquilin 1. *J Biol Chem.* 2009;284:8083–92.
298. Jantrapirom S, Lo Piccolo L, Yoshida H, Yamaguchi M. Depletion of Ubiquilin induces an augmentation in soluble ubiquitinated Drosophila TDP-43 to drive neurotoxicity in the fly. *Biochim Biophys Acta Mol Basis Dis.* 2018;1864 9 Pt B:3038–49.
299. Picher-Martel V, Renaud L, Bareil C, Julien J-P. Neuronal Expression of UBQLN2P497H Exacerbates TDP-43 Pathology in TDP-43G348C Mice through Interaction with Ubiquitin. *Mol Neurobiol.* 2019;56:4680–96.
300. Buratti E. TDP-43 post-translational modifications in health and disease. *Expert Opin Ther Targets.* 2018;22:279–93.
301. Akimov V, Barrio-Hernandez I, Hansen SVF, Hallenborg P, Pedersen A-K, Bekker-Jensen DB, et al. UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nat Struct Mol Biol.* 2018;25:631–40.
302. Yasuda S, Tsuchiya H, Kaiho A, Guo Q, Ikeuchi K, Endo A, et al. Stress- and ubiquitylation-dependent phase separation of the proteasome. *Nature.* 2020;578:296–300.
303. Farrawell NE, McAlary L, Lum JS, Chisholm CG, Warraich ST, Blair IP, et al. Ubiquitin Homeostasis Is Disrupted in TDP-43 and FUS Cell Models of ALS. *iScience.* 2020;23:101700.
304. Shodai A, Morimura T, Ido A, Uchida T, Ayaki T, Takahashi R, et al. Aberrant Assembly of RNA Recognition Motif 1 Links to Pathogenic Conversion of TAR DNA-

- binding Protein of 43 kDa (TDP-43)\*. *J Biol Chem.* 2013. <https://doi.org/10.1074/jbc.M113.451849>.
305. Bozzo F, Salvatori I, Iacovelli F, Mirra A, Rossi S, Cozzolino M, et al. Structural insights into the multi-determinant aggregation of TDP-43 in motor neuron-like cells. *Neurobiol Dis.* 2016;94:63–72.
306. Zhao M-J, Yao X, Wei P, Zhao C, Cheng M, Zhang D, et al. O-GlcNAcylation of TDP-43 suppresses proteinopathies and promotes TDP-43's mRNA splicing activity. *EMBO Rep.* 2021. <https://doi.org/10.15252/embr.202051649>.
307. Pirie E, Oh C, Zhang X, Han X, Cieplak P, Scott HR, et al. S-nitrosylated TDP-43 triggers aggregation, cell-to-cell spread, and neurotoxicity in hiPSCs and in vivo models of ALS/FTD. *Proc Natl Acad Sci.* 2021;118.
308. Grundke-Iqbal I, Iqbal K, Tung YC, Quinlan M, Wisniewski HM, Binder LI. Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc Natl Acad Sci U S A.* 1986;83:4913–7.
309. Despres C, Byrne C, Qi H, Cantrelle F-X, Huvent I, Chambraud B, et al. Identification of the Tau phosphorylation pattern that drives its aggregation. *Proc Natl Acad Sci.* 2017;114:9080–5.
310. Suárez-Calvet M, Neumann M, Arzberger T, Abou-Ajram C, Funk E, Hartmann H, et al. Monomethylated and unmethylated FUS exhibit increased binding to Transportin and distinguish FTLD-FUS from ALS-FUS. *Acta Neuropathol (Berl).* 2016;131:587–604.
311. Guerrero EN, Wang H, Mitra J, Hegde PM, Stowell SE, Liachko NF, et al. TDP-43/FUS in Motor Neuron Disease: Complexity and Challenges. *Prog Neurobiol.* 2016;145–146:78–97.
312. Klim JR, Pintacuda G, Nash LA, Guerra San Juan I, Eggan K. Connecting TDP-43 Pathology with Neuropathy. *Trends Neurosci.* 2021;44:424–40.
313. Buratti E. Chapter One - Functional Significance of TDP-43 Mutations in Disease. In: Friedmann T, Dunlap JC, Goodwin SF, editors. *Advances in Genetics.* Academic Press; 2015. p. 1–53.
314. White MA, Kim E, Duffy A, Adalbert R, Phillips BU, Peters OM, et al. TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in mouse model of ALS-FTD. *Nat Neurosci.* 2018;21:552–63.
315. Fratta P, Sivakumar P, Humphrey J, Lo K, Ricketts T, Oliveira H, et al. Mice with endogenous TDP-43 mutations exhibit gain of splicing function and characteristics of amyotrophic lateral sclerosis. *EMBO J.* 2018;37:e98684.

316. Jiang T, Handley E, Brizuela M, Dawkins E, Lewis KEA, Clark RM, et al. Amyotrophic lateral sclerosis mutant TDP-43 may cause synaptic dysfunction through altered dendritic spine function. *Dis Model Mech*. 2019;12:dmm038109.
317. Chen H-J, Topp SD, Hui HS, Zacco E, Katarya M, McLoughlin C, et al. RRM adjacent TARDBP mutations disrupt RNA binding and enhance TDP-43 proteinopathy. *Brain*. 2019;142:3753–70.
318. Chiang C-H, Grauffel C, Wu L-S, Kuo P-H, Doudeva LG, Lim C, et al. Structural analysis of disease-related TDP-43 D169G mutation: linking enhanced stability and caspase cleavage efficiency to protein accumulation. *Sci Rep*. 2016;6:21581.
319. Maurel C, Madji-Hounoum B, Thepault R-A, Marouillat S, Brulard C, Danel-Brunaud V, et al. Mutation in the RRM2 domain of TDP-43 in Amyotrophic Lateral Sclerosis with rapid progression associated with ubiquitin positive aggregates in cultured motor neurons. *Amyotroph Lateral Scler Front Degener*. 2018;19:149–51.
320. Winton MJ, Van Deerlin VM, Kwong LK, Yuan W, Wood EM, Yu C-E, et al. A90V TDP-43 variant results in the aberrant localization of TDP-43 in vitro. *FEBS Lett*. 2008;582:2252–6.
321. Shang Y, Huang EJ. Mechanisms of FUS Mutations in Familial Amyotrophic Lateral Sclerosis. *Brain Res*. 2016;1647:65–78.
322. Qiu H, Lee S, Shang Y, Wang W-Y, Au KF, Kamiya S, et al. ALS-associated mutation FUS-R521C causes DNA damage and RNA splicing defects. *J Clin Invest*. 2014;124:981–99.
323. Sun S, Ling S-C, Qiu J, Albuquerque CP, Zhou Y, Tokunaga S, et al. ALS-causative mutations in FUS/TLS confer gain- and loss-of-function by altered association with SMN and U1-snRNP. *Nat Commun*. 2015. <https://doi.org/10.1038/ncomms7171>.
324. Chhangani D, Martín-Peña A, Rincon-Limas DE. Molecular, functional, and pathological aspects of TDP-43 fragmentation. *iScience*. 2021;24:102459.
325. Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun*. 2006;351:602–11.
326. Igaz LM, Kwong LK, Xu Y, Truax AC, Uryu K, Neumann M, et al. Enrichment of C-terminal fragments in TAR DNA-binding protein-43 cytoplasmic inclusions in brain but not in spinal cord of frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Am J Pathol*. 2008;173:182–94.

327. Li Q, Yokoshi M, Okada H, Kawahara Y. The cleavage pattern of TDP-43 determines its rate of clearance and cytotoxicity. *Nat Commun.* 2015;6:6183.
328. Zhang Y-J, Xu Y-F, Cook C, Gendron TF, Roettges P, Link CD, et al. Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity. *Proc Natl Acad Sci U S A.* 2009;106:7607–12.
329. Caccamo A, Majumder S, Oddo S. Cognitive Decline Typical of Frontotemporal Lobar Degeneration in Transgenic Mice Expressing the 25-kDa C-Terminal Fragment of TDP-43. *Am J Pathol.* 2012;180:293–302.
330. Walker AK, Tripathy K, Restrepo CR, Ge G, Xu Y, Kwong LK, et al. An insoluble frontotemporal lobar degeneration-associated TDP-43 C-terminal fragment causes neurodegeneration and hippocampus pathology in transgenic mice. *Hum Mol Genet.* 2015;24:7241–54.
331. Wang Y-T, Kuo P-H, Chiang C-H, Liang J-R, Chen Y-R, Wang S, et al. The truncated C-terminal RNA recognition motif of TDP-43 protein plays a key role in forming proteinaceous aggregates. *J Biol Chem.* 2013;288:9049–57.
332. Che M-X, Jiang L-L, Li H-Y, Jiang Y-J, Hu H-Y. TDP-35 sequesters TDP-43 into cytoplasmic inclusions through binding with RNA. *FEBS Lett.* 2015;589:1920–8.
333. Nishimoto Y, Ito D, Yagi T, Nihei Y, Tsunoda Y, Suzuki N. Characterization of alternative isoforms and inclusion body of the TAR DNA-binding protein-43. *J Biol Chem.* 2010;285:608–19.
334. Shenoy J, El Mammeri N, Dutour A, Berbon M, Saad A, Lends A, et al. Structural dissection of amyloid aggregates of TDP-43 and its C-terminal fragments TDP-35 and TDP-16. *FEBS J.* 2020;287:2449–67.
335. Waibel S, Neumann M, Rosenbohm A, Birve A, Volk AE, Weishaupt JH, et al. Truncating mutations in FUS/TLS give rise to a more aggressive ALS-phenotype than missense mutations: a clinico-genetic study in Germany. *Eur J Neurol.* 2013;20:540–6.
336. Bosco DA, Lemay N, Ko HK, Zhou H, Burke C, Kwiatkowski TJ Jr, et al. Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. *Hum Mol Genet.* 2010;19:4160–75.
337. DeJesus-Hernandez M, Kocerha J, Finch N, Crook R, Baker M, Desaro P, et al. De Novo Truncating FUS Gene Mutation as a Cause of Sporadic Amyotrophic Lateral Sclerosis. *Hum Mutat.* 2010;31:E1377–89.
338. Huang C, Xia PY, Zhou H. Sustained Expression of TDP-43 and FUS in Motor Neurons in Rodent's Lifetime. *Int J Biol Sci.* 2010;6:396–406.

339. Koza P, Beroun A, Konopka A, Górkiewicz T, Bijoch L, Torres JC, et al. Neuronal TDP-43 depletion affects activity-dependent plasticity. *Neurobiol Dis.* 2019;130:104499.
340. Sephton CF, Tang AA, Kulkarni A, West J, Brooks M, Stubblefield JJ, et al. Activity-dependent FUS dysregulation disrupts synaptic homeostasis. *Proc Natl Acad Sci U S A.* 2014;111:E4769-4778.
341. Ayala YM, Zago P, D'Ambrogio A, Xu Y-F, Petrucelli L, Buratti E, et al. Structural determinants of the cellular localization and shuttling of TDP-43. *J Cell Sci.* 2008;121 Pt 22:3778–85.
342. Wang I-F, Wu L-S, Chang H-Y, Shen C-KJ. TDP-43, the signature protein of FTLD-U, is a neuronal activity-responsive factor. *J Neurochem.* 2008;105:797–806.
343. Gregory RI, Yan K-P, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, et al. The Microprocessor complex mediates the genesis of microRNAs. *Nature.* 2004;432:235–40.
344. Andersson MK, Ståhlberg A, Arvidsson Y, Olofsson A, Semb H, Stenman G, et al. The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. *BMC Cell Biol.* 2008;9:37.
345. Wang D, Eraslan B, Wieland T, Hallström B, Hopf T, Zolg DP, et al. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Mol Syst Biol.* 2019;15:e8503.
346. Tomkins JE, Ferrari R, Vavouraki N, Hardy J, Lovering RC, Lewis PA, et al. PINOT: an intuitive resource for integrating protein-protein interactions. *Cell Commun Signal.* 2020;18:92.
347. Hu B, Arpag S, Zuchner S, Li J. A novel missense mutation of CMT2P alters transcription machinery. *Ann Neurol.* 2016;80:834–45.
348. Haenig C, Atias N, Taylor AK, Mazza A, Schaefer MH, Russ J, et al. Interactome Mapping Provides a Network of Neurodegenerative Disease Proteins and Uncovers Widespread Protein Aggregation in Affected Brains. *Cell Rep.* 2020;32:108050.
349. Cacciottolo R, Ciantar J, Lanfranco M, Borg RM, Vassallo N, Bordonné R, et al. SMN complex member Gemin3 self-interacts and has a functional relationship with ALS-linked proteins TDP-43, FUS and Sod1. *Sci Rep.* 2019;9:18666.
350. Ishiguro T, Sato N, Ueyama M, Fujikake N, Sellier C, Kanegami A, et al. Regulatory Role of RNA Chaperone TDP-43 for RNA Misfolding and Repeat-Associated Translation in SCA31. *Neuron.* 2017;94:108-124.e7.

351. Honda D, Ishigaki S, Iguchi Y, Fujioka Y, Udagawa T, Masuda A, et al. The ALS/FTLD-related RNA-binding proteins TDP-43 and FUS have common downstream RNA targets in cortical neurons. *FEBS Open Bio*. 2013;4:1–10.
352. Hergesheimer RC, Chami AA, de Assis DR, Vourc'h P, Andres CR, Corcia P, et al. The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain*. 2019;142:1176–94.
353. Vance C, Rogelj B, Hortobágyi T, De Vos KJ, Nishimura AL, Sreedharan J, et al. Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science*. 2009;323:1208–11.
354. Munoz DG, Neumann M, Kusaka H, Yokota O, Ishihara K, Terada S, et al. FUS pathology in basophilic inclusion body disease. *Acta Neuropathol (Berl)*. 2009;118:617–27.
355. Armstrong R. Neuronal cytoplasmic inclusions in tau, TDP-43, and FUS molecular subtypes of frontotemporal lobar degeneration share similar spatial patterns. *Folia Neuropathol*. 2017;55:185–92.
356. Mackenzie IRA, Neumann M, Bigio EH, Cairns NJ, Alafuzoff I, Kril J, et al. Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: an update. *Acta Neuropathol (Berl)*. 2010;119:1–4.
357. Keller BA, Volkening K, Droppelmann CA, Ang LC, Rademakers R, Strong MJ. Co-aggregation of RNA binding proteins in ALS spinal motor neurons: evidence of a common pathogenic mechanism. *Acta Neuropathol (Berl)*. 2012;124:733–47.
358. Chou C-C, Zhang Y, Umoh ME, Vaughan SW, Lorenzini I, Liu F, et al. TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. *Nat Neurosci*. 2018;21:228–39.
359. Dhakal S, Wyant CE, George HE, Morgan SE, Rangachari V. Prion-like C-Terminal Domain of TDP-43 and  $\alpha$ -Synuclein Interact Synergistically to Generate Neurotoxic Hybrid Fibrils. *J Mol Biol*. 2021;433:166953.
360. Takanashi K, Yamaguchi A. Aggregation of ALS-linked FUS mutant sequesters RNA binding proteins and impairs RNA granules formation. *Biochem Biophys Res Commun*. 2014;452:600–7.
361. Baron DM, Kaushansky LJ, Ward CL, Sama R, Chian R-J, Boggio KJ, et al. Amyotrophic lateral sclerosis-linked FUS/TLS alters stress granule assembly and dynamics. *Mol Neurodegener*. 2013;8:30.

362. Lenzi J, De Santis R, de Turrís V, Morlando M, Laneve P, Calvo A, et al. ALS mutant FUS proteins are recruited into stress granules in induced pluripotent stem cell-derived motoneurons. *Dis Model Mech*. 2015;8:755–66.
363. Trieu EP, Gross JK, Targoff IN. Immunoprecipitation-western blot for proteins of low abundance. *Methods Mol Biol Clifton NJ*. 2009;536:259–75.
364. Louche A, Salcedo SP, Bigot S. Protein-Protein Interactions: Pull-Down Assays. *Methods Mol Biol Clifton NJ*. 2017;1615:247–55.
365. Brymora A, Valova VA, Robinson PJ. Protein-protein interactions identified by pull-down experiments and mass spectrometry. *Curr Protoc Cell Biol*. 2004;Chapter 17:Unit 17.5.
366. Gould KL, Ren L, Feoktistova AS, Jennings JL, Link AJ. Tandem affinity purification and identification of protein complex components. *Methods San Diego Calif*. 2004;33:239–44.
367. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Séraphin B. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol*. 1999;17:1030–2.
368. Du X, Li Y, Xia Y-L, Ai S-M, Liang J, Sang P, et al. Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci*. 2016;17:144.
369. Homola J, Yee SS, Gauglitz G. Surface plasmon resonance sensors: review. *Sens Actuators B Chem*. 1999;54:3–15.
370. Patching SG. Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. *Biochim Biophys Acta*. 2014;1838 1 Pt A:43–55.
371. Miura K. An Overview of Current Methods to Confirm Protein-Protein Interactions. *Protein Pept Lett*. 2018;25:728–33.
372. Rao VS, Srinivas K, Sujini GN, Kumar GNS. Protein-Protein Interaction Detection: Methods and Analysis. *Int J Proteomics*. 2014;2014:147648.
373. O’Connell MR, Gamsjaeger R, Mackay JP. The structural analysis of protein-protein interactions by NMR spectroscopy. *Proteomics*. 2009;9:5224–32.
374. Landrieu I, Leroy A, Smet-Nocca C, Huvent I, Amniai L, Hamdane M, et al. NMR spectroscopy of the neuronal tau protein: normal function and implication in Alzheimer’s disease. *Biochem Soc Trans*. 2010;38:1006–11.
375. Sekar RB, Periasamy A. Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations. *J Cell Biol*. 2003;160:629–33.

376. Jares-Erijman EA, Jovin TM. FRET imaging. *Nat Biotechnol.* 2003;21:1387–95.
377. Alam MS. Proximity Ligation Assay (PLA). *Curr Protoc Immunol.* 2018;123:e58.
378. Wang P, Yang Y, Hong T, Zhu G. Proximity ligation assay: an ultrasensitive method for protein quantification and its applications in pathogen detection. *Appl Microbiol Biotechnol.* 2021;105:923–35.
379. Rodríguez-Negrete E, Bejarano ER, Castillo AG. Using the yeast two-hybrid system to identify protein-protein interactions. *Methods Mol Biol Clifton NJ.* 2014;1072:241–58.
380. Boca M, Kretov DA, Desforges B, Mephon-Gaspard A, Curmi PA, Pastré D. Probing protein interactions in living mammalian cells on a microtubule bench. *Sci Rep.* 2015;5:17304.
381. Béganton B, Coyaud E, Mangé A, Solassol J. Approches nouvelles pour l'étude des interactions protéine-protéine. *MS Médecine Sci ISSN Pap 0767-0974 ISSN Numér 1958-5381 Vol 35 N° 3 P 223-231.* 2019. <https://doi.org/10.1051/medsci/2019035>.
382. Remy I, Michnick SW. A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways. *Methods San Diego Calif.* 2004;32:381–8.
383. Trugnan G, Fontanges P, Delautier D, Ait-Slimane T. FRAP, FLIP, FRET, BRET, FLIM, PRIM... De nouvelles techniques pour voir la vie en couleur! *Medecine/Sciences.* 2004;20:1027–34.
384. Lin Z, Chen S, Lin C. Sensitivity Improvement of a Surface Plasmon Resonance Sensor Based on Two-Dimensional Materials Hybrid Structure in Visible Region: A Theoretical Study. *Sensors.* 2020;20:2445.
385. Weimann M, Grossmann A, Woodsmith J, Özkan Z, Birth P, Meierhofer D, et al. A Y2H-seq approach defines the human protein methyltransferase interactome. *Nat Methods.* 2013;10:339–42.
386. Maucuer A, Desforges B, Joshi V, Boca M, Kretov DA, Hamon L, et al. Microtubules as platforms for probing liquid-liquid phase separation in cells - application to RNA-binding proteins. *J Cell Sci.* 2018;131:jcs214692.
387. Pankivskiy S, Pastré D, Steiner E, Joshi V, Rynditch A, Hamon L. ITS1N1 regulates SAM68 solubility through SH3 domain interactions with SAM68 proline-rich motifs. *Cell Mol Life Sci.* 2021;78:1745–63.



388. Kamelgarn M, Chen J, Kuang L, Arenas A, Zhai J, Zhu H, et al. Proteomic analysis of FUS interacting proteins provides insights into FUS function and its role in ALS. *Biochim Biophys Acta*. 2016;1862:2004–14.
389. Lee W, Lee SW, Lee G, Yoon DS. Atomic Force Microscopy Analysis of EPPS-Driven Degradation and Reformation of Amyloid- $\beta$  Aggregates. *J Alzheimers Dis Rep*. 2:41–9.
390. Ikegawa M, Nirasawa T, Kakuda N, Miyasaka T, Kuzuhara Y, Murayama S, et al. Visualization of Amyloid  $\beta$  Deposits in the Human Brain with Matrix-assisted Laser Desorption/Ionization Imaging Mass Spectrometry. *J Vis Exp JoVE*. 2019. <https://doi.org/10.3791/57645>.
391. Bigio EH, Wu JY, Deng H-X, Bit-Ivan EN, Mao Q, Ganti R, et al. Inclusions in frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) and amyotrophic lateral sclerosis (ALS), but not FTLD with FUS proteinopathy (FTLD-FUS), have properties of amyloid. *Acta Neuropathol (Berl)*. 2013;125:463–5.
392. Reber S, Jutzi D, Lindsay H, Devoy A, Mechttersheimer J, Levone BR, et al. The phase separation-dependent FUS interactome reveals nuclear and cytoplasmic function of liquid–liquid phase separation. *Nucleic Acids Res*. 2021;49:7713–31.
393. Chi B, O’Connell JD, Yamazaki T, Gangopadhyay J, Gygi SP, Reed R. Interactome analyses revealed that the U1 snRNP machinery overlaps extensively with the RNAP II machinery and contains multiple ALS/SMA-causative proteins. *Sci Rep*. 2018;8:8755.
394. Elden AC, Kim H-J, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, et al. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*. 2010;466:1069–75.
395. Yamazaki T, Chen S, Yu Y, Yan B, Haertlein TC, Carrasco MA, et al. FUS-SMN protein interactions link the motor neuron diseases ALS and SMA. *Cell Rep*. 2012;2:799–806.
396. Ryan VH, Dignon GL, Zerze GH, Chabata CV, Silva R, Conicella AE, et al. Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol Cell*. 2018;69:465-479.e7.
397. Sun Z, Diaz Z, Fang X, Hart MP, Chesi A, Shorter J, et al. Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS. *PLoS Biol*. 2011;9:e1000614.
398. Farrowell N, Lambert-Smith IA, Warraich S, Blair I, Saunders D, Hatters D, et al. Distinct partitioning of ALS associated TDP-43, FUS and SOD1 mutants into cellular inclusions. *Sci Rep*. 2015. <https://doi.org/10.1038/srep13416>.

399. Che M-X, Jiang Y-J, Xie Y-Y, Jiang L-L, Hu H-Y. Aggregation of the 35-kDa fragment of TDP-43 causes formation of cytoplasmic inclusions and alteration of RNA processing. *FASEB J.* 2011;25:2344–53.
400. Knowles TPJ, Vendruscolo M, Dobson CM. The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol.* 2014;15:384–96.

**Titre :** Caractérisation de l'interaction entre FUS et TDP-43, deux protéines de liaison à l'ARN impliquées dans les maladies neurodégénératives

**Mots clés :** Maladie neurodégénératives, Protéine FUS de liaison à l'ARN, Protéine TDP-43 de liaison à l'ARN, Ribonucléoprotéines, Complexes multiprotéiques, séparation de phase liquide-liquide

**Résumé :** FUS et TDP-43 sont deux protéines de liaison à l'ARN qui partagent une structure similaire. Elles sont aussi impliquées dans un nombre de fonctions semblables autour de la prise en charge de l'ARN, de son transport, de l'épissage, ou encore de la formation de condensats par le processus de séparation de phases liquide-liquide. Toutefois, elles ont des comportements distincts et des modes de prise en charge de l'ARN qui diffèrent. FUS se lie à l'ARN sans spécificité de séquence, tandis que TDP-43 fait preuve d'une grande spécificité pour des motifs GU. Ces protéines présentent un intérêt majeur puisque leur présence au sein d'agrégats cytoplasmiques est une caractéristique de certaines maladies neurodégénératives, dont les caractéristiques cliniques et génétiques s'entrecroisent, comme la SLA, la DFT ou encore la LATE. La présence de ces protéines à la fois au sein de mêmes agrégats, mais aussi de mêmes complexes cellulaires, suppose une convergence dans les mécanismes dans lesquels elles sont impliquées. Par ailleurs, l'interaction entre ces deux protéines a déjà été mise en évidence.

Ce travail de thèse s'intéresse à mieux comprendre et caractériser cette interaction, ainsi que leur relation avec l'ARN grâce à des techniques originales. Les expériences menées au cours de ce travail ont permis de confirmer cette interaction, en apportant des éléments nouveaux sur la miscibilité de ces protéines et leur organisation en condensats. Nous avons montré que le RRM de TDP-43 était particulièrement important dans cette interaction, et que les formes pathologiques issues du clivage de TDP-43 perdaient la capacité d'interagir avec FUS, du fait de leur liaison altérée à l'ARN. Enfin, les acquisitions effectuées en AFM à l'échelle de la molécule unique, permettent de mettre en évidence le comportement coopératif de ces deux protéines dans la prise en charge de l'ARN.

**Title:** *Characterization of the interaction between FUS and TDP-43, two RNA-binding proteins involved in neurodegenerative diseases*

**Keywords:** *Neurodegenerative diseases, FUS RNA-binding protein, TDP-43 RNA binding protein, Ribonucleoproteins, Multiprotein complexes, Liquid-liquid phase separation*

**Abstract:** *FUS and TDP-43 are two RNA binding proteins sharing similar structures. They are involved in several functions of RNA processing, such as RNA transport, splicing, transcription, as well as in cell compartmentalization. Despite these similarities, they exhibit distinct behaviours in their RNA-binding modes. FUS can bind RNA non-specifically, while TDP-43 is very specific for GU motif. These two proteins have been found aggregated in cytoplasmic inclusions of neurons from patients with overlapped neurodegenerative diseases as ALS, FTLD, and LATE. The fact that these two proteins were co-aggregated in the same inclusions, and are involved in the same cellular complexes, suggests that they are implicated in common pathways. Thus, the interaction between these two proteins has already been highlighted.*

*This thesis report focuses on a better understanding and characterization of this interaction, in particular on the influence on mRNA, using original approach. Experiments conducted during this work have confirmed the interaction, with a focus on protein miscibility and the compartmentalization mechanisms. Besides, we demonstrated that RRM1 of TDP-43 are particularly important for the interaction, while pathological truncated forms of TDP-43 disrupt the interaction with FUS by inducing a change in their RNA binding. Finally, data obtained by high resolution imaging by AFM, highlight a cooperative binding of TDP-43 and FUS on RNA.*