



HAL
open science

Caractérisations biochimiques des protéines à répétitions de dipeptides et de leurs assemblages

Laurent Brasseur

► **To cite this version:**

Laurent Brasseur. Caractérisations biochimiques des protéines à répétitions de dipeptides et de leurs assemblages. Biochimie [q-bio.BM]. Université Paris Saclay (COMUE), 2019. Français. NNT : 2019SACLS538 . tel-03505883

HAL Id: tel-03505883

<https://theses.hal.science/tel-03505883>

Submitted on 1 Jan 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.

Caractérisations biochimiques des protéines à répétitions de dipeptides et de leurs assemblages

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°568 Signalisations et réseaux
intégratifs en biologie (BIOSIGNE)

Spécialité de doctorat: Aspects moléculaires et cellulaires de la
biologie

Thèse présentée et soutenue à Fontenay-aux-Roses, le 17/12/19, par

Mr. Laurent BRASSEUR

Composition du Jury :

Myriam TAVERNA Professeur, Université Paris-Saclay (Institut Galien Paris Sud CNRS)	Président
Anja BÖCKMANN Directeur de recherche, Université de Lyon (Institut de Biologie et Chimie des Protéines)	Rapporteur
Séverine BOILLÉE Chargé de Recherche, Sorbonne Université (ICM, INSERM, CNRS)	Rapporteur
Luc BOUSSET Chargé de recherche, Université Paris-Saclay (MIRGen - CEA)	Co-Directeur de thèse



« Un jour j'arracherai l'ancre qui tient mon navire loin des mers. »

Henri Michaux (1899 – 1984)

Remerciements

En premier lieu, je tiens à remercier Anja Böckmann, Séverine Boillée et Myriam Taverna d'avoir accepté de prendre le temps de lire ma thèse et juger mon travail.

Je tiens ensuite à remercier Ronald Melki et Luc Bousset. Ronald pour m'avoir accepté au sein de ton équipe, m'avoir enseigné le métier de chercheur et partagé ton savoir. Je remercie également Luc de m'avoir partagé ses connaissances théoriques et expérimentales. Sans votre patience et vos conseils, je n'aurais jamais pu faire le travail que je présente ici.

Evidemment, je remercie aussi chaque membre de l'équipe. Tous, ont contribué à ce travail, que ce soit en prodiguant des conseils ou tout simplement en ayant participé à l'ambiance de camaraderie de l'équipe. Ainsi, je tiens particulièrement à remercier Alexis Fenyi, Maud Landureau, Stéphane Tchankouo, Maya Bendifallah, Nolwen Rey et Emilie Caroux avec lesquels j'ai eu des conversations passionnantes et avec qui j'ai aussi partagé de bons moments d'amitié. Je remercie également Tracy Bellande pour la bonne humeur et la bienveillance angélique qu'elle apporte à tous.

Je n'oublie pas Elodie Monsellier, Virginie Redeker, Karine Madiona, Laura Pieri, Mehdi Kabani, Amulya Shrivastava et Jimmy Savistchenko qui m'ont accompagné dans mon travail et dans ses trois marquantes années de ma vie.

Je remercie également Mimoun Azzouz et Paolo Marchi de l'Université de Sheffield avec qui j'ai travaillé en collaboration.

Je remercie également les services administratifs, techniques et d'hygiène de MIRCCen et de Neuro-PSI. Tous participent à faire avancer nos connaissances en assurant le bon fonctionnement des instituts de recherche.

Enfin, je tiens à remercier chaleureusement ma famille et mes amis. Sans le soutien et l'amour qu'ils me portent, je n'aurais jamais pu arriver au bout des 8 années d'études

qui se clôturent avec cette thèse. J'adresse ainsi des sentiments de profondes reconnaissances à toute ma famille. Je remercie ma mère dont le courage et la résilience inspirent l'humilité. Je remercie mon père, qui m'a transmis la soif de connaissances et dont l'intégrité morale force au respect. Je remercie mes grands-parents qui sont des modèles de probité et de dévouement. Je remercie mon grand frère Julien qui me connaît mieux que quiconque et mon petit frère Vincent – la mémoire de la famille. Nous partageons tous une seule et même âme.

SOMMAIRE

Index des illustrations	10
Abréviations	11
Glossaire des symptômes	12
Résumé	13
Introduction	17
I – Aspects cliniques de la Sclérose Latérale Amyotrophique et des Démences Fronto-Temporale	19
A – La Sclérose Latérale Amyotrophique	19
B – Les Démences Fronto-Temporales	24
C – La SLA et la DFT, les deux pôles d’un même spectre pathologique ?	29
II – Les mécanismes pathologiques et les déterminants moléculaires majeurs des SLA/DFT	33
A – Comorbidités et facteurs environnementaux dans les SLA/DFT.	33
B – Aspects génétiques majeurs des SLA/DFT	34
1 – La mutation du gène SOD1 dans la SLA.....	34
2 – La mutation du gène MAPT dans les DFT	35
3 – La mutation des gènes TARDBP et Fus dans la SLA/DFT	35
4 – Plusieurs gènes impliqués dans la protéostasie sont retrouvés mutés dans les SLA/DFT	36
C – Diversité de l’agrégation protéique dans les SLA/DFT	37
1) L’agrégation de la protéine SOD1	37
2) L’agrégation de la protéine Tau	39
3) L’agrégation de la protéine TDP-43	40
4) L’agrégation de la protéine FUS/TLS	41
D – Les SLA/DFT, des maladies de type prion et du dérèglement des ARN	42

1) L'agrégation protéique dans les pathologies neurodégénératives et le concept de maladies de type prion.....	42
2) Les SLA/DFT sont des protéinopathies multiples appartenant à la famille des maladies de type prion	46
3) La dérégulation du métabolisme des ARN dans les SLA/DFT.....	48
III – La présence de répétitions de l'hexanucleotide GGGGCC au niveau du gène C9orf72 et la cause génétique principale des SLA/DFT.	50
A – Le gène C9orf72 et sa mutation.....	50
B – Trois hypothèses sur la pathogénicité de la mutation C9orf72.....	51
1) La perte de fonction du gène C9orf72	52
2) La toxicité des ARN sens et antisens	54
3) Les Protéines à Répétitions de Dipeptides.....	56
IV – Mécanismes de toxicité des PRD.....	60
A – L'agrégation des PRD	60
1) Poly-GA est capable de former des fibres amyloïdes <i>in vitro</i> ayant des caractéristiques de type prion	60
2) Hypothèses sur l'agrégation des PRD.....	61
B – Les différentes voies de toxicité des PRD.....	64
1) Les cibles multiples des PRD.....	64
2) L'agrégation des PRD engendre-t-elle celles des autres protéines retrouvées sous forme insoluble dans les SLA/DFT ?	65
Conclusions.....	68
Travaux de thèse	72
I – Dipeptide Repeat derived from C9orf72 Hexanucleotide Expansions Forms Amyloids or natively unfolded structures in vitro	74
II – Propriétés cellulaires de protéines à répétitions de dipeptides et de leurs assemblages.....	109
Conclusions et perspectives	115
I – Conclusions	117
A – Production de PRD et caractérisation de leur structure en solution.	117

B – Conditions d'agrégation et caractérisations des assemblages de PRD.....	118
C – Marquage des assemblages de PRD dans le cadre d'une collaboration.....	120
II – Perspectives	122
Bibliographie.....	125

Index des illustrations

Figure 1 : Section d'une moelle épinière.....	20
Figure 2 : Représentation du spectre pathologique de la Sclérose Latérale Amyotrophique.....	24
Figure 3 : Localisation des Aires de Broca et de Wernicke.....	28
Figure 4 : Contributions génétiques dans la SLA et la DFT..	37
Figure 5 : Proportions des protéines formant des inclusions insolubles dans la SLA et la DFT.....	38
Figure 6 : Schéma de l'agrégation amyloïde d'une protéine.....	45
Figure 7 : Conséquences de la présence de répétitions de l'hexanucléotide GGGGCC au niveau du gène C9orf72 et hypothèses de toxicité.	53
Figure 8 : Les brins sens et antisens d'hexanucléotides sont traduits pour chaque cadre de lecture ouvert.....	57

Abréviations

AMP : Atrophie Musculaire Progressive

ALYREF ou Aly/REF export factor : ALY; THO complex subunit 4, Tho4; RNA and export factor binding protein 1, Refbp1

APP : Aphasie Primaire Progressive

BMAA : Beta-methylamino-L-alanine

C9orf72 : Chromosome 9 open reading frame 72

DCB : Démence cortico-basale

DENN : Differentially Expressed in Normal and Neoplasia

DFT : Démence Fronto-Temporale

EWSR1 : Ewing Sarcoma breakpoint Region 1

FUS / TLS : Fused in Sarcoma / Translocated in Sarcoma

GEF : Guanine Exchange Factor

hnRNP : heterogenous nuclear ribonucleoproteins

NLS : Nuclear Localization Sequence

PRD : Protéine à Répétitions de Dipeptides

PSP : Paralysie Supra-nucléaire Progressive

SCB : Syndrome Cortico-Basal

SLA : Sclérose Latérale Amyotrophique

SLP : Sclérose Latérale Progressive

SOD1 : Superoxyde Dismutase 1

SRSF : Serine-rich-splicing factor

TAF15 : TATA-binding protein-associated factor 2N

TARDBP ou TDP-43 : TAR DNA-binding protein 43

UBQLN2 : Ubiquiline 2

UPS : Ubiquitin-Proteasom System

vc-DFT : Variant comportemental de la Démence Fronto-Temporale

VCP : Valosin-Containing Protein

Glossaire des symptômes

Apathie : Perte de motivation et d'intérêt

Bradykinésie : Ralentissement et perte de contrôle de la finesse des mouvements

Dysarthrie : Trouble articulaire de la parole

Dyslexie : Trouble de la lecture et de l'écriture

Dysphagie : Trouble de la déglutition

Dysphorie : Trouble de l'humeur caractérisé par un sentiment d'inconfort psychique. Opposé à l'euphorie.

Fasciculation : Contractions brèves et involontaires de faisceaux musculaires

Hypertonie : Tonus musculaire trop élevé

Paraphasie : Trouble de la parole lié à la substitution ou l'altération des mots/phonèmes

Persévération : Répétitions incontrôlées de mêmes réponses verbales ou gestuelles sans tenir compte d'un changement de situation.

Stéréotypie : Répétitions incontrôlées de gestes ou paroles qui ne servent aucune intention

Résumé

La Sclérose Latérale Amyotrophique (SLA) et la Démence Fronto-temporale (DFT) sont deux pathologies neurodégénératives complexes. La première se manifeste essentiellement par des symptômes moteurs tandis que la deuxième par des troubles cognitifs et comportementaux. Bien que ces deux maladies fussent initialement décrites séparément, de nombreux aspects symptomatiques, anatomiques et biochimiques les rapprochent.

Récemment, la découverte de la mutation du gène C9orf72 dans les deux maladies a soutenu l'hypothèse que la SLA et la DFT constituaient les deux pôles d'un même spectre pathologique. Cette mutation est la première cause génétique des deux maladies. Elle consiste en la répétition d'un hexanucléotide (GGGGCC) pouvant aller jusqu'à plusieurs milliers d'exemplaires. Trois hypothèses ont été avancées pour rendre compte de la pathogénicité de ces expansions :

- (1) L'haplo-insuffisance du gène C9orf72 induit par la séquence d'hexanucléotides répétés.
- (2) Les ARN issus de la transcription de ces séquences seraient toxiques.
- (3) L'agrégation de Protéines à Répétitions de Dipeptides (PRD) issues de la traduction des ARN d'hexanucléotides constitue également une voie de toxicité possible. Ces PRD sont au nombre de 5 : poly-GA (Glycine-Alanine), poly-GP (Glycine-Proline), poly-GR (Glycine-Arginine), poly-PR (Proline-Arginine) et poly-PA (Proline-Alanine).

Parmi ces trois hypothèses, il semble que la dernière soit la plus à même de reproduire, dans des modèles animaux, des marqueurs essentiels de la SLA et de la DFT,

tel que la dégénérescence neuronale.

Les mécanismes de toxicité de ces protéines sont en cours d'investigation. Seulement, peu d'études se sont intéressées au comportement *in vitro* des PRD. Pourtant, explorer la manière dont s'agrègent ces protéines dans ces maladies est essentiel pour les comprendre davantage. En effet, l'agrégation protéique est un mécanisme fondamental responsable de la mort cellulaire dans de nombreuses maladies neurodégénératives telles que la maladie d'Alzheimer, de Parkinson ou encore celle de Huntington.

L'objectif de cette thèse a été de produire *in vitro* certains de ces PRD afin de déterminer leur structure en solution, ainsi que les conditions et la nature de leur agrégation. Pour cela, Audrey Coens et Luc Bousset ont sous-cloné des séquences d'ADN codant pour les PRD GAx34 et GPx24 dans un vecteur d'expression bactérien. Nous nous sommes également procuré deux plasmides codant pour les peptides GAx50 et PAx50. En exprimant ces PRD dans des bactéries, j'ai pu les produire puis les purifier par FPLC.

J'ai d'abord montré que GAx34, GPx24 et PAx50 étaient des protéines très peu structurées. Dans nos conditions expérimentales, seul poly-GA s'assemble pour former des agrégats fibrillaires de nature amyloïde. Ces résultats sont cohérents avec la littérature disponible sur les PRD. Nous avons aussi montré que ces agrégats ne liaient ni la Thioflavin T, ni la Primuline : deux molécules fluorescentes habituellement utilisées pour marquer les assemblages amyloïdes. Quant aux PRD GPx24 et PAx50, ils forment peu d'agrégats à la morphologie amorphe.

En parallèle de ces caractérisations biochimiques, j'ai produit des échantillons de protéines oligomériques et fibrillaires marquées covalamment avec des molécules fluorescentes. Ces préparations ont été réalisées dans le cadre d'une collaboration avec

l'équipe de Pr. Mimoun Azzouz de l'université de Sheffield. Ce travail est en cours et a pour objectif d'étudier l'internalisation, la toxicité et la propagation de ces PRD dans des modèles cellulaires.

Introduction

I – Aspects cliniques de la Sclérose Latérale Amyotrophique et des Démences

Fronto-Temporale

A – La Sclérose Latérale Amyotrophique

La Sclérose Latérale Amyotrophique (SLA), également appelée Maladie de Charcot ou encore Maladie de Lou Gehrig, est une maladie neurodégénérative caractérisée par des symptômes moteurs comme la paralysie, une atrophie des muscles squelettiques et dans certains cas des symptômes cognitifs et comportementaux (Hardiman et al. 2017). La SLA est la forme la plus courante des maladies du motoneurone avec une incidence de 2-3/100 000 et une prévalence de 3-6/100 000 (Grad et al., 2017).

L'anatomiste Charles Bell (1774-1842) fut le premier à s'intéresser à cette pathologie (Bell, 1830). Après avoir différencié la fonction motrice des racines antérieures de la moelle épinière de celle sensorielle des racines postérieures, il rechercha des patients ayant une atteinte purement motrice (Fig 1.). Bell et les neurologues de son époque ne réussirent pourtant pas à définir si les symptômes moteurs de leurs patients étaient d'origine musculaire ou nerveuse. Ce n'est qu'à partir de travaux effectués en 1865 par l'un des pères fondateurs de la neurologie moderne, Jean-Martin Charcot (1825-1893) et son élève Alix Joffroy (1844-1908), qu'on a commencé à caractériser la SLA et la séparer d'autres maladies semblables – en associant les aspects cliniques de la pathologie à des lésions anatomiques définies. Pour Charcot, la SLA se caractérise par une atteinte duale du système moteur. Ainsi, des lésions observées dans la colonne latérale de la moelle épinière entraînent une paralysie chronique progressive ainsi que des contractions musculaires, tandis que des lésions de la corne antérieure sont associées à une atrophie musculaire (Rowland., 2001 ; Goetz, 2000). Le terme Sclérose Latérale

Amyotrophique n'apparaît sous la plume de Charcot qu'en 1874. Les termes « Sclérose Latérale » se réfèrent aux lésions tissulaires et « Amyotrophique » à l'atrophie musculaire résultante.

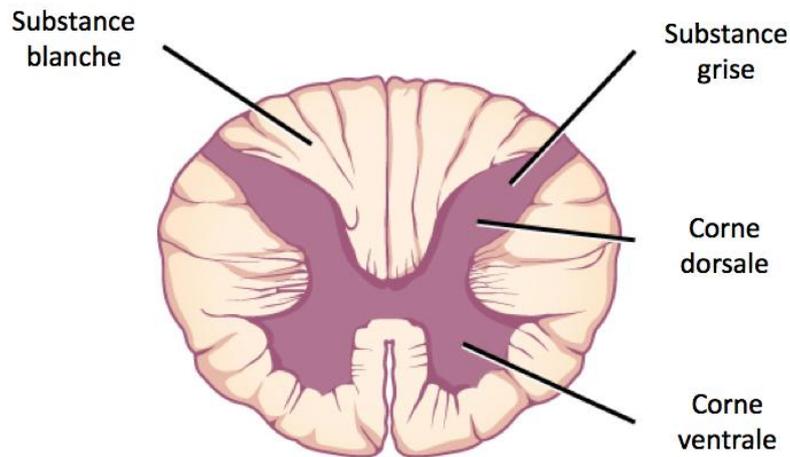


Figure 1 : Section d'une moelle épinière. La corne dorsale est située sur la partie supérieure de la substance grise, tandis que la corne ventrale se trouve dans la partie inférieure.

Depuis les découvertes de Charcot, nos connaissances sur la SLA se sont nettement enrichies. Malheureusement, dans de nombreux cas, le diagnostic reste encore délicat à établir. Cela dit, une dégénérescence des motoneurones au niveau central et périphérique, ainsi qu'une apparition sérielle de symptômes moteurs sont les critères nécessaires pour pouvoir dire d'une pathologie qu'elle correspond à une Maladie de Charcot (Hardiman et al., 2017). Le diagnostic se base essentiellement sur les signes cliniques et les antécédents familiaux du patient.

La plupart des malades sont atteints d'une faiblesse et d'une atrophie musculaire, d'une paralysie progressive, d'hyperréflexie, d'hypertonie, de fasciculation (contraction involontaire de faisceaux musculaires) et de dysphagie (difficulté à déglutir). A cela

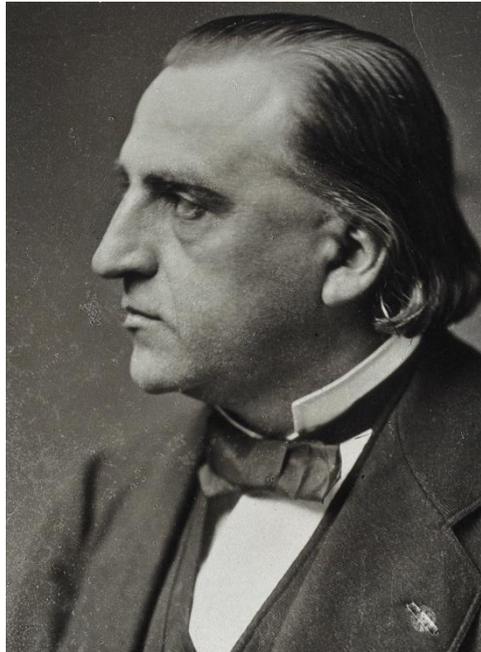
s'ajoute que près de la moitié des patients développent des symptômes cognitifs et/ou comportementaux (Renton et al., 2014).

Dans de nombreux cas, la pathologie progresse par contiguïté. Ce qui permet de différencier différentes formes phénotypiques de la pathologie par l'origine de sa propagation (Ravits et al., 2009 ; Swinnen et al. 2014). Les formes les plus courantes se propagent depuis la moelle épinière ou depuis le bulbe rachidien.

Dans le premier cas, les premiers symptômes apparaissant consistent en une atrophie, une faiblesse musculaire, une fasciculation, une hyperréflexie et une hypertonie. Ces manifestations cliniques sont cohérentes avec un dysfonctionnement des fonctions normales de la moelle épinière.

Dans les formes « bulbaires », les premiers symptômes sont une dysarthrie (troubles articulaires de la parole), de la dysphagie, une fasciculation des muscles linguaux ainsi que des rires et des pleurs incontrôlés. Il existe également une forme rare de SLA qui se déclare avec des symptômes respiratoires.

Bien qu'on ne comprenne pas pourquoi il existe ces différentes formes de SLA, certaines mutations génétiques semblent partiellement en rendre compte. Ainsi, la mutation C9orf72 est associée à la forme bulbaire tandis que les mutations de la Superoxyde Dismutase 1 (SOD1) sont retrouvées chez des patients ayant une forme spinale (Swinnen et al., 2014).



Jean-Martin Charcot (1825-1893), un des pères fondateurs de la neurologie moderne

L'âge moyen de l'apparition de la maladie se situe entre 50 et 60 ans, mais il existe aussi des formes juvéniles se déclarant avant 25 ans (Renton et al., 2014). Ces dernières sont généralement à progression lente. Des mutations touchant les gènes ALS2, SETX et FUS sont retrouvées chez ces patients (Sabatelli et al., 2008 ; Pupillo et al., 2014). Pour 60% des malades âgés entre 20 et 60 ans, on remarque surtout une dégénérescence du système moteur central. Les formes plus tardives sont plus virulentes et ont une probabilité plus élevée d'être d'origine bulbaire (Zoccolella et al., 2008).

La rapidité de la maladie est variable. La valeur médiane de survie des patients se situe entre 2 et 4 ans après l'apparition des premiers signes cliniques. Seulement 5 à 10 % des sujets survivent au-delà de 10 ans (Del Aguila et al., 2003 ; Chi et al., 2009). Chez ces derniers, on n'observe pourtant pas de différence anatomique majeure (Iwanaga et al., 1997). Cependant, les patients ayant développé des formes juvéniles de la maladie et ceux ayant une lésion majoritairement centrale, ont plus de chances de dépasser l'espérance de vie moyenne (Zoccolella et al., 2008 ; Pupillo et al., 2014). Cependant, au

sein d'une même famille, différents individus portant pourtant la même mutation voient leurs maladies avoir des évolutions variables (Kim et al., 2007 ; Penco et al., 2011). Il existe donc des facteurs, génétiques ou environnementaux, qui moduleraient l'espérance de vie des patients SLA.

D'autres maladies du motoneurone sont proches de la SLA tant par leurs manifestations cliniques que par leurs déterminants génétiques. C'est le cas de l'Atrophie Musculaire Progressive (AMP). Cette pathologie neurodégénérative n'atteint que les motoneurones périphériques et se traduit par une amyotrophie, une fasciculation des muscles squelettiques et une faiblesse musculaire. Cependant, l'AMP ne partage aucun des signes cliniques de la SLA associés à une dégénérescence motrice centrale (Swinnen et al., 2014). Il est pourtant notable qu'on retrouve des lésions neuronales centrales extra-motrices partagées entre les deux pathologies (Ince et al., 2003 ; Prudlo et al., 2012). De plus, des mutations communes du gène SOD1 sont impliquées dans la SLA et l'AMP (Cudkowicz et al., 1998 ; Cervenakova et al., 2000). Ainsi, il apparaît que la séparation de ces deux maladies ne repose que sur un critère, la lésion motrice centrale, alors que d'autres aspects cliniques, anatomiques et génétiques les rapprochent.

D'une façon complémentaire à l'AMP, la Sclérose Latérale Primaire (SLP) est une maladie du motoneurone centrale, non-périphérique, qui partage des similitudes avec la SLA. Par ailleurs, les deux pathologies sont tant similaires que les patients SLP ne peuvent être véritablement diagnostiqués que tardivement. On retrouve d'ailleurs ces mêmes lésions après l'autopsie de sujets qui étaient pourtant diagnostiqués SLP (D'Amico et al., 2013). Il est aussi notable que plusieurs des mutations soient communes aux deux maladies (Renton et al., 2014).

Les convergences entre la SLP ou l'AMP avec la SLA ont conduit à élargir la définition de cette dernière en un spectre pathologique regroupant toutes ces formes. Les

malades n'ayant que des lésions motrices centrales (SLP) se situeraient à une extrémité du spectre, tandis que ceux n'ayant que l'atteinte médullaire (AMP) seraient à l'autre extrémité (Fig. 2).



Figure 2 : Représentation du spectre pathologique de la Sclérose Latérale Amyotrophique. La composante médullaire de la pathologie est représentée en rouge, tandis que la composante centrale est en bleu.

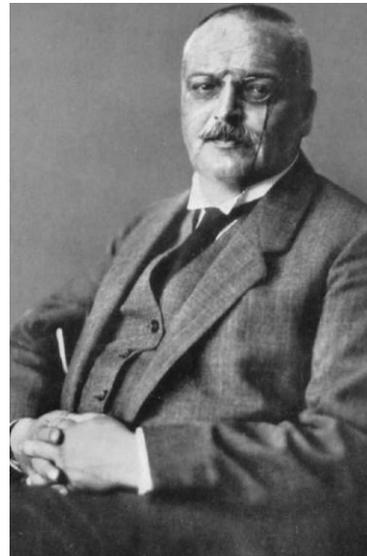
La diversité symptomatique et anatomopathologique de la SLA en fait une pathologie neurodégénérative complexe.

B – Les Démences Fronto-Temporales

Les Démences Fronto-Temporales (DFT) correspondent à un groupe de pathologies neurodégénératives progressives qui entraînent des troubles du comportement et du langage associés à des lésions des cortex frontaux et temporaux. Les DFT sont la deuxième cause de démence après la Maladie d'Alzheimer, avec une incidence de 3-4/100 000 et une prévalence de 15-22/100 000 (Onike et al., 2013).

En 1892, le psychiatre Arnold Pick (1851-1924) décrit le cas médical d'un homme de 71 ans, August H, frappé d'un déclin cognitif progressif. A la fin de sa vie, ce patient avait du mal à comprendre le langage écrit et oral, une tendance à répéter de façon

incontrôlée certaines phrases et à altérer des mots (paraphasie) (Pick et al., 1901). C'est le premier cas étudié de Maladie de Pick. Le professeur d'anatomie Hans Chiari (1851-1916) effectua une autopsie sur le corps d'August H et mit en évidence une atrophie cérébrale, notamment au niveau du gyrus temporal supérieur gauche – dont Pick avait postulé qu'elle était responsable de l'aphasie de son patient. En 1911, le neurologue et psychiatre allemand Alois Alzheimer (1864-1915), célèbre pour la description qu'il fit de la démence qui porte son nom, rapporte la présence d'inclusions cytoplasmiques globulaires, appelées Corps de Pick, dans les neurones de ces patients (Alzheimer, 1991).



A gauche, Arnold Pick (1851-1924) et, à droite, Alois Alzheimer (1864-1915), deux pionniers de la recherche sur les DFT

Les recherches qui suivirent vinrent enrichir les connaissances sur cette pathologie qu'on préférera appeler Démence Fronto-Temporale (DFT) ou Démence Fronto-Temporale Lobaire (DFTL) – les corps de Pick n'étant pas retrouvés chez tous les patients partageant pourtant des signes cliniques identiques. La dénomination Maladie de Pick sera réservée à une forme particulière.

La DFT s'avère être une neuropathologie au moins autant hétérogène que l'est la SLA. Aujourd'hui, on classifie la DFT en 3 catégories cliniques principales : la variante comportementale (vc-DFT), l'Aphasie Primaire Progressive (APP) et la Démence Fronto-Temporale associée à une maladie du motoneurone (Finger, 2016). Parmi les pathologies assimilées à la DFT, il existe deux Tauopathies : la Dégénérescence Cortico-Basale (DCB) et la Paralyse Supranucléaire Progressive (PSP).

La vc-DFT se diagnostique chez des patients ayant au moins trois des symptômes suivants : apathie, manque d'empathie, troubles alimentaires, comportements désinhibés ou compulsifs et des troubles de la fonction cognitive exécutive (Neary et al., 1998). Deux profils comportementaux dont les symptômes peuvent parfois se recouper ont été décrits :

- 1) Un profil apathique (DFT-A) caractérisé par une baisse de la volition, de la motivation, une tendance à l'isolement ainsi qu'une augmentation du temps de réponse à la douleur
- 2) Un profil désinhibé (DFT-D) caractérisé par une hyperoralité, une préférence pour la nourriture sucrée, une stéréotypie motrice (répétition incontrôlée de gestes ou paroles qui ne servent aucune intention) et une persévération (répétition incontrôlée de réponses verbales ou gestuelles sans tenir compte d'un changement de situation) (Snowden et al., 2001).

Pour que le diagnostic clinique soit véritablement confirmé, il est nécessaire d'observer des lésions anatomiques au niveau des lobes frontaux et temporaux.

L'Aphasie Primaire Progressive (APP), initialement dénommée maladie de

Medulam, est une forme de DFT linguistique comportant 3 variantes : sémantique (vsAPP), non-fluente/agrammatique (naAPP) et logopénique (lpAPP) (Gorno-Tempini et al., 2011).

Le variant sémantique se caractérise par une impossibilité à nommer les choses et une difficulté à comprendre certains mots. En dehors de cette perte de la mémoire des mots, d'autres symptômes peuvent être observés comme une difficulté à former certains gestes d'écriture (dysgraphie), des troubles pour reconnaître les individus, des difficultés à manipuler certains objets, à comprendre le sens de concepts abstraits, une tendance à substituer ou altérer des mots par similitude grammaticale (erreurs paragrammatiques) et une difficulté à lire et écrire des mots dont l'écriture est atypique par rapport à la prononciation (dyslexie de surface).

Comme son nom l'indique, le variant agrammatique (ou non-fluent) de l'APP se caractérise en premier lieu par des difficultés à construire des phrases. Les phrases produites par les patients sont dans un style télégraphique et peu ou pas structurées. La difficulté à articuler correctement les mots est le symptôme le plus courant et c'est souvent le premier à apparaître. Il est nécessaire que l'un de ces deux symptômes soit présent pour poser le diagnostic. Parmi les autres manifestations cliniques, on retrouve une tendance à déformer des mots à l'oral par substitution, insertion ou suppression de sons (paraphasie phonologique). Contrairement à d'autres démences comme le variant sémantique de l'APP ou la Maladie d'Alzheimer, la naAPP n'affecte pas la compréhension sémantique des mots et des objets, ce qui facilite le diagnostic en cas d'ambiguïté. Celui-ci peut être confirmé par imagerie, des lésions étant présentes au niveau du gyrus frontal inférieur, de l'insula et des aires motrices et pré-motrices (Gorno-Tempini et al., 2011). Il est à noter que le gyrus frontal inférieur inclut l'Aire de Broca (Fig. 3), impliquée dans la production et la compréhension du langage notamment en étant le support de la mémoire de travail syntaxique et phonologique (Dronkers et al., 2004).

Les patients atteints de la forme logopénique de l'APP développent deux symptômes principaux qui permettent de poser le diagnostic. Ces malades ont une difficulté à répéter des phrases et des pertes d'accès au vocabulaire. Cette pathologie fait suite à des lésions cérébrales au niveau du carrefour temporo-pariétal qui contient, entre autres, l'Aire de Wernicke, une région cérébrale impliquée dans la compréhension du langage écrit et oral (Gorno-Tempini et al., 2011). Cette aire est fondamentale dans la représentation mentale des phonèmes et non dans la compréhension sémantique du langage comme il a été avancé pendant de nombreuses années (Fig. 3.). Ainsi, des lésions de cette structure cérébrale rendent compte des troubles phonologiques qui caractérisent la lpAPP (Mesulam et al., 2014 Binder et al., 2017).

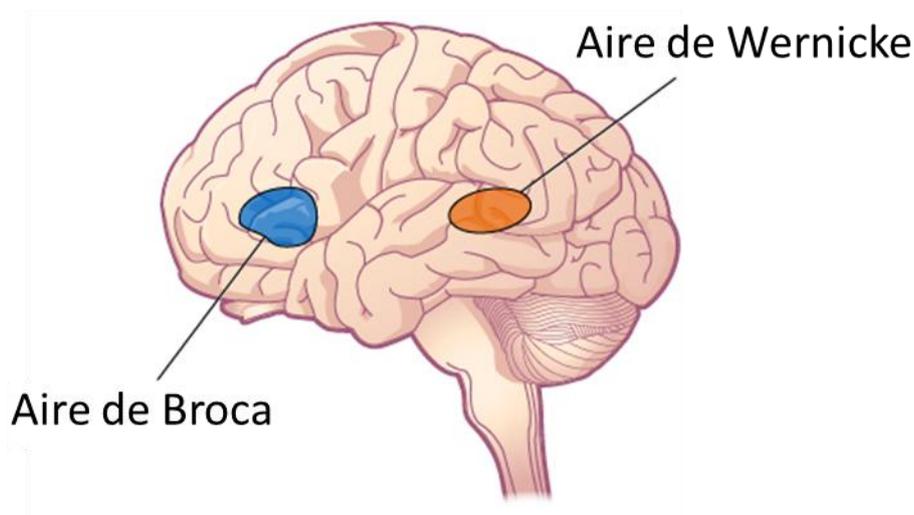


Figure 3 : Localisation des Aires de Broca et de Wernicke deux aires fondamentales dans la production et la compréhension du langage. Dans les Démences Fronto-Temporales, ces aires sont lésées, ce qui explique les aphasies présentes chez ces patients.

De même que des symptômes cognitifs et comportementaux de type DFT se retrouvent chez de nombreux sujets SLA, on estime à 15% le nombre de patients DFT développant des symptômes moteurs correspondant à une SLA (Lomen-Hoerth et al.,

2011). On parle de Démence Fronto-temporale associée à une Sclérose Latérale Amyotrophique (DFT/SLA ou SLA/DFT).

C'est l'étiologie de la SLA/DFT qui constituera le cadre clinique de cette thèse.

C – La SLA et la DFT, les deux pôles d'un même spectre pathologique ?

Quelques décennies après la définition de la SLA par Charcot, plusieurs études décrivent l'existence de symptômes cognitifs et/ou comportementaux chez ces patients (Pilcz et al., 1906 ; Cullerre et al., 1906 ; Resegotti et al., 1907). Des observations post-mortem ont même montré des lésions des lobes frontaux et temporaux (Tronconi et al., 1941 ; Delay et al., 1959 ; Van Reeth et al., 1961). Ces découvertes furent d'abord interprétées comme l'association de deux pathologies distinctes chez un même patient ou assimilées à des lésions cérébrales extra-motrices non spécifiques.

Malgré ces observations, ce n'est que récemment que les troubles cognitifs et comportementaux s'ajoutèrent au tableau des manifestations cliniques de la Maladie de Charcot (Brooks et al., 2000). Parallèlement à cet élargissement de la définition de la maladie, on commença à parler de DFT associée à une maladie du motoneurone (Massman et al., 1996). Le recoupement entre ces deux pathologies resta sujet à débat jusqu'à l'établissement de critères consensus pour parler de syndromes cognitifs et comportementaux de type DFT dans la SLA (Brooks et al., 2000). Afin de clarifier le statut de ces symptômes, 5 catégories différentes ont été établies (Massman et al., 1996):

1. SLA sans présence de troubles cognitifs et comportementaux
2. SLA présentant certains troubles cognitifs
3. SLA présentant certains troubles comportementaux
4. SLA associée à une DFT

5. SLA associée à une démence autre que la DFT

Les catégories 2 et 3 ont été séparées de la SLA associée à une DFT, car tous les symptômes cognitifs et comportementaux ne sont pas dus à une dégénérescence des lobes frontaux ou temporaux. De nombreux patients SLA ayant ces profils ne remplissent pas les critères d'une démence. Il a été estimé par deux larges études épidémiologiques que 36 à 51% des patients SLA développeront des symptômes cognitifs (Neary et al., 2000 ; Ringholz et al., 2005). Le nombre de malades qui développeront une démence correspondant aux critères de la DFT a été estimé entre 3 et 10 % (Lomen-Hoerth et al. Al., 2003). Une autre étude parle de 52% (Neary et al., 1998).

L'observation complémentaire a également été reportée. En effet, de nombreux patients diagnostiqués avec une DFT ont des lésions au niveau de la corne antérieure de la moelle épinière (Jackson et al., 1996). Des inclusions cytoplasmiques caractéristiques de la SLA ont été retrouvées dans ces motoneurones – quand bien même certains patients ne montrèrent aucun trouble moteur lors de leur vie (Lomen-Hoerth et al., 2002). On a estimé à 15% le nombre de patients atteints de DFT qui remplissent les critères d'une SLA (Lomen-Hoerth et al., 2011). Egalement, 36% des patients DFT montrent des symptômes communs avec la SLA, tels que la fasciculation de muscles ou de la dysphagie, sans pour autant remplir tous les critères suffisants pour leur diagnostiquer cette pathologie (Neary et al., 2005).

Ainsi l'idée d'un continuum entre ces deux maladies s'imposa progressivement. Durant les dix dernières années, les chercheurs ont tenté de clarifier l'entrecouplement clinique de la maladie de Charcot avec les Démences fronto-temporales. Notamment, la fréquence des différentes manifestations cliniques de la DFT au sein des populations SLA

a été étudiée. Les malades atteints de SLA partagent ainsi des symptômes avec le variant comportemental de la DFT ou la forme sémantique et agrammatique de l'APP (Ringholz et al., 2005 ; Neary et al., 2005 ; Rakowicz et al., 1998). Il est à noter, qu'aucun patient SLA n'a été pour l'instant diagnostiqué avec le variant logopénique.

Plusieurs études tendent à montrer que les symptômes cognitifs et comportementaux précèdent souvent ceux moteurs chez les patients SLA/DFT (Guedj et al. 2007 ; Murphy et al., 2007). Ainsi, 58% d'entre eux développent une DFT plus de 3 ans avant l'apparition d'une SLA et 38% développeraient les deux maladies au même moment (Murphy et al., 2007). En moyenne, les premiers symptômes cognitifs et/ou comportementaux apparaîtraient entre 7 et 8 ans avant les symptômes moteurs (Kato et al., 1994). Certains ont tenté d'expliquer ce phénomène en comparant la sévérité des deux pathologies. En effet, si on observe davantage de patients atteints de démence développer une SLA que l'inverse, ce pourrait être parce que les patients SLA ont une espérance de vie plus faible (Del Aguila et al., 2003 ; Chio et al., 2009). D'après cette hypothèse, si les malades SLA vivaient plus longtemps, ils seraient plus nombreux à avoir une pathologie mixte.

Parmi les malades SLA, ceux ayant une forme bulbaire auraient davantage de risques de développer une démence que ceux ayant une forme spinale (Jackson et al., 1996 ; Neary et al., 1998). Cette observation pourrait s'expliquer par la distance plus grande entre les neurones spinaux que bulbaires avec ceux corticaux impliqués dans les démences. Comme ces deux maladies neurodégénératives progressent par contiguïté, il apparaît cohérent que les formes bulbaires soient plus souvent associées à une DFT que les formes spinales.

L'atteinte cérébrale des patients SLA ne se restreint pas aux zones motrices, mais affecte de nombreuses autres régions, dont les lobes frontaux et temporaux (Frank et al.,

1997). La différence entre les patients qui développent des symptômes cognitifs et/ou comportementaux avec ceux ayant des formes purement motrices de SLA s'observe par le degré d'atrophie des lobes frontaux et corticaux (Chang et al., 2005).

II – Les mécanismes pathologiques et les déterminants moléculaires majeurs des

SLA/DFT

A – Comorbidités et facteurs environnementaux dans les SLA/DFT.

La SLA apparait majoritairement comme une maladie sporadique, avec seulement 10% d'héritabilité autosomique dominante. Seulement, la présence de mutations de novo, ainsi que la découverte d'une origine oligogénétique de la pathologie, observée chez des frères jumeaux, ont permis d'estimer que près de 60% des formes sporadiques seraient d'origine génétique (Al-Chalabi et al., 2013).

Toutefois, la répartition géographique des patients forme des clusters, témoignant de la possible existence de facteurs environnementaux qui interviendraient dans la pathogénicité de la SLA. Par exemple, il a été observé que sur l'île de Guam, dans le Pacifique, le nombre de cas de SLA était nettement supérieur à la moyenne, suggérant une influence de l'environnement. La cyanotoxine Beta-méthylamino-L-alanine (BMAA), présente à de fortes doses dans le régime alimentaire des habitants de l'île, a été incriminée (Bradley WG et al., 2009). Egalement, un haut niveau de pesticides a été retrouvé dans le sang de patients SLA au Michigan, suggérant un possible lien avec la maladie (Su et al., 2016). On voit ainsi que des toxines pourraient être à l'origine de certains cas SLA.

D'autres facteurs de risque ont également été constatés : les traumatismes crâniens, la consommation d'alcool, le tabagisme chez les femmes, un IMC faible, une forte activité athlétique, des traumas crâniens répétés, les chocs électriques ou encore l'exposition à des métaux lourds. Nombre de ces facteurs sont communs à d'autres maladies neurodégénératives. Il existe également des comorbidités entre la SLA et

certaines maladies, comme le diabète de type I ou des maladies auto-immunes (Ingre et al., 2015).

Concernant la DFT, on retrouve des facteurs communs avec la SLA comme les maladies auto-immunes ou les traumatismes crâniens. L'obésité pourrait aussi être corrélée avec ces maladies (Rasmussen et al., 2018). Parmi les facteurs à risque, certains sont spécifiques à des formes particulières de démence. Ainsi, il a été observé une prévalence significative de maladies cardiovasculaires parmi les patients atteints du variant comportemental de la DFT, tandis que les maladies auto-immunes seraient plus souvent associées au variant sémantique de l'APP. Plus surprenant, un haut niveau d'éducation est constaté chez les patients atteints de ces démences (Chen et al. 2009).

B – Aspects génétiques majeurs des SLA/DFT

Nombre des formes sporadiques de SLA/DFT sont d'origine génétique. A ce jour, plus de 100 gènes ont été identifiés comme étant impliqués dans ces pathologies, témoignant de leur forte composante héréditaire (Cirulli et al., 2015 ; Ciani et al 2019). La plupart de ces gènes ont une contribution mineure et une faible pénétrance. Pour l'instant, nous ne savons pas si ces gènes, combinés avec des facteurs environnementaux, déclenchent les SLA/DFT ou bien s'ils constituent un modèle polygénique de la maladie. Dans cette partie, nous ne discuterons que des gènes majeurs impliqués dans la SLA et la DFT.

1 – La mutation du gène SOD1 dans la SLA

En 1993, des mutations du gène SOD1, codant pour la Superoxyde dismutase-1,

ont été mises en évidence au sein d'une famille où la SLA se transmettait héréditairement (Rosen et al., 1993). Les mutations dominantes de SOD1 sont retrouvées dans 13% des formes familiales de la SLA et dans 1 % des formes sporadiques. Plus de 170 différentes mutations ont été identifiées au sein de ce gène qui code pourtant pour une protéine de seulement 153 acides aminés. Certaines mutations sont connues pour provoquer des formes rapides et précoces de la maladie comme A4V, tandis que D90A induit une forme lente (Ji et al., 2017).

Il est important de noter que les mutations de ce gène ne sont impliquées que dans la SLA et non dans la DFT.

2 – La mutation du gène MAPT dans les DFT

De même que le gène SOD1 n'est impliqué que dans la SLA, les mutations qui affectent le gène MAPT, codant pour la protéine Tau, conduisent uniquement à des formes pures de DFT. Initialement, ces formes furent appelées DFTL-17 – le gène étant situé sur le chromosome 17 – mais on privilégie maintenant la dénomination DFT-Tau. Ces formes sont le plus souvent associées à un parkinsonisme. Les mutations de ce gène sont autosomiques dominantes. On en a dénombré 44 (Clark et al., 1998 ; Hutton et al., 1998 ; Spillantini et al 1998).

3 – La mutation des gènes TARDBP et Fus dans la SLA/DFT

Plus de 40 mutations du gène TARDBP (TAR DNA Binding Protein), codant pour la protéine TDP-43, ont été retrouvées dans les formes sporadiques (3 %) et familiales (1,5 %) de SLA, ainsi que dans de rares cas de DFT (Lattante et al., 2013). La plupart de ces mutations se trouvent au niveau d'un domaine prion-like – ainsi nommé car la séquence

protéique est similaire à celles des prions de levure.

Les mutations du gène Fus sont retrouvées dans 4% des formes familiales de la SLA et moins de 1% des formes sporadiques (Lattante et al., 2013). Souvent, les mutations de ce gène sont associées à des formes juvéniles (Vance et al., 2009 ; Kwiatkowski et al., 2009).

De même que pour TARDBP, les mutations de Fus sont peu retrouvées chez les patients DFT. Par ailleurs, les troubles cognitifs et comportementaux sont très rarement observés chez des patients atteints de SLA portant des mutations de ce gène.

4 – Plusieurs gènes impliqués dans la protéostasie sont retrouvés mutés dans les SLA/DFT

VCP, UBQLN2 et SQSTM1 sont trois gènes codants pour des protéines qui jouent un rôle régulateur dans l'homéostasie protéique. Ainsi, VCP est une enzyme qui cible des substrats destinés à être dégradés par le système ubiquitine-protéasome (UPS) (Xu et al., 2011) tandis que UBQLN2 code pour l'ubiquilin-2, un autre régulateur essentiel de ce système de dégradation (Ko et al., 2004). SQSTM1 marque des substrats à digérer par l'UPS (Seibenhener et al., 2004) mais permet aussi de diriger les agrégats protéiques vers l'autophagie (Bjørkøy et al., 2006).

Comparativement aux mutations que nous avons vues précédemment, celles affectant la protéostasie sont moins fréquentes dans les SLA/DFT. Le dérèglement de la protéostasie dans ces pathologies n'est pas le mécanisme principal de la dégénérescence neuronale, mais il est un facteur favorisant un phénomène majeur observé dans les SLA/DFT : l'agrégation protéique.

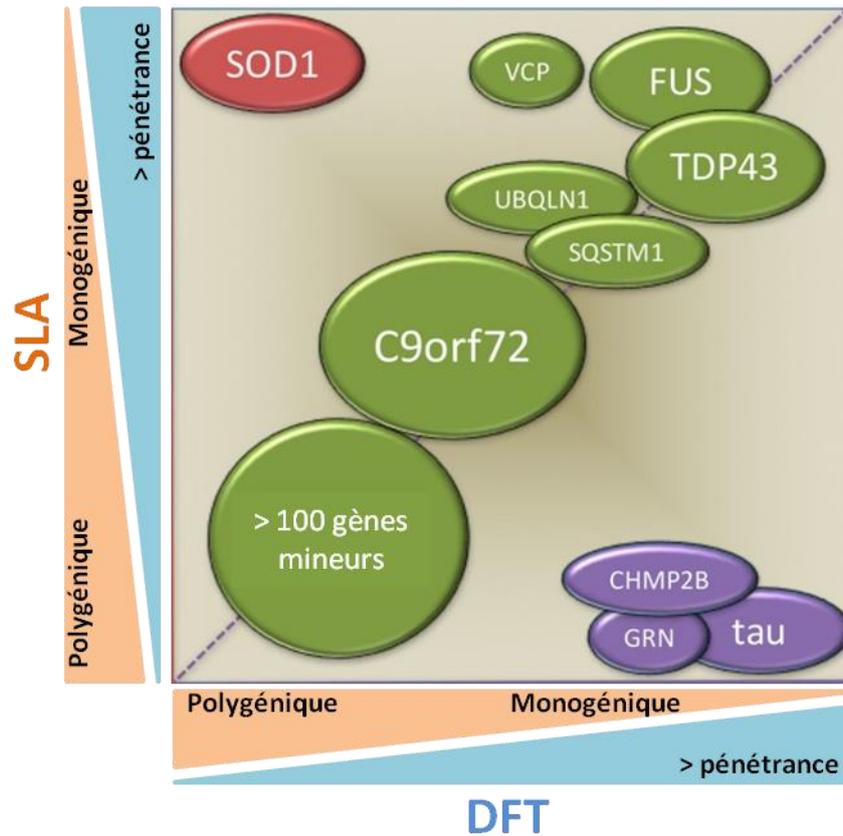


Figure 4 : Contributions génétiques dans la SLA et la DFT. En abscisse et en ordonnée sont respectivement indiqués les phénotypes SLA et DFT. Un gradient (orange) indique si les gènes mutés participent à des pathologies polygéniques ou monogéniques. Un deuxième gradient (bleu) indique la pénétrance des mutations. Adaptée de Ji et al., 2017.

C – Diversité de l'agrégation protéique dans les SLA/DFT

1) L'agrégation de la protéine SOD1

SOD1 est une protéine ubiquitaire qui a principalement un rôle d'antioxydant. Composée de 153 acides aminés, elle fonctionne en homodimère et convertit les radicaux superoxyde en dioxygène et peroxyde d'hydrogène (Sea et al., 2015).

Bien que nous ne connaissions pas le mécanisme par lequel chacune des

centaines de mutations affectant le gène SOD1 engendrent une pathologie SLA, il semblerait que cela ne soit pas dû à une perte de fonction du gène. En effet, la délétion totale de SOD1 n'engendre aucun symptôme moteur chez la souris, alors qu'on les observe chez des animaux transgéniques possédant des mutations du gène retrouvées chez des malades (Reaume et al., 1996 ; Gurney et al., 1994 ; Tu et al., 1996). Il apparaît que cette toxicité par gain-de-fonction soit due à l'agrégation de la protéine SOD1, observable dans le cerveau des patients SLA/DFT. Ces agrégats de SOD1 se retrouvent essentiellement au niveau de la moelle épinière qui est l'une des deux régions principalement lésées dans la SLA (Johnston et al., 2000 ; Jonsson et al., 2004 ; Zetterström et al 2007). Ces observations suggèrent que l'assemblage de la protéine SOD1 en agrégats serait responsable de la destruction des neurones moteurs.

Seulement, ce mécanisme ne suffit pas à expliquer toutes les formes de la SLA, les agrégats de SOD1 n'étant retrouvés que chez 2% des patients (Fig. 5).

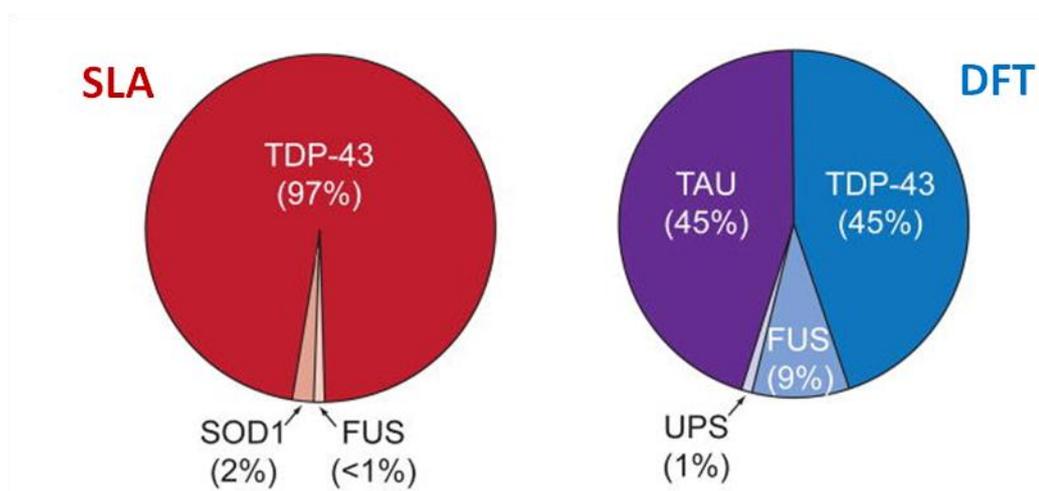


Figure 5 : Proportions des protéines formant des inclusions insolubles dans la SLA et la DFT (Ling et al., 2013).

2) L'agrégation de la protéine Tau

Tau est une protéine exprimée essentiellement au niveau des axones des neurones et, dans des proportions nettement plus faibles, dans les astrocytes et les oligodendrocytes (Schoenfeld et al., 1994 ; M Chin et al., 2005). Son gène comporte 16 exons et permet, par épissage alternatif, la production de 6 isoformes différents : 0N3R, 0N4R, 1N3R, 1N4R, 2N3R et 2N4R (Himmler et al., 1989 ; Andreadis et al., 1995).

Deux types d'épissage ont lieu lors de la maturation des ARNm. Un épissage au niveau amino-terminal produira des protéines de longueurs différentes : 0N, 1N (+29 acides aminés) ou 2N (+ 58). De même, une région, comportant des séquences de liaison aux microtubules, est aussi la cible d'un épissage conduisant à la production de protéines avec 3 ou 4 domaines répétés : 3R et 4R. Chez un individu sain, le ratio protéique 3R:4R est de 1:1 (Goedert et al., 1991). En se liant aux microtubules par ces séquences répétées, la protéine Tau assure l'assemblage et la stabilité du cytosquelette (Cleaveland et al., 1977 ; Harada et al., 1994). Cette fonction est régulée par de nombreux sites de phosphorylation (Billingsley et al., 1997).

Les mutations affectant le gène MAPT peuvent avoir un impact sur l'épissage des domaines répétés ou sur l'état de phosphorylation de la protéine. Les mutations favorisant la forme 4R entraînent l'agrégation de la protéine. Des mutations favorisant l'état de phosphorylation auront le même effet (Clark et al., 1998 ; Hutton et al., 1998 ; Spillantini et al 1998).

L'agrégation de cette protéine dans le système nerveux est responsable de nombreuses pathologies regroupées sous le nom de Tauopathies. Parmi celles-ci, on retrouve la maladie d'Alzheimer mais aussi près de la moitié des DFT (Fig. 5). Ces pathologies sont subdivisées en fonction du ratio 3R:4R des corps d'inclusions.

La DCB ainsi que la PSP sont des pathologies où l'isoforme 4R est dominant. Pour

les DFT où les protéines 3R sont les plus abondantes, on parle de Maladie de Pick, bien que cette dénomination regroupe des pathologies aux manifestations cliniques diverses (Irwin et al., 2015).

3) L'agrégation de la protéine TDP-43

L'étude histologique de cerveaux de patients atteints de DFT a permis de mettre en lumière la présence d'inclusions cytoplasmiques négatives pour Tau et positives pour la protéine TDP-43 (Arai et al., 2006). Contrairement aux agrégats de SOD1 ou de Tau, ceux de TDP-43 se retrouvent tant chez des patients SLA que DFT. Cette observation renforça l'idée d'un continuum entre ces pathologies.

Dans la quasi-totalité des cas de SLA (97%), on retrouve TDP-43 sous forme agrégée. Dans la DFT, les assemblages de cette protéine sont aussi courants que ceux de Tau (45%) (Fig.5). Ils sont présents dans les formes familiales et sporadiques des deux maladies.

TDP-43 est une protéine ubiquitaire de 414 acides aminés. Elle est majoritairement retrouvée au niveau du noyau, mais peut se délocaliser au niveau du cytoplasme pour assurer de nombreuses fonctions essentielles au métabolisme des ARN : l'épissage, le transport et la traduction d'ARNm, la synthèse de micro-ARN et la formation de granules de stress en interagissant avec des complexes de la classe hnRNPs (heterogenous nuclear ribonucleoproteins) ainsi que d'autres facteurs (Lagier-Tourenne et al., 2010 ; Buratti et al., 2012). Cette protéine possède un site de localisation nucléaire (NLS) dans sa partie N-terminale, deux sites de reconnaissance à l'ARN, un domaine d'export nucléaire et, en C-terminal, un domaine prion-like (Lattante et al. 2013).

La majorité des mutations se situe dans la région C-terminale et favorise

l'agrégation de TDP-43. Il a été démontré que la délocalisation de la protéine depuis le noyau jusqu'au cytoplasme induisait aussi son agrégation (Winton et al., 2008). A90V, une mutation située au niveau du site de localisation nucléaire, est la seule mutation de TDP-43 qui induise son accumulation cytoplasmique (Barmada et al., 2010). Cependant, la plupart des patients ayant une pathologie TDP-43 n'ont pas de mutation directe sur le gène TARDBP. Par ailleurs, ces formes sont indiscernables de celles où le gène est muté (Pamphlett et al., 2009). Les raisons de la déplétion nucléaire de TDP-43 restent mal comprises.

4) L'agrégation de la protéine FUS/TLS

FUS (Fused in Sarcoma), aussi appelé TLS (Translocated in Sarcoma), possède de nombreux aspects fonctionnels et pathologiques communs avec TDP-43. En effet, cette protéine nucléaire qui appartient à un complexe de la classe hnRNPs (Calvio et al., 1995) est impliquée dans la transcription, l'épissage et le transport d'ARN (Lagier-Tourenne et al., 2012 ; Colombrita et al., 2012). On retrouve également dans son domaine C-terminal la plupart des mutations. Cette séquence contient un site de localisation nucléaire qui, lorsqu'il est muté, conduit à une translocation cytoplasmique de FUS et à son agrégation (Dormann et al., 2010). D'autres mutations, situées au niveau d'un domaine prion-like entraînent aussi l'agrégation de la protéine (Vance et al., 2009 ; King et al., 2012). Plusieurs de ces mutants sont retrouvés chez des patients atteints de formes juvéniles de la SLA (Baümer et al., 2010 ; Belzil et al., 2012 ; Yamashita et al., 2012).

D – Les SLA/DFT, des maladies de type prion et du dérèglement des ARN

1) L'agrégation protéique dans les pathologies neurodégénératives et le concept de maladies de type prion

L'agrégation protéique est indubitablement un mécanisme biochimique majeur ayant lieu dans la SLA/DFT. Ce phénomène est commun à plusieurs maladies neurodégénératives, parmi lesquelles on trouve les maladies de Creutzfeld-Jakob, Parkinson, Alzheimer ou encore Huntington (Renner et Melki, 2014). Dans ces pathologies, certaines protéines forment des agrégats qui constituent des plaques extracellulaires ou des inclusions cytoplasmiques dans le système nerveux (Tab. 1). Cette agrégation est un phénomène délétère et rendrait compte de la dégénérescence neuronale dans les pathologies évoquées.

Pathologie	Protéine agrégée	Localisation des dépôts
Maladies à prions	PrP	Extracellulaire
Synucléinopathies (ex : Maladie de Parkinson)	α -synucléine	Intracellulaire
Maladie d'Alzheimer	peptide amyloïde A β	Extracellulaire
Tauopathies	Tau	Intracellulaire
Maladie d'Huntington	Htt+PolyQ	Intracellulaire

Tableau 1 : Les principales protéinopathies, les protéines agrégées et la localisation de leurs dépôts. Adapté de Renner et Melki, 2014.

Dans un contexte non-pathologique, les protéines adoptent spontanément la structure essentielle à leur fonction ou par interaction avec d'autres partenaires cellulaires. Ceci dit, une protéine peut adopter de multiples conformations (Fersht, 1999). Leur état de stabilité dépend de plusieurs facteurs. Les contacts intra- et inter-moléculaires ainsi que

les conditions physico-chimiques, telles que la température, la force ionique ou encore l'acidité du milieu, sont autant de paramètres décidant de la stabilité et des équilibres entre les conformations protéiques. Ces états sont dynamiques et réversibles.

Lorsque des mal-conformations protéiques persistent, des protéines appelées chaperons moléculaires les prennent en charge (Hartl et al., 2011). Leur rôle consiste à replier correctement la protéine mais elles peuvent également agir dès la synthèse protéique. Si ce système n'est pas en mesure de corriger la structure protéique, un système de dégradation prend le relais, le système UPS (Ubiquitin-Proteasome System) (Goldberg, 2003). Ce système fait intervenir différents complexes de protéines. En premier lieu, des ubiquitines ligases vont marquer la protéine à dégrader en ajoutant covalamment des chaînes de poly-ubiquitine aux lysines. Ces étiquettes sont des signaux reconnus par la cellule pour orienter le substrat marqué vers un complexe enzymatique appelé protéasome. L'activité catalytique de ce complexe dégrade la protéine dont les acides aminés sont ensuite recyclés (Glickman et Ciechanover, 2002).

Malgré ces systèmes prévenant l'intégrité des structures protéiques, il arrive que la cellule n'arrive pas à prendre en charge certains polypeptides mal-repliés. Dans des contextes pathologiques, des protéines ayant la même séquence et la même conformation peuvent s'assembler en formant des feuilletts β entre chaque monomère (Fig. 6). On appelle ces assemblages très stables des agrégats amyloïdes (Barrow et al., 1992 ; Pan et al., 1993 ; Conway et al., 2000 ; Chen et al., 2002). Leur morphologie est le plus souvent fibrillaire. Ces assemblages s'amplifient en incorporant d'autres monomères à leurs extrémités. Plus les agrégats sont de taille importante, plus ils sont stables et résistants. (Brundin et al., 2011). Ils peuvent se fragmenter en de plus petits assemblages qui se propagent et s'amplifient.

Ces amyloïdes sont considérés comme toxiques soit parce qu'ils entraînent la perte de fonction de la protéine séquestrée dans l'agrégat, soit par des gains de fonctions inhérents à leur nouvelle structure. Si les assemblages sont extracellulaires, ils peuvent se lier à la surface des cellules. Cette liaison s'accompagne de changements délétères pour la cellule comme la perméabilisation de la membrane, la réorganisation de ses constituants lipidiques et protéiques par regroupement autour des agrégats, mais aussi la déformation de la cellule (Kayed et al., 2004 ; Smith et al., 2009 ; Shrivastava et al., 2015 ; Shrivastava et al., 2019). Une fois internalisés, les assemblages ont montré qu'ils échappaient aux compartiments endosomal-lysosomal. En effet, les fibres sont capables de recruter la protéine soluble de la cellule (Ren et al., 2009 ; Bousset et al., 2013 ; Frost et al., 2009). La rupture du lysosome induit par le passage des agrégats peut libérer au passage des composants lysosomaux dangereux pour la cellule, comme la Cathepsin B qui produit des dérivés réactifs de l'oxygène (Freeman et al., 2013). Egalement, les assemblages présents dans le cytosol peuvent perturber l'ensemble de l'homéostasie protéique en saturant les systèmes de chaperons moléculaires et UPS.

Aujourd'hui ces assemblages sont largement considérés comme responsables de la destruction des cellules nerveuses dans les maladies neurodégénératives. On retrouve ainsi l' α -synucléine présente sous forme fibrillaire dans la maladie de Parkinson, mais aussi le peptide A β et la protéine Tau dans la maladie d'Alzheimer (Tab. 1).

Un débat persiste quant à savoir si les dépôts composés de ces assemblages sont des espèces toxiques, ou bien si elles constituent un moyen de défense de la cellule contre des oligomères de plus petites tailles.

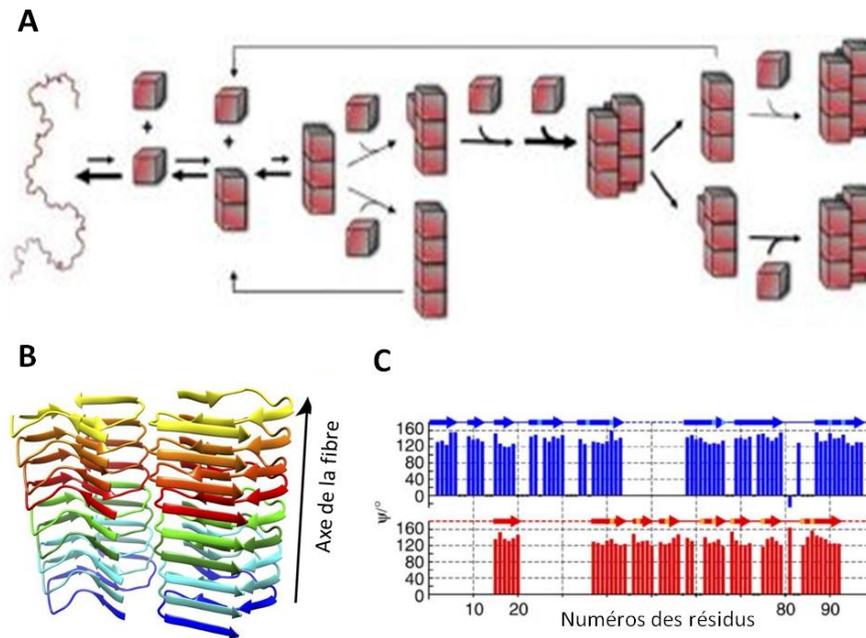


Figure 6 : A : Schéma de l'agrégation amyloïde d'une protéine. Cette protéine est représentée dans un état intermédiaire, non structurée, à gauche. Elle peut acquies parmi ses différentes conformations possibles, une capable de s'agréger. Si la protéine en rencontre d'autres de même séquence et conformation, elles s'assemblent par l'intermédiaire d'interactions inter-moléculaires de type feuillet β , formant ainsi des oligomères de nature amyloïde. Des monomères peuvent s'en détacher, mais plus les oligomères sont allongés, plus les assemblages deviennent stables. La stabilité de ces agrégats fibrillaires est d'autant plus forte que des liaisons latérales se forment entre eux. Les assemblages de longues tailles peuvent également se fragmenter, créant ainsi de plus petites fibres pouvant aussi s'allonger et ainsi propager l'amyloïde dans l'organisme. Adaptée de Melki, 2018. B : Représentation 3D basée sur l'analyse par RMN du solide d'assemblages formés par le peptide $A\beta(1-42)$ qu'on retrouve sous forme agrégée dans la maladie d'Alzheimer. Dans cette structure, deux molécules sont présentes à chaque niveau de la fibre. Les brins β sont représentés par des flèches. Adapté de Wälti et al., 2016. C : Différences structurales entre deux souches d' α -synucléine : les fibres (rouge) et les rubans (bleu). Ces modèles ont été obtenus à partir d'analyse en RMN du solide des assemblages d' α -synucléine. Les angles Ψ de chaque résidu ont été déterminés par le logiciel de prédiction TALOS. Les feuillets β sont représentés par des flèches au-dessus du graphique. Adaptée de Bousset et al., 2013.

Les propriétés de propagation et de toxicité des agrégats protéiques dans ces pathologies neurodégénératives ont permis de les classer comme des maladies de type prion (Brundin et al., 2011). Le concept du prion fut initialement introduit pour désigner le comportement infectieux de la protéine PrP dans la maladie de la tremblante du mouton puis dans d'autres pathologies associées, parmi lesquelles la maladie de Creutzfeld-Jakob, du Kuru, l'encéphalopathie spongiforme bovine ou encore l'encéphalopathie des cervidés (Prusiner, 1982 ; Bolton et al., 1982 ; Prusiner 1998). Depuis ce concept s'est élargi pour désigner d'autres maladies neurodégénératives régies par les mêmes mécanismes, mais impliquant d'autres protéines. Il faut toutefois préciser que la protéine PrP peut se transmettre d'un organisme à un autre – parfois même entre deux espèces différentes. Cette propriété infectieuse est absente dans le cas des autres protéines de type prion. Leur propagation se limite aux cellules d'un même organisme.

Il est à noter qu'une même protéine peut former des assemblages de différentes conformations (Fig. 6C). On parle alors de souches (Melki, 2015). Ces souches n'ont pas les mêmes propriétés biochimiques et cellulaires. En effet, les résidus présents à la surface de ces assemblages diffèrent. La résistance à la protéolyse et les partenaires cellulaires sont différents d'une souche à l'autre, bien que celles-ci soient obtenues à partir de la même protéine. Cela pourrait expliquer pourquoi une seule protéine, comme l' α -synucléine, en formant des agrégats de différentes conformations, est impliquée dans des maladies distinctes, les synucléinopathies : la Maladie de Parkinson, l'Atrophie Multisystématisée et la Démence à Corps de Lewy (Peelarets et al., 2015).

2) Les SLA/DFT sont des protéinopathies multiples appartenant à la famille des maladies de type prion

Qu'il s'agisse de Tau (Kirschner et al., 1989), SOD1 (Elam et al., 2003), ou TDP-43

(Shimonaka et al., 2016), tous peuvent former des agrégats de nature amyloïde. Concernant Fus, seul un mutant retrouvé en clinique a été montré comme appartenant à cette famille d'assemblage, mais ce n'est pas le cas des agrégats obtenus à partir de la protéine sauvage (Nomura et al., 2014).

Ces assemblages se comportent-ils à la manière de prions?

L'ajout de fibres de Tau à des cultures cellulaires a montré que ces agrégats recrutent la protéine monomérique et se propageaient de cellule en cellule, témoignant de propriétés de type prion (Guo et al., 2011, Frost et al., 2009 ; Kfoury et al., 2012). Ces observations ont été validées *in vivo* par injection de fibres de Tau dans le cerveau de souris transgéniques exprimant la forme humaine de la protéine (Clavaguera et al., 2009).

Les mutants de la protéine SOD1 retrouvés chez des patients atteints de SLA forment des agrégats fibrillaires amyloïdes *in vitro* (Chattopadhyay et al., 2008). Ces fibres ont également montré leur capacité d'amplification et de propagation *in cellulo* (Münch et al., 2011) et dans des modèles murins transgéniques (Chia et al., 2010 ; Deng et al., 2006).

Contrairement à Tau qu'on ne trouve agrégée que dans la DFT et SOD1 dans la SLA, TDP-43 est présente sous forme insoluble dans les deux pathologies. Les propriétés de type prion de TDP-43 ont été démontrées dans des cultures cellulaires (Furukawa et al., 2011). En particulier, l'une de ces études a montré que des agrégats purifiés à partir de cerveaux de patients décédés de SLA/DFT reproduisaient ces propriétés (Nonaka et al., 2013).

Des modèles cellulaires ont pu démontrer l'agrégation de la protéine FUS en conditions pathologiques, mais pour l'instant la capacité de propagation de ces assemblages n'a pas été démontrée (Kryndushkin et al., 2011 ; Nomura et al., 2014).

Rappelons enfin, que TDP-43 et FUS ont toutes deux un domaine prion-like dans

leur séquence.

A l'exception de FUS dont le statut reste incertain, les protéines que l'on trouve agrégées dans le cerveau des patients SLA/DFT ont des propriétés de type prion. Malgré la diversité des acteurs mis en jeu, ce phénomène est un mécanisme central pour comprendre l'étiologie de ces pathologies complexes.

3) La dérégulation du métabolisme des ARN dans les SLA/DFT

La forte implication des protéines TDP-43 et FUS dans les SLA/DFT suggère une hypothèse supplémentaire à celle de maladie de type prion. En effet, chez presque tous les patients atteints de SLA (98%) ainsi que dans plus de la moitié des cas de DFT, on retrouve l'une de ces deux protéines agrégée (Fig. 5). Et, comme nous l'avons vu, les deux protéines possèdent de nombreuses analogies structurales et fonctionnelles. Les fonctions majeures qu'occupent ces deux protéines dans le métabolisme des ARN indiquent que la dérégulation de cette voie pourrait être un des mécanismes pathogènes à l'œuvre dans les SLA/DFT. En effet, FUS cible à peu près de 5,500 ARN dans le cerveau humain tandis que TDP-43 en a plus de 6,000 (Lagier-Tourenne et al 2012 ; Sephton et al., 2011). La diminution des taux de TDP-43 et de FUS affecte chacun plus de 600 ARNm (Lagier-Tourenne et al 2012 ; Rogelj et al., 2012). La perte de leur fonction a donc un impact majeur dans l'homéostasie des ARN.

Appuyant cette hypothèse, une étude a montré que l'agrégation de TDP-43 induisait un appauvrissement en ARN (Pesiridis et al., 2011). De plus, une perte de fonction du gène dans un modèle murin engendre une dégénérescence neuronale et des symptômes moteurs typiques de la SLA (Yang et al., 2014). Deux études *in vivo* similaires ont été réalisées sur le gène *Fus*. Cependant sa délétion complète dans des souris n'a

pas pu établir un lien clair entre la perte de fonction et un phénotype rappelant les SLA/DFT (Kuroda et al., 2000 ; Hicks et al., 2000).

Enfin, de nombreuses mutations mineures affectant des gènes ayant des rôles similaires à TARDBP et Fus ont été retrouvées chez des patients SLA/DFT. C'est le cas des gènes codant pour TAF15 (TATA-binding protein-associated factor 15), EWSR1 (Ewing's sarcoma breakpoint region 1), l'angiogenin, la senataxin et l'ataxin-2 (Couthouis et al., 2011 ; Couthouis et al., 2012 ; Greenway et al., 2006 ; Chen et al., 2004 ; Elden et al., 2010).

L'ensemble de ces données appuie l'idée que deux mécanismes de toxicité sont à l'œuvre dans les SLA/DFT : l'agrégation de protéines de type prion et la dérégulation de l'homéostasie des ARN.

III – La présence de répétitions de l'hexanucleotide GGGGCC au niveau du gène

C9orf72 et la cause génétique principale des SLA/DFT.

A – Le gène C9orf72 et sa mutation.

En 2011, deux équipes différentes ont mis à jour la mutation autosomale dominante la plus communément retrouvée dans les formes familiales de SLA/DFT (DeJesus-Hernandez et al., 2011 ; Renton et al., 2011). Cette mutation consiste en la répétition d'un hexanucléotide GGGGCC situé au niveau du promoteur ou du premier intron du gène C9orf72, appelé d'après son locus (Chromosome 9 open reading frame 72). Le nombre de répétitions peut aller jusqu'à des milliers. Il est à noter qu'on retrouve des répétitions de cet hexanucleotide d'une longueur inférieure à 30 chez des sujets sains, signifiant que le nombre des répétitions est déterminant pour qu'une pathologie se déclare. Par ailleurs, dans le cadre des DFT, une étude a montré que plus le nombre de répétitions était grand, plus la maladie se déclarait tôt (Gijssels et al., 2016). Une autre étude a également suggéré que la longueur de ces extensions est positivement corrélée avec la progression de la maladie (Nordin et al., 2014).

Cette mutation se retrouve respectivement dans ~20 et ~16% des formes familiales de SLA et de DFT. Selon les études, on la retrouve également dans 6 à 8% des formes sporadiques des deux pathologies (DeJesus-Hernandez et al., 2011 ; Renton et al., 2011 ; Majouine et al., 2012 ; Smith et al., 2013). On constate aussi une grande variabilité géographique de la fréquence de cette mutation dans les SLA (Smith et al., 2013).

La fonction du gène C9orf72 est encore peu comprise. Chez l'homme, ce gène donne lieu à trois transcrits issus d'épissages alternatifs qui sont par la suite traduits en deux isoformes: un long de 481 acides aminés et un court de 221. La séquence de ces

protéines possède une haute similitude avec les DENN (Differentially Expressed in Normal and Neoplasia) (Levine et al., 2013). Cette famille de protéines appartient aux facteurs d'échange de nucléotide guanine, ou GEF (Guanine Exchange Factors), suggérant que C9orf72 aurait la même fonction (Marat et al 2011). Depuis, l'appartenance de C9orf72 à ce groupe de protéines a été confirmée (Farg et al., 2014 ; Iyer et al., 2018). De récents travaux montrent que C9orf72 serait ainsi impliqué dans différents processus cellulaires tel que la régulation du trafic endosomal (Aoki et al., 2017 ; Iyer et al., 2018), l'activité lysosomale (Armick et al., 2016) ou encore l'autophagie (Sellier et al., 2016 ; Ho et al., 2019).

Les rôles spécifiques des isoformes longs et courts de C9orf72 restent mal compris. Des travaux montrent cependant que leurs localisations subcellulaires diffèrent. Ainsi, l'isoforme long est présent de façon diffuse dans le cytoplasme, tandis que l'isoforme court semble agir au niveau des pores nucléaires (Xiao et al., 2015).

B – Trois hypothèses sur la pathogénicité de la mutation C9orf72

Un nombre grandissant d'études tentent d'expliquer comment la mutation du gène C9orf72 rendrait compte des maladies SLA/DFT. A ce jour, trois hypothèses ont été avancées (Fig. 7) :

- 1) La perte de fonction du gène C9orf72 induit par la présence des répétitions de l'hexanucléotide serait toxique pour la cellule.
- 2) Les ARN transcrits à partir de ces séquences répétées seraient délétères en séquestrant des protéines nucléaires.

- 3) Les ARN seraient traduits en Protéines à Répétitions de Dipeptides (PRD), toxiques pour les cellules, présentes dans une forme agrégée.

1) La perte de fonction du gène C9orf72

Une forte diminution des transcrits C9orf72 a été remarquée chez les patients porteurs de répétitions GGGGCC (Renton et al., 2011), ainsi qu'une diminution des quantités de la protéine aux niveaux frontal et temporal (Gijssels et al., 2016). Ces observations pourraient suggérer que la mutation serait délétère en abaissant le niveau d'expression du gène C9orf72. Une étude a montré que c'est l'expression de l'isoforme long qui est diminué dans le cerveau des patients (Xiao et al., 2015). Deux mécanismes ont été proposés pour expliquer cette diminution : l'hyperméthylation de l'ADN induite par les hexanucléotides répétés et la formation de structures de type G-quadruplex et boucle-R de ces séquences.

Les répétitions d'hexanucléotides engendrent une hyperméthylation des îlots CpG situés au niveau du promoteur du gène C9orf72 (Xi et al., 2013 ; Xi et al., 2014). Cette régulation épigénétique pourrait rendre compte de l'haplo-insuffisance du gène. De plus, l'état de méthylation est proportionnel à la longueur des répétitions (Gijssels et al., 2016), ce qui pourrait expliquer l'importance de cette dernière dans l'apparition de la maladie.

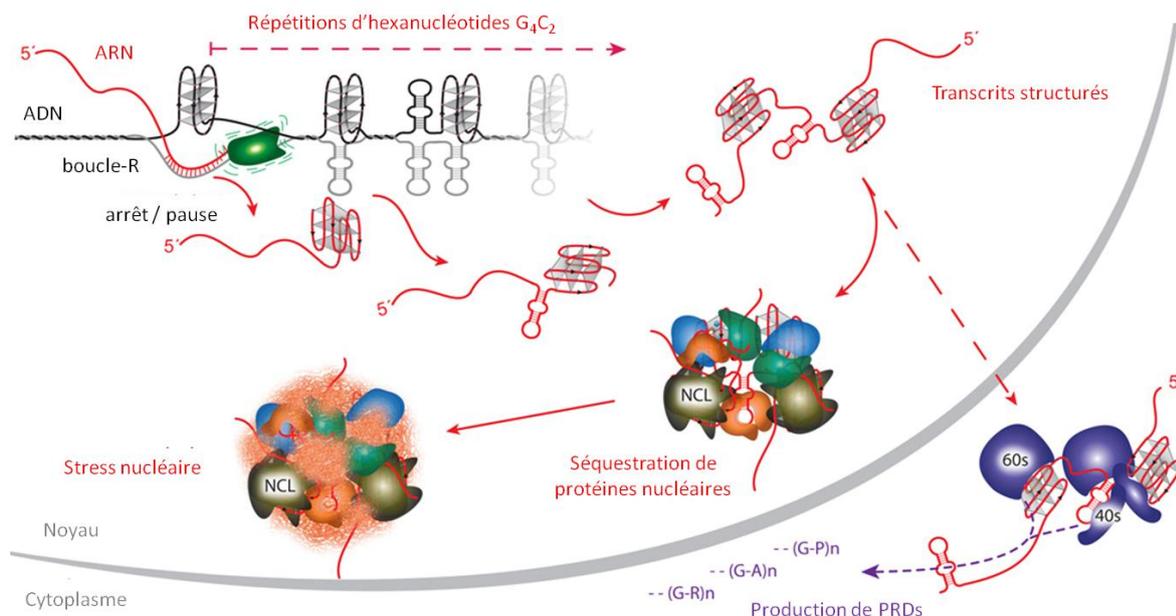


Figure 7 : Conséquences de la présence de répétitions de l'hexanucléotide GGGGCC au niveau du gène C9orf72 et hypothèses de toxicité. Les séquences d'ADN répétées forment des structures en épingle et en G-quadruplex qui interrompent la transcription du gène C9orf72, conduisant à son haplo-insuffisance. Les transcrits avortés peuvent s'hybrider avec les répétitions GGGGCC d'ADN pour former des boucles-R qui inhibent également la transcription de C9orf72. Ces ARN forment aussi des structures en épingle et en G-quadruplex. Ces transcrits structurés vont recruter et séquestrer des protéines nucléaires, engendrant ainsi de nombreuses défaillances de transcription et de traduction. Certains ARN pourront également être transloqués jusqu'au cytoplasme où ils seront traduits par un mécanisme indépendant d'un codon d'initiation en protéines à répétitions de dipeptides, ces polypeptides sont toxiques pour la cellule. Adaptée de Haeusler et al., 2014.

Cela dit, une autre hypothèse a été avancée pour expliquer la diminution d'expression de C9orf72. Les répétitions d'hexanucléotides forment des structures d'ADN en épingle ou en G-quadruplex qui sont capables d'interrompre la transcription (Haeusler et al., 2014) (Fig. 7). Les G-quadruplex se présentent suivant deux topologies différentes : parallèle et anti-parallèle (Zhou et al., 2018). A cela s'ajoute que les brins d'ADN et d'ARN contenant les répétitions sont capables de s'hybrider pour former des boucles-R. Celles-ci

ont des répercussions sur la transcription des séquences voisines (Haeusler et al., 2014 ; Walker et al., 2017).

L'invalidation du gène C9orf72 dans des modèles de poisson zèbre et *C. elegans* a reproduit des symptômes moteurs de type SLA (Ciura et al 2013 ; Therrien et al., 2013). Ces résultats semblent appuyer l'hypothèse selon laquelle l'haplo-insuffisance de C9orf72 serait un élément déterminant dans l'étiologie des pathologies SLA/DFT. Seulement la version humaine du gène a une homologie faible avec celles du poisson zèbre ou de *C. elegans*. Elle est bien plus proche de la version murine. Plusieurs modèles de souris invalidées pour le gène, ou diminuées dans son expression, n'ont pourtant montré aucun des symptômes réminiscents de la SLA/DFT (Koppers et al., 2015 ; Atanasio et al., 2016 ; Sudria-Lopez et al., 2016). Seuls des troubles immunitaires ont pu être constatés.

Il semble que d'autres déterminants moléculaires soient mis en jeu pour expliquer la pathogénicité de la mutation C9orf72.

2) La toxicité des ARN sens et antisens

Des ARN issus de la transcription sens et antisens des hexanucléotides répétés ont été détectés dans le cerveau de patients SLA/DFT (Mori et al., 2013a). On a retrouvé des agrégats nucléaires, ou foci, de ces ARN chez des patients porteurs de la mutation C9orf72 (DeJesus-Hernandez et al., 2011). Ces agrégats sont à la fois composés de transcrits sens et antisens (Mizielinska et al., 2013 ; Lee et al., 2013 ; Gendron et al., 2013). La présence de ce matériel agrégé dans les neurones est toxique. En effet, de nombreuses protéines nucléaires sont piégées dans ces foci d'ARN. La toxicité résultante dépend de la nature des protéines séquestrées par les transcrits agrégés. Par exemple, la nucléoline qui est une protéine du nucléole essentielle à la synthèse et à la maturation des

ribosomes est séquestrée dans ces foci d'ARN (Haeusler et al., 2014). S'ensuit une dérégulation globale de la machinerie transcriptionnelle. On observe ainsi une accumulation d'ARNm non-traduits. Egalement, les protéines SRSF1 et SRSF2 (Sérine-rich-splicing factor), impliquées dans l'épissage d'ARNm, sont aussi piégées dans ces agrégats et n'assurent plus leurs fonctions (Cooper-Knock et al., 2014 ; Hautbergue et al., 2017). D'autres protéines sont piégées par ces foci, tel ALYREF, impliquée dans le transport jusqu'au cytoplasme d'ARNm mature, ou les hnRNP (heterogenous nuclear ribonucleoproteins), qui ont pour principal rôle la rétention dans le noyau des ARNm non-épissés (Lee et al., 2013 ; Cooper-Knock et al., 2014 ; Haeusler et al., 2014).

Bien que ces agrégats d'hexanucléotides aient un impact majeur sur la machinerie transcriptionnelle et traductionnelle des cellules, leur implication dans les SLA/DFT reste encore mal comprise. Une étude a pu corrélérer l'accumulation de ces foci dans le cortex de patients DFT avec une apparition précoce de la maladie (Mizielinska et al., 2013), mais aucun lien n'a pu être établi avec la dégénérescence neuronale. Ainsi, plusieurs modèles de drosophiles reproduisant uniquement la composante ARN des patients ont été réalisés, mais ces travaux n'ont montré aucune neurodégénérescence, ni symptôme de type SLA/DFT (Mizielinska et al., 2014 ; Moens et al., 2018). Seuls les modèles de drosophiles (Mizielinska et al., 2014 ; Solomon et al., 2018) et de souris (Jiang et al., 2016 ; Liu et al., 2016 ; Herranz-Martin et al., 2017) qui présentent la composante ARN et qui expriment aussi les protéines PRD reproduisent la dégénérescence des neurones et des symptômes de type SLA/DFT. Ces souris développent des troubles moteurs, cognitifs et comportementaux tels que des déficits de mémoire ou de l'anxiété.

Les PRD semblent donc avoir un rôle plus déterminant que les foci d'ARN dans les pathologies liées à la mutation de C9orf72.

3) Les Protéines à Répétitions de Dipeptides

Les transcrits d'hexanucléotides sont capables de traverser le noyau pour se retrouver dans le cytoplasme. Il a été montré que la liaison des ARN à la protéine nucléaire SRSF1 pouvait permettre cette translocation (Hautbergue et al., 2017). Par ailleurs, l'inhibition de cette liaison induit une diminution de la neurodégénérescence et des symptômes moteurs dans des modèles de drosophiles et d'iPSCs (induced pluripotent stem cells) issues de patients. L'inhibition de cette voie pourrait donc constituer une potentielle cible thérapeutique.

Une fois présents dans le cytoplasme, les transcrits d'hexanucléotides sont traduits suivant un mécanisme indépendant d'un codon d'initiation ATG appelé RANT (Repeat associated non-ATG translation) (Mori et al., 2013a ; Ash et al., 2013 ; Zu et al., 2013 ; Gendron et al., 2013). Ce mécanisme de traduction rare a lieu en la présence de séquences microsatellites répétées formant des structures en épingles (Zu et al., 2011). Il a d'abord été mis à jour dans d'autres pathologies, comme l'ataxie cérébrospinale de type 8 ou la dystrophie myotonique de type 1 (Zu et al., 2011).

La traduction des ARN sens et antisens d'hexanucléotides étant indépendante de codon d'initiation, elle s'effectue dans les 3 cadres de lecture possibles (Fig. 8). Cette traduction permet la production de protéines de séquences différentes mais étant toutes composées d'une répétition de deux acides aminés, d'où leur nom : Protéines à Répétitions de Dipeptides (PRD). Elles sont au nombre de 5. La traduction des ARN sens produit les protéines : poly-GA (Glycine-Alanine), poly-GP (Glycine-Proline) et poly-GR (Glycine-Arginine) (Mori et al., 2013a). Celle des ARN antisens produit également des poly-GP, mais aussi des poly-PA (Proline-Alanine) et poly-PR (Proline-Arginine) (Mori et al., 2013b ; Zu et al., 2013 ; Gendron et al., 2013) (Fig. 8).

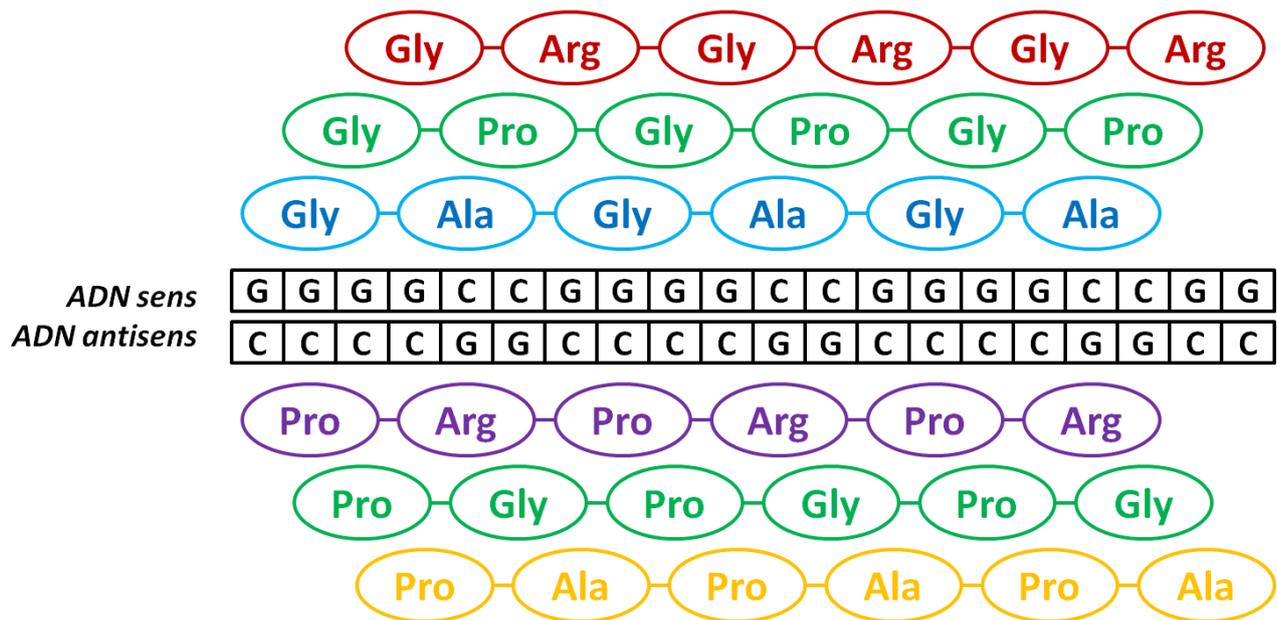


Figure 8 : Les brins sens et antisens d'hexanucléotides sont traduits pour chaque cadre de lecture ouvert. Les lectures du brin sens donnent les PRD : poly-GA, poly-GP et poly-PR, tandis que les lectures du brin antisens donnent ; poly-PR, poly-GP et poly-PA.

A l'aide d'anticorps, les PRD ont été détectés sous forme d'inclusions cytoplasmiques dans le cerveau de patients portant la mutation du gène C9orf72 (Mori et al., 2013a ; Mori et al., 2013b ; Ash et al., 2013 ; Zu et al., 2013 ; Gendron et al., 2013). Ces agrégats sont présents dans les neurones, mais absents des astrocytes. Il est à noter que des anticorps dirigés contre la partie C-terminale des PRD, qui suit les séquences répétées, marquent une portion des agrégats (Gendron et al., 2013). En l'absence d'un codon stop, la traduction se poursuit donc, pour une partie des transcrits, au-delà des répétitions d'hexanucléotides.

Les PRD sont retrouvés dans de nombreuses structures cérébrales, mais dans des proportions variables. Les régions qui sont le plus affectées sont : le cortex frontal, temporal, moteur, le cervelet, l'hippocampe, l'amygdale et le thalamus (Ash et al., 2013). On retrouve également ces protéines en plus faible quantité au niveau du cortex

entorhinal, du tronc cérébral, du putamen, de la medulla et de la moelle épinière. Il est intéressant de remarquer que nombre de ces régions sont lésées dans les SLA/DFT. Toutefois, les régions les plus affectées par la dégénérescence dans ces pathologies ne sont pas tout le temps celles qui sont le plus marquées par les anticorps dirigés contre les PRD. Par exemple, la moelle épinière est une région où on retrouve peu de ces protéines, bien qu'elle soit l'une des structures les plus affectées dans la SLA. Trois hypothèses sont alors envisageables :

1. Les PRD ne sont pas responsables, ou pas entièrement responsables, de la dégénérescence neuronale dans les SLA/DFT.
2. Les travaux histologiques étant réalisés sur les cerveaux de patients décédés, certaines régions cérébrales sont déjà détruites, ce qui pourrait expliquer un marquage plus faible.
3. Une faible quantité de PRD hautement toxiques pourrait suffire à détruire les neurones des régions les moins marquées.
4. Certaines cellules pourraient également être plus résistantes ou plus vulnérables que d'autres à ces protéines.

Il faut préciser que la quantification des PRD la plus exhaustive effectuée jusqu'ici a été réalisée avec des anticorps polyclonaux qui reconnaissent tous les PRD (Ash et al., 2013). Pourtant, il est possible que les 5 protéines n'aient pas la même toxicité et soient présentes à des taux différents d'une région à une autre.

Bien que nous ne connaissions pas la répartition de chacun d'entre eux pour toutes les structures cérébrales affectées, des études, qui ne se sont concentrées que sur les régions les plus marquées, ont révélé que les concentrations entre PRD étaient significativement différentes. Les plus abondants sont les poly-GA et poly-GP. Les poly-

PA, poly-PR et poly-GR sont détectés à des niveaux plus faibles (Mori et al., 2013a ; Zu et al., 2013 ; Gendron et al., 2013). Malgré cette tendance générale, la concentration relative des PRD varie d'une région à l'autre. De plus, des ratios de PRD différents d'un patient à l'autre ont déjà été observés (Zu et al., 2013).

Ces observations restent à être généralisées à tout le système nerveux.

La toxicité des PRD a été confirmée par de nombreuses études. En effet, que ce soit dans des modèles cellulaires (May et al., 2014 ; Yamakawa et al., 2015 Chang et al., 2016 ; Flores et al., 2016), de drosophiles (Solomon et al., 2018 ; Wu et al., 2018) ou de souris (Zhang et al., 2018 ; Hao et al., 2019), les PRD seuls suffisent à reproduire la mort cellulaire et les symptômes observés dans les SLA/DFT.

Ainsi, parmi les trois hypothèses évoquées plus haut, il apparaît que celle concernant la toxicité des PRD soit la plus pertinente pour rendre compte de la pathogénicité de la mutation C9orf72. L'haplo-insuffisance du gène et la toxicité des ARN d'hexanucléotides prennent certainement part à la maladie, mais ne sont pas suffisantes pour en expliquer les principaux aspects.

IV – Mécanismes de toxicité des PRD

A – L'agrégation des PRD

- 1) Poly-GA est capable de former des fibres amyloïdes *in vitro* ayant des caractéristiques de type prion

L'analyse histologique de cerveaux de patients portants la mutation C9orf72 a révélé que les PRD se présentaient sous forme d'inclusions cytoplasmiques de forme sphérique ou en étoile (Mori et al., 2013a ; Ash et al., 2013 ; Zu et al., 2013 ; Gendron et al., 2013). Ces observations rappellent les caractéristiques des autres protéines agrégées dans les SLA/DFT. Les PRD appartiennent-ils aussi à la famille des protéines de type prion ?

Rares sont les travaux ayant caractérisé le comportement des PRD *in vitro*. L'une de ces études a montré que des peptides poly-GA contenant 15 répétitions s'assemblent pour former des agrégats fibrillaires de nature amyloïde (Chang et al., 2016). Une autre équipe a montré que parmi les PRD poly-GA, poly-GP et poly-GR ayant 3 ou 6 répétitions, les poly-GA ayant 6 répétitions s'assemblent en agrégats amyloïdes non fibrillaires (Flores et al., 2016). Les autres peptides formaient peu d'agrégats non-amyloïdiens et de morphologies sphériques. De plus, de récents travaux de Cryo-EM ont révélé que l'expression de poly-GA contenant 175 répétitions dans des cultures de neurones primaires conduisait à la formation d'agrégats plans, semblables à des rubans, de largeur variable et ramifiés (Guo et al., 2018).

Les fibres de poly-GA sont amyloïdes et sont également capables d'être internalisées, de se propager de cellule à cellule, de recruter la protéine monomérique et

de provoquer la mort des neurones (Chang et al., 2016). Ces caractéristiques correspondent à celles des protéines de type prion. Toutefois, ces résultats préliminaires doivent encore être reproduits et complétés.

Les études réalisées jusqu'ici suggèrent que la protéine poly-GA est celle qui s'agrège le plus parmi les PRD. Pour l'instant, c'est aussi la seule de ces protéines qui a été montrée comme se comportant à la manière d'un prion. Cette observation pourrait expliquer pourquoi c'est ce PRD qu'on retrouve le plus abondamment présent sous forme insoluble dans les neurones de patients SLA/DFT. Le peu de résultats obtenus sur les autres PRD ne permet pas de conclure sur leur propension à pouvoir s'agréger d'eux-mêmes. Toutefois, dans les analyses histologiques réalisées sur les cerveaux de patients, les inclusions cytoplasmiques sont souvent positives pour plusieurs PRD (Mori et al 2013a ; Mori et al., 2013b ; Yamakawa et al., 2015). Ainsi, il est possible que certaines de ces protéines co-agrégent entre elles ou avec d'autres partenaires.

2) Hypothèses sur l'agrégation des PRD

Le peu d'informations obtenues sur la nature et la morphologie des agrégats formés par les PRD devra être complété. Toutefois, en considérant leur séquence certains points essentiels de l'agrégation protéique peuvent nous éclairer quant à l'assemblage de telles protéines.

Il existe deux classes d'agrégats protéiques :

- 1) Les agrégats amorphes qui ont une morphologie variable. L'incorporation de monomères dans ces agrégats n'est pas ordonnée (Stanks et al., 2009).

2) Les agrégats amyloïdes sont le plus souvent fibrillaires. Les monomères, assemblés en feuillets β , sont disposés de façon perpendiculaire le long de l'axe de la fibre (voir Fig. 6).

Contrairement à l'agrégation amorphe, celle amyloïde est organisée. Elle est soumise à différentes contraintes. Notamment, la nature chimique des acides aminés est déterminante dans la formation des feuillets β .

L'interface formée par les feuillets β ne contient pas de molécules d'eau. Les brins sont légèrement décalés dans le feuillet de façon à ce que les chaînes latérales des acides aminés puissent se disposer par interdigitation (Sawaya et al., 2007). Cette organisation des assemblages amyloïdes est favorisée par les acides aminés hydrophobes (Tartaglia et al., 2008).

Il a aussi été montré que les résidues prolines, de part la contrainte stérique qu'induit les coudes qu'ils forment dans les protéines, sont défavorables à la formation de feuillets β (Wood et al., 1994 ; Williams et al., 2004). Egalement, les résidus aromatiques, en créant des interactions π , favoriseraient la formation d'agrégats amyloïdes (Gazit E., 2002).

Lorsqu'on considère les séquences des 5 PRD, il s'avère que peu d'entre eux possèdent une séquence favorisant la formation d'assemblages amyloïdes. Leur séquence est peu hydrophobe et n'est composée d'aucun résidu aromatique. De plus, les poly-GP, poly-PA et poly-PR sont composés pour la moitié de leur séquence de prolines, qui interrompent les feuillets β . Enfin, comme les poly-PR et poly-GR présentent un résidu d'arginine tous les deux acides aminés, leur charge électrique est fortement positive dans des conditions de pH physiologique. La répulsion de charge entre les chaînes latérales d'arginines pourrait également contrevenir à la formation d'agrégats.

Ainsi, il n'est pas surprenant que pour l'instant seul le poly-GA, dont la séquence est hydrophobe, ait été décrit comme formant des assemblages amyloïdes. Il est possible que les assemblages des autres PRD soient d'une autre nature.

Egalement, il est possible que la particularité des séquences des PRD, composées uniquement de répétitions d'un dipeptide, puisse avoir un impact sur leur agrégation. En effet, les protéines formant des agrégats amyloïdes, comme le peptide A β ou l' α -synucléine, ont leurs monomères alignés face-à-face le long de l'axe de la fibre (Fig. 6B). Seulement, la séquence des PRD n'impose pas que les monomères s'alignent face-à-face.

La nature répétée de la séquence des PRD rappelle celle de la protéine HTT mutée dans la maladie de Huntington. Dans cette pathologie, la répétition d'un codon CAG dans la séquence d'ADN de HTT conduit à la production de protéines contenant une séquence de polyglutamine. Au-delà de 36 répétitions, ces séquences répétées induisent l'agrégation de la protéine conduisant à la pathologie. La longueur de la séquence de polyglutamine est déterminante dans l'assemblage de la protéine et dans l'âge d'apparition de la maladie (Duyao et al., 1993 ; Monsellier et al., 2015). Ce déterminisme rappelle l'impact du nombre de répétitions de l'hexanucléotide GGGGCC au niveau du gène C9orf72 dans les SLA/DFT. Il est possible que les similitudes partagés entre HTT-polyQ et les PRD se retrouvent aussi au niveau de la structure de leurs agrégats.

La structure des assemblages de polyglutamine reste encore indéterminée. Une récente étude a montré que ces assemblages étaient structurés par des feuillets β et que son cœur était organisé par des épingles β (Hoop et al., 2016). Que de telles caractéristiques structurales puissent être partagées avec les assemblages de PRD

restent à être établi.

B – Les différentes voies de toxicité des PRD

1) Les cibles multiples des PRD

Différents travaux ont montré l'impact délétère des agrégats de poly-GA sur la protéostase (Zhang et al., 2014 ; May et al., 2014 ; Yamakawa et al., 2015 ; Guo et al., 2018). En effet, les assemblages de poly-GA recrutent massivement la sous-unité 26S du protéasome (Guo et al., 2018), ainsi que d'autres éléments de l'UPS comme l'Ubiquiline 1 et 2 ou encore PSMB6 (May et al., 2014). Cette dérégulation a également été observée pour les PRD poly-GP et poly-GR (Yamakawa et al., 2015). Un dérèglement de la protéostase engendre de nombreuses conséquences qui sont délétères pour la cellule.

On a aussi observé que la protéine Unc119 était fortement enrichie au niveau des agrégats de poly-GA (May et al., 2014). Ce polypeptide assure le développement et le maintien des axones. Sa perte de fonction est létale pour les neurones et pourrait expliquer la toxicité des fibres. Il est intéressant de noter qu'Unc119 reconnaît un motif GAGASA sur la transducine, l'un de ses partenaires (Zhang et al., 2011). Ce motif ayant une forte homologie avec la séquence des poly-GA, il est fort probable que cela explique la co-agrégation des deux protéines.

Parmi les PRD, poly-GR et poly-PR forment un groupe à part. Leurs mécanismes de toxicité divergent des autres. Dans des modèles cellulaires, ils forment des agrégats nucléolaires toxiques qui perturbent la synthèse et la régulation des ARN (Wen et al.,

2014 ; Kwon et al., 2014 ; Yamakawa et al., 2015). Les PRD riches en arginine lient directement les ARN, probablement par leurs résidus hautement basiques.

Il a aussi été montré que ces deux PRD co-localisent avec des composants des pores nucléaires comme la karyophérine, l'importine ou encore l'exportine – perturbant ainsi le transport nucléocytoplasmique (Jovičić et al., 2015 ; Boeynaems et al., 2016). Un modèle de souris n'exprimant que le PRD poly-GR a également manifesté une dérégulation de la machinerie traductionnelle, notamment en séquestrant des sous-unités du ribosome et des facteurs d'initiation. Cette souris développait une neurodégénérescence progressive, des troubles moteurs et de la mémoire (Zhang et al., 2018). Ces dérégulations ont été également constatées dans des modèles de drosophiles et d'iPSCs humains pour poly-PR (Moens et al., 2019).

Ces travaux s'entendent pour affirmer que les PRD poly-GR et poly-PR sont les espèces les plus toxiques. Seulement, ces résultats sont difficilement transposables à la clinique. En effet, on retrouve de faibles quantités de ces protéines chez les malades. Également, la présence intranucléaire de ces PRD n'a pas été observée dans le cerveau de ces patients.

2) L'agrégation des PRD engendre-t-elle celles des autres protéines retrouvées sous forme insoluble dans les SLA/DFT ?

On retrouve de nombreuses inclusions neuronales chez les patients SLA/DFT portant la mutation de C9orf72. Dans la plupart des cas, les inclusions sont positives pour les PRD ou pour TDP-43 (Mori et al., 2013a). La localisation des agrégats de TDP-43 correspond davantage aux lésions neurologiques observées que celle des inclusions de PRD (Mackenzie et al., 2013). Bien que la pathologie TDP-43 apparait donc être la plus déterminante, il semble que la présence de PRD soit bien corrélée avec plusieurs aspects

cliniques de la maladie (Schludi et al., 2015). Par ailleurs, certains patients développent une pathologie en présence de ces protéines bien qu'aucun agrégat de TDP-43 ne soit retrouvé dans leurs neurones (Mori et al., 2013a ; Mann et al., 2014). Mais dans la plupart des cas, les deux pathologies co-existent. Plusieurs études tentent d'expliquer ce phénomène.

Ainsi, il a été montré que chez des patients DFT, l'agrégation des PRD apparaissait avant celle de TDP-43 (Baborie et al., 2014). De plus, on retrouve des inclusions de PRD et de TDP-43 dans les mêmes cellules pour certaines régions du cerveau comme l'hippocampe, le cortex fronto-temporal et le cervelet. On a également remarqué que des inclusions de poly-GA sont parfois entourées d'agrégats de TDP-43 (Mori et al., 2013a). Ces observations suggèrent que l'accumulation de PRD sous forme insoluble précède, voire provoque, celle de TDP-43.

Différentes équipes ont montré que la présence d'agrégats de poly-GA dans une cellule induisait une délocalisation de TDP-43 depuis le noyau vers le cytoplasme, provoquant ainsi son agrégation (Khosravi et al., 2017 ; Nonaka et al., 2018 ; Solomon et al., 2018). Ce phénomène est dépendant de la séquence de localisation nucléaire NLS de la protéine (Khosravi et al., 2017) et du nombre de dipeptides des poly-GA (Nonaka et al., 2018). Bien que toutes ces études s'accordent à dire que les agrégats de poly-GA entraînent une agrégation de TDP-43, des résultats contradictoires concernant les autres PRD ne permettent pas de conclure sur leur possible implication.

Le mécanisme responsable de cette co-agrégation reste incompris. Néanmoins, comme nous l'avons vu les PRD, et particulièrement poly-GA, provoquent de multiples dérèglements de la protéostasie. On peut supposer que l'agrégation de TDP-43 en résulte.

Peu de travaux ont été réalisés concernant la possibilité de dialogue entre les PRD

et les autres protéines agrégées dans les SLA/DFT.

Une étude a montré que la présence d'hexanucléotides répétés suffisait à induire la phosphorylation et l'agrégation de Tau – sans qu'il ne soit précisé si ce sont les répétitions d'ADN, d'ARN ou les PRD qui sont responsables de ces évènements (He et al., 2018). Pour l'instant, aucune observation similaire n'a été rapportée pour les protéines Fus ou SOD1 – bien qu'on puisse retrouver cette dernière agrégée chez des patients ayant une SLA/DFT associée à une répétition d'hexanucléotides (Forsberg et al., 2019).

D'autres travaux doivent être réalisés pour comprendre davantage l'existence de protéinopathie mixte dans les SLA/DFT. Il est néanmoins possible que les PRD puissent avoir un effet toxique indirect en déclenchant l'agrégation d'autres protéines, comme TDP-43 – dont l'importance dans l'étiologie des SLA/DFT n'est plus à démontrer.

Conclusions

Bien qu'initialement décrites séparément, la SLA et la DFT se ressemblent sous bien des aspects. Plusieurs symptômes, lésions anatomiques ou marqueurs moléculaires sont communs aux deux maladies. L'idée que SLA et DFT ne soient plus des pathologies séparées, mais les deux pôles d'un même spectre pathologique s'impose de plus en plus.

La grande diversité de manifestations cliniques de ces maladies se reflète au niveau moléculaire. De nombreux acteurs génétiques et protéiques ont été incriminés pour expliquer l'origine de ces pathologies. Ces différents acteurs expliquent en partie la diversité phénotypique des SLA/DFT. Deux mécanismes principaux se dégagent de ces déterminants : l'agrégation de protéines se comportant à la manière de prions et la dérégulation de l'homéostasie des ARN.

Récemment, la découverte de la mutation la plus communément retrouvée dans les formes génétiques de ces maladies a permis de formuler de nouvelles hypothèses sur les mécanismes délétères à l'œuvre dans les SLA/DFT. La répétition d'un hexanucléotide (GGGGCC) au niveau du gène C9orf72 est une mutation dominante retrouvée tant chez les patients SLA que DFT. La découverte de cette mutation renforça l'idée d'un continuum entre les deux maladies. Trois hypothèses ont été avancées pour rendre compte de la part que prenait cette mutation dans la maladie :

- (1) Les modifications structurales de l'ADN induites par ces hexanucléotides diminuent fortement le nombre de transcrits de la protéine C9orf72. Cette protéine, encore peu connue, interviendrait dans différents processus cellulaires comme l'autophagie

ou la régulation du trafic endosomal. Seulement, sa délétion dans des modèles animaux n'a pas permis de reproduire des marqueurs essentiels de SLA/DFT.

(2) Les ARN issus de ces répétitions d'hexanucléotides ont montré une forte toxicité dans de nombreuses études. En formant des agrégats séquestrant des protéines nucléaires, ces ARN perturbent les machineries transcriptionnelle et traductionnelle de la cellule. Pourtant, ce mécanisme n'est pas suffisant pour reproduire le phénomène de dégénérescence neuronale des SLA/DFT et les symptômes qui en découlent.

(3) La traduction de ces ARN permet la production de 5 Protéines à Répétitions de Di-peptides (PRD). Ces protéines sont retrouvées sous forme d'inclusions cytoplasmiques dans de nombreuses structures cérébrales. Ces agrégats en perturbant la protéostasie ainsi que la synthèse, le transport et la traduction des ARN sont toxiques pour les cellules. Seulement, la corrélation entre les régions marquées pour ces protéines et les lésions observées chez les patients n'est que partielle. Par ailleurs, la nature des agrégats de PRD reste à élucider. Déterminer l'impact de la spécificité de séquence de ces protéines sur la formation d'agrégats représente un intérêt particulier pour l'étude biochimique des assemblages de protéines.

Afin d'unifier les principales voies de toxicité à l'œuvre dans les SLA/DFT, un modèle de protéinopathies croisées a été proposé. Il semble ainsi que les PRD participeraient à la dégénérescence par plusieurs voies, dont une indirecte mais potentiellement dramatique. En effet, l'agrégation de ces protéines induit aussi celle de TDP-43, qui est massivement retrouvée sous forme insoluble dans le cerveau des patients SLA et DFT. Les inclusions cytoplasmiques de cette protéine sont fortement corrélées aux lésions tissulaires. C'est un acteur central de la mort neuronale dans la plupart des cas de SLA/DFT.

Ainsi, si la présence de la mutation du gène C9orf72 est impliquée dans l'étiologie des SLA/DFT, ce serait, non seulement par les 3 voies évoquées plus haut, mais aussi en déclenchant une autre protéinopathie – plus déterminante encore dans le cours des maladies.

Cependant, ce phénomène ne peut rendre compte de toutes les différentes formes de la maladie et de nombreuses autres questions restent en suspend.

Travaux de thèse

**I – Dipeptide Repeat derived from C9orf72
Hexanucleotide Expansions Forms
Amyloids or natively unfolded structures
in vitro**

PLOS ONE

Dipeptide Repeat derived from C9orf72 Hexanucleotide Expansions Forms Amyloids or natively unfolded structures in vitro

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research Article
Full Title:	Dipeptide Repeat derived from C9orf72 Hexanucleotide Expansions Forms Amyloids or natively unfolded structures in vitro
Short Title:	Dipeptide Repeat derived from C9orf72 Forms Amyloids or natively unfolded structures in vitro
Corresponding Author:	luc bousset CNRS Fontenay aux roses, FRANCE
Keywords:	C9orf72; hexanucleotide repeat; Amyotrophic Lateral Sclerosis; Dipeptide Repeat Proteins; amyloid; Fronto-temporal dementia
Abstract:	The abnormal repetition of the hexanucleotide GGGGCC within the C9orf72 gene is the most common genetic cause of both Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD). Different hypothesis have been proposed to explain the pathogenicity of this mutation. Among them, the production of aberrant proteins called Dipeptide Repeat Proteins (DPR) from the repeated sequence. Those proteins are of interest, as they are toxic and form insoluble deposits in patient brains. In this study, we characterize the structural features of three different DPR encoded by the hexanucleotide repeat GGGGCC, namely poly-GA, poly-GP and poly-PA. We show that DPR are natively unstructured proteins but that only poly-GA forms in vitro fibrillary aggregates. Poly-GA fibrils are of amyloid nature as revealed by their high content in beta sheets. They neither bind Thioflavin T nor Primuline, the commonly used amyloid fluorescent dyes. Remarkably, not all of the poly-GA primary structure was part of fibrils amyloid core.
Order of Authors:	Laurent Brasseur Audrey Coens Jehan Waeytens Ronald Melki Luc Bousset
Opposed Reviewers:	Human Rezaei INRA Centre de Jouy-en-Josas Conflict of interest for the research project
Additional Information:	
Question	Response

<p>Financial Disclosure</p> <p>Enter a financial disclosure statement that describes the sources of funding for the work included in this submission. Review the submission guidelines for detailed requirements. View published research articles from PLOS ONE for specific examples.</p> <p>This statement is required for submission</p>	<p>LB, AC, RM, AC were supported by grants from the Agence Nationale de la Recherche Scientifique, European Commission Joint Programme on Neurodegenerative Diseases (JPND-TransPathND). LB and RM were supported by Equipe FRM (Fondation pour la Recherche Médicale) 2016 (DEQ2016033489). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.</p>
<p>and will appear in the published article if the submission is accepted. Please make sure it is accurate.</p> <p>Unfunded studies Enter: <i>The author(s) received no specific funding for this work.</i></p> <p>Funded studies Enter a statement with the following details:</p> <ul style="list-style-type: none"> • Initials of the authors who received each award • Grant numbers awarded to each author • The full name of each funder • URL of each funder website • Did the sponsors or funders play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript? • NO - Include this sentence at the end of your statement: <i>The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.</i> • YES - Specify the role(s) played. <p>* typeset</p>	

Competing Interests

Use the instructions below to enter a competing interest statement for this submission. On behalf of all authors, disclose any [competing interests](#) that could be perceived to bias this work—acknowledging all financial support and any other relevant financial or non-financial competing interests.

This statement **will appear in the published article** if the submission is accepted. Please make sure it is accurate. View published research articles from [PLOS ONE](#) for specific examples.

The authors have declared that no competing interests exist.

NO authors have competing interests

Enter: *The authors have declared that no competing interests exist.*

Authors with competing interests

Enter competing interest details beginning with this statement:

I have read the journal's policy and the authors of this manuscript have the following competing interests: [insert competing interests here]

* typeset

Ethics Statement

N/A

Enter an ethics statement for this submission. This statement is required if the study involved:

- Human participants
- Human specimens or tissue
- Vertebrate animals or cephalopods
- Vertebrate embryos or tissues
- Field research

Write "N/A" if the submission does not require an ethics statement.

General guidance is provided below. Consult the [submission guidelines](#) for detailed instructions. **Make sure that all information entered here is included in the Methods section of the manuscript.**

Format for specific study types

Human Subject Research (involving human participants and/or tissue)

- Give the name of the institutional review board or ethics committee that approved the study
- Include the approval number and/or a statement indicating approval of this research
- Indicate the form of consent obtained (written/oral) or the reason that consent was not obtained (e.g. the data were analyzed anonymously)

Animal Research (involving vertebrate animals, embryos or tissues)

- Provide the name of the Institutional Animal Care and Use Committee (IACUC) or other relevant ethics board that reviewed the study protocol, and indicate whether they approved this research or granted a formal waiver of ethical approval
- Include an approval number if one was obtained
- If the study involved *non-human primates*, add *additional details* about animal welfare and steps taken to ameliorate suffering
- If anesthesia, euthanasia, or any kind of animal sacrifice is part of the study, include briefly which substances and/or methods were applied

Field Research

Include the following details if this study involves the collection of plant, animal, or other materials from a natural setting:

- Field permit number
- Name of the institution or relevant body that granted permission

<p>Data Availability</p> <p>Authors are required to make all data underlying the findings described fully available, without restriction, and from the time of publication. PLOS allows rare exceptions to address legal and ethical concerns. See the PLOS Data Policy and FAQ for detailed information.</p>	<p>Yes - all data are fully available without restriction</p>
<p>A Data Availability Statement describing where the data can be found is required at submission. Your answers to this question constitute the Data Availability Statement and will be published in the article, if accepted.</p> <p>Important: Stating 'data available on request from the author' is not sufficient. If your data are only available upon request, select 'No' for the first question and explain your exceptional situation in the text box.</p> <p>Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?</p>	

<p>Describe where the data may be found in full sentences. If you are copying our sample text, replace any instances of XXX with the appropriate details.</p> <ul style="list-style-type: none"> • If the data are held or will be held in a public repository, include URLs, accession numbers or DOIs. If this information will only be available after acceptance, indicate this by ticking the box below. For example: <i>All XXX files are available from the XXX database (accession number(s) XXX, XXX).</i> • If the data are all contained within the manuscript and/or Supporting Information files, enter the following: <i>All relevant data are within the manuscript and its Supporting Information files.</i> • If neither of these applies but you are able to provide details of access elsewhere, with or without limitations, please do so. For example: <p><i>Data cannot be shared publicly because of [XXX]. Data are available from the XXX Institutional Data Access / Ethics Committee (contact via XXX) for researchers who meet the criteria for access to confidential data.</i></p> <p><i>The data underlying the results presented in the study are available from (include the name of the third party</i></p>	<p>All relevant data are within the manuscript and its Supporting Information files</p>
<p><i>and contact information or URL).</i></p> <ul style="list-style-type: none"> • This text is appropriate if the data are owned by a third party and authors do not have permission to share the data. <p>* typeset</p>	
<p>Additional data availability information:</p>	

Dear Editor,

My coworkers Laurent Brasseur, Audrey Coens, Jehan Waeytens, Ronald Melki and myself would like you to consider for publication in Plos One Please our manuscript entitled: "*Dipeptide Repeat derived from C9orf72 Hexanucleotide Expansions Forms Amyloids or natively unfolded structures in vitro*".

Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD) are linked to abnormal repetition of the hexanucleotide GGGGCC within the C9orf72 gene. This mutation account for the most common genetic factor. Aberrant proteins synthesis from hexanucleotide repeats produces Dipeptide Repeat Proteins (DPR) which are toxic and form insoluble deposits in patient brains.

In this study, we investigated the structural features of three different DPR proteins encoded by the hexanucleotide repeat GGGGCC, namely poly-Gly/Ala, poly- Gly/Pro and poly-Pro/Ala. We show that DPR are natively unstructured proteins but that only poly-GA forms in vitro fibrillary aggregates of amyloid nature as revealed by infra- red spectroscopy. Remarkably, these assemblies bind neither Thioflavin T nor Primuline, the commonly used amyloid fluorescent dyes. Our findings suggest that, not all of the poly- GA primary structure is part of fibrils amyloid core.

We consider that this work is of particular interest to the fields of ALS and FTD as it brings new insights on the toxicity of mutated C9orf72 gene product. We confirmed the amyloidogenic nature of poly-Gly/Ala peptides in vitro, widely considered to be at the origin of early onset of ALS and FTD in vivo.

Sincerely yours,

Dipeptide Repeat derived from C9orf72 Hexanucleotide Expansions Forms

Amyloids or natively unfolded structures *in vitro*

Laurent Brasseur¹, Audrey Coens¹, Jehan Waeytens^{2,3}, Ronald Melki¹, Luc Bousset¹

¹Institut de Biologie François Jacob, Molecular Imaging Research Center (MIRCent), Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Direction de la Recherche Fondamentale (DRF), Laboratoire des Maladies Neurodégénératives, Centre National de la Recherche Scientifique (CNRS), Paris, Fontenay-aux-Roses, F-92265, France

² Laboratoire de Chimie Physique, CNRS UMR 8000 Université Paris-Sud Orsay France

³ Structure et Fonction des Membranes Biologiques, Université libre de Bruxelles, Bruxelles, Belgique

Author email addresses:

Laurent Brasseur (lbrasseur@gmail.com)

Audrey Coens (aud_coens@msn.com)

Jehan Waeytens (jehan.waeytens@ulb.ac.be)

Ronald Melki (ronald.melki@cnrs.fr)

Luc Bousset (luc.bousset@cnrs.fr)

Corresponding Author:

Luc Bousset, Institut de Biologie François Jacob, Molecular Imaging Research Center (MIRCent), Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Direction de la Recherche Fondamentale (DRF), Laboratoire des Maladies Neurodégénératives, Centre National de la Recherche Scientifique (CNRS), Paris, Fontenay-aux-Roses, F-92265, France. luc.bousset@cnrs.fr

Abstract

The abnormal repetition of the hexanucleotide GGGGCC within the C9orf72 gene is the most common genetic cause of both Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD). Different hypothesis have been proposed to explain the pathogenicity of this mutation. Among them, the production of aberrant proteins called Dipeptide Repeat Proteins (DPR) from the repeated sequence. Those proteins are of interest, as they are toxic and form insoluble deposits in patient brains. In this study, we characterize the structural features of three different DPR encoded by the hexanucleotide repeat GGGGCC, namely poly-GA, poly-GP and poly-PA. We show that DPR are natively unstructured proteins but that only poly-GA forms *in vitro* fibrillary aggregates. Poly-GA fibrils are of amyloid nature as revealed by their high content in beta sheets. They neither bind Thioflavin T nor Primuline, the commonly used amyloid fluorescent dyes. Remarkably, not all of the poly-GA primary structure was part of fibrils amyloid core.

Introduction

Amyotrophic Lateral Sclerosis (ALS) is a progressive neurodegenerative disease that affects both the central and peripheral nervous system leading to impaired motor functions and a progressive paralysis of the skeletal muscles. Frontotemporal dementia (FTD), the second cause of dementia after Alzheimer's disease, is also a neurodegenerative disease diagnosed by cognitive and behavioral symptoms. These pathologies share numerous clinical and pathological aspects and are now considered as the two extremities of the same pathological spectrum, rather than distinct entities.

Mutation within the open reading frame 72 of chromosome 9 (C9orf72) is a genetic factor shared by both pathologies [1,2]. The mutation consists of hexanucleotide (GGGGCC) repetition from 25 to 1000 times in the promoter or intron 1 sequence. It is the most common genetic cause for these diseases that account for 16% of familial ALS and 20% of familial FTD patients. It is also reported in 6 - 8 % of ALS sporadic cases [3].

The role of this mutation in the etiology of ALS/FTD is poorly understood. So far, three mechanisms have been proposed to explain it: (1) Haploinsufficiency of the gene C9orf72, possibly encoding for a GTP Exchange Factor [4–7], (2) Toxicity of the hexanucleotide sense and antisense RNA that sequester RNA-binding proteins and other factors into insoluble RNA foci [8–10], (3) Unconventional translation of repeated hexanucleotide into toxic proteins called Dipeptide Repeat Proteins (DPR) [11–16]. These proteins have been shown to be produced independently of any start codon by a rare mechanism called repeat associated non-ATG (RAN) translation [11,13–15]. As GGGGCC and CCGGGG RNA are translated in all alternate reading frames: poly-GA (Glycine-Alanine), poly-GP (Glycine-Proline) and poly-GR (Glycine-Arginine) for sense transcripts or poly-PA (Proline-Alanine), poly-PR (Proline-Arginine) and also poly-GP for antisense transcripts [11,12,14]. Antibodies directed against several DPR stain neuronal inclusions

that are p62-positive and TDP-43-negative. Poly-GA seems to be the most abundant DPR in the inclusions [13]. Nonetheless, the distribution of DPR aggregates does not completely correlate with lesions observed in patient brains [12,13,15,17,18]. It remains possible that small amounts of highly toxic DPR may be sufficient to kill cells. Another hypothesis is that DPR could trigger TDP-43 aggregation, a major pathological mechanism occurring in ALS/FTD [19–23].

So far, very few studies have been done to characterize the biochemical properties of DPR [24,25]. Here, we show that poly-GA, poly-GP and poly-PA DPR are mostly unstructured proteins. Only poly-GA form fibrillary aggregates in-vitro. We show, using FTIR spectroscopy that GAx34 fibrillar assemblies are of amyloid nature albeit they neither bind Thioflavin T nor Primuline, the fluorescent dyes that are considered as markers of amyloid aggregates. Interestingly, not all of the poly-GA repeated sequence was involved in fibrils amyloid core.

Results

The DPR proteins were expressed in *E. coli* and purified by affinity for a 6xhistidine Tag on a Cobalt column (see Materials and Methods section). We produced four different DPR proteins of distinct composition: PAX50, GAX50, GAX34 and GPX24. The different DPR proteins were expressed as fusion proteins with different tags for purification and immunodetection purposes (Fig S1).

The oligomeric state of the DPR proteins, was determined by assessment of their apparent molecular weights by size-exclusion chromatography (SEC). DPR proteins freshly solubilized in HFIP were dried and the resulting film was resuspended in PBS buffer. All DPR were found soluble after this treatment, except GAX50 protein that forms immediately fibrillary aggregates. GAX34 protein in PBS displays two species with apparent molecular weights of 21 and 6,4 kDa (Fig 1A, 1G). GAX34 protein presence in each peak was confirmed by Western-blot analysis using an anti-V5 antibody (Fig 1D). GPX24 protein shows three species (Fig 1B) all of which contain GPX24 protein as confirmed by Western blot analysis (Fig 1E). The principal specie has an apparent molecular weight of 30 kDa, whereas that of the second is 63 kDa. The third species corresponds to aggregated GPX24 protein with an apparent molecular weight over 2000 kDa (Fig 1G). The species corresponding to peak 3 represents monomeric GPX24 protein. Peaks 1 and 2 correspond to oligomeric GPX24 protein. PAX50 protein presents an elution profile similar that of GPX24 protein with the major species exhibiting an apparent molecular weight of 105 and 32 kDa and a third species corresponding to aggregated PAX50 protein (Fig 1C). PAX50 protein species with apparent molecular weight 105 and 32 kDa were recognized by an anti-FLAG antibody (FIG 1F). Altogether, our results shows that all three DPR we produced are of mostly monomeric in nature with varying amounts of low and high molecular weight species corresponding to small or larger oligomeric species

(Fig S1).

We next measured the secondary structure content of the DPR proteins by Circular Dichroism (CD). The different DPR proteins were prepared as for SEC (Fig 2A). GAx34, GPx24 and PAX50 proteins possess random coil structures (Fig 2B). Indeed, deconvolution of the CD spectra indicates that over 78% of the polypeptide consists of random coil (Fig 2B).

We next investigated the capacity of the DPR we produced to assemble into higher molecular weight species. We explored various assembly conditions e.g. buffers, temperature and agitation. In all explored conditions, GPx24 and PAX50 proteins remained mostly soluble with a fraction ranging from 10 to 20% of the proteins forming aggregates of amorphous nature (data not shown). In contrast, GAx34 and GAx50 proteins assembled into fibrillar aggregates that bundle together in purification buffer or PBS at 4°C, without shaking, as observed by transmission electron microscopy after negative staining with Uranyl acetate (Fig 3A and 3B). While GAx50 protein assembled within minutes to hours into fibrils, GAx34 protein (100µM) assembly into fibrils was slow. Assembly of the poly-GA proteins was followed using a sedimentation assay and SDS-PAGE analysis (Fig 3C). GAx34 protein (100 µM) assembly into fibrils was complete after 30 days (Fig 3A). Interestingly, GAx34 fibrils do not bind the fluorescent dyes Thioflavin T or Primuline that are known to bind in a quantitative manner to amyloid fibrils such as those α -synuclein forms (Fig 3B).

We next performed FTIR measurements to determine whether fibrillar GAx34 protein is of amyloid nature. Fibrillar GAx34 protein FTIR spectrum clearly shows a shoulder at 1620 cm^{-1} , indicative of the amyloid nature of the aggregates (Fig 4A). Fourier deconvolution of the spectra indicates that 42 % of GAx34 amino acids residues are involved in the β -sheet structures within the fibrillar form of the protein. The 58% remaining amino acid residues are disordered (Fig 4B). The peak centered at 1628 cm^{-1} indicates

that GAx34 fibrils have cross- β -sheet content and are therefore of amyloid nature. Surprisingly, our data clearly indicate that only 40 out of the 68 GA residues are involved into the amyloid core of fibrillar GAx34.

Discussion

DPR have been proposed to be key contributors in C9orf72+ pathologies. Little is known about their biochemical properties and the deleterious effect of these proteins is still under debate. Here we have characterized the properties of three DPR proteins of different lengths: GAx34, GPx24 and PAX50. We have shown that, in PBS, these proteins have apparent molecular weights compatible with monomeric or small sized oligomeric forms. The three DPR proteins exhibit the characteristics of disordered proteins as assessed by CD measurements.

We did not succeed to assemble GPx24 and PAX50 into fibrils under any of the numerous experimental conditions we tested. Amorphous aggregates were obtained instead. In contrast, GAx34 and GAx50 proteins readily assembled into fibrils at 4°C. This is in agreement with a previous study where synthetic peptides GAX3 and 6 were demonstrated to exhibit the highest propensity to aggregate [24]. This is of particular interest given that most DPR inclusions in patient brains are GA positive [12,13]. FTIR measurements reveal that the fibrillar GAx34 proteins we generated are of amyloid nature, in agreement with previous observations made with DPR with little number of repeats (25). Altogether, the present findings and previous results suggest that GA DPR assemblies of amyloid nature may play a role in neuronal degeneration in ALS/FTD.

The FTIR measurements we performed indicate that only part of GA repeat is involved in the amyloid core of GAx34 fibrils. Whether this observation is relevant to pathology remains to be established given that 400 and over GA repeats are associated to ALS/FTD in normal population. Still, our result might mean that GA aggregates elongate in a way where monomers stack with offset. The repeated nature of their sequence can allow such phenomenon, contrary to other aggregation-prone proteins such as Tau or α -Synuclein [27–31]. This could also account for the poly-GA assemblies with variable width

and bifurcation reported previously [26].

Contrary to previous observations, we neither observed Thioflavin T nor Primuline binding to poly-GA fibrils [24,25]. These fluorescent dyes are often used as markers of amyloids assemblies, despite some controversies [32]. The inability of DPR to bind canonical amyloid dyes may be due to the lack of key amino acids within primary structure, in agreement with previous studies showing a requirement for Tyrosine and Leucine for Thioflavin T binding [33,34].

As indicated above, we were not able to generate fibrillar poly-GP and poly-PA. Those DPR are nonetheless detected within neuronal inclusions [12–14]. Thus, they either form under conditions we did not explore or co-assemble with other DPR or aggregate into amorphous structures such as those we generated that are relocated in neurons, as many other proteins, to inclusions containing filamentous poly-GA. The ability of poly-GP and poly-PA at different concentrations to co-assemble into fibrils with poly-GA at different molar ratio was assessed. Amorphous aggregates together with fibrils were obtained as when poly-GP or poly-PA amorphous assemblies are diluted in the presence of fibrillar poly-GA (not shown). This suggests that poly-GP and poly-PA do not co-assemble into fibrils with poly-GA.

Acknowledgments

This work was supported by grants from the Agence Nationale de la Recherche Scientifique, European Commission Joint Programme on Neurodegenerative Diseases (JPND-TransPathND). The present work has benefited from Imagerie-Gif core facility, supported by l'Agence Nationale de la Recherche (ANR-11-EQPX-0029/Morphoscope, ANR-10-INBS-04/FranceBioImaging; ANR-11-IDEX-0003-02/ Saclay Plant Sciences) for access to Electron Microscopes. LB and RM were supported by Equipe FRM (Fondation pour la Recherche Médicale) 2016 (DEQ2016033489).

Materials and methods

Dipeptide Repeat Proteins (DPR) cloning and purification

DNA sequences encoding a V5 tag followed by 34 repeats of GA or 24 repeats of GP were subcloned in a bacterial expression vector containing an N-terminal 6xHis-Tag and a TEV protease cleavage site (pETM-11 vector, EMBL). Plasmids pAG416-Gal encoding for PAX50 and GAX50 were obtained from Addgene. Proteins were expressed in *E. coli* BL21, and purified on 5 mL Talon column (Clontech®) loaded with Cobalt. The proteins were eluted with a linear gradient of 12 ml from buffer A (20mM Tris pH 7.5, 250mM NaCl, 5mM Imidazole, 1mM β MerCaptoethanol, Glycerol 10%, 0.1 mM PMSF) to buffer B (20mM Tris pH 7.5, 250mM NaCl, 250mM Imidazole, 1mM β MerCaptoethanol, Glycerol 10%, 0.1 mM PMSF). Eluted fraction were analysed by SDS PAGE, and proteins were quantified spectrophotometrically using a extinction coefficient of $\epsilon_{GP/GA} = 2980 \text{ M}^{-1}\text{cm}^{-1}$ and $\epsilon_{PA} = 4470 \text{ M}^{-1}\text{cm}^{-1}$. Proteins were assembled into fibrils at 4°C without shaking for at least 30 days.

Size-Exclusion Chromatography (SEC) and Western-Blot analysis

DPR species of different sizes were separated by SEC on a Superose 6 (10/300) column (GE Healthcare) equilibrated in 0.1X PBS pH 7.4 at a flow rate of 0.4 ml/min. Elution was monitored by measuring absorbance at 280 nm. During elution, fractions were collected for Western-Blot analysis using antibodies directed against V5 tag for GAX34 or GPx24 and FLAG tag for PAX50. The Superose 6 column was calibrated using Dextran blue (2000 kDa), Thyroglobulin (669 kDa), Alcohol Deshydrogenase (150 kDa), BSA (66 kDa), and carbonic anhydrase (29 kDa) standards (Sigma).

Circular dichroism

CD spectra of GAx34, GPx24 and PAX50 were recorded at 20°C in a Jasco J810 dichrograph using 0.1 mm pathlength quartz cuvettes (Hellma) containing 20 µL of the protein solutions. Proteins were solubilized in HFIP, dried under an N₂ stream before being dissolved in PBS buffer and analyzed. All spectra were normalized to the mean residue weight ellipticity (θ_{MRW}) [deg cm² /dmol] using the equation $\theta(\lambda)_{MRW} = \theta(\lambda)_{deg}/10cnd$ where $\theta(\lambda)_{deg}$ is the recorded spectra in degrees, d is the path length of the cuvette in centimeters, n is the number of amino acid residues, and c is the sample concentration in moles per liter.

Electron microscopy

The morphology of DPR assemblies was examined by transmission electron microscopy (TEM) in a Jeol 1400 transmission electron microscope, following adsorption onto carbon-coated 200 mesh grids and negative staining with 1% uranyl acetate. The images were recorded with a Gatan Orius CCD camera (Gatan Inc., Pleasanton, CA, USA).

Fibrillization kinetics by SDS-PAGE analysis

Protein fibrillization was followed by disappearance over time of monomeric DPR (100 µM) from the supernatant of aliquots withdrawn at different time points from the assembling reaction at 4°C after sedimentation at 100,000 g for 30 min using SDS-PAGE following Coomassie Blue staining. Quantification was done with ImageLab. Kinetic points are the mean of three different experiments ± SEM.

Fourier transform infrared spectroscopy

GA fibrils were centrifuged for 30 min at 100,000 g at 4°C then extensively washed through three cycles of resuspension and sedimentation with D₂O. The spectra were recorded on a Bruker Vertex 70 Fourier transform infrared spectroscopy (FTIR)

spectrometer equipped with a liquid nitrogen-cooled MCT detector. The background consisted of D₂O and water vapour. A total of 100 interferograms were collected with a resolution of 1 cm⁻¹. The sample were loaded on a single reflection ATR chamber (PikeTech). All the spectra were baseline-corrected, smoothed and normalized prior to further data processing. The amide I (1,600–1,700 cm⁻¹) band of the spectra was fitted using a Gaussian species model centered at 1628, 1648,5, 1667 and 1680 cm⁻¹. During the fitting procedure the peak height was free, whereas the width at half-height was maintained at <20 cm⁻¹.

Thioflavin T and Primuline fluorescence in presence of GA and α -Synuclein fibrils

The fibrillary samples were spun at 100,000g for 30 min at 4°C for GA and at 37°C for α -Synuclein. The pellets were resuspended in PBS at 150 μ M and then diluted at different concentrations (100, 50, 25 and 10 μ M). Samples were incubated with Thioflavin T or Primuline (10 μ M). Fluorescence was recorded with a Cary Eclipse spectrofluorimeter (Varian Medical Systems Inc.) using excitation and emission wavelengths set at 440 and 480 nm for Thioflavin T and 400 and 480 nm for Primuline.

Legends

Figure 1. Apparent molecular weights of solubilized GAx34, GPx24 and PAx50.

(A – C) Size-exclusion chromatograms on Superose 6 (10/300) column of GAx34, GPx24 and PAx50 (150 μ M). Samples were solubilized in HFIP and filtered through 0,2 μ m pore filters and loaded onto the column. Peptides elution was followed by absorbancy at 280 nm. Column calibration is shown for Dextran blue (>2000 kDa), thyroglobulin (670 kDa), β -amylase (200 kDa), BSA (66 kDa) and carbonic anhydrase (29 kDa). All spectra were normalized to the highest value. **(D – F)** Western-Blot analysis of collected fractions using antibodies directed against V5 tag for GAx34 or GPx24 and FLAG tag for PAx50. **(G)** Apparent molecular weight of GAx34, GPx24 and PAx50 peaks based on their V_e/V_0 ratio (see supplemental 1).

Figure 2. Secondary structure determination of solubilized GAx34, GPx24 and PAx50.

A Circular dichroism spectra of GAx34 (red line), GPx24 (blue line) and PAx50 (green line). DPR were all solubilized in HFIP before resuspension in PBS and analysis. All spectra were normalized to the mean residue weight ellipticity (θ_{MRW}) [deg cm² /dmol] **B** Secondary structure estimation of solubilized GAx34, GPx24 and PAx50 obtained by deconvolution of the CD spectroscopic measurements. Deconvolution was performed on Dichroweb server using CONTIN algorithm (reference set 7).

Figure 3. DPR made of poly-GA form fibrillar assemblies that do not bind Thioflavin T nor Primuline

(A, B) Negatively stained TEM of GAx34 and GAx50 fibrillar assemblies obtained by incubation at 4°C without shaking. Scale bar, 200 nm **(C)** poly-GAx34 aggregation followed by disappearance over time of monomeric protein (100 μ M) from the SDS-PAGE gel. Aliquots of GAx34 were taken during assembly and analyzed by SDS-PAGE followed by

Coomassie Blue staining (inset). Quantification was done by comassie signal integration using ImageLab. Quantification is shown as the average and SEM of three independent experiments. **(D)** Thioflavin T and Primuline (10 μ M) binding assay with increasing concentration of GAx34 fibrils (●) and control amyloid fibrils made of alpha synuclein aSyn (X). Fluorescence intensity was recorded using excitation and emission wavelengths set at 440 and 480 nm for Thioflavin T and 400 and 480 nm for Primuline. Quantification are shown as the average and SEM of three independent experiments.

Figure 4. Deconvoluted FT-IR spectrum of GA fibrils

(A) GAx34 fibrils deconvoluted FT-IR spectrum. GAx34 fibrils were washed 3 times in D₂O with repeated cycles of spinning/resuspension. FTIR spectra (thick line), Fourier deconvolution (thin lines) and curve fit (dashed line) data are presented. FT-IR spectrum was fitted using a Gaussian species model centered at 1628, 1680 (β -sheets), 1649 (disordered and α helix) and 1667 (loops) cm^{-1} . **(B)** Secondary structure estimation for GAx34 fibrils obtained by deconvolution of FTIR spectrum. The percent of the Area and the associated number of peptide bonds are listed for each structure assignment.

Supplemental Figure 1. Sequence DPR use in this study

The His tag and TEV cleavage site are common to all DPR studied. An V5 TAG is present upstream to GAx34 and GPx24, while a Myc and FLAG tag are present downstream to GAx50 and PAX50 sequences.

Supplemental Figure 2. Calibration of Superose 6 (10/300) column

A Size-exclusion chromatograms of molecular weights markers (1 : Blue Dextran, 2 : Thyroglobulin, 3 : Alcohol deshydrogenase, 4 : BSA, 5 : Carbonic deshydrogenase). **B** Calibration curve derived from molecular weight standards. The ratio V_e/V_0 relate to the elution volume of the proteins divide by the dead volume (elution volume of Blue Dextran). Molecular weights are presented on a log scale. Exponential regression's equation is presented on the graph.

References

1. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*. 2011. doi:10.1016/j.neuron.2011.09.011
2. Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. 2011. doi:10.1016/j.neuron.2011.09.010
3. Marogianni C, Rikos D, Provatas A, Dadouli K, Ntellas P, Tsitsi P, et al. The role of C9orf72 in neurodegenerative disorders: a systematic review, an updated meta-analysis, and the creation of an online database. *Neurobiol Aging*. 2019. doi:10.1016/j.neurobiolaging.2019.04.012
4. Levine TP, Daniels RD, Gatta AT, Wong LH, Hayes MJ. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics*. 2013. doi:10.1093/bioinformatics/bts725
5. Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, et al. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann Neurol*. 2013. doi:10.1002/ana.23946
6. Therrien M, Rouleau GA, Dion PA, Parker JA. Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in *C. elegans*. *PLoS One*. 2013. doi:10.1371/journal.pone.0083450
7. Haeusler AR, Donnelly CJ, Periz G, Simko EAJ, Shaw PG, Kim MS, et al. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature*. 2014. doi:10.1038/nature13124
8. Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S, et al. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron*.

2013. doi:10.1016/j.neuron.2013.10.015

9. Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Štalekar M, et al. Hexanucleotide repeats in ALS/FTD form length-dependent RNA Foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* 2013. doi:10.1016/j.celrep.2013.10.049
10. Cooper-Knock J, Walsh MJ, Higginbottom A, Highley JR, Dickman MJ, Edbauer D, et al. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain.* 2014. doi:10.1093/brain/awu120
11. Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PEA, Caulfield T, et al. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol.* 2013. doi:10.1007/s00401-013-1192-8
12. Mori K, Arzberger T, Grässer FA, Gijssels I, May S, Rentzsch K, et al. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* 2013. doi:10.1007/s00401-013-1189-3
13. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science (80-).* 2013. doi:10.1126/science.1232927
14. Zu T, Liu Y, Bañez-Coronel M, Reid T, Pletnikova O, Lewis J, et al. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A.* 2013. doi:10.1073/pnas.1315438110
15. Ash PEA, Bieniek KF, Gendron TF, Caulfield T, Lin WL, DeJesus-Hernandez M, et al. Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron.* 2013. doi:10.1016/j.neuron.2013.02.004
16. Mann DMA, Rollinson S, Robinson A, Bennion Callister J, Thompson JC, Snowden JS, et al. Dipeptide repeat proteins are present in the p62 positive inclusions in patients with

- frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol Commun.* 2014. doi:10.1186/20515960168
17. Gomez-Deza J, Lee YB, Troakes C, Nolan M, Al-Sarraj S, Gallo JM, et al. Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. *Acta Neuropathol Commun.* 2015. doi:10.1186/s40478-015-0218-y
 18. Davidson YS, Barker H, Robinson AC, Thompson JC, Harris J, Troakes C, et al. Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol Commun.* 2014. doi:10.1186/2051-5960-2-70
 19. Chew J, Gendron TF, Prudencio M, Sasaguri H, Zhang YJ, Castanedes-Casey M, et al. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science (80-)*. 2015. doi:10.1126/science.aaa9344
 20. Khosravi B, Hartmann H, May S, Möhl C, Ederle H, Michaelson M, et al. Cytoplasmic poly-GA aggregates impair nuclear import of TDP-43 in C9orf72 ALS/FTLD. *Hum Mol Genet.* 2017. doi:10.1093/hmg/ddw432
 21. Solomon DA, Stepto A, Au WH, Adachi Y, Diaper DC, Hall R, et al. A feedback loop between dipeptide-repeat protein, TDP-43 and karyopherin- α mediates C9orf72-related neurodegeneration. *Brain.* 2018. doi:10.1093/brain/awy241
 22. Baborie A, Griffiths TD, Jaros E, Perry R, McKeith IG, Burn DJ, et al. Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. *Neuropathol Appl Neurobiol.* 2015. doi:10.1111/nan.12178
 23. Nonaka T, Masuda-Suzukake M, Hosokawa M, Shimozawa A, Hirai S, Okado H, et al. C9ORF72 dipeptide repeat poly-GA inclusions promote intracellular aggregation of phosphorylated TDP-43. *Hum Mol Genet.* 2018. doi:10.1093/hmg/ddy174

24. Flores BN, Dulchavsky ME, Krans A, Sawaya MR, Paulson HL, Todd PK, et al. Distinct c9orf72-associated dipeptide repeat structures correlate with neuronal toxicity. *PLoS One*. 2016. doi:10.1371/journal.pone.0165084
25. Chang YJ, Jeng US, Chiang YL, Hwang IS, Chen YR. The glycine-alanine dipeptide repeat from C9 or f72 hexanucleotide expansions forms toxic amyloids possessing cell-to-cell transmission properties. *J Biol Chem*. 2016. doi:10.1074/jbc.M115.694273
26. Guo Q, Lehmer C, Martínez-Sánchez A, Rudack T, Beck F, Hartmann H, et al. In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. *Cell*. 2018. doi:10.1016/j.cell.2017.12.030
27. Fitzpatrick AWP, Falcon B, He S, Murzin AG, Murshudov G, Garringer HJ, et al. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature*. 2017. doi:10.1038/nature23002
28. Falcon B, Zhang W, Schweighauser M, Murzin AG, Vidal R, Garringer HJ, et al. Tau filaments from multiple cases of sporadic and inherited Alzheimer's disease adopt a common fold. *Acta Neuropathol*. 2018. doi:10.1007/s00401-018-1914-z
29. Zhang W, Falcon B, Murzin AG, Fan J, Crowther RA, Goedert M, et al. Heparin-induced tau filaments are polymorphic and differ from those in alzheimer's and pick's diseases. *Elife*. 2019. doi:10.7554/eLife.43584
30. Guerrero-Ferreira R, Taylor NMI, Mona D, Ringler P, Lauer ME, Riek R, et al. Cryo-EM structure of alpha-synuclein fibrils. *Elife*. 2018. doi:10.7554/eLife.36402
31. Li B, Ge P, Murray KA, Sheth P, Zhang M, Nair G, et al. Cryo-EM of full-length α -synuclein reveals fibril polymorphs with a common structural kernel. *Nat Commun*. 2018. doi:10.1038/s41467-018-05971-2
32. Sabaté R, Lascu I, Saupe SJ. On the binding of Thioflavin-T to HET-s amyloid fibrils assembled at pH 2. *J Struct Biol*. 2008. doi:10.1016/j.jsb.2008.02.002

33. Biancalana M, Makabe K, Koide A, Koide S. Molecular Mechanism of Thioflavin-T Binding to the Surface of β -Rich Peptide Self-Assemblies. *J Mol Biol.* 2009. doi:10.1016/j.jmb.2008.11.006
34. Wu C, Biancalana M, Koide S, Shea JE. Binding Modes of Thioflavin-T to the Single-Layer β -Sheet of the Peptide Self-Assembly Mimics. *J Mol Biol.* 2009. doi:10.1016/j.jmb.2009.09.056

Figure 1

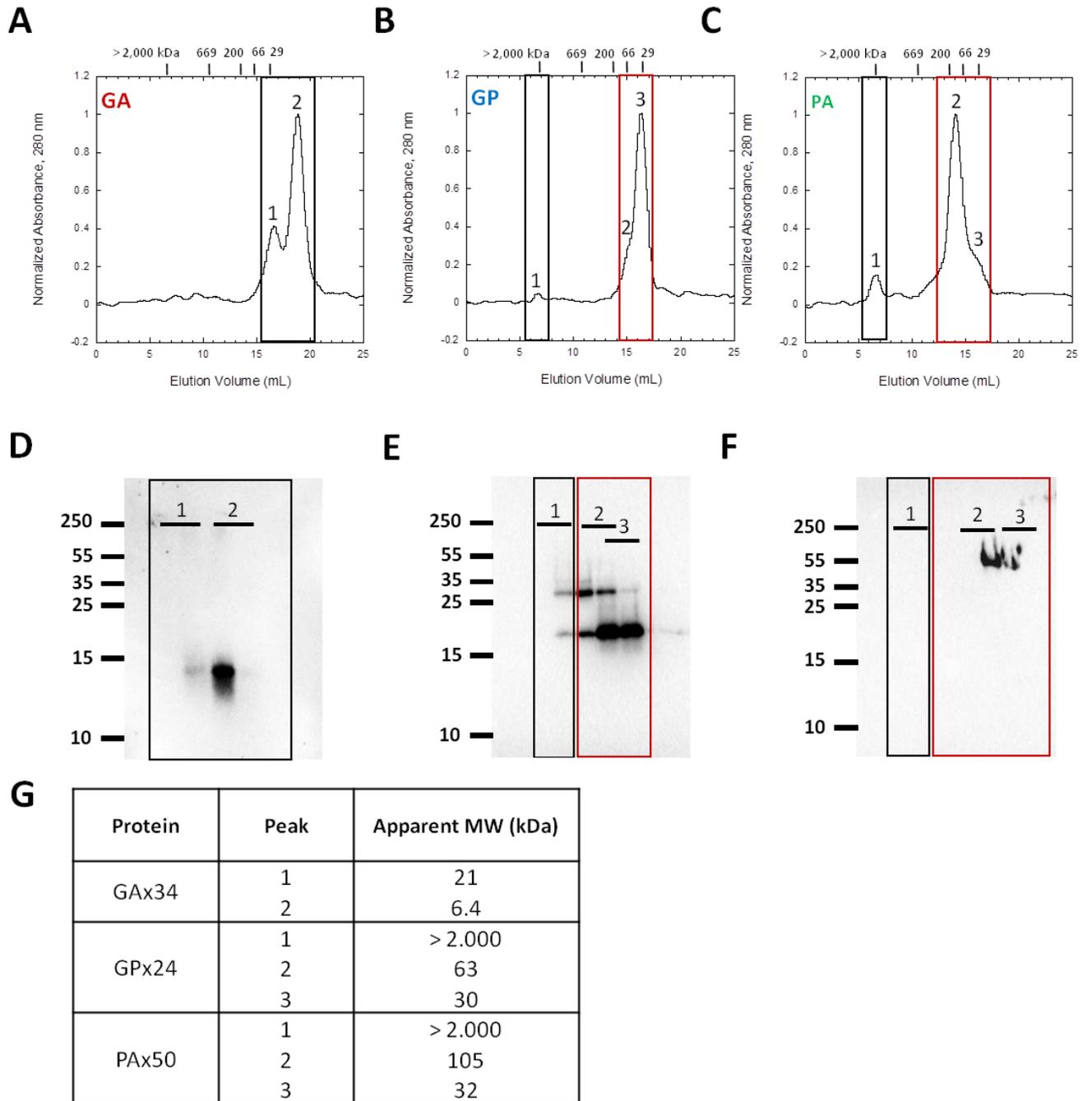
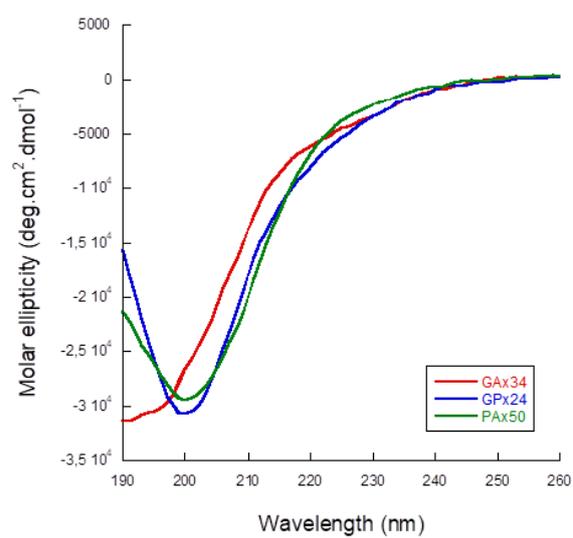


Figure 2

A



B

	α -helix (%)	β -sheet (%)	Other (%)
GAx34	10	1	89
GPx24	12	2	86
PAX50	20	2	78

Figure 3

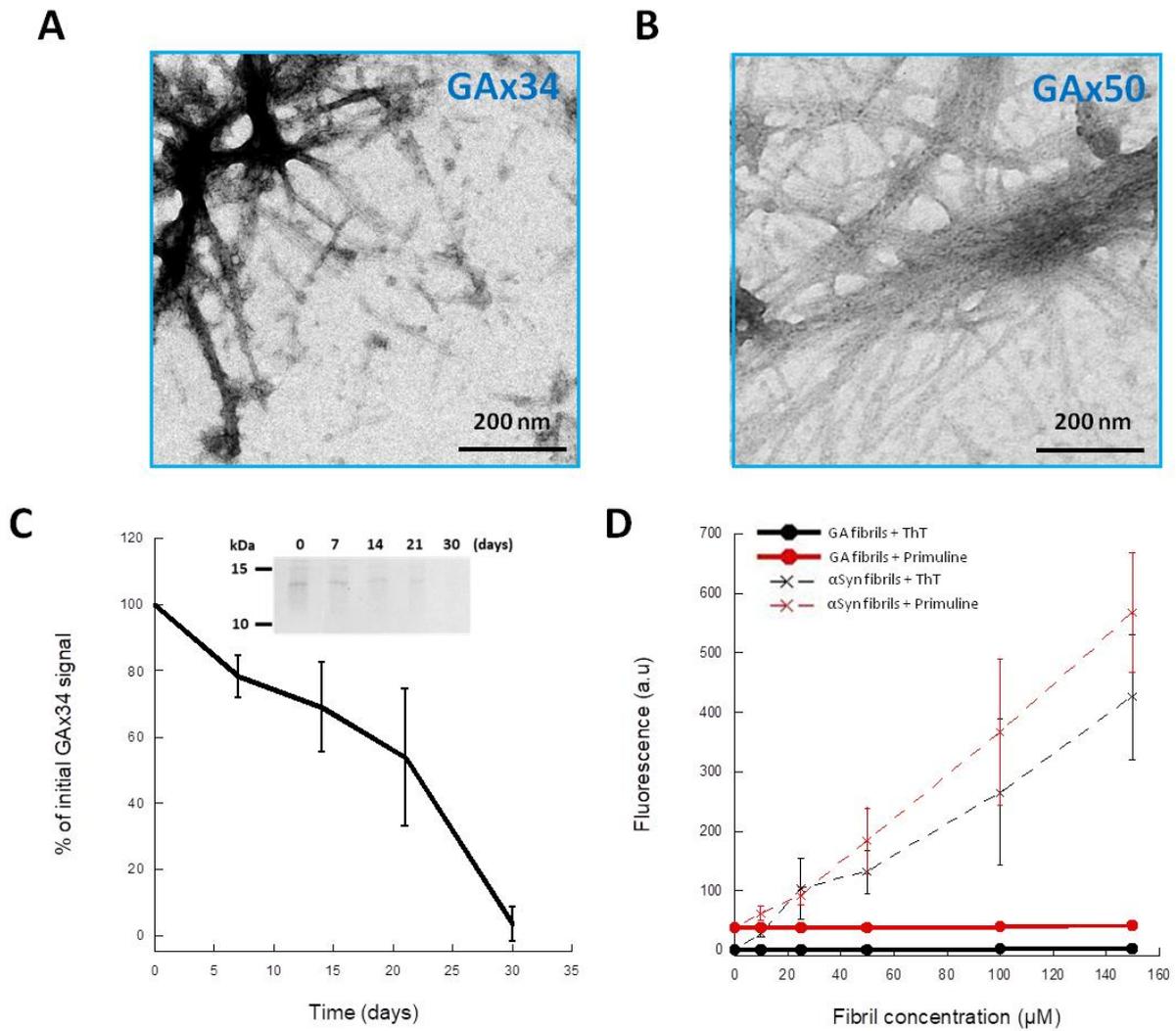
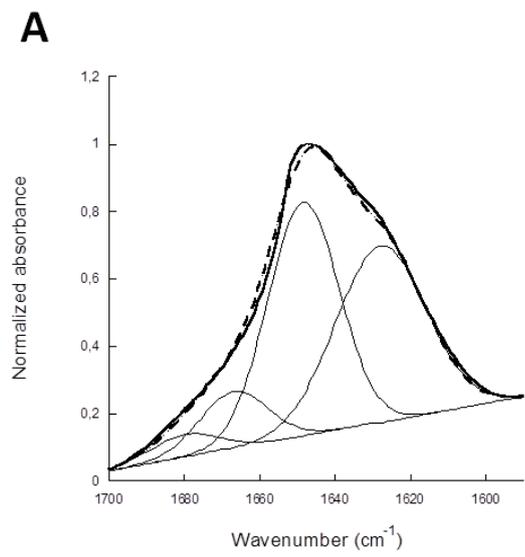


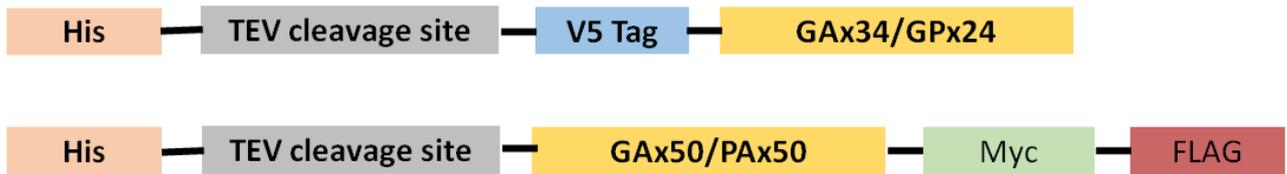
Figure 4



B

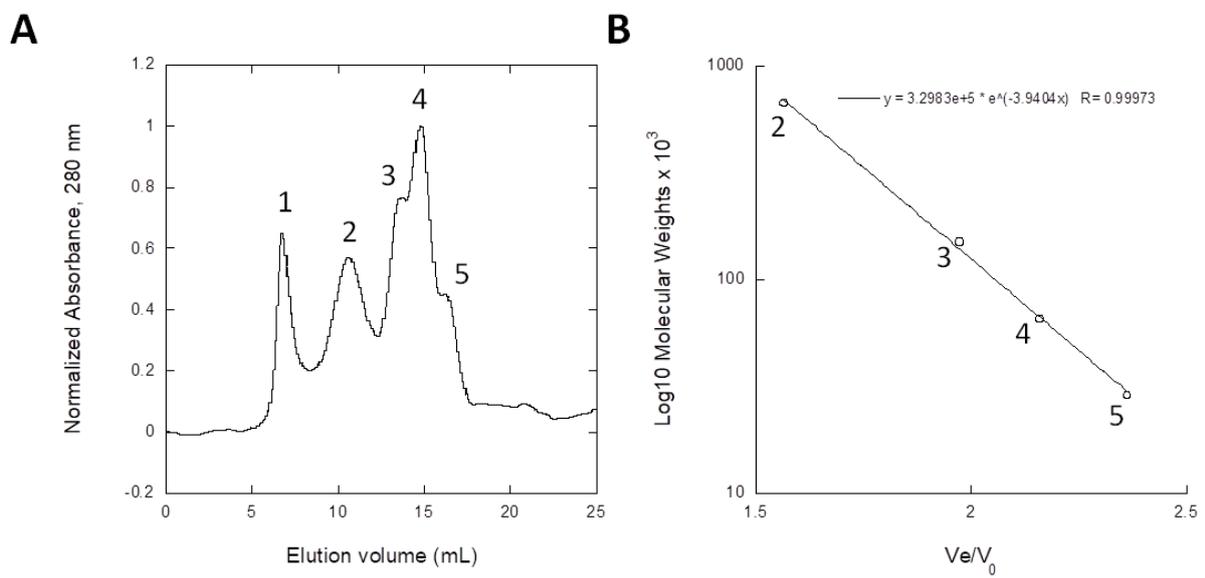
	Structure assignment	Area (%)	Number of peptide bonds
1628 cm ⁻¹	β-sheets	38	40
1680 cm ⁻¹	β-sheets	4	4
1667 cm ⁻¹	Turns	10	11
1649 cm ⁻¹	Disordered and α-helix	48	52

Supplemental Figure 1



DPR	Mw (kDa)
GPx24	9,2
GAx34	9,0
GAx50	12,3
PAx50	14,3

Supplemental Figure 2



II – Propriétés cellulaires de protéines à répétitions de dipeptides et de leurs assemblages

Ce travail est le fruit d'une collaboration avec l'équipe du Professeur Mimoun Azzouz. Les expériences ont été réalisées par Paolo Marchi, étudiant en thèse, à partir d'échantillons protéiques que j'ai préparé. Ce travail est en cours, nous ne présenterons ici qu'une partie des résultats obtenus.

Objectif de l'étude

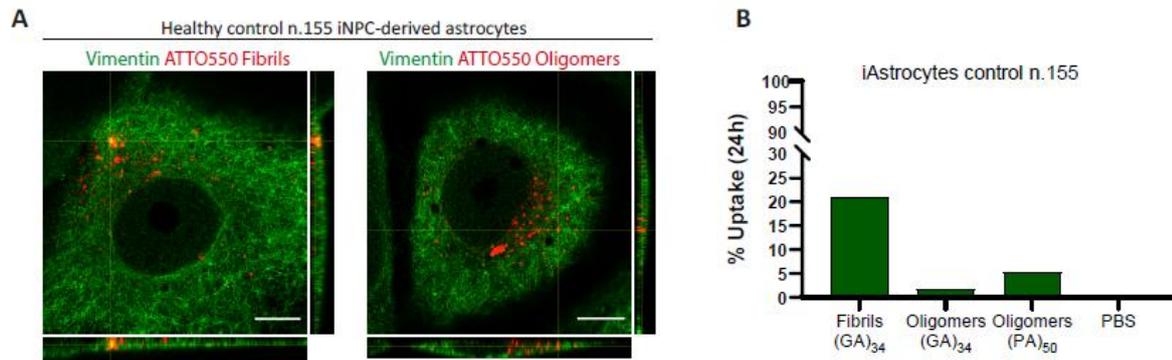
La part de mon travail dans ce projet a été de produire différentes Protéines à Répétitions de Dipeptides (PRD) et de réaliser leurs assemblages. J'ai marqué ces protéines avec l'Atto-550, un composé fluorescent qui se lie aux amines primaires. A partir de ce matériel, Paolo Marchi explore les différentes propriétés cellulaires de ces PRD. L'internalisation, la localisation intracellulaire et la toxicité de ces protéines sont testées dans différents modèles cellulaires.

Préparation et marquage des Protéines à Répétitions de Dipeptides (PRD) et de leurs assemblages

Une séquence d'ADN encodant pour une protéine possédant un tag V5 et 34 répétitions du dipeptide GA a été sous clonée par Audrey Coens et Luc Bousset dans un vecteur d'expression bactérien possédant un tag 6xHis, suivi d'un site de clivage à la protéase TEV en N-terminal (pETM-11 vector, EMBL). Un plasmide pAG416-Gal encodant pour une protéine contenant 50 répétitions du dipeptide PA, un TAG 6xHis et un site de clivage à la TEV en N-terminal, les TAG Myc et FLAG en C-terminal (Addgene) a également été utilisé. Les deux protéines ont été exprimées dans *E. coli*, puis purifiées sur une colonne Talon de 5 mL (Clontech®) chargé avec du Cobalt. Les protéines ont été éluées avec un gradient linéaire depuis 12 mL d'un tampon A (20mM Tris pH 7.5, 250mM NaCl, 5mM Imidazole, 1mM β Mercaptoethanol, Glycerol 10% PMSF 0.1 mM) jusqu'à un

tampon B (20mM Tris pH 7.5, 250mM NaCl, 250mM Imidazole, 1mM β Mercaptoethanol, Glycerol 10% PMSF 0.1 mM). Les fractions éluées ont été analysées sur gel SDS-PAGE et la quantité de protéines a été déterminée par spectrophotométrie en utilisant les coefficients suivants : $\epsilon_{\text{GAX34}} = 2980 \text{ M}^{-1}\text{cm}^{-1}$ et $\epsilon_{\text{PAX50}} = 4470 \text{ M}^{-1}\text{cm}^{-1}$. Le peptide GAX34 a été assemblé en fibres pendant 30 jours à 4°C sans agitation. Les fibres ont ensuite été centrifugées à 100 000 g, 30 min à 4°C, puis lavées avant d'être marquées à l'Atto-550 (Atto-Tec #AD550-35) pendant 1h à température ambiante à un ratio molaire Atto:GA = 1:1. Le marqueur non-lié a été retiré par centrifugation à 100 000 g, 30 min à 4°C. Les assemblages marqués ont été soniqués pendant 5 min et congelés à l'azote liquide avant d'être stockés à -80°C. Les préparations non-fibrillaires de GAX34 et de PAX50 ont été solubilisées 24h en HFIP. L'HFIP a ensuite été évaporé avant de reprendre les peptides dans un tampon PBS contenant l'Atto-550. Le marquage a été réalisé dans les mêmes conditions que les fibres. L'Atto-550 non-lié a été neutralisé par ajout de 1mM de Tris. Ces préparations sont considérées comme des oligomères de petites tailles (voir article).

Les protéines à répétitions de dipeptides sont internalisées par des cellules souches pluripotentes reprogrammées en astrocytes (iAstrocytes)



Les iAstrocytes ont été exposés pendant 24h à 1 μ M d'assemblages oligomériques ou fibrillaires des protéines à répétitions de dipeptides (PRD) GAx34 et PAx50. Après plusieurs lavages avec du PBS pour retirer les assemblages du milieu, les cellules ont été fixées et le réseau de vimentine ou CD44 décoré par des anticorps spécifiques (vert). (A) L'internalisation des PRD est montrée par microscopie super-résolutive (Zeiss LSM 880 Airyscan) en utilisant des vues orthogonales d'images en Z-stacking de cellules marquées pour la vimentine (n = 2). (B) L'internalisation a ensuite été quantifiée par microscopie confocale (Perkin Elmer Opera Phenix) en utilisant le logiciel Columbus sur des cellules marquées pour CD44. Sur le graphique est montré le pourcentage d'iAstrocytes ayant internalisés au moins un agrégat de PRD (n = 2). Echelle, 10 μ m sur chaque image.

Conclusions et perspectives

Ces données préliminaires indiquent que les assemblages de PAx50 et de GAx34 sont internalisés dans des iAstrocytes. Parmi les différents agrégats testés, les fibres de GAx34 sont ceux qui s'internalisent le plus, alors que les oligomères de ce même PRD ne pénètrent que très peu dans les cellules. Cette observation pourrait représenter l'une des voies possibles parmi lesquelles les PRD se propagent à travers le système nerveux des patients atteints de Sclérose Latérale Amyotrophique ou de Démence Fronto-Temporale.

Ces résultats devront être comparés avec ceux obtenus sur d'autres types

cellulaires, comme les neurones.

Les oligomères et les fibres de GAx34 sont observables au niveau du cytoplasme et à la périphérie du noyau. Paolo Marchi a montré qu'une partie des assemblages internalisés sont retrouvés dans les lysosomes. Dans quel autre compartiment intracellulaire ils sont localisés reste à être établi.

La toxicité de ces PRD et de leurs assemblages est en cours d'étude dans différents modèles cellulaires.

Conclusions et perspectives

I – Conclusions

Durant ma thèse, j'ai pu :

- (1) Produire plusieurs PRD (GPx24, GAx34, GAx50, et PAX50) et déterminer leur état d'oligomérisation et leur structure en solution.
- (2) Etablir leurs conditions d'agrégation et démontrer la nature amyloïde des fibres de GAx34.
- (3) Marquer les oligomères de ces PRD et des fibres de GAx34 dans le cadre d'une collaboration avec l'équipe du Pr. Mimoun Azzouz de l'Université de Sheffield.

A – Production de PRD et caractérisation de leur structure en solution.

De nombreuses études de toxicité ont déjà été réalisées sur les PRD dans des modèles cellulaires ou animaux. Mais à ce jour, seulement deux équipes se sont intéressées aux propriétés *in vitro* des PRD (Chang et al., 2016 ; Flores et al., 2016).

Le peu de données obtenues sur ce sujet a motivé le travail de ma thèse. Notre objectif était de produire des PRD *in vitro* afin d'en caractériser les propriétés biochimiques.

Audrey Coens et Luc Bousset ont produit par sous-clonage dans un vecteur bactérien deux plasmides contenant les séquences codantes pour les protéines GAx34 et GPx24. Parallèlement, nous avons obtenu d'autres plasmides commerciaux contenant les séquences des PRD GAx50 et PAX50.

Ces protéines ont été exprimées dans des bactéries *E. coli*. Leur purification a été réalisée par FPLC.

Seules les propriétés biochimiques des protéines GAx34, GPx24 et PAX50 ont pu

être explorées. La protéine GAx50 a une solubilité trop faible.

L'état d'oligomérisation des PRD GAx34, GPx24 et PAx50 a pu être déterminé par chromatographie d'exclusion de taille. Les valeurs déterminées expérimentalement pour les trois protéines sont cohérentes avec des protéines monomériques et des oligomères de petites tailles (inférieur à 10 monomères). Nous avons également déterminé la structure secondaire de ces PRD en étudiant leur spectre de dichroïsme circulaire. Ces spectres correspondent à ceux de protéines dont la séquence est majoritairement désorganisée.

B – Conditions d'agrégation et caractérisations des assemblages de PRD

En faisant varier la composition du tampon, la température et l'agitation, j'ai réussi à établir des conditions d'incubation pour lesquelles les PRD s'agrégeaient. La formation d'agrégats a été vérifiée par sédimentation. Leur morphologie a été observée par microscopie électronique.

Dans notre étude, les protéines poly-GA s'assemblent pour former des agrégats fibrillaires. Les feuilletts β qui se forment dans ces agrégats n'impliquent qu'une partie des dipeptides répétés. Sa séquence étant uniforme, il est possible que les peptides poly-GA s'assemblent avec un décalage entre chaque monomère le long de la fibre. De plus, pour la même concentration, l'assemblage du peptide GAx50 s'effectue en moins d'une heure, alors que GAx34 forme des fibres après 30 jours d'incubation. La vitesse d'agrégation de poly-GA est donc proportionnelle au nombre de dipeptides. Cette observation est complémentaire de précédentes observations réalisées à propos de l'importance du nombre de répétitions sur l'apparition de la maladie et l'agrégation des PRD (Gijssels et

al., 2016 ; Herranz-Martin et al., 2017 Nonaka et al., 2018 ; Chang et al., 2016 ; Flores et al., 2016).

Les agrégats de GAx34 montrent une signature amyloïde par spectroscopie infrarouge. De façon surprenante, ces fibres ne lient pas la Thioflavin T, ni la Primuline – des molécules fluorescentes pourtant utilisées pour marquer les assemblages amyloïdes. Une étude montrant l'importance des résidus leucine et tyrosine dans la fixation de ces ligands aux amyloïdes pourrait expliquer ce résultat (Biancalana et al., 2009 ; Wu et al., 2009).

Sur la liaison de ces composés, nos résultats diffèrent de ceux obtenus précédemment sur des poly-GA contenant 3, 6 ou 15 répétitions (Chang et al., 2016 ; Flores et al., 2016).

Il est important de souligner que les protéines poly-GA que nous avons produites ont un nombre de répétitions de dipeptides supérieur à celles des protéines étudiées dans les autres études *in vitro*. De plus, la température à laquelle nous les avons assemblé est de 4° et non 37°C. Les assemblages que nous avons produits ont peut-être une structure différente de celles des autres études. Cela pourrait expliquer les différences de marquage. Cette hypothèse est appuyée par le fait que les agrégats obtenus à partir de 3 et 6 répétitions de GA ne sont pas fibrillaires (Flores et al., 2016), contrairement à ceux de notre étude.

Dans les conditions que nous avons testées, les protéines GPx24 et PAX50 forment peu d'assemblages. Les agrégats observés par microscopie électronique ont une morphologie amorphe. Ainsi, parmi les PRD étudiés, seul poly-GA s'assemble pour former des fibres amyloïdes. Ce résultat est cohérent avec les précédents travaux réalisés (Chang et al., 2016). Ces observations pourraient expliquer pourquoi le poly-GA est le PRD qu'on retrouve le plus abondamment présent sous forme d'inclusions dans les

neurones de patients SLA et DFT (Mori et al., 2013a ; Zu et al., 2013 ; Gendron et al., 2013).

Poly-GP et poly-PA sont aussi retrouvés sous forme insoluble dans le cerveau des patients. Nos résultats *in vitro* ne permettent pas de véritablement conclure sur cette observation. Nous pouvons seulement avancer les hypothèses suivantes :

- (1) Les protéines poly-GP et poly-PA ne peuvent pas former d'assemblages fibrillaires amyloïdes. Cette hypothèse est appuyée par la présence de prolines dans leur séquence. Ces résidus sont connus pour interrompre la formation des feuillets β .
- (2) D'autres conditions physico-chimiques d'incubation sont peut-être requises pour l'agrégation de ces PRD.
- (3) Dans notre étude le nombre de répétitions des dipeptides est égal ou inférieure à 50, alors que dans un contexte pathologique il peut atteindre plusieurs milliers. Il est possible qu'en deçà d'un certain nombre de répétitions, des protéines comme poly-GP ou poly-PA ne s'agrègent pas ou trop lentement. C'est le cas pour l'assemblage des polyglutamines. En effet, en dessous de 37 résidus l'assemblage de ces peptides est incomplet et nettement plus lent que ceux contenant plus de 40 glutamines (Monsellier et al., 2015).
- (4) Dans le cerveau, plusieurs PRD co-localisaient parfois. De plus, ces inclusions contiennent toujours d'autres protéines. On peut ainsi supposer que certains PRD ne peuvent s'agréger seul.

C – Marquage des assemblages de PRD dans le cadre d'une collaboration

Au cours de ma thèse, j'ai également préparé des échantillons de PRD dans un état

oligomérique ou fibrillaire que j'ai marqué par l'Atto-550 – fluorophore qui se lie aux amines primaires. Ces préparations ont été réalisées dans le cadre d'une collaboration avec l'équipe du Pr. Mimoun Azzouz de l'Université de Sheffield. Cette étude a pour objectif de déterminer les propriétés des PRD telles que l'internalisation, la propagation et la toxicité dans des modèles cellulaires. Paolo Marchi, doctorant, réalise ce travail.

Il a notamment pu montrer que les fibres de GAx34 étaient davantage internalisées que les oligomères de GAx34 et de PAX50 dans des cellules souches pluripotentes reprogrammées en astrocytes. Cette observation pourrait représenter l'une des voies de propagation des PRD dans les maladies SLA et DFT. Il a également observé qu'une partie de ces assemblages sont localisés dans les lysosomes de la cellule.

II – Perspectives

Nos travaux montrent que poly-GA forme des assemblages fibrillaires de nature amyloïde. La résolution de la structure tridimensionnelle de ces agrégats par cryo-EM pourrait nous renseigner davantage sur l'assemblage de ces fibres.

Dans notre étude, poly-GP et poly-PA ne formaient que peu d'agrégats, d'aspects amorphes. Nous ne pouvons pas conclure sur la nature des inclusions cytoplasmiques marquées avec des anticorps dirigés contre ces deux PRD. Il est possible que ces protéines s'assemblent avec d'autres partenaires, comme c'est le cas des poly-PR et poly-GR qui co-agrégent avec des protéines nucléaires. L'analyse de la composition des inclusions cytoplasmiques contenant poly-GP et poly-PA pourrait nous renseigner sur ce point.

Afin de compléter nos travaux, il faudrait pouvoir produire et caractériser les PRD contenant des arginines, qui sont les seuls absents de notre étude. Au cours de cette thèse, j'ai produit un peptide PRx50. Ce peptide n'était pas reconnu par son anticorps et ne formait aucun agrégat pour les différentes conditions d'incubation explorées.

Le travail effectué par Paolo Marchi de l'équipe du Pr. Mimoun Azzouz a pu mettre en évidence l'internalisation des fibres de poly-GA et des oligomères de poly-PA. Ces observations effectuées sur les iAstrocytes seront comparées à des expériences réalisées sur d'autres modèles cellulaires. La détermination de la toxicité, de la propagation et de la dégradation des PRD et de leurs assemblages est en cours.

Bibliographie

1. Al-Chalabi, A., & Hardiman, O. (2013). The epidemiology of ALS: A conspiracy of genes, environment and time. *Nature Reviews Neurology*. Nature Publishing Group.
2. Alzheimer, A. (1991). Über eigenartige Krankheitsfälle des späteren Alters: (On certain peculiar diseases of old age. *History of Psychiatry*, 2(5), 74–101.
3. Amick, J., Rocznik-Ferguson, A., & Ferguson, S. M. (2016). C9orf72 binds SMCR8, localizes to lysosomes, and regulates mTORC1 signaling. *Molecular Biology of the Cell*, 27(20), 3040–3051.
4. Andreadis, A., Broderick, J. A., & Kosik, K. S. (1995). Relative exon affinities and suboptimal splice site signals lead to non-equivalence of two cassette exons. *Nucleic Acids Research*, 23(17), 3585–3593.
5. Aoki, Y., Manzano, R., Lee, Y., Dafinca, R., Aoki, M., Douglas, A. G. L., ... Wood, M. J. A. (2017). C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia. *Brain*, 140(4), 887–897.
6. Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., ... Oda, T. (2006). TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochemical and Biophysical Research Communications*, 351(3), 602–611.
7. Ash, P. E. A., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W. L., DeJesus-Hernandez, M., ... Petrucelli, L. (2013). Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron*, 77(4), 639–646.
8. Atanasio, A., Decman, V., White, D., Ramos, M., Ikiz, B., Lee, H. C., ... Lai, K. M. V. (2016). C9orf72 ablation causes immune dysregulation characterized by leukocyte

- expansion, autoantibody production, and glomerulonephropathy in mice. *Scientific Reports*, 6.
9. Baborie, A., Griffiths, T. D., Jaros, E., Perry, R., McKeith, I. G., Burn, D. J., ... Mann, D. M. A. (2015). Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. *Neuropathology and Applied Neurobiology*, 41(5), 601–612.
 10. Barmada, S. J., Skibinski, G., Korb, E., Rao, E. J., Wu, J. Y., & Finkbeiner, S. (2010). Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 30(2), 639–49.
 11. Barrow, C. J., Yasuda, A., Kenny, P. T. M., & Zagorski, M. G. (1992). Solution conformations and aggregational properties of synthetic amyloid β -peptides of Alzheimer's disease. Analysis of circular dichroism spectra. *Journal of Molecular Biology*, 225(4), 1075–1093.
 12. Bäumer, D., Hilton, D., Paine, S. M. L., Turner, M. R., Lowe, J., Talbot, K., & Ansorge, O. (2010). Juvenile ALS with basophilic inclusions is a FUS proteinopathy with FUS mutations. *Neurology*, 75(7), 611–618.
 13. Bell C. *The nervous system of the human body*. London: Longman ; 1830. P. 132-6 (Special Edition, copyright Classics of Neurology and Neurosurgery Library, Birmingham Alabama 1985).
 14. Belzil, V. V., Langlais, J. S., Daoud, H., Dion, P. A., Brais, B., & Rouleau, G. A. (2012). Novel FUS deletion in a patient with juvenile amyotrophic lateral sclerosis. *Archives of Neurology*, 69(5), 653–656.

15. Biancalana, M., Makabe, K., Koide, A., & Koide, S. (2009). Molecular Mechanism of Thioflavin-T Binding to the Surface of β -Rich Peptide Self-Assemblies. *Journal of Molecular Biology*, 385(4), 1052–1063.
16. Billingsley, M. L., Kincaid, R. L. (1997). Regulated phosphorylation and dephosphorylation of tau protein: effects on microtubule interaction, intracellular trafficking and neurodegeneration. *Biochemical Journal*, 323(3), 577–591.
17. Binder, J. R. (2017, August 1). Current Controversies on Wernicke's Area and its Role in Language. *Current Neurology and Neuroscience Reports*. Current Medicine Group LLC 1.
18. Boeynaems, S., Bogaert, E., Michiels, E., Gijssels, I., Sieben, A., Jovičić, A., ... Van Den Bosch, L. (2016). Drosophila screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Scientific Reports*, 6.
19. Bolton, D.C., McKinley, M.P., and Prusiner, S.B. (1982). Identification of a protein that purifies with the scrapie prion. *Science* 218, 1309–1311.
20. Bjørkøy, G., Lamark, T., & Johansen, T. (2006). p62/SQSTM1: A missing link between protein aggregates and the autophagy machinery. *Autophagy*, 2(2), 138–139.
21. Bousset, L., Pieri, L., Ruiz-Arlandis, G., Gath, J., Jensen, P. H., Habenstein, B., ... Melki, R. (2013). Structural and functional characterization of two alpha-synuclein strains. *Nature Communications*, 4.
22. Bradley, W. G., & Cox, P. A. (2009). Beyond Guam: Cyanobacteria, BMAA and sporadic amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis*, 10(sup2), 5–6.
23. Brooks B.R., Miller R.G., Swash M. Munsat T.L. (2000) World Federation of Neurology Research Group on Motor Neuron Diseases : El Escorial revisited : revised criteria for diagnosis of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Other Motor Neuron Disord.*, 1 : 293-239.

24. Brundin, P., Melki, R., & Kopito, R. (2010, April). Prion-like transmission of protein aggregates in neurodegenerative diseases. *Nature Reviews Molecular Cell Biology*.
25. Buratti, E., & Baralle, F. E. (2012, June). TDP-43: Gumming up neurons through protein-protein and protein-RNA interactions. *Trends in Biochemical Sciences*. Calvio, C., Lamond, A. I., Neubauer, G., & Mann, M. (1995). Identification of hnRNP P2 as TLS/FUS using electrospray mass spectrometry. *RNA*, 1(7), 724–733.
26. Cervenakova, L., Protas, I. I., Hirano, a, Votiakov, V. I., Nedzved, M. K., Kolomiets, N. D., ... Goldfarb, L. G. (2000). Progressive muscular atrophy variant of familial amyotrophic lateral sclerosis (PMA/ALS). *Journal of the Neurological Sciences*, 177(2), 124–130.
27. Chang, J. L., Lomen-Hoerth, C., Murphy, J., Henry, R. G., Kramer, J. H., Miller, B. L., & Gorno-Tempini, M. L. (2005). A voxel-based morphometry study of patterns of brain atrophy in ALS and ALS/FTLD. *Neurology*, 65(1), 75–80.
28. Chattopadhyay, M., Durazo, A., Sohn, S. H., Strong, C. D., Gralla, E. B., Whitelegge, J. P., & Valentine, J. S. (2008). Initiation and elongation in fibrillation of ALS-linked superoxide dismutase. *Proceedings of the National Academy of Sciences*, 105(48), 18663–18668.
29. Chen, S., Berthelie, V., Hamilton, J. B., O’Nuallain, B., & Wetzel, R. (2002). Amyloid-like features of polyglutamine aggregates and their assembly kinetics. *Biochemistry*, 41(23), 7391–7399.
30. Chen, J. H., Lin, K. P., & Chen, Y. C. (2009). Risk factors for dementia. *Journal of the Formosan Medical Association*. Elsevier B.V.
31. Chiò, A., Logroscino, G., Hardiman, O., Swingler, R., Mitchell, D., Beghi, E., & Traynor, B. G. (2009). Prognostic factors in ALS: A critical review. *Amyotrophic Lateral Sclerosis*.

32. Ciani, M., Benussi, L., Bonvicini, C., & Ghidoni, R. (2019). Genome Wide Association Study and Next Generation Sequencing: A Glimmer of Light Toward New Possible Horizons in Frontotemporal Dementia Research. *Frontiers in Neuroscience*, 13(MAI).
33. Cirulli, E. T., Lasseigne, B. N., Petrovski, S., Sapp, P. C., Dion, P. A., Leblond, C. S., ... Muñoz-Blanco, J. L. (2015). Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science*, 347(6229), 1436–1441.
34. Ciura, S., Lattante, S., Le Ber, I., Latouche, M., Tostivint, H., Brice, A., & Kabashi, E. (2013). Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Annals of Neurology*, 74(2), 180–187.
35. Chang, Y. J., Jeng, U. S., Chiang, Y. L., Hwang, I. S., & Chen, Y. R. (2016). The glycine-alanine dipeptide repeat from C9 or f72 hexanucleotide expansions forms toxic amyloids possessing cell-to-cell transmission properties. *Journal of Biological Chemistry*, 291(10), 4903–4911.
36. Chen, H. J., Anagnostou, G., Chai, A., Withers, J., Morris, A., Adhikaree, J., ... De Bellerocche, J. S. (2010). Characterization of the properties of a novel mutation in VAPB in familial amyotrophic lateral sclerosis. *Journal of Biological Chemistry*, 285(51), 40266–40281.
37. Chia, R., Tattum, M. H., Jones, S., Collinge, J., Fisher, E. M. C., & Jackson, G. S. (2010). Superoxide dismutase 1 and tgSOD1G93A mouse spinal cord seed fibrils, suggesting a propagative cell death mechanism in amyotrophic lateral sclerosis. *PLoS ONE*, 5(5).
38. Chin, S. S., & Goldman, J. E. (1996). Glial Inclusions in CNS Degenerative Diseases. *Journal of Neuropathology and Experimental Neurology*, 55(5), 499–508.
39. Clark, L. N., Poorkaj, P., Wszolek, Z., Geschwind, D. H., Nasreddine, Z. S., Miller, B., ... Wilhelmsen, K. C. (1998). Pathogenic implications of mutations in the tau

- gene in pallido-ponto-nigral degeneration and related neurodegenerative disorders linked to chromosome 17. *Proceedings of the National Academy of Sciences of the United States of America*, 95(22), 13103–7.
40. Clavaguera, F., Bolmont, T., Crowther, R. A., Abramowski, D., Frank, S., Probst, A., ... Tolnay, M. (2009). Transmission and spreading of tauopathy in transgenic mouse brain. *Nature Cell Biology*, 11(7), 909–913.
41. Cleveland, D. W., Hwo, S. Y., & Kirschner, M. W. (1977). Purification of tau, a microtubule-associated protein that induces assembly of microtubules from purified tubulin. *Journal of Molecular Biology*, 116(2), 207–225.
42. Colombrita, C., Onesto, E., Megjorni, F., Pizzuti, A., Baralle, F. E., Buratti, E., ... Ratti, A. (2012). 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. *Journal of Biological Chemistry*, 287(19), 15635–15647.
43. Cooper-Knock, J., Walsh, M. J., Higginbottom, A., Highley, J. R., Dickman, M. J., Edbauer, D., ... Shaw, P. J. (2014). Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain*, 137(7), 2040–2051.
44. Conway, K. A., Harper, J. D., & Lansbury, P. T. (2000). Fibrils formed in vitro from α -synuclein and two mutant forms linked to Parkinson's disease are typical amyloid. *Biochemistry*, 39(10), 2552–2563.
45. Couthouis, J., Hart, M. P., Shorter, J., DeJesus-Hernandez, M., Erion, R., Oristano, R., ... Gitler, A. D. (2011). A yeast functional screen predicts new candidate ALS disease genes. *Proceedings of the National Academy of Sciences*, 108(52), 20881–20890.
46. Couthouis, J., Hart, M. P., Erion, R., King, O. D., Diaz, Z., Nakaya, T., ... Gitler, A. D. (2012). Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. *Human Molecular Genetics*, 21(13), 2899–2911.

47. Cudkovicz, M. E., McKenna-Yasek, D., Chen, C., Hedley-Whyte, E. T., & Brown, R. H. (1998). Limited corticospinal tract involvement in amyotrophic lateral sclerosis subjects with the A4V mutation in the copper/zinc superoxide dismutase gene. *Annals of Neurology*, 43(6), 703–710.
48. Culler, A. (1906) Troubles Mentaux dans la sclérose latérale amyotrophique. *Arch, de Neurol*, 21 : 433-450.
49. D'Amico, E., Pasmantier, M., Lee, Y. W., Weimer, L., & Mitsumoto, H. (2013). Clinical evolution of pure upper motor neuron disease/dysfunction (PUMMD). *Muscle and Nerve*, 47(1), 28–32.
50. DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., ... Rademakers, R. (2011). Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*, 72(2), 245–256
51. Del Aguila, M. A., Longstreth, W. T., McGuire, V., Koepsell, T. D., & Van Belle, G. (2003). Prognosis in amyotrophic lateral sclerosis: A population-based study. *Neurology*, 60(5), 813–819.
52. Delay J., Brion S., Escourolle R., Marty R. (1959) Sclérose latérale amyotrophique et démence (a propos de deux cas anatomo-cliniques). *Rev. Neurol.*, 100 :191-204.
53. Deng, H.-X., Shi, Y., Furukawa, Y., Zhai, H., Fu, R., Liu, E., ... Siddique, T. (2006). Conversion to the amyotrophic lateral sclerosis phenotype is associated with intermolecular linked insoluble aggregates of SOD1 in mitochondria. *Proceedings of the National Academy of Sciences*, 103(18), 7142–7147.
54. Dormann, D., Rodde, R., Edbauer, D., Bentmann, E., Fischer, I., Hruscha, A., ... Haass, C. (2010). ALS-associated fused in sarcoma (FUS) mutations disrupt transportin-mediated nuclear import. *EMBO Journal*, 29(16), 2841–2857.

55. Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ (2004). "Lesion analysis of the brain areas involved in language comprehension". *Cognition*. 92 (1-2): 145–77.
56. Duyao, M., Ambrose, C., Myers, R., Novelletto, A., Persichetti, F., Frontali, M., ... Macdonald, M. (1993). Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nature Genetics*, 4(4), 387–392.
57. Elam, J. S., Taylor, A. B., Strange, R., Antonyuk, S., Doucette, P. A., Rodriguez, J. A., ... Hart, P. J. (2003). Amyloid-like filaments and water-filled nanotubes formed by SOD1 mutant proteins linked to familial ALS. *Nature Structural Biology*, 10(6), 461–467
58. Elden, A. C., Kim, H. J., Hart, M. P., Chen-Plotkin, A. S., Johnson, B. S., Fang, X., ... Gitler, A. D. (2010). Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*, 466(7310), 1069–1075.
59. Farg, M. A., Sundaramoorthy, V., Sultana, J. M., Yang, S., Atkinson, R. A. K., Levina, V., ... Atkin, J. D. (2014). C9ORF72, implicated in amyotrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Human Molecular Genetics*, 23(13), 3579–3595.
60. Fersht, A. (1999). Structure and mechanism in protein science. Lavoisier.Fr (Vol. 13409, pp. 54–101). Freeman, New York.
61. Ferguson, J. H., & Boller, F. (1977). A different form of agraphia: Syntactic writing errors in patients with motor speech and movement disorders. *Brain and Language*, 4(3), 382–389.
62. Finger, E. C. (2016, April 1). Frontotemporal dementias. *CONTINUUM Lifelong Learning in Neurology*. Lippincott Williams and Wilkins.

63. Flores, B. N., Dulchavsky, M. E., Krans, A., Sawaya, M. R., Paulson, H. L., Todd, P. K., ... Ivanova, M. I. (2016). Distinct c9orf72-associated dipeptide repeat structures correlate with neuronal toxicity. *PLoS ONE*, 11(10).
64. Forsberg, K., Graffmo, K., Pakkenberg, B., Weber, M., Nielsen, M., Marklund, S., ... Andersen, P. M. (2019). Misfolded SOD1 inclusions in patients with mutations in C9orf72 and other ALS/FTD-associated genes. *Journal of Neurology, Neurosurgery and Psychiatry*.
65. Frank, B., Haas, J., Heinze, H. J., Stark, E., & Münte, T. F. (1997). Relation of neuropsychological and magnetic resonance findings in amyotrophic lateral sclerosis: Evidence for subgroups. *Clinical Neurology and Neurosurgery*, 99(2), 79–86.
66. Freeman, D., Cedillos, R., Choyke, S., Lukic, Z., McGuire, K., Marvin, S., ... Campbell, E. M. (2013). Alpha-Synuclein Induces Lysosomal Rupture and Cathepsin Dependent Reactive Oxygen Species Following Endocytosis. *PLoS ONE*, 8(4).
67. Frick, P., Sellier, C., Mackenzie, I. R. A., Cheng, C. Y., Tahraoui-Bories, J., Martinat, C., ... Neumann, M. (2018). Novel antibodies reveal presynaptic localization of C9orf72 protein and reduced protein levels in C9orf72 mutation carriers. *Acta Neuropathologica Communications*, 6(1), 72.
68. Frost, B., Jacks, R. L., & Diamond, M. I. (2009). Propagation of Tau misfolding from the outside to the inside of a cell. *Journal of Biological Chemistry*, 284(19), 12845–12852.
69. Furukawa, Y., Kaneko, K., Watanabe, S., Yamanaka, K., & Nukina, N. (2011). A seeding reaction recapitulates intracellular formation of sarkosyl-insoluble transactivation response element (TAR) DNA-binding protein-43 inclusions. *Journal of Biological Chemistry*, 286(21), 18664–18672.
70. Gazit, E. (2002). A possible role for π -stacking in the self-assembly of amyloid fibrils. *FASEB Journal*.

71. Gendron, T. F., Bieniek, K. F., Zhang, Y. J., Jansen-West, K., Ash, P. E. A., Caulfield, T., ... Petrucelli, L. (2013). Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathologica*, 126(6), 829–844.
72. Gijssels, I., Van Mossevelde, S., Van Der Zee, J., Sieben, A., Engelborghs, S., De Bleecker, J., ... Van Broeckhoven, C. (2016). The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Molecular Psychiatry*, 21(8), 1112–1124.
73. Glickman, M. H., & Ciechanover, A. (2002). The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction.
74. Goedert, M., & Jakes, R. (1990). Expression of separate isoforms of human tau protein: correlation with the tau pattern in brain and effects on tubulin polymerization. *The EMBO Journal*, 9(13), 4225–30
75. Goetz, C. G. (2000). Amyotrophic lateral sclerosis: Early contributions of Jean-Martin Charcot. *Muscle and Nerve*
76. Goldberg, A.L. (2003). Protein degradation and protection against misfolded or damaged proteins. *Nature* 426, 895–899.
77. Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., ... Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006–1014.
78. Grad, L. I., Rouleau, G. A., Ravits, J., & Cashman, N. R. (2017, August 1). Clinical Spectrum of Amyotrophic Lateral Sclerosis (ALS). *Cold Spring Harbor Perspectives in Medicine*.
79. Greenway, M. J., Andersen, P. M., Russ, G., Ennis, S., Cashman, S., Donaghy, C., ... Hardiman, O. (2006). ANG mutations segregate with familial and “sporadic” amyotrophic lateral sclerosis. *Nature Genetics*, 38(4), 411–413.

80. Guedj E., Le Ber I., Lacomblez L., Dubois B., Vepillat P., Didic M., Salachas F., Vera P., Hannequin D., Lotterie J.A., Puel M., Decousus M., Thomas-Antérion C., Magne C., Verceletto M., Bernard A. M., Golfier V., Pasquier J., Volteau M., Brice A., Meininger V. (2007) French Research Network on FTD/FTD-MND, Habert M.O. Brain spect perfusion of frontotemporal dementia associated with motor neuron disease. *Neurology*, 69 : 488-490.
81. Guo, J. L., & Lee, V. M. Y. (2011). Seeding of normal tau by pathological tau conformers drives pathogenesis of Alzheimer-like tangles. *Journal of Biological Chemistry*, 286(17), 15317–15331.
82. Guo, Q., Lehmer, C., Martínez-Sánchez, A., Rudack, T., Beck, F., Hartmann, H., ... Fernández-Busnadiego, R. (2018). In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. *Cell*, 172(4), 696-705.e12.
83. Gurney, M. E., Pu, H., Chiu, A. Y., Dal Canto, M. C., Polchow, C. Y., Alexander, D. D., ... Siddique, T. (1994). Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. *Science*, 264(5166), 1772–1775.
84. Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A. J., Shaw, P. G., Kim, M. S., ... Wang, J. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature*, 507(7491), 195–200.
85. Hao, Z., Liu, L., Tao, Z., Wang, R., Ren, H., Sun, H., ... Wang, G. (2019). Motor dysfunction and neurodegeneration in a C9orf72 mouse line expressing poly-PR. *Nature Communications*, 10(1), 2906.
86. Harada, A., Oguchi, K., Okabe, S., Kuno, J., Terada, S., Ohshima, T., ... Hirokawa, N. (1994). Altered microtubule organization in small-calibre axons of mice lacking tau protein. *Nature*, 369(6480), 488–491.

87. Hardiman, O., Al-Chalabi, A., Chio, A., Corr, E. M., Logroscino, G., Robberecht, W., ... Van Den Berg, L. H. (2017, October 5). Amyotrophic lateral sclerosis. *Nature Reviews Disease Primers*. Nature Publishing Group.
88. Hartl, F. U., Bracher, A., & Hayer-Hartl, M. (2011, July 21). Molecular chaperones in protein folding and proteostasis. *Nature*.
89. Hautbergue, G. M., Castelli, L. M., Ferraiuolo, L., Sanchez-Martinez, A., Cooper-Knock, J., Higginbottom, A., ... Shaw, P. J. (2017). SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. *Nature Communications*, 8.
90. He, H., Huang, W., Wang, R., Lin, Y., Guo, Y., Deng, J., ... Duan, R. (2019). Amyotrophic Lateral Sclerosis-associated GGGGCC repeat expansion promotes Tau phosphorylation and toxicity. *Neurobiology of Disease*, 130.
91. Herranz-Martin, S., Chandran, J., Lewis, K., Mulcahy, P., Higginbottom, A., Walker, C., ... Azzouz, M. (2017). Viral delivery of C9orf72 hexanucleotide repeat expansions in mice leads to repeat-length-dependent neuropathology and behavioural deficits. *DMM Disease Models and Mechanisms*, 10(7), 859–868.
92. Hicks, G. G., Singh, N., Nashabi, A., Mai, S., Bozek, G., Klewes, L., ... Ruley, H. E. (2000). *Fus* deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nature Genetics*, 24(2), 175–179.
93. Himmler, A., Drechsel, D., Kirschner, M. W., & Martin, D. W. (1989). Tau consists of a set of proteins with repeated C-terminal microtubule-binding domains and variable N-terminal domains. *Molecular and Cellular Biology*, 9(4), 1381–8.
94. Ho, W. Y., Tai, Y. K., Chang, J. C., Liang, J., Tyan, S. H., Chen, S., ... Ling, S. C. (2019). The ALS-FTD-linked gene product, C9orf72, regulates neuronal morphogenesis via autophagy. *Autophagy*, 15(5), 827–842.

95. Hoop, C. L., Lin, H. K., Kar, K., Magyarfalvi, G., Lamley, J. M., Boatz, J. C., ... Van Der Wel, P. C. A. (2016). Huntingtin exon 1 fibrils feature an interdigitated β -hairpin-based polyglutamine core. *Proceedings of the National Academy of Sciences of the United States of America*, 113(6), 1546–1551.
96. Hutton, M., Lendon, C. L., Rizzu, P., Baker, M., Froelich, S., Houlden, H. H., ... Heutink, P. (1998). Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*, 393(6686), 702–704.
97. Ichikawa, H., Koyama, S., Ohno, H., Ishihara, K., Nagumo, K., & Kawamura, M. (2008). Writing errors and anosognosia in amyotrophic lateral sclerosis with dementia. *Behavioural Neurology*, 19(3), 107–116.
98. Igaz, L. M., Kwong, L. K., Chen-Plotkin, A., Winton, M. J., Unger, T. L., Xu, Y., ... Lee, V. M. Y. (2009). Expression of TDP-43 C-terminal fragments in vitro recapitulates pathological features of TDP-43 proteinopathies. *Journal of Biological Chemistry*, 284(13), 8516–8524.
99. Ince, P. G., Evans, J., Knopp, M., Forster, G., Hamdalla, H. H. M., Wharton, S. B., & Shaw, P. J. (2003). Corticospinal tract degeneration in the progressive muscular atrophy variant of ALS. *Neurology*, 60(8), 1252–1258.
100. Ingre, C., Roos, P. M., Piehl, F., Kamel, F., & Fang, F. (2015, February 12). Risk factors for amyotrophic lateral sclerosis. *Clinical Epidemiology*. Dove Medical Press Ltd
101. Irwin, D. J., Cairns, N. J., Grossman, M., McMillan, C. T., Lee, E. B., Van Deerlin, V. M., ... Trojanowski, J. Q. (2015, April 1). Frontotemporal lobar degeneration: defining phenotypic diversity through personalized medicine. *Acta Neuropathologica*. Springer Verlag.
102. Iwanaga, K., Hayashi, S., Oyake, M., Horikawa, Y., Hayashi, T., Wakabayashi, M., ... Takahashi, H. (1997). Neuropathology of sporadic amyotrophic lat-

- eral sclerosis of long duration. *Journal of the Neurological Sciences*, 146(2), 139–143.
103. Iyer, S., Subramanian, V., & Acharya, K. R. (2018). C9orf72, a protein associated with amyotrophic lateral sclerosis (ALS) is a guanine nucleotide exchange factor. *PeerJ*, 6, e5815.
104. Jackson, M., Lennox, G., & Lowe, J. (1996). Motor neurone disease-inclusion dementia. *Neurodegeneration*, 5(4), 339–350.
105. Ji, A. L., Zhang, X., Chen, W. W., & Huang, W. J. (2017, March 1). Genetics insight into the amyotrophic lateral sclerosis/frontotemporal dementia spectrum. *Journal of Medical Genetics*.
106. Jiang, J., Zhu, Q., Gendron, T. F., Saberi, S., McAlonis-Downes, M., Seelman, A., ... Lagier-Tourenne, C. (2016). Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron*, 90(3), 535–550.
107. Johnston, J. A., Dalton, M. J., Gurney, M. E., & Kopito, R. R. (2000). Formation of high molecular weight complexes of mutant Cu,Zn-superoxide dismutase in a mouse model for familial amyotrophic lateral sclerosis. *Proceedings of the National Academy of Sciences*, 97(23), 12571–12576.
108. Jonsson, P. A., Ernhill, K., Andersen, P. M., Bergemalm, D., Brännström, T., Gredal, O., ... Marklund, S. L. (2004). Minute quantities of misfolded mutant superoxide dismutase-1 cause amyotrophic lateral sclerosis. *Brain*, 127(1), 73–88.
109. Jovičić, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S. B., ... Gitler, A. D. (2015). Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nature Neuroscience*, 18(9), 1226–1229.

110. Kato, S., Oda, M., Hayashi, H., Kawata, A., & Shimizu, T. (1994). Participation of the limbic system and its associated areas in the dementia of amyotrophic lateral sclerosis. *Journal of the Neurological Sciences*, 126(1), 62–69.
111. Kaye, R., Sokolov, Y., Edmonds, B., McIntire, T. M., Milton, S. C., Hall, J. E., & Glabe, C. G. (2004). Permeabilization of lipid bilayers is a common conformation-dependent activity of soluble amyloid oligomers in protein misfolding diseases. *Journal of Biological Chemistry*, 279(45), 46363–46366.
112. Kfoury, N., Holmes, B. B., Jiang, H., Holtzman, D. M., & Diamond, M. I. (2012). Trans-cellular propagation of Tau aggregation by fibrillar species. *Journal of Biological Chemistry*, 287(23), 19440–19451.
113. Khosravi, B., Hartmann, H., May, S., Möhl, C., Ederle, H., Michaelson, M., ... Edbauer, D. (2017). Cytoplasmic poly-GA aggregates impair nuclear import of TDP-43 in C9orf72 ALS/FTLD. *Human Molecular Genetics*, 26(4), 790–800.
114. Kim, W., Kim, J. S., Lee, K. S., Gwoun, Y. J., Kim, J. M., & Lee, K. H. (2007). Anticipation and phenotypic heterogeneity in Korean familial amyotrophic lateral sclerosis with superoxide dismutase 1 gene mutation. *Journal of Clinical Neurology*, 3(1), 38–44.
115. King, O. D., Gitler, A. D., & Shorter, J. (2012, June 26). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Research*.
116. Kipps, C. M., Hodges, J. R., & Hornberger, M. (2010, December). Nonprogressive behavioural frontotemporal dementia: Recent developments and clinical implications of the “bvFTD phenocopy syndrome.” *Current Opinion in Neurology*.
117. Kirschner, D. A., Abraham, C., & Selkoe, D. J. (1986). X-ray diffraction from intraneuronal paired helical filaments and extraneuronal amyloid fibers in Alzheimer

- disease indicates cross- β conformation. *Proceedings of the National Academy of Sciences of the United States of America*, 83(2), 503–507.
118. Ko, H. S., Uehara, T., Tsuruma, K., & Nomura, Y. (2004). Ubiquilin interacts with ubiquitylated proteins and proteasome through its ubiquitin-associated and ubiquitin-like domains. *FEBS Letters*, 566(1–3), 110–114.
119. Koppers, M., Blokhuis, A. M., Westeneng, H. J., Terpstra, M. L., Zundel, C. A. C., Vieira De Sá, R., ... Pasterkamp, R. J. (2015). C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Annals of Neurology*, 78(3).
120. Kryndushkin, D., Wickner, R. B., & Shewmaker, F. (2011). FUS/TLS forms cytoplasmic aggregates, inhibits cell growth and interacts with TDP-43 in a yeast model of amyotrophic lateral sclerosis. *Protein and Cell*, 2(3), 223–236.
121. Kuroda, M. (2000). Male sterility and enhanced radiation sensitivity in TLS-/- mice. *The EMBO Journal*, 19(3), 453–462.
122. Kwiatkowski, T. J., Bosco, D. A., LeClerc, A. L., Tamrazian, E., Vanderburg, C. R., Russ, C., ... Brown, R. H. (2009). Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science*, 323(5918), 1205–1208.
123. Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T., ... McKnight, S. L. (2014). Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science*, 345(6201), 1139–1145.
124. Lagier-Tourenne, C., Polymenidou, M., & Cleveland, D. W. (2010). TDP-43 and FUS/TLS: Emerging roles in RNA processing and neurodegeneration. *Human Molecular Genetics*, 19(R1).
125. Lagier-Tourenne, C., Polymenidou, M., Hutt, K. R., Vu, A. Q., Baughn, M., Huelga, S. C., ... Yeo, G. W. (2012). Divergent roles of ALS-linked proteins

- FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nature Neuroscience*, 15(11), 1488–1497.
126. Lattante, S., Rouleau, G. A., & Kabashi, E. (2013). TARDBP and FUS Mutations Associated with Amyotrophic Lateral Sclerosis: Summary and Update. *Human Mutation*, 34(6), 812–826.
127. Lee, Y. B., Chen, H. J., Peres, J. N., Gomez-Deza, J., Attig, J., Štalekar, M., ... Shaw, C. E. (2013). Hexanucleotide repeats in ALS/FTD form length-dependent RNA Foci, sequester RNA binding proteins, and are neurotoxic. *Cell Reports*, 5(5), 1178–1186.
128. Levine, T. P., Daniels, R. D., Gatta, A. T., Wong, L. H., & Hayes, M. J. (2013). The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics*, 29(4), 499–503.
129. Ling, S. C., Polymenidou, M., & Cleveland, D. W. (2013, August 7). Converging mechanisms in als and FTD: Disrupted RNA and protein homeostasis. *Neuron*.
130. Liu, Y., Pattamatta, A., Zu, T., Reid, T., Bardhi, O., Borchelt, D. R., ... Ranum, L. P. W. (2016). C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. *Neuron*, 90(3), 521–534.
131. Lomen-Hoerth, C., Anderson, T., & Miller, B. (2002). The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology*, 59(7), 1077–1079.
132. Lomen-Hoerth, C., Murphy, J., Langmore, S., Kramer, J. H., Olney, R. K., & Miller, B. (2003). Are amyotrophic lateral sclerosis patients cognitively normal? *Neurology*, 60(7), 1094–7.
133. Lomen-Hoerth, C. (2011). Clinical Phenomenology and Neuroimaging Correlates in ALS-FTD. *Journal of Molecular Neuroscience*, 45(3), 656–662.

134. MacKenzie, I. R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., ... Neumann, M. (2013). Dipeptide repeat protein pathology in C9ORF72 mutation cases: Clinico-pathological correlations. *Acta Neuropathologica*, 126(6), 859–879.
135. Majounie, E., Renton, A. E., Mok, K., Dopper, E. G. P., Waite, A., Rollinson, S., ... Logroscino, G. (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: A cross-sectional study. *The Lancet Neurology*, 11(4), 323–330.
136. Mann, D. M. A., Rollinson, S., Robinson, A., Bennion Callister, J., Thompson, J. C., Snowden, J. S., ... Pickering-Brown, S. (2014). Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathologica Communications*, 2(1).
137. Marat, A. L., Dokainish, H., & McPherson, P. S. (2011, April 22). DENN domain proteins: Regulators of Rab GTPases. *Journal of Biological Chemistry*.
138. May, S., Hornburg, D., Schludi, M. H., Arzberger, T., Rentzsch, K., Schwenk, B. M., ... Edbauer, D. (2014). C9orf72 FTL/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathologica*, 128(4), 485–503.
139. Massman PJ, Sims J, Cooke N Haverkamp L.J., Appel V., Appel S.H. (1996) Prevalence and correlates of neuropsychological deficits in amyotrophic lateral sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* 61:450-455.
140. Melki, R. (2015, June 1). Role of different alpha-synuclein strains in synucleinopathies, similarities with other neurodegenerative diseases. *Journal of Parkinson's Disease*. IOS

141. Melki, R. (2018). How the shapes of seeds can influence pathology. *Neurobiology of Disease*, 109, 201–208.
142. Mesulam, M. M., Rogalski, E. J., Wieneke, C., Hurley, R. S., Geula, C., Bigio, E. H., ... Weintraub, S. (2014, January 1). Primary progressive aphasia and the evolving neurology of the language network. *Nature Reviews Neurology*. Nature Publishing Group.
143. Mizielska, S., Lashley, T., Norona, F. E., Clayton, E. L., Ridler, C. E., Fratta, P., & Isaacs, A. M. (2013). C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathologica*, 126(6), 845–857.
144. Mizielska, S., Grönke, S., Niccoli, T., Ridler, C. E., Clayton, E. L., Devoy, A., ... Isaacs, A. M. (2014). C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science*, (6201), 1192–1194.
145. Moens, T. G., Mizielska, S., Niccoli, T., Mitchell, J. S., Thoeng, A., Ridler, C. E., ... Isaacs, A. M. (2018). Sense and antisense RNA are not toxic in *Drosophila* models of C9orf72-associated ALS/FTD. *Acta Neuropathologica*, 135(3), 445–457.
146. Moens, T. G., Niccoli, T., Wilson, K. M., Atilano, M. L., Birsa, N., Gittings, L. M., ... Isaacs, A. M. (2019). C9orf72 arginine-rich dipeptide proteins interact with ribosomal proteins in vivo to induce a toxic translational arrest that is rescued by eIF1A. *Acta Neuropathologica*, 137(3), 487–500.
147. Monsellier, E., Redeker, V., Ruiz-Arlandis, G., Bousset, L., & Melki, R. (2015). Molecular interaction between the chaperone Hsc70 and the N-terminal flank of huntingtin exon 1 modulates aggregation. *Journal of Biological Chemistry*, 290(5), 2560–2576.

148. Mori, K., Weng, S. M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., ... Edbauer, D. (2013a). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science*, 339(6125), 1335–1338.
149. Mori, K., Arzberger, T., Grässer, F. A., Gijssels, I., May, S., Rentzsch, K., ... Edbauer, D. (2013b). Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathologica*, 126(6), 881–893.
150. Münch, C., O'Brien, J., & Bertolotti, A. (2011). Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells. *Proceedings of the National Academy of Sciences*, 108(9), 3548–3553.
151. Murphy, J. M., Henry, R. G., Langmore, S., Kramer, J. H., Miller, B. L., & Lomen-Hoerth, C. (2007). Continuum of frontal lobe impairment in amyotrophic lateral sclerosis. *Archives of Neurology*, 64(4), 530–534
152. Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., ... Benson, D. F. (1998). Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. *Neurology*, 51(6), 1546–1554.
153. Neary, D., Snowden, J. S., & Mann, D. M. A. (2000). Cognitive change in motor neurone disease/amyotrophic lateral sclerosis (MND/ALS). *Journal of the Neurological Sciences*
154. Neary D., Snowden J., and Mann D. (2005), Frontotemporal dementia. *Lancet Neurology*, 4 : 711-780.
155. Nomura, T., Watanabe, S., Kaneko, K., Yamanaka, K., Nukina, N., & Furukawa, Y. (2014). Intranuclear aggregation of mutant FUS/TLS as a molecular pathomechanism of amyotrophic lateral sclerosis. *Journal of Biological Chemistry*, 289(2), 1192–1202.

156. Nonaka, T., Masuda-Suzukake, M., Arai, T., Hasegawa, Y., Akatsu, H., Obi, T., ... Hasegawa, M. (2013). Prion-like Properties of Pathological TDP-43 Aggregates from Diseased Brains. *Cell Reports*, 4(1), 124–134.
157. Nonaka, T., Masuda-Suzukake, M., Hosokawa, M., Shimozawa, A., Hirai, S., Okado, H., & Hasegawa, M. (2018). C9ORF72 dipeptide repeat poly-GA inclusions promote intracellular aggregation of phosphorylated TDP-43. *Human Molecular Genetics*, 27(15), 2658–2670.
158. Nordin, A., Akimoto, C., Wuolikainen, A., Alstermark, H., Jonsson, P., Birve, A., ... Andersen, P. M. (2014). Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Human Molecular Genetics*, 24(11), 3133–3142.
159. Onyike, C. U., & Diehl-Schmid, J. (2013). The epidemiology of frontotemporal dementia. *International Review of Psychiatry*, 25(2), 130–137.
160. Pamphlett, R., Luquin, N., McLean, C., Jew, S. K., & Adams, L. (2009). TDP-43 neuropathology is similar in sporadic amyotrophic lateral sclerosis with or without TDP-43 mutations. *Neuropathology and Applied Neurobiology*, 35(2), 222–225.
161. Pan, K. M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., ... Prusiner, S. B. (1993). Conversion of α -helices into β -sheets features in the formation of the scrapie prion proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 90(23), 10962–10966.
162. Peelaerts, W., Bousset, L., Van Der Perren, A., Moskalyuk, A., Pulizzi, R., Giugliano, M., ... Baekelandt, V. (2015). α -Synuclein strains cause distinct synucleinopathies after local and systemic administration. *Nature*, 522(7556), 340–344.
163. Penco, S., Lunetta, C., Mosca, L., Maestri, E., Avemaria, F., Tarlarini, C., ... Corbo, M. (2011). Phenotypic heterogeneity in a SOD1 G93D Italian ALS family: An

- example of human model to study a complex disease. *Journal of Molecular Neuroscience*, 44(1), 25–30.
164. Pesiridis, G. S., Tripathy, K., Tanik, S., Trojanowski, J. Q., & Lee, V. M. Y. (2011). A “two-hit” hypothesis for inclusion formation by carboxyl-terminal fragments of TDP-43 protein linked to RNA depletion and impaired microtubule-dependent transport. *Journal of Biological Chemistry*, 286(21), 18845–18855
165. Pick, A., Girling, D. M., & Marková, I. S. (1901). Senile Hirnatrophie als Grundlage von Herderscheinungen. *Wien Klin Wochensehr*, 14, 403-404.
166. Pilcz, A (1906): Über einen Fall von amyotrophischer Lateralsklerose und Paranoia. *Jahrb. für Psychiat. und Neurol.*, 17:221, 1898.
167. Prudlo, J., Bißbort, C., Glass, A., Grossmann, A., Hauenstein, K., Benecke, R., & Teipel, S. J. (2012). White matter pathology in ALS and lower motor neuron ALS variants: A diffusion tensor imaging study using tract-based spatial statistics. *Journal of Neurology*, 259(9), 1848–1859.
168. Prusiner, S.B. (1982). Novel proteinaceous infectious particles cause scrapie. *Science* 216, 136–144.
169. Prusiner, S. B. (1998). Prions, Nobel Lecture. *Pnas*, 95(November), 13363–13383.
170. Pupillo, E., Messina, P., Logroscino, G., Beghi, E., Micheli, A., Rosettani, P., ... Mazzini, L. (2014). Long-term survival in amyotrophic lateral sclerosis: A population-based study. *Annals of Neurology*, 75(2), 287–297.
171. Rakowicz, W. P., & Hodges, J. R. (1998). Dementia and aphasia in Motor Neuron Disease: An underrecognised association? *Journal of Neurology Neurosurgery and Psychiatry*, 65(6), 881–889.
172. Rasmussen Eid, H., Rosness, T. A., Bosnes, O., Salvesen, Ø., Knutli, M., & Stordal, E. (2018). Smoking and Obesity as Risk Factors in Frontotemporal Demen-

- tia and Alzheimer's Disease: The HUNT Study. *Dementia and Geriatric Cognitive Disorders Extra*, 1–10.
173. Ravits, J. M., & La Spada, A. R. (2009). ALS motor phenotype heterogeneity, focality, and spread: Deconstructing motor neuron degeneration. *Neurology*, 73(10), 805.
174. Reaume, A. G., Elliott, J. L., Hoffman, E. K., Kowall, N. W., Ferrante, R. J., Siwek, D. F., ... Snider, W. D. (1996). Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury. *Nature Genetics*, 13(1), 43–47.
- Ren, P. H., Lauckner, J. E., Kachirskaya, I., Heuser, J. E., Melki, R., & Kopito, R. R. (2009). Cytoplasmic penetration and persistent infection of mammalian cells by polyglutamine aggregates. *Nature Cell Biology*, 11(2), 219–225.
175. Renner, M., & Melki, R. (2014). Protein aggregation and prionopathies. *Pathologie Biologie*. Elsevier Masson SAS.
176. Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J. R., ... Traynor, B. J. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*, 72(2), 257–68.
177. Renton, A. E., Chio, A., & Traynor, B. J. (2014). State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci*, 17(1), 17–23.
178. Resegotti E. (1907) Un caso di sclerosi laterale amiotrofica osservato nella clinica neuropatologica di Pavia. *Il Morgani*, 49 : 586-590.
179. Ringholz, G. M., Appel, S. H., Bradshaw, M., Cooke, N. A., Mosnik, D. M., & Schulz, P. E. (2005). Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology*, 65(4), 586–590.

180. Rogelj, B., Easton, L. E., Bogu, G. K., Stanton, L. W., Rot, G., Curk, T., ... Ule, J. (2012). Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Scientific Reports*, 2.
181. Rosen, D. R., Siddique, T., Patterson, D., Figlewicz, D. A., Sapp, P., Hentati, A., ... Brown, R. H. (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature*, 362(6415).
182. Rowland, L. P. (2001). How amyotrophic lateral sclerosis got its name: the clinical-pathologic genius of Jean-Martin Charcot. *Archives of Neurology*, 58(3), 512–515.
183. Sabatelli, M., Madia, F., Conte, A., Luigetti, M., Zollino, M., Mancuso, I., ... Tonali, P. (2008). Natural history of young-adult amyotrophic lateral sclerosis. *Neurology*, 71(12), 876–881.
184. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., ... Eisenberg, D. (2007). Atomic structures of amyloid cross- β spines reveal varied steric zippers. *Nature*, 447(7143), 453–457.
185. Schoenfeld, T. A., & Obar, R. A. (1994). Diverse Distribution and Function of Fibrous Microtubule-Associated Proteins in the Nervous System. *International Review of Cytology*, 151(C), 67–137.
186. Schludi, M. H., May, S., Grässer, F. A., Rentzsch, K., Kremmer, E., Küpper, C., ... Roeber, S. (2015). Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta Neuropathologica*, 130(4), 537–555.
187. Sea, K., Sohn, S. H., Durazo, A., Sheng, Y., Shaw, B. F., Cao, X., ... Valentine, J. S. (2015). Insights into the role of the unusual disulfide bond in copper-zinc superoxide dismutase. *Journal of Biological Chemistry*, 290(4), 2405–2418.

188. Seibenhener, M. L., Babu, J. R., Geetha, T., Wong, H. C., Krishna, N. R., & Wooten, M. W. (2004). Sequestosome 1/p62 Is a Polyubiquitin Chain Binding Protein Involved in Ubiquitin Proteasome Degradation. *Molecular and Cellular Biology*, 24(18), 8055–8068.
189. Sellier, C., Campanari, M., Julie Corbier, C., Gaucherot, A., Kolb-Cheynel, I., Oulad-Abdelghani, M., ... Charlet-Berguerand, N. (2016). Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *The EMBO Journal*, 35(12), 1276–1297.
190. Sephton, C. F., Cenik, C., Kucukural, A., Dammer, E. B., Cenik, B., Han, Y., ... Yu, G. (2011). Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes. *Journal of Biological Chemistry*, 286(2), 1204–1215.
191. Shimonaka, S., Nonaka, T., Suzuki, G., Hisanaga, S. I., & Hasegawa, M. (2016). Templated aggregation of TAR DNA-binding protein of 43 kDa (TDP-43) by seeding with TDP-43 peptide fibrils. *Journal of Biological Chemistry*, 291(17), 8896–8907.
192. Shrivastava, A. N., Redeker, V., Fritz, N., Pieri, L., Almeida, L. G., Spolidoro, M., ... Triller, A. (2015). α -synuclein assemblies sequester neuronal β -Na⁺/K⁺-ATPase and impair Na⁺ gradient. *The EMBO Journal*, 34(19), 2408–2423.
193. Shrivastava, A. N., Redeker, V., Pieri, L., Bousset, L., Renner, M., Madiona, K., ... Melki, R. (2019). Clustering of Tau fibrils impairs the synaptic composition of α -Na⁺/K⁺-ATPase and AMPA receptors. *The EMBO Journal*, 38(3), e99871.
194. Smith, B. N., Newhouse, S., Shatunov, A., Vance, C., Topp, S., Johnson, L., ... Shaw, C. E. (2013). The C9ORF72 expansion mutation is a common cause of ALS⁺/FTD in Europe and has a single founder. *European Journal of Human Genetics*, 21(1), 102–108.

195. Smith, P. E. S., Brender, J. R., & Ramamoorthy, A. (2009). Induction of negative curvature as a mechanism of cell toxicity by amyloidogenic peptides: The case of islet amyloid polypeptide. *Journal of the American Chemical Society*, 131(12), 4470–4478.
196. Snowden, J. S., Bathgate, D., Varma, A., Blackshaw, A., Gibbons, Z. C., & Neary, D. (2001). Distinct behavioural profiles in frontotemporal dementia and semantic dementia. *Journal of Neurology Neurosurgery and Psychiatry*, 70(3), 323–332.
197. Solomon, D. A., Stepto, A., Au, W. H., Adachi, Y., Diaper, D. C., Hall, R., ... Hirth, F. (2018). A feedback loop between dipeptide-repeat protein, TDP-43 and karyopherin- α mediates C9orf72-related neurodegeneration. *Brain*, 141(10), 2908–2924.
198. Spillantini, M. G., Murrell, J. R., Goedert, M., Farlow, M. R., Klug, A., & Ghetti, B. (2002). Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proceedings of the National Academy of Sciences*, 95(13), 7737–774.
199. Stranks, S. D., Ecroyd, H., Van Sluyter, S., Waters, E. J., Carver, J. A., & Von Smekal, L. (2009). Model for amorphous aggregation processes. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 80(5).
200. Su, F. C., Goutman, S. A., Chernyak, S., Mukherjee, B., Callaghan, B. C., Batterman, S., & Feldman, E. L. (2016). Association of environmental toxins with amyotrophic lateral sclerosis. *JAMA Neurology*, 73(7), 803–811.
201. Sudria-Lopez, E., Koppers, M., de Wit, M., van der Meer, C., Westeneng, H. J., Zundel, C. A. C., ... Pasterkamp, R. J. (2016, July 1). Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. *Acta Neuropathologica*. Springer Verlag.

202. Sullivan, P. M., Zhou, X., Robins, A. M., Paushter, D. H., Kim, D., Smolka, M. B., & Hu, F. (2016). The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathologica Communications*, 4(1), 51.
203. Swinnen, B., & Robberecht, W. (2014, November 5). The phenotypic variability of amyotrophic lateral sclerosis. *Nature Reviews Neurology*. Nature Publishing Group.
204. Tartaglia, G. G., Pawar, A. P., Campioni, S., Dobson, C. M., Chiti, F., & Vendruscolo, M. (2008). Prediction of Aggregation-Prone Regions in Structured Proteins. *Journal of Molecular Biology*, 380(2), 425–436.
205. Therrien, M., Rouleau, G. A., Dion, P. A., & Parker, J. A. (2013). Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in *C. elegans*. *PLoS ONE*, 8(12).
206. Tronconi V (1941), Sindrome demenziale presenile e sclerosi laterale amiotrofica. *Riv Sper. Freniatr. Med. Leg.*, 65 : 811-813.
207. Tu, P. H., Raju, P., Robinson, K. A., Gurney, M. E., Trojanowski, J. Q., & Lee, V. M. (1996). Transgenic mice carrying a human mutant superoxide dismutase transgene develop neuronal cytoskeletal pathology resembling human amyotrophic lateral sclerosis lesions. *Proceedings of the National Academy of Sciences of the United States of America*, 93(7), 3155–60.
208. Van Reeth P. C., Perier O., Coers C., Van B.L. (1961) Pick's dementia associated with atypical amyotrophic lateral sclerosis. (Anatomoclinical study), *Acta Neurol. Belg.*, 61 : 309-325.
209. Vance, C., Rogelj, B., Hortobágyi, T., De Vos, K. J., Nishimura, A. L., Sreedharan, J., ... Shaw, C. E. (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science*, 323(5918), 1208–1211.

210. Vercelletto, M., Ronin, M., Huvet, M., Magne, C., & Feve, J. R. (1999). Frontal type dementia preceding amyotrophic lateral sclerosis: A neuropsychological and SPECT study of five clinical cases. *European Journal of Neurology*, 6(3), 295–299.
211. Walker, C., Herranz-Martin, S., Karyka, E., Liao, C., Lewis, K., Elsayed, W., ... El-Khamisy, S. F. (2017). C9orf72 expansion disrupts ATM-mediated chromosomal break repair. *Nature Neuroscience*, 20(9), 1225–1235.
212. Wälti, M. A., Ravotti, F., Arai, H., Glabe, C. G., Wall, J. S., Böckmann, A., ... Riek, R. (2016). Atomic-resolution structure of a disease-relevant A β (1-42) amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America*, 113(34), E4976–E4984.
213. Wen, X., Tan, W., Westergard, T., Krishnamurthy, K., Markandaiah, S. S., Shi, Y., ... Trotti, D. (2014). Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate invitro and invivo neuronal death. *Neuron*, 84(6), 1213–1225.
214. Williams, A. D., Portelius, E., Kheterpal, I., Guo, J. T., Cook, K. D., Xu, Y., & Wetzel, R. (2004). Mapping A β amyloid fibril secondary structure using scanning proline mutagenesis. *Journal of Molecular Biology*, 335(3), 833–842.
215. Winton, M. J., Igaz, L. M., Wong, M. M., Kwong, L. K., Trojanowski, J. Q., & Lee, V. M. Y. (2008). Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. *Journal of Biological Chemistry*, 283(19), 13302–13309.
216. Wood, S. J., Wetzel, R., Martin, J. D., & Hurle, M. R. (1995). Prolines and Amyloidogenicity in Fragments of the Alzheimer's Peptide β /A4. *Biochemistry*, 34(3), 724–730.

217. Wu, C., Biancalana, M., Koide, S., & Shea, J. E. (2009). Binding Modes of Thioflavin-T to the Single-Layer β -Sheet of the Peptide Self-Assembly Mimics. *Journal of Molecular Biology*, 394(4), 627–633
218. Xu S, Peng G, Wang Y, Fang S, Karbowski M (February 2011). "The AAA-ATPase p97 is essential for outer mitochondrial membrane protein turnover". *Molecular Biology of the Cell*. 22 (3): 291–300.
219. Xu, W., & Xu, J. (2018). C9orf72 dipeptide repeats cause selective neurodegeneration and cell-autonomous excitotoxicity in *Drosophila* glutamatergic neurons. *Journal of Neuroscience*, 38(35), 7741–7752.
220. Xi, Z., Zinman, L., Moreno, D., Schymick, J., Liang, Y., Sato, C., ... Rogaeva, E. (2013). Hypermethylation of the CpG island near the G 4 C 2 repeat in ALS with a C9orf72 expansion. *American Journal of Human Genetics*, 92(6), 981–989.
221. Xi, Z., Rainero, I., Rubino, E., Pinessi, L., Bruni, A. C., Maletta, R. G., ... Rogaeva, E. (2014). Hypermethylation of the CpG-island near the C9orf72 G 4 C 2 - repeat expansion in FTLD patients. *Human Molecular Genetics*, 23(21), 5630–5637.
222. Xiao, S., MacNair, L., McGoldrick, P., McKeever, P. M., McLean, J. R., Zhang, M., ... Robertson, J. (2015). Isoform-specific antibodies reveal distinct sub-cellular localizations of C9orf72 in amyotrophic lateral sclerosis. *Annals of Neurology*, 78(4), 568–583.
223. Yamakawa, M., Ito, D., Honda, T., Kubo, K. I., Noda, M., Nakajima, K., & Suzuki, N. (2015). Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS. *Human Molecular Genetics*, 24(6), 1630–1645.
224. Yamashita, S., Mori, A., Sakaguchi, H., Suga, T., Ishihara, D., Ueda, A., ... Hirano, T. (2012). Sporadic juvenile amyotrophic lateral sclerosis caused by mutant

- FUS/TLS: Possible association of mental retardation with this mutation. *Journal of Neurology*, 259(6), 1039–1044.
225. Yang, C., Wang, H., Qiao, T., Yang, B., Aliaga, L., Qiu, L., ... Xu, Z. (2014). Partial loss of TDP-43 function causes phenotypes of amyotrophic lateral sclerosis. *Proceedings of the National Academy of Sciences*, 111(12), E1121–E112.
226. Zago, S., Poletti, B., Corbo, M., Adobbati, L., & Silani, V. (2008). Dysgraphia in patients with primary lateral sclerosis: a speech-based rehearsal deficit? *Behavioural Neurology*, 19(4), 169–75.
227. Zetterstrom, P., Stewart, H. G., Bergemalm, D., Jonsson, P. A., Graffmo, K. S., Andersen, P. M., ... Marklund, S. L. (2007). Soluble misfolded subfractions of mutant superoxide dismutase-1s are enriched in spinal cords throughout life in murine ALS models. *Proceedings of the National Academy of Sciences*, 104(35), 14157–1416.
228. Zhang, H., Constantine, R., Vorobiev, S., Chen, Y., Seetharaman, J., Huang, Y. J., ... Baehr, W. (2011). UNC119 is required for G protein trafficking in sensory neurons. *Nature Neuroscience*, 14(7), 874–880.
229. Zhang, Y. J., Jansen-West, K., Xu, Y. F., Gendron, T. F., Bieniek, K. F., Lin, W. L., ... Petrucelli, L. (2014). Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathologica*, 128(4), 505–524.
230. Zhang, Y. J., Gendron, T. F., Ebbert, M. T. W., O'Raw, A. D., Yue, M., Jansen-West, K., ... Petrucelli, L. (2018, August 1). Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nature Medicine*. Nature Publishing Group.

231. Zhou, B., Geng, Y., Liu, C., Miao, H., Ren, Y., Xu, N., ... Zhu, G. (2018). Characterizations of distinct parallel and antiparallel G-quadruplexes formed by two-repeat ALS and FTD related GGGGCC sequence. *Scientific Reports*, 8(1).
232. Zoccolella, S., Beghi, E., Palagano, G., Fraddosio, A., Guerra, V., Samarelli, V., ... Logroscino, G. (2008). Analysis of survival and prognostic factors in amyotrophic lateral sclerosis: A population based study. *Journal of Neurology, Neurosurgery and Psychiatry*, 79(1), 33–37.
233. Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., ... Ranum, L. P. W. (2011). Non-ATG-initiated translation directed by microsatellite expansions. *Proceedings of the National Academy of Sciences of the United States of America*, 108(1), 260–265.
234. Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., ... Ranum, L. P. W. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proceedings of the National Academy of Sciences of the United States of America*, 110(51).

Title : Biochemical characterisations of dipeptide repeat proteins and their assemblies

Keywords : Amyotrophic Lateral Sclerosis ; Fronto-temporal Dementia ; C9orf72 ; Dipeptide Repeat Proteins ; Protein aggregation

Abstract : The most common genetic factor between familial Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD) is the repetition of (GGGGCC)_n in the open reading frame 72 of chromosome 9 (C9orf72). FTD is characterized by personality, behavior and language disorders due to neuronal degeneration in the cortex, thalamus, hippocampus and cerebellum. ALS patients, on the other hand, exhibit motor symptoms primarily consisting of paralysis of the skeletal muscles caused by motoneuronal loss in the spinal cord and the central motor system. These pathologies have significant genetic and clinic overlaps suggesting common mechanisms.

Large amounts of cytoplasmic inclusions of the Dipeptide Repeat Proteins (DRP), translated from C9orf72 RNA, were found in brains of patients who died from ALS or FTD. These proteins are constituted of repetitions of Dipeptides : glycine-alanine (poly-GA), glycine-proline (poly-GP), glycine-arginine (poly-GR), proline-arginine (poly-PR) and proline-alanine (poly-PA). The toxicity of the aggregated DRP has been shown in various cellular models as well as in drosophila and mice.

Nevertheless, few studies have described the biochemistry of the DRP. The aim of my thesis is the *in vitro* characterization of the DRP. I have characterized the biochemistry of these proteins and in particular their self-assembly abilities in Ronald Melki's team « Proteins folding and conformational diseases ».

Recombinant DRP of different lengths and nature have been produced to determined the *in vitro* conditions in which they assemble. The morphology of the related aggregates has been studied by transmission electron microscopy. Structure informations of DPR were obtained with Circular Dichroism spectrometry for monomers and Infrared spectroscopy for assemblies.

Lastly, the toxicity and internalization abilities of these proteins and their assemblies are tested on culture cells in collaboration with Pr. Mimoun Azzouz of the University of Sheffield.

Titre : Caractérisations biochimiques des protéines à répétitions de dipeptides et de leurs assemblages

Mots clés : Sclérose Latérale Amyotrophique ; Démences Fronto-Temporale ; C9orf72 ; Protéines à Répétitions de Dipeptides ; Agrégation protéique

Résumé : La répétition de l'hexanucléotide GGGGCC est le facteur génétique le plus présent chez les patients atteints de démence fronto-temporale (DFT) et de sclérose latérale amyotrophique (SLA). Elle apparaît au niveau du cadre de lecture ouvert 72 du chromosome 9 (C9orf72). D'une part, la DFT se caractérise par des troubles de la personnalité, du langage et du comportement causés principalement par une dégénérescence de neurones corticaux, thalamiques, hippocampiques et cérébelleux. D'autre part, les patients SLA sont atteints de paralysie des muscles squelettiques causée par la mort de motoneurons au niveau de la moelle épinière et du système moteur central. Ces deux pathologies partagent de nombreuses caractéristiques cliniques et génétiques, suggérant qu'il existe des mécanismes communs entre elles.

D'abondantes inclusions cytoplasmiques de protéines issues de la traduction des ARN d'hexanucléotides ont été retrouvées dans le cerveau de patients décédés de SLA et de DFT. On appelle ces polypeptides : Protéines à Répétitions de Dipeptides (PRD), car ils ne sont composés que d'une suite de deux résidus : glycine-alanine (GA), glycine-proline (GP), glycine-arginine (GR), proline-arginine (PR) et proline-alanine (PA). La toxicité de ces PRD a été mise en évidence dans des modèles cellulaires, chez la drosophile mais aussi la souris.

Cependant, peu d'études se sont concentrées sur la caractérisation biochimique de ces PRD.

Le travail de cette thèse a consisté en la production et la caractérisation *in vitro* de PRD. J'ai réalisé la caractérisation biochimique de ces protéines et étudié leur capacité à s'auto-assembler au sein de l'équipe « Repliement des protéines *in vitro* et maladies conformationnelles » dirigée par Ronald Melki.

Les PRD recombinants poly-GA, poly-PA et poly-GP ont été produits. Nous avons déterminé les conditions *in vitro* dans lesquelles ces protéines s'assemblaient. La morphologie des agrégats a été étudiée par microscopie électronique, tandis que la structure des protéines a été déterminée par spectromètre de dichroïsme circulaire pour la forme monomérique et par spectroscopie à infrarouge pour les assemblages.

La toxicité, la propagation et la capacité de ces PRD à être internalisé sont testées dans des modèles cellulaires en collaboration avec l'équipe du Pr. Mimoun Azzouz de l'Université de Sheffield.