



**HAL**  
open science

# Déterminants structuraux d'agrégats de Tau distincts : vers de nouveaux outils moléculaires pour discriminer les tauopathies

Emilie Caroux

## ► To cite this version:

Emilie Caroux. Déterminants structuraux d'agrégats de Tau distincts : vers de nouveaux outils moléculaires pour discriminer les tauopathies. Biochimie [q-bio.BM]. Université Paris Saclay (COmUE), 2019. Français. NNT : 2019SACLS592 . tel-03494708

**HAL Id: tel-03494708**

**<https://theses.hal.science/tel-03494708>**

Submitted on 20 Dec 2021

**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.

# Déterminants structuraux d'agrégats de Tau distincts : vers de nouveaux outils moléculaires pour discriminer les tauopathies

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

École doctorale n°568 Biosigne  
Signalisations et réseaux intégratifs en biologie

Spécialité de doctorat: Aspects moléculaires et cellulaires de la biologie

Thèse présentée et soutenue à Fontenay-aux-Roses, le 19 décembre 2019, par

**Emilie Caroux**

Composition du Jury :

Florence GONNET <i>Professeure, Université d'Evry Val d'Essonne (LAMBE, Université d'Evry)</i>	Président
Marie-Claude POTIER <i>Directrice de recherche, CNRS (ICM, UMR7225 CNRS, U112 INSERM)</i>	Rapporteur
Julia CHAMOT-ROOKE <i>Directrice de recherche, CNRS (Institut Pasteur, UtechS MS Bio)</i>	Rapporteur
Noëlle POTIER <i>Chargée de recherche, CNRS (LSMIS, UMR 7140 CNRS)</i>	Examineur
Virginie REDEKER <i>Chargée de recherche, INSERM (CNRS UMR 9199, MIRGen, CEA)</i>	Directrice de thèse
Ronald MELKI <i>Directeur de recherche, CNRS (CNRS UMR 9199, MIRGen, CEA)</i>	Co-Directeur de thèse



## Remerciements

J'exprime toute ma gratitude à la Fondation Vaincre Alzheimer qui a financé cette thèse.

Je remercie chaleureusement Mme Julia Chamot-Rooke et Mme Marie-Claude Potier d'avoir accepté d'être rapporteurs de ce manuscrit. J'adresse également mes remerciements à Mesdames Florence Gonnet et Noëlle Potier pour s'être rendues disponibles pour examiner mon travail de thèse. Enfin je remercie, à nouveau, Mme Chamot-Rooke, et M Buée pour avoir suivi, tout au long de ma thèse, l'évolution de mon projet et pour nos enrichissants échanges.

Je souhaite bien sûr remercier Ronald Melki de m'avoir accueillie dans son équipe. Ronald, je te remercie pour nos nombreuses discussions scientifiques autour de mon projet et tes précieux conseils.

Ensuite, je tiens évidemment à remercier Virginie Redeker de m'avoir encadrée pendant cette thèse. Virginie, merci de m'avoir offert l'opportunité de travailler sur un sujet aussi intéressant. Notre collaboration a renforcé mon envie de continuer dans le monde de la spectrométrie de masse.

Merci également à Karine qui m'a formée lors de mon arrivée dans l'équipe, qui est à l'origine de ce projet et pour tous ses conseils tout au long de ma thèse.

Merci à Luc pour s'être rendu disponible pour répondre à chacune de mes questions.

Merci à Tracy pour faire de notre quotidien, à la paillasse et en dehors, un plaisir de chaque instant.

Merci à tous les membres de l'équipe ou anciens collègues que je n'ai pas encore cités : Alexis, Amu, Audrey, Elodie, Hugo, Jimmy, Kai, Laura, Laurent, Maud, Maya, Mehdi, Nolwen, Pauline, Stéphane. J'ai également eu la chance de travailler aux côtés d'Elodie, Marion et Margaux<sup>2</sup>. Je suis très heureuse d'avoir pu vous rencontrer et d'avoir passé autant de temps avec vous au laboratoire, et en dehors, pour la plupart d'entre vous.

Merci à David, Laïla et Sofia pour nos nombreux moments à la plateforme ; cette thèse n'aurait jamais été possible sans vous.

Merci à tous les membres de MIRCen de nous avoir si gentiment accueillis lors de ma dernière année de thèse.

Je tiens à remercier chaleureusement toutes les personnes qui ont accepté de relire ce manuscrit.

Je profite de cette occasion pour remercier ma famille et nos amis pour leur amour et leur soutien indéfectible : évidemment mes merveilleux parents, ma grand-mère, Benjamin, Gabi, Antoine, Samira et Régis, Valérie et Bertrand, Sandrine et Christophe, Claire et Philippe, Aline, Magalie, Oriane, Ferdinand, Cécile et Isabelle. Merci à mes oncles et tantes, et plus particulièrement Godeleine. Merci à toute ma famille, Caroux et Le Bail (dont l'énumération pourrait constituer un chapitre entier de cette thèse).

Enfin, je tiens à exprimer ma gratitude pour toutes les personnes que j'ai eu la chance de rencontrer et qui ont éclairé mon quotidien, notamment pour leur soutien après mes performances trampolinesques : ma famille de tous les jours (Sergi et Kévin), Guillaume, Anne-Cécile, Audrey, Chris, Eddy, Fahd, Florian, Humberto, Julie, Léo, Marianne, Merouann, Meryem, Mich, Michel, Nico, Robin, Sham, Tristan, Valérie, Venkat, Yanis, mes anciens collègues de l'INRA et tous les autres ! Pour un autre type de soutien, je remercie Ed S, la NPR et Cadbury.

## Table des matières

<b>Abréviations</b> .....	7
<b>Avant-Propos</b> .....	8
<b>I. Introduction</b> .....	10
I. Maladies à prions et protéinopathies .....	10
I.1 Bref historique des maladies à prions et généralités sur les protéinopathies .....	10
Perte de fonction normale ou gain de fonction toxique ? .....	12
I.2 Mécanismes moléculaires de l'agrégation .....	14
L'agrégation à l'origine des protéinopathies ? .....	14
Mécanisme d'agrégation .....	14
I.3 Caractère prion-like des protéinopathies .....	16
Transmission prion-like des agrégats .....	16
Notion de souches .....	18
II. Les tauopathies .....	20
II.1 Historique et généralités .....	20
II.2 Clinique et Anatomopathologie .....	21
II.2.1 Observations cliniques .....	21
II.2.2 Anatomopathologie et évolution de la pathologie Tau .....	23
II.2.2.1 Observations histologiques .....	23
II.2.2.2 Evolution de la pathologie .....	25
II.3 Caractéristiques moléculaires et biochimiques .....	27
II.4 Déterminants génétiques .....	30
II.5 Méthodes diagnostiques .....	33
II.5.1 Examen clinique et neuropsychologique .....	33
II.5.2 Imagerie .....	33
II.5.3 Analyse du Liquide Céphalo-Rachidien (LCR) – Biomarqueurs biochimiques .....	35
II.5.4 Difficultés associées au diagnostic différentiel des tauopathies .....	36
II.6 Essais cliniques et prise en charge thérapeutique des patients .....	38
Maladie d'Alzheimer .....	38
Autres tauopathies .....	39
III. Tau .....	41
III.1 Le gène MAPT, épissage alternatif et expression des isoformes .....	41
III.2 Domaines et fonctions .....	47
III.3 Modifications post-traductionnelles .....	50
III.3.1 Addition de groupements fonctionnels .....	50
III.3.2 Ponts disulfure .....	54
III.3.3 Clivages protéolytiques .....	55
III.3.4 Isomérisations .....	56
III.4 Agrégation .....	57
III.4.1 Formation des agrégats .....	57
III.4.2 Structure de Tau monomérique et de ses agrégats .....	62
III.4.2.1 Prédiction de structures secondaires .....	62
III.4.2.2 Etude RMN de la structure de Tau monomérique .....	63
III.4.2.3 Structure de Tau fibrillaire .....	63
III.4.3 Transmission « prion-like » .....	70
IV. Approches structurales applicables à l'étude d'assemblages de haut poids moléculaire insolubles .....	75
IV.1 Méthodes de caractérisation structurales classiques .....	75

IV.1.1	Diffraction aux rayons X .....	75
IV.1.2	Spectroscopie infrarouge à transformée de Fourier (FTIR) .....	76
IV.1.3	Résonance Magnétique Nucléaire (RMN) .....	77
IV.2	Cryo-Microscopie .....	77
IV.3	Protéolyse ménagée .....	79
IV.4	Méthodes associant la spectrométrie de masse .....	81
IV.4.1	Marquage non covalent : Echanges Hydrogène-Deutérium .....	81
IV.4.2	Marquage covalent .....	82
IV.4.3	Pontage covalent .....	84
<b>II.</b>	<b>Objectifs de la thèse</b> .....	<b>86</b>
<b>III.</b>	<b>Résultats</b> .....	<b>88</b>
III.1	Mise en place de méthodes dédiées à l'analyse des surfaces accessibles des fibres de Tau .	88
A.	Identification de surfaces accessibles et de régions protégées par protéolyse ménagée .....	88
A.1	Présentation des enzymes disponibles et de leurs sites de clivage .....	88
A.2	Choix des conditions de protéolyse ménagée .....	90
A.3	Identification des fragments protéolytiques de haut poids moléculaire .....	94
A.3.1	Analyse par immunodétection (Western-Blots) (crédit : Karine Madiona) .....	95
A.3.2	Analyse du recouvrement de séquence des fragments par une approche protéomique Bottom-up .....	97
A.3.3	Identification du résidu N-terminal des fragments protéolytiques .....	99
a.	Par dérivatisation du N-terminal au TMPP et analyse par spectrométrie de masse .....	99
b.	Par séquençage d'Edman .....	108
A.3.4	Déduction du résidu C-terminal des fragments protéolytiques .....	110
a.	Fragments protéolytiques GluC : d'après des mesures de masse entière .....	110
b.	Fragments chymotryptiques : d'après le recouvrement de séquence .....	113
A.4	Analyse de peptides protéolytiques issus d'une protéolyse avancée .....	115
B.	Marquage covalent des surfaces accessibles par la NHS-Biotine .....	119
B.1	Optimisation des conditions de marquage .....	119
B.1.1	Choix du marquage covalent par la NHS-Biotine .....	119
B.1.2	Mesure du niveau moyen de marquage par MALDI-MS .....	121
B.1.3	Détermination des conditions de marquage optimales .....	124
B.2	Choix des conditions de digestion des protéines biotinylées .....	127
B.3	Identification et quantification des peptides biotinylés .....	128
B.3.1	Identification des peptides biotinylés .....	128
B.3.2	Quantification des peptides biotinylés .....	129
III.2	Application aux fibres de Tau (article) .....	134
<b>IV.</b>	<b>Conclusion et perspectives</b> .....	<b>203</b>
	<b>Références</b> .....	<b>209</b>
	<b>Matériel et méthodes</b> .....	<b>225</b>
	Expression et purification de Tau 1N3R et 1N4R recombinante humaine .....	225
	Séquences primaires de Tau 1N3R et 1N4R humaines .....	225
	Assemblage de Tau 1N3R et 1N4R .....	226
	Suivi de la polymérisation par ultracentrifugation suivie d'un SDS-PAGE .....	226
	Suivi de la polymérisation par liaison à la Thioflavine T .....	227
	Protéolyses ménagées et identification des fragments protéolytiques .....	228
	Protéolyses ménagées .....	228
	Cinétique de protéolyses ménagées suivie par SDS-PAGE .....	228
	Test de sédimentation des fragments protéolytiques .....	229

Suivi de la protéolyse ménagée des fibres de Tau par la GluC par microscopie électronique à transmission.....	229
Analyse des fragments protéolytiques associant la spectrométrie de masse.....	229
Analyse Bottom-up sans marquage par LC-MS/MS des fragments protéolytiques GluC et Chymotrypsine.....	229
Analyse Bottom-up avec marquage spécifique du résidu N-terminal par LC-MS/MS .....	230
Analyse en masse entière par MALDI-MS .....	231
Marquage par la NHS-Biotine .....	231
Analyse du marquage sur protéine entière par MALDI-MS .....	231
Gamme de ratios molaires [K/NHS-Biotine] et cinétique de marquage .....	231
<b>Annexes</b> .....	233
Annexe 1. Liste des tauopathies actuellement répertoriées .....	233
Annexe 2. Formes tronquées de Tau couramment utilisées en laboratoire .....	234
Annexe 3. La spectrométrie de masse.....	235
Annexe 4. Recouvrement de séquence obtenu sur la protéine entière Tau 1N3R et Tau 1N4R après digestion en gel par la trypsine et analyse LC-MS/MS .....	242
Annexe 5. Identification du peptide N-terminal dérivatisé au TMPP des fragments protéolytiques	243

## Liste des figures

<b>Figure 1.</b> Mécanisme général de formation des fibres amyloïdes proposé et cinétique de formation.....	15
<b>Figure 2.</b> Représentation schématique de différents conformères d'assemblages protéiques ou souches.....	18
<b>Figure 3.</b> Observations histologiques de coupes hippocampales humaines de patients atteints de tauopathies .....	23
<b>Figure 4.</b> Schématisation de la pathologie amyloïde et de la pathologie Tau dans la maladie d'Alzheimer. ....	26
<b>Figure 5.</b> Catégories 3R et 4R des isoformes de Tau. ....	27
<b>Figure 6.</b> Comparaison des déterminants biochimiques d'AD et de PART.....	28
<b>Figure 7.</b> Schématisation du code-barre moléculaire caractéristique des différentes tauopathies.....	29
<b>Figure 8.</b> Images par IRM du cerveau de patients atteints de tauopathies.....	34
<b>Figure 9.</b> Traceurs PET et suivi de la pathologie Alzheimer.....	35
<b>Figure 10.</b> Illustration du recouvrement des spectres cliniques présentés par différentes tauopathies.....	37
<b>Figure 11.</b> Le gène MAPT, régulation de l'épissage alternatif de l'exon 10 et isoformes de Tau.....	42
<b>Figure 12.</b> Expression du gène MAPT. ....	45
<b>Figure 13.</b> Domaines de Tau et son interaction avec les microtubules.....	48
<b>Figure 14.</b> Phosphorylation de Tau chez des individus sains et des patients atteints de la maladie d'Alzheimer. ....	51
<b>Figure 15.</b> Sites de clivage et protéases associées sur la séquence de Tau 2N4R.....	56
<b>Figure 16.</b> Schéma proposé du processus de polymérisation des fibres de Tau.....	58
<b>Figure 17.</b> Structures secondaires, hélices $\alpha$ et brins $\beta$ , prédites le long de la séquence de Tau 2N4R monomérique. ....	62
<b>Figure 18.</b> Observation par MET de fibres de Tau issues de patients et assemblées in vitro .....	65
<b>Figure 19.</b> Détermination du coeur des fibres par protéolyse ménagée. ....	66
<b>Figure 20.</b> Etudes structurales du coeur des fibres de Tau par cryomicroscopie électronique après digestion à la pronase.....	68
<b>Figure 21.</b> Mécanismes possibles de propagation cellule à cellule des agrégats de Tau .....	70
<b>Figure 22.</b> Principe de l'étude structurale par cryo-microscopie .....	78
<b>Figure 23.</b> Principe de l'étude structurale de souches par protéolyse ménagée suivie par SDS-PAGE.....	80
<b>Figure 24.</b> Présentation des enzymes disponibles et de leurs sites de clivage .....	89
<b>Figure 25.</b> Cinétique des protéolyses ménagées suivie par SDS-PAGE de fibres 1N3R et fibres 1N4R .....	90
<b>Figure 26.</b> Cinétique des protéolyses ménagées suivie par SDS-PAGE des monomères et des fibres de Tau 1N3R et 1N4R.....	92
<b>Figure 27.</b> Les fragments protéolytiques de haut poids moléculaire générés par la GluC sont fibrillaires.....	93
<b>Figure 28.</b> Codes-barres moléculaires des fibres de Tau 1N3R et 1N4R protéolysées.....	95
<b>Figure 29.</b> Identification des fragments issus de la protéolyse ménagée à la GluC des fibres de Tau 1N3R et de Tau 1N4R par Western Blot.....	96
<b>Figure 30.</b> Stratégie d'identification des fragments protéolytiques par spectrométrie de masse et approche Bottom-up.....	97
<b>Figure 31.</b> Recouvrement de séquence des fragments générés par protéolyse ménagée de Tau 1N3R et 1N4R.....	98

<b>Figure 32.</b> Réaction de dérivation du groupement amine primaire spécifique du résidu N-terminal de la protéine par le TMPP à pH 8,2. ....	100
<b>Figure 33.</b> Dérivation du résidu N-terminal : preuve de faisabilité.....	101
<b>Figure 34.</b> Aspécificité du marquage par le TMPP. ....	102
<b>Figure 35.</b> Chromatogramme de Tau 1N3R entière digérée en gel par la trypsine après marquage par le TMPP .....	103
<b>Figure 36.</b> XIC, spectres MS et MS <sup>2</sup> du peptide N-terminal de la protéine Tau entière marqué par le TMPP. ....	104
<b>Figure 37.</b> XIC, spectres MS et MS <sup>2</sup> d'un peptide interne de la protéine Tau entière marqué par le TMPP. ....	105
<b>Figure 38.</b> Identification des extrémités N-terminales des fragments protéolytiques GluC des fibres de Tau 1N3R et 1N4R par marquage au TMPP et analyse LC-MS/MS .....	106
<b>Figure 39.</b> Identification des extrémités N-terminales des fragments protéolytiques chymotryptiques des fibres de Tau 1N3R et 1N4R par marquage au TMPP et analyse LC-MS/MS .....	108
<b>Figure 40.</b> Mécanisme de la dégradation d'Edman .....	109
<b>Figure 41.</b> Electrotransfert des fragments protéolytiques GluC des fibres 1N3R et 1N4R générés après 45 minutes de digestion ménagée. ....	109
<b>Figure 42.</b> Identification des extrémités N-terminales des fragments protéolytiques GluC des fibres de Tau 1N3R et 1N4R par séquençage d'Edman .....	110
<b>Figure 43.</b> Spectre MALDI-MS des produits de protéolyse ménagée des fibres 1N3R et 1N4R par la GluC .....	112
<b>Figure 44.</b> Extrémités N et C-terminales des fragments protéolytiques générés après 45 minutes de digestion par la GluC.....	113
<b>Figure 45.</b> Extrémités N et C-terminales des fragments protéolytiques générés après 2 minutes de digestion par la chymotrypsine. ....	115
<b>Figure 46.</b> Intensité du peptide 77-88 choisi pour la normalisation.....	117
<b>Figure 47.</b> Box-plots représentant les trois catégories de peptides associés à différents « fold-changes » entre formes fibrillaire et monomérique de Tau 1N4R. ....	118
<b>Figure 48.</b> Principe du marquage par la NHS-Biotine et sites putatifs le long de la séquence de 1N3R et 1N4R. ....	120
<b>Figure 49.</b> Schématisation de la stratégie de marquage de surface de Tau 1N3R et 1N4R monomérique et fibrillaire et son analyse par spectrométrie de masse.....	121
<b>Figure 50.</b> Influence de la matrice sur la qualité de l'analyse de masse de molécules de haut poids moléculaire .....	122
<b>Figure 51.</b> Influence du tampon de reprise de matrice HCCA sur la qualité spectrale.....	123
<b>Figure 52.</b> Mesure de l'efficacité de biotinylation de Tau 1N3R et 1N4R monomérique et fibrillaire en fonction d'une gamme de ratios molaires (K/NHS-Biotine) .....	125
<b>Figure 53.</b> Mesure de l'efficacité de biotinylation de Tau 1N3R et 1N4R monomérique et fibrillaire en fonction d'une cinétique de marquage. ....	126
<b>Figure 54.</b> XIC, spectres MS et MS <sup>2</sup> d'un peptide biotinylé. ....	129
<b>Figure 55.</b> Intensité du peptide standard 77-88 dans différents échantillons et intensité normalisée d'un peptide biotinylé en fonction des différents ratios [K/NHS-Biotine]......	131
<b>Figure 56.</b> Box-plots représentant les trois catégories de peptides biotinylés associés à différents « fold-changes » entre formes fibrillaire et monomérique de Tau 1N4R.....	132

## Abréviations

<b>ACN</b>	Acétonitrile
<b>AD</b>	Maladie d'Alzheimer
<b>AF</b>	Acide Formique
<b>AGD</b>	Démence à Grains Argyrophiles
<b>APP</b>	Protéine Précurseur de l'Amyloïde
<b>BDNF</b>	Brain-Derived Neurotrophic Factor
<b>CBD</b>	Dégénérescence Cortico-Basale
<b>CR</b>	Extrémité C-terminale de Tau
<b>EGTA</b>	Acide EGTazique
<b>DTT</b>	DiThioThréitol
<b>FTDP-17</b>	Démence Fronto-Temporale et Parkinsonisme associée au chromosome 17
<b>FDA</b>	Food and Drug Administration
<b>GEE-EDC</b>	Glycine Ethyl Ester - 1-Ethyl-3-(3-Diméthylaminopropyl)-Carbodiimide
<b>HFIP</b>	Hexafluoropropan-2-ol
<b>HCCA</b>	Acide $\alpha$ -Cyano-4-hydroxycinnamique
<b>IRM</b>	Imagerie par Résonance Magnétique
<b>LC-MS</b>	Chromatographie Liquide – couplée à de la Spectrométrie de Masse
<b>LCR</b>	Liquide Céphalo-Rachidien
<b>MALDI-TOF</b>	Matrix Assisted Laser Desorption Ionization – Time Of Flight
<b>MES</b>	acide 2-(N-Morpholino)EthanoSulfonique
<b>MET</b>	Microscopie Electronique à Transmission
<b>MTBR</b>	Domaines de liaison aux microtubules de Tau
<b>NPF</b>	Narrow Pick Filament
<b>PART</b>	Primary-Age Related Tauopathy
<b>PEP</b>	Parkinsonisme post-encéphalique
<b>Peptide A<math>\beta</math></b>	Peptide Amyloïde Bêta
<b>PHF</b>	Paired Helical Filament
<b>PiD</b>	Maladie de Pick
<b>PMSF</b>	Fluorure de PhénylMéthylSulfonyle
<b>PR</b>	Domaine de projection de Tau
<b>PRR</b>	Région riche en prolines de Tau
<b>PSP</b>	Paralyse Supranucléaire Progressive
<b>PrP<sup>C</sup></b>	Protéine Prion normale
<b>PrP<sup>Sc</sup></b>	Protéine Prion scrapie
<b>RMN</b>	Résonance Magnétique Nucléaire
<b>ROS</b>	Espèces Réactives de l'Oxygène
<b>SA</b>	Acide Sinapinique
<b>SDD-AGE</b>	Semi-Denaturing Detergent Agarose Gel Electrophoresis
<b>SDS-PAGE</b>	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
<b>SF</b>	Straight Filament
<b>SNU</b>	Small Unit
<b>STEM</b>	Microscope électronique en transmission à balayage
<b>TEP</b>	Tomographie par Emission de Positons
<b>TFA</b>	Acide trifluoroacétique
<b>WPF</b>	Wide Pick Filament

## Avant-Propos

Les protéinopathies sont des pathologies caractérisées par l'agrégation anormale d'une ou de plusieurs protéines, retrouvées sous forme d'agrégats intracellulaires ou extracellulaires. L'observation de caractéristiques communes suggère un mécanisme pathologique partagé pour ces différentes pathologies. Parmi elles, les tauopathies constituent une famille de pathologies caractérisées par l'agrégation de la protéine Tau et l'observation d'enchevêtrements intracellulaires d'agrégats de la protéine Tau. Ces pathologies présentent des similitudes et des dissemblances cliniques, anatomopathologiques et biochimiques. Comme nous le développerons en introduction, le diagnostic définitif différentiel des tauopathies reste difficile, voir impossible à ce jour; seul l'examen histologique post-mortem permet de confirmer ou d'infirmer le diagnostic posé dans les derniers stades de la pathologie. L'identification de déterminants moléculaires permettant de distinguer les tauopathies permettraient de concevoir des outils pour élaborer un diagnostic différentiel robuste des tauopathies, une prise en charge adaptée des patients, et d'ouvrir la voie vers le développement de nouvelles thérapies.

Parmi les cibles possibles, la protéine Tau semble être la plus prometteuse, non seulement parce que la protéine Tau est constitutive des fibres de Tau qui propagent de cellule en cellule, mais aussi parce qu'il a été montré que la composition en isoformes de Tau différait selon les populations cellulaires affectées et selon la tauopathie considérée.

C'est dans ce cadre que s'inscrit la bourse décernée à l'équipe de Ronald Melki par la Fondation Vaincre Alzheimer et ainsi mon projet de thèse dont l'objectif a été de mettre en place des approches capables de cartographier la surface accessible au solvant de fibres de Tau distinctes dans le but d'identifier des déterminants moléculaires structuraux permettant de discriminer des agrégats de Tau distincts.

Les problématiques qui ont conduit à mon projet de thèse reposent sur les observations suivantes:

- (1) Chez les patients atteints de tauopathies sont retrouvées des inclusions cytoplasmiques de Tau agrégée, corrélées à une dégénérescence cellulaire
- (2) La chronologie des régions touchées par l'agrégation de la protéine Tau et les symptômes apparents sont différents en fonction des tauopathies
- (3) Dans certaines populations cellulaires, l'isoforme de Tau, 3R et/ou 4R (3 séquences répétées et/ou 4 séquences répétées dans la partie C-terminale de la protéine), qui

compose les agrégats diffère en fonction de la tauopathie considérée et semble constituer une marque biochimique caractéristique de chaque pathologie

- (4) La protéine Tau se propage de façon prion-like dans le système nerveux central des patients atteints de tauopathies
- (5) Dans le cas des tauopathies et des synucléinopathies, il a été mis en évidence qu'à chaque souche d'agrégats correspondait une pathologie ; les propriétés conformationnelles des différents assemblages macroprotéiques seraient donc spécifiques, voire causales, de chaque tauopathie.

Ainsi, en introduction, je présenterai les caractéristiques principales qui définissent aujourd'hui les protéinopathies. Ensuite, je dresserai un tableau épidémiologique, clinique, anatomopathologique et biochimique des principales tauopathies. Par la suite, l'état des connaissances sur la protéine Tau, du gène à la structure de ses conformères fera l'objet d'une discussion. Enfin, je présenterai succinctement les méthodes disponibles pour l'étude structurale d'assemblages insolubles de haut poids moléculaire.

L'objectif de ma thèse était de caractériser structurellement des polymorphes fibrillaires de Tau 1N3R et 1N4R humaine pour identifier des déterminants conformationnels et de comparer ces cartographies structurales pour mettre en évidence de potentielles différences entre ces assemblages.

Pour ce faire, j'ai développé deux approches dont la mise au point est présentée dans la première partie des résultats. J'ai ensuite appliqué ces méthodes à l'analyse des fibres de Tau 1N3R et 1N4R. L'ensemble des avancées est présenté dans la deuxième partie des résultats sous forme d'un article actuellement en cours de finalisation.

Je finirai ce manuscrit en reprenant les conclusions principales et les perspectives de ce travail.

## I. Introduction

### I. Maladies à prions et protéinopathies

Les protéinopathies sont des pathologies pour lesquelles il est observé une agrégation anormale d'une ou plusieurs protéines dans le milieu intra- ou extra- cellulaire. Ainsi, certaines protéines adoptent une conformation « anormale », différente de celle observée dans un contexte physiologique sain. Ce variant structural engendre une agrégation de la protéine qui sera transmise aux structures avoisinantes. Ces assemblages protéiques sont associés à des propriétés conformationnelles, biochimiques et pathologiques particulières.

#### I.1 Bref historique des maladies à prions et généralités sur les protéinopathies

Dès le XVII<sup>ème</sup> siècle, la tremblante ovine, encéphalopathie spongiforme transmissible, est décrite (Field. 1976). Elle se manifeste par des troubles du comportement et des troubles locomoteurs chez les petits ruminants. Dans les années 1930, Cuillé et Chelle démontrent le caractère transmissible de la tremblante en inoculant des homogénats de cerveaux de brebis mortes des suites de la maladie à des animaux sains qui seront à leur tour affectés (Cuillé and Chelle. 1936). Apparu dans une tribu Fore en Papouasie-Nouvelle Guinée, le Kuru est une encéphalopathie humaine associée à un syndrome cérébelleux. Un long travail de recherche effectué durant les années 1960 conduit Gajdusek et ses collaborateurs à justifier le caractère transmissible du Kuru chez le chimpanzé et à en déterminer la transmission, assurée par des rites anthropophagiques mortuaires (Liberski 2013).

C'est en 1982 que Stanley Prusiner édicte une définition des protéines prions dans le journal Science (Prusiner. 1982) qui le conduira à l'obtention du prix Nobel de médecine en 1997. Les prions ( $\text{PrP}^{\text{Sc}}$ ) sont donc des particules protéiques infectieuses ne contenant pas d'acide nucléique et adoptant une structure tridimensionnelle particulière « scrapie » qui peut se propager à une protéine prion normale ( $\text{PrP}^{\text{C}}$ ).  $\text{PrP}^{\text{Sc}}$  est transmissible de cellule à cellule, d'individu à individu et peut passer la barrière d'espèce. L'illustration la plus célèbre de ce phénomène est certainement la maladie de Creutzfeldt-Jakob (MCJ) et son nouveau variant ( $\text{vMCJ}$ ) (Hill *et al.* 1997). On sait aujourd'hui que la MCJ peut avoir une cause génétique ou sporadique ; cette dernière représentant 85% de l'ensemble des Encéphalopathies Subaiguës Spongiformes Transmissibles (ESST) diagnostiquées chaque année d'après l'INSERM. Dans les années 80, certaines personnes ayant reçu de l'hormone de croissance provenant de

l'hypophyse de patients décédés ou encore des greffes de tissus (cornée, dure-mère), avaient développé une MCJ prouvant la possible transmission inter-individuelle de la MCJ. On parle alors de MCJ iatrogène (Brown *et al.* 2012).

Ainsi, cette même protéine Prion, PrP, est impliquée dans un certain nombre de neuropathologies humaines telles que la Maladie de Creutzfeld-Jakob (MCJ), la variante de la Maladie de Creutzfeld-Jakob (vMCJ), le Kuru, le Syndrome de Gerstmann-Sträussler-Scheinker (SGSS) ou encore l'Insomnie Fatale Familiale (IFF).

Les protéinopathies, dont font partie les maladies à prions, regroupent un certain nombre de pathologies du système nerveux, enjeux majeurs de santé publique, dont une liste non exhaustive est présentée dans la **Table 1**.

Protéinopathies	Protéines agrégées
<b>Maladie d'Alzheimer</b>	Peptide A $\beta$ et Protéine Tau
<b>Maladie de Parkinson</b> <b>Atrophie Multi-Systématisée (AMS)</b> <b>Démence à corps de Lewy (DLB)</b>	$\alpha$ -Synucléine
<b>Autres Tauopathies</b>	Protéine Tau
<b>Maladie de Huntington</b>	Huntingtine avec queue polyglutamine
<b>Sclérose Latérale Amyotrophique familiale</b>	Superoxide dismutase 1 (SOD1)
<b>Encéphalopathies spongiformes (Kuru, Insomnie Fatale familiale,...)</b>	Protéine Prion
<b>Neuropathie amyloïde familiale</b>	Transthyréline mutée

**Table 1.** Protéinopathies du système nerveux et protéines agrégées associées

Adaptée de Sweeney *et al.* 2017

Le lien entre les maladies à prions et les autres protéinopathies, est basé sur l'observation de plusieurs caractéristiques partagées. En effet, dans un contexte pathologique, on observe une conversion de la protéine « normale » vers une protéine mépliée. Ce conformère est alors capable de s'agrèger. De plus, ces éléments mépliés agrégés sont capables de se répliquer en recrutant la protéine endogène. Enfin, il a été démontré que ces espèces agrégées étaient capables de se propager de cellule à cellule et, plus largement, d'une région cérébrale à une

autre. Toutes ces similitudes sous-tendent un éventuel mécanisme commun d'agrégation et de toxicité résultante. Toutefois, il manque à ce jour un élément essentiel à la définition *stricto sensu* de prion : celle d'infectivité inter-individuelle et inter-espèce. En effet, bien qu'une infectivité inter-cellulaire et inter-régionale a été mise en évidence, il n'a pas encore été mis en évidence un phénomène de transmission entre individus ou un franchissement de la barrière d'espèce dans le cas des autres protéinopathies. Ainsi, en regard de ces nombreuses propriétés partagées, la communauté scientifique tend à s'accorder pour associer à ces pathologies le terme « prion-like » (Aguzzi. 2009, Scialò *et al.* 2019). Nous discuterons plus loin de ces caractéristiques « prion-like ».

Dans le cadre de ces différentes pathologies, il est donc observé une forte corrélation entre présence de particules protéiques agrégées et toxicité cellulaire. Il en résulte la question de toxicité des agrégats. Le caractère pathologique de cette agrégation protéique anormale peut s'expliquer à la fois par le gain d'une fonction toxique, par la perte de la fonction normale de la protéine ou un effet des deux.

#### *Perte de fonction normale ou gain de fonction toxique ?*

D'une part, ces protéines retrouvées agrégées dans le cadre des protéinopathies listées dans la **Table 1** possèdent de multiples fonctions dans un contexte non pathologique. Ainsi, il a par exemple été suggéré que l' $\alpha$ -synucléine, retrouvée agrégée dans le cadre des synucléinopathies, jouait un rôle dans la plasticité synaptique, le relargage de neurotransmetteurs et, par le biais d'interactions, pourrait même jouer un rôle de chaperon (Bendor *et al.* 2014 ; Burré. 2015). La protéine Huntingtine, retrouvée agrégée dans la maladie de Huntington, a quant à elle un rôle essentiel dans le développement embryonnaire et possède une activité anti-apoptotique chez l'adulte (Cattaneo *et al.* 2005). Le peptide A $\beta$ , impliqué dans la maladie d'Alzheimer, dérive du clivage protéolytique de la glycoprotéine transmembranaire APP (Protéine Précurseur de l'Amyloïde) qui détient une activité trophique pour les neurones et les synapses (Müller *et al.* 2012). Enfin, la protéine Tau, qui fera l'objet du chapitre III, régule la polymérisation des microtubules et est impliquée dans leur stabilité (Weingarten *et al.* 1975). Ainsi, la perte de ces différentes fonctions physiologiques, qu'elles soient directes ou indirectes (interactions interrompues avec des partenaires moléculaires et cellulaires), pourrait être à l'origine de la dégénérescence neuronale observée dans le contexte pathologique.

D'autre part, ces fibres pourraient elle-mêmes constituer des espèces toxiques pour leur hôte. En effet, la présence de ces agrégats pathologiques engendre une réponse au niveau

cellulaire : le système HSR (Heat Shock Response) dans le cytosol et le système UPR (Unfolded Protein Response) dans le réticulum endoplasmique (Lim and Yue. 2015). Ces processus font notamment intervenir les protéines chaperons et la voie du protéasome afin de rétablir l'homéostasie cellulaire et, *in fine*, d'assurer la survie neuronale. L'établissement au cours de l'évolution de ces machineries tend à prouver que les protéines mal repliées et leur agrégation résultante constituent intrinsèquement un réel danger pour la cellule hôte. Pour appuyer cette hypothèse, des déficiences de la voie Ubiquitine-Protéasome ont été largement décrites dans différentes protéinopathies telles que la démence fronto-temporale, la maladie d'Alzheimer ou encore la maladie de Parkinson. Dans des cellules exprimant le mutant du gène Préséniline 1, prédisposant à la maladie d'Alzheimer, il a été observé des défauts du contrôle qualité opéré sur les protéines dans le réticulum endoplasmique (Katayama *et al.* 1999 ; Sherman *et al.* 2001). Récemment, il a pu être mis en évidence une fragmentation de fibres matures de Tau et d' $\alpha$ -synucléine par le protéasome de façon ATP-dépendante (Cliffe *et al.* 2019). De plus, la toxicité intrinsèque des agrégats pourrait reposer sur leur capacité à séquestrer les protéines chaperons et à inhiber le protéasome comme semblent le suggérer les travaux réalisés sur SOD1, retrouvée agrégée dans la Sclérose Latérale Amyotrophique familiale (Renner and Melki. 2014). Par ailleurs, de nombreuses études ont été menées sur les perturbations de l'équilibre redox intracellulaire et la quantité de  $Ca^{2+}$  libre dans les cellules exposées aux agrégats. Le déséquilibre redox est associé à une augmentation de la quantité d'espèces réactives de l'oxygène (ROS), conjointement à des défauts du métabolisme mitochondrial, qui pourraient impacter négativement la survie cellulaire (Lefaki *et al.* 2017). Des travaux ont montré que le vieillissement et le stress oxydatif avaient des effets directs sur le taux d'expression et l'activité du protéasome et donc sur la survie neuronale. Cette présence accrue de ROS pourrait être la conséquence de l'accumulation intracellulaire du  $Ca^{2+}$  que la cellule tente d'éliminer par le biais des pompes membranaires, grandes consommatrices d'ATP dont la synthèse est un processus oxydatif (Squier. 2001). Une autre piste pouvant expliquer le rôle toxique de ces agrégats est la formation de pores membranaires (observés notamment pour A $\beta$ ,  $\alpha$ -synucléine et polyQ par microscopie électronique et AFM), qui pourrait être à l'origine du déséquilibre calcique. De façon réciproque, l'effet toxique de l'expansion PolyQ peut être supprimé grâce à l'expression des Gluthation peroxidases qui jouent un rôle prépondérant dans le catabolisme des ROS (Stefani and Dobson 2003 ; Holmes *et al.* 2014). Enfin, dans un modèle de souris Alzheimer, il a été suggéré un effet collaboratif de Tau et d'A $\beta$  sur l'activité mitochondriale, à l'origine d'une accumulation de ROS (Rhein *et al.* 2009). D'autre part, il a récemment été démontré que les fibres de Tau pouvaient former des agglomérats à la membrane neuronale et piègeraient ainsi la diffusion des récepteurs AMPA au niveau des synapses excitatrices et modulerait donc la vulnérabilité des neurones à l'influx de  $Ca^{2+}$  qui suit le potentiel d'action (Shrivastava *et al.* 2019).

Pour conclure, considérant l'ensemble des arguments développés en faveur de la perte de fonctions normales et du gain de fonctions toxiques liés à ces agrégats, il semblerait raisonnable de penser que la toxicité observée puisse résulter de l'effet combiné des deux phénomènes. Le poids de chacun pourrait d'ailleurs différer en fonction de la protéine agrégée et même des arrangements conformationnels inhérents à cette agrégation (séquestration différentielle de partenaires).

## I.2 Mécanismes moléculaires de l'agrégation

### *L'agrégation à l'origine des protéinopathies ?*

La présence de dépôts protéiques étant la caractéristique commune à toutes les maladies neurodégénératives listées ci-dessus, ceci suggère un lien de causalité entre la formation d'agrégats et le développement de la pathologie. La nature spécifique des espèces pathogéniques (monomères, oligomères ou fibres multimériques matures) et les mécanismes par lesquels ces dernières endommagent les cellules hôtes restent débattus et sujets de nombreux travaux de recherche (Golde *et al.* 2013 ; Mudher *et al.* 2017). Une hypothèse prévaut toutefois : le dysfonctionnement cellulaire résulte de l'interaction de ces agrégats avec des acteurs cellulaires majeurs. D'un point de vue biochimique, les espèces agrégantes peuvent prendre différentes formes en fonction de la protéinopathie considérée (Marsh. 2019) : protéine entière, peptides biologiques, fragments protéolytiques résultant d'un processing finement établi ou d'une dégradation plus large. La séquence des protéines considérées peut être sauvage (forme sporadique de la protéinopathie) ou mutée (formes familiales). Les formes familiales de protéinopathies sont généralement à l'origine d'une apparition précoce des symptômes et de phénotypes plus sévères. Une transmission génétique de certaines maladies neurodégénératives (e.g Maladie de Huntington) est parfois observée. L'existence de ces formes familiales et la capacité accrue des protéines mutées à s'agréger *in vitro* renforcent l'hypothèse selon laquelle que le processus d'agrégation est à l'origine même de la pathologie.

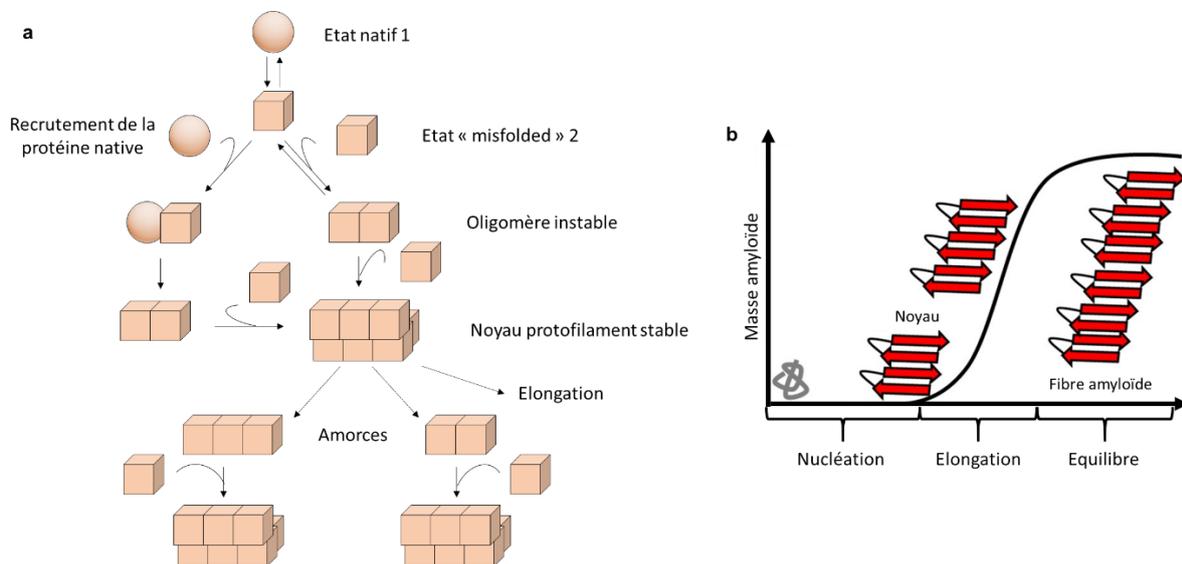
### *Mécanisme d'agrégation*

La formation de structures amyloïdes (i.e agrégats) est rendue possible par le caractère méplié de protéines ; la structure tridimensionnelle est alors différente de celle dans laquelle elle assure normalement ses fonctions *in vivo*. Ceci peut être dû à la présence de mutations, de

défaut de processing, d'interactions aberrantes avec des ions métalliques, de changements environnementaux (pH, température), de modifications chimiques. Ces paramètres peuvent perturber l'équilibre résidant entre les états 1 (« normal ») et 2 (« méplié ») de la protéine et ainsi enrichir la population 2 dans un contexte pathologique, favorisant ainsi la formation d'agrégats par effet de concentration. Cet enrichissement peut être influencé par un défaut de clairance ou par une production exacerbée de la protéine. La présence d'agrégats matures semble favoriser la formation d'agrégats *de novo*.

Ces assemblages matures seraient issus de l'agrégation d'espèces plus petites, oligomères ou protofibrilles d'apparence ronde, agrégats amorphes ou micelles, au microscope électronique, dont la structure diffère de celle des fibres matures, comme semblent l'indiquer des expériences d'immunodétection par des anticorps conformationnels dirigés contre des agrégats matures. La nature toxique des oligomères contre celle des fibres matures est à l'heure actuelle sujet de vifs débats au sein de la communauté (Stefani and Dobson 2003).

Le mécanisme de nucléation et d'élongation de ces assemblages protéiques a été, et reste, au cœur de nombreuses études ; un schéma consensus et la courbe sigmoïde associée sont présentés en **Figure 1**.



**Figure 1.** Mécanisme général de formation des fibres amyloïdes proposé et cinétique de formation

(a) Mécanisme général de formation des fibres amyloïdes proposé avec recrutement de la protéine native ou mépliée par la protéine mépliée conduisant à la formation d'un oligomère instable qui donnera naissance à un noyau de protofilament stable pouvant, soit s'allonger pour former des fibres matures, soit se dissocier partiellement et servir d'amorces à la formation de néoprotofilaments. Issu de Brundin *et al.* 2010 (b) Représentation sigmoïde de la formation de fibres amyloïdes divisée en trois phases : (1) Nucléation/Latence (2) Elongation (3) Etat stationnaire. Issu de Chatani & Yamamoto 2018.

Certains éléments semblent influencer ce processus d'agrégation. En effet, l'exposition au solvant transitoire de régions riches en acides aminés hydrophobes favorisent les interactions intermoléculaires et donc la formation d'oligomères puis des fibres matures. De plus, la concentration en amorces pré-fibrillaires associées à des formes natives de la protéine monomérique favorise et accélère l'agrégation jusqu'à supprimer la phase de latence observée lors de la nucléation initiale ; phénomène largement observable dans les expériences de RT-Quick ou PMCA (Brundin *et al.* 2010 ; Fenyi *et al.* 2019). Ce phénomène a d'ailleurs été aussi mis en évidence pour Tau (Ramachandran and Udgaonkar. 2013). Plus récemment, la communauté scientifique s'interroge sur un possible phénomène de séparation de phase de la protéine, notamment de Tau, qui conduirait à augmenter localement sa concentration et participerait ainsi au processus d'agrégation (Wegmann *et al.* 2018). Il s'agit d'un processus de compartimentation sans membrane, une séparation de deux phases coexistantes induite par des interactions intermoléculaires et qui serait dépendante des domaines de faible complexité (intrinsèquement désordonnés et enrichis en acides aminés polaires non chargés Gln et Asn) (Scheckel & Aguzzi 2018).

### I.3 Caractère prion-like des protéinopathies

A partir de la définition de « prion » établie par Prusiner en 1982, la communauté scientifique s'est rapidement intéressée au caractère dit « prion-like » des protéinopathies humaines mettant en jeu d'autres protéines agrégeantes. De très nombreux groupes ont ainsi pu mettre en évidence une transmission de cellule à cellule et d'une région cérébrale à une autre.

#### *Transmission prion-like des agrégats*

Le phénomène de transmission de cellule à cellule des agrégats pathologiques démontré par Prusiner pour la Prp et sa potentielle transposition aux maladies neurodégénératives constitue un challenge scientifique prépondérant. Les conditions nécessaires à cette transmission de cellule à cellule qui ont été proposées sont les suivantes (Brundin *et al.* 2010) :

- (1) Les agrégats doivent être capables de recruter la protéine native pour procéder à leur élongation et de générer des amorces nécessaires à l'amplification de l'agrégation (voir Figure 1).
- (2) La cellule infectée doit continuer à produire des formes natives de la protéine nécessaire à son recrutement par les espèces agrégées.

(3) Les espèces agrégées transmissibles doivent être relarguées par la cellule hôte (par voie passive ou exocytose).

(4) Les espèces agrégées doivent se lier à la membrane (ou ses composants protéiques/lipidiques) pour entrer dans la cellule réceptrice.

(Alternative 3/4) Les espèces agrégées sont transmises par des nanotubes établis entre la cellule hôte et la cellule réceptrice.

Ainsi, de nombreuses études ont tenté de répondre à la question de transmission cellule à cellule des agrégats retrouvés dans les différentes pathologies neurodégénératives ; en guise d'exemple, nous développerons brièvement ci-après certains arguments en faveur de la propagation de ces assemblages. L'examen histologique de cerveaux de patients atteints de la maladie de Parkinson décédés tardivement après avoir reçu des greffes de neurones embryonnaires sains, a révélé la présence de corps de Lewy au sein des neurones greffés (Kordower *et al.* 2008 ; Kurowska *et al.* 2011). Ceci suggère que les espèces présentes dans le cerveau d'un patient parkinsonien seraient capables de convertir la forme native de l' $\alpha$ -synucléine vers sa forme mépliée pro-agrégante et de se propager aux régions cérébrales avoisinantes. Des études ont ensuite montré que de l' $\alpha$ -synucléine fibrillaire, soniquée et marquée puis ajoutée à des cultures cellulaires, était retrouvée internalisée puis transférée aux cellules voisines. De plus, ces fibres exogènes induisaient la formation de corps d'inclusion dans les cellules réceptrices. Ces fibres pourraient donc jouer le rôle d'amorces et recruter la protéine endogène pour former des agrégats (Luk *et al.* 2009). Enfin, l'injection de fibres synthétiques d' $\alpha$ -synucléine dans le striatum de souris sauvages engendre une agrégation endogène, une dégénérescence des neurones dopaminergiques et l'apparition de troubles moteurs (Luk *et al.* 2012). A l'image de ces fibres générées *in vitro*, l'injection intracérébrale de fractions sarkosyl-insolubles issues de patients atteints de démence à corps de Lewy induit une synucléinopathie chez la souris sauvage (Masuda-Suzukake *et al.* 2013). De la même façon, l'injection d'homogénats de cerveaux de patients atteints de la maladie d'Alzheimer et d'agrégats d'A $\beta$  synthétiques est à l'origine d'une  $\beta$ -amyloïdose et du développement ultérieur de la pathologie chez la souris transgénique APP23<sup>1</sup> ; cependant, l'expérience n'a pas été fructueuse chez la souris sauvage (Meyer-Luehmann *et al.* 2006 ; Stöhr *et al.* 2012). Concernant la pathologie Tau, c'est seulement très récemment, en 2016, qu'il a été mis en évidence que l'injection d'agrégats de Tau purifiés (i.e PHF) du cerveau d'un patient atteint de la maladie d'Alzheimer induisait une pathologie argyrophile 4R chez la souris

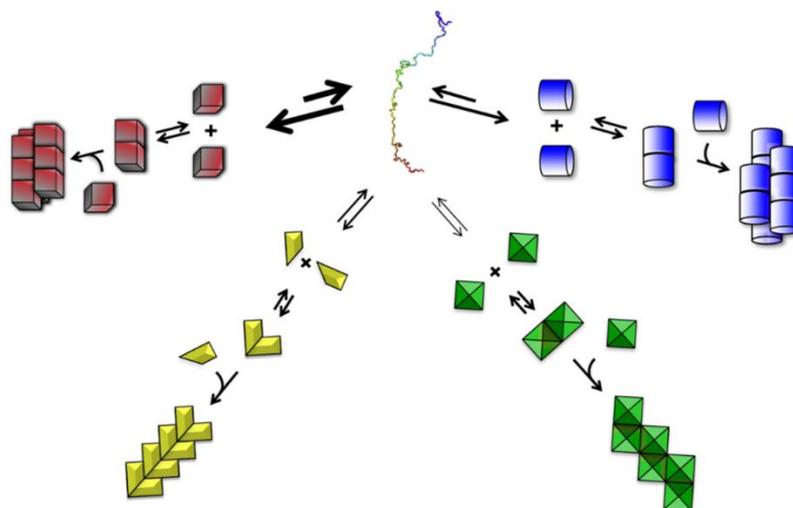
---

<sup>1</sup> Le modèle transgénique de souris APP23 comprend l'expression de l'APP751 humaine portant une double mutation K670N/M671L sous le contrôle du promoteur neural murin Thy-1. Le transgène est exprimé 7 fois plus que l'APP endogène, de façon prévalente au sein du néocortex et de l'hippocampe (Van Dam *et al.* 2005).

sauvage. (Audouard *et al.* 2016 ; Guo *et al.* 2016). Alors que la capacité d'ensemencement des fibres de Tau formées *in vitro* (ici, K18 P301L) chez la souris exprimant une forme mutée de Tau humaine a été validée (Peeraer *et al.* 2015), elle reste encore à démontrer chez la souris sauvage.

### Notion de souches

La notion de souches a émergé avec l'identification de la protéine prion PrP et son implication dans différentes pathologies communément regroupées sous le nom d'encéphalopathies spongiformes. Ainsi, la protéine PrP peut s'agréger sous plusieurs conformations qui (1) seront associées différemment aux pathologies et (2) peuvent parfois coexister. Ces conformères, aussi appelés « souches » (**Figure 2**) présentent des surfaces accessibles différentes qui conduisent à des profils de protéolyse à la protéinase K distincts. De plus, des différences de contenu en structures secondaires de ces conformères résultent des profils spectraux FTIR différents. Les caractéristiques conformationnelles de ces souches, notamment de Tau, sont stables et se perpétuent lors d'expériences répétées d'ensemencement (ou « seeding » en anglais) (Frost and Diamond 2010). Elles sont caractérisées par des temps d'incubation, des conséquences histologiques et cliniques différents. Certains conformères pourraient d'ailleurs être à l'origine du franchissement de la barrière d'espèce. Ainsi, le polymorphisme de la protéine prion influence sa capacité à s'agréger, son potentiel infectieux et son impact sur le phénotype de l'hôte (Morales *et al.* 2007).



**Figure 2.** Représentation schématique de différents conformères d'assemblages protéiques ou souches. Issu de Melki 2018.

Cette hétérogénéité conformationnelle que présente la protéine PrP, indépendante de la séquence de la protéine, a également été identifiée dans d'autres maladies neurodégénératives. En effet, l' $\alpha$ -synucléine et le peptide A $\beta$ 1-40 par exemple peuvent adopter *in vitro* différentes conformations en fonction des conditions d'assemblage (Bousset *et al.* 2013 ; Wood *et al.* 1996). De plus, lorsque ces différentes souches sont injectées à des rats, différents phénotypes sont observés, associés à des synucléinopathies distinctes (Peelaerts *et al.* 2015). Concernant Tau, différentes souches de Tau issues de patients atteints de tauopathies distinctes ont pu être isolées. De plus, lorsque des clones HEK293<sup>2</sup> exprimant différentes souches de Tau ont été injectés chez la souris transgénique P301S<sup>3</sup> ; ils ont été à l'origine de phénotypes distincts, stables entre les générations. (Sanders *et al.* 2014). Enfin, des injections d'homogénats de cerveaux de patients atteints de différentes tauopathies ont été réalisées sur des souris non transgéniques et ont généré des phénotypes cellulaires-spécifiques différents (Narasimhan *et al.* 2017) et des inclusions de Tau spécifiques de la pathologie humaine pour chacune des tauopathies (Clavaguera *et al.* 2013).

Pour conclure, à l'image des maladies à prions, il a été identifié un certain nombre de protéinopathies humaines au sein du système nerveux mettant en jeu différentes protéines toutes retrouvées agrégées. La capacité des espèces assemblées à recruter la protéine endogène normale et à en potentialiser l'agrégation a pu être démontrée par de nombreux travaux pour différentes maladies dégénératives. La transmission de cellule à cellule et entre régions cérébrales a également été mise en évidence. *In vivo* et *in vitro*, ces espèces assemblées peuvent adopter, pour une même protéine, différentes conformations, appelées souches, qui seront à l'origine de phénotypes pathologiques distincts (Frost and Diamond. 2010). Parmi ces protéinopathies, nous nous sommes intéressés particulièrement aux pathologies pour lesquelles la protéine Tau est retrouvée mépliée et agrégée : les tauopathies.

---

<sup>2</sup> Cellules rénales embryonnaires exprimant ici une forme tronquée de Tau (MTBR) portant ou non différentes mutations

<sup>3</sup> Le modèle transgénique de souris P301S comprend l'expression de Tau humaine 0N4R portant une mutation P301S sous le contrôle du promoteur neural murin Thy-1.2. Le transgène est exprimé 2 fois plus que la Tau endogène de souris (Allen *et al.* 2002)

## II. Les tauopathies

### II.1 Historique et généralités

En 1892, Arnold Pick décrit pour la première fois en clinique le cas d'une patiente présentant une atrophie cérébrale associée à une démence présénile et une dysphasie (Dickson. 1998). Plus tard, en 1906, Alois Alzheimer découvre les symptômes de la maladie, conséquence d'une dégénérescence corticale, maladie qui portera son nom. Il y associe alors des observations histologiques aujourd'hui caractéristiques de la maladie : les plaques amyloïdes et les dégénérescences neurofibrillaires (Stelzmann *et al.* 1995). Il établira également que ces traits histopathologiques ne sont pas retrouvés dans le cas de la maladie de Pick pour laquelle il observait des neurones gonflés et des inclusions argyrophiles (Constantinidis *et al.* 1974).

La maladie d'Alzheimer est aujourd'hui la plus répandue des maladies neurodégénératives (60 à 70% des cas de démences (Rapport Mondial Alzheimer 2015 de l'ADI (Alzheimer's Disease International)) et appartient, tout comme la maladie de Pick, à une famille de pathologies appelée tauopathies. Ces dernières partagent comme caractéristique commune une accumulation intracellulaire de la protéine Tau (Lee *et al.* 2001 ; De Silva *et al.* 2003; Kovacs. 2015). Il existe plus d'une vingtaine de tauopathies dont les tableaux cliniques sont très variés et se recouvrent partiellement. Certaines tauopathies, comme la maladie d'Alzheimer, présentent, en plus de l'agrégation intracellulaire de Tau, une accumulation extracellulaire du peptide A $\beta$  agrégé. Ces tauopathies sont désignées sous le terme de « tauopathies secondaires », et sont à opposer avec les « tauopathies primaires » pour lesquelles seule la protéine Tau est retrouvée agrégée. Cette agrégation de la protéine Tau peut concerner différents types cellulaires du système nerveux central : neurones, astrocytes et oligodendrocytes en fonction de la pathologie considérée. Une liste des tauopathies répertoriées à ce jour est présentée en **Annexe 1**.

Les données d'épidémiologie associées aux différentes tauopathies ne sont pas toutes connues ; ceci s'explique d'une part par la difficulté rencontrée par les praticiens d'apposer un diagnostic certain et d'autre part par la rareté des pathologies (comme la maladie de Pick par exemple). Nous en donnerons un bref aperçu dans la **Table 2**. D'un point de vue global, on estime d'une part la prévalence de l'ensemble des démences à 6,4 % de la population de plus de 65 ans ; elle augmente drastiquement entre 65 et 90 ans jusqu'à atteindre 28,5% ; l'âge constitue donc un facteur de risque majeur. D'autre part, l'incidence des démences est de 19,4/1000 personnes chaque année (Ankri. 2009).

	<b>Incidence</b>	<b>Prévalence</b>	<b>Références</b>
<b>AD</b>	1,2/1000 pers/an entre 65-69 ans 53,5/1000 pers/an après 90 ans	4,4/100 des plus de 65 ans	Ankri. 2009
<b>PSP</b>	1,7/100000 pers/an à 60 ans 14,7/100000 pers/an entre 80 et 89 ans	6/100000	Agarwal and Gilbert. 2019
<b>CBD</b>	0,62-0,92/100000	4,9-7,3/100000	Mahapatra <i>et al.</i> 2004
<b>AGD</b>	n.d	6/1000 des plus de 65 ans et 187/1000 des centenaires	Rodriguez and Grinberg. 2015
<b>PiD</b>	n.d	1/100000 des plus de 60 ans	Williams. 2006

**Table 2.** Données épidémiologiques des différentes tauopathies AD maladie d'Alzheimer, PSP Paralyse Supranucléaire Progressive, CBD Dégénérescence cortico-basale, AGD Maladie aux Grains Argyrophiles et PiD Maladie de Pick.

Ainsi, la maladie d'Alzheimer constitue de loin la tauopathie la plus représentée au sein de la population. La maladie aux grains argyrophyles est elle aussi particulièrement répandue chez le sujet âgé. Aujourd'hui, les tauopathies constituent un enjeu de santé publique majeur.

## II.2 Clinique et Anatomopathologie

Les caractéristiques cliniques et anatomopathologiques développées ci-après sont résumées en **Table 3** (p31).

### II.2.1 Observations cliniques

De manière générale, la démence est définie par l'OMS comme « un syndrome, généralement chronique ou évolutif, dans lequel on observe une altération de la fonction cognitive [...], plus importante que l'on pourrait attendre du vieillissement normal. Elle affecte la mémoire, le raisonnement, l'orientation, la compréhension, le calcul, la capacité d'apprentissage, le langage et le jugement. La conscience n'est pas touchée. Une détérioration du contrôle émotionnel, du comportement social ou de la motivation accompagne souvent, et parfois précède, les troubles de la fonction cognitive. ».

Ainsi, la maladie d'Alzheimer est caractérisée par un déclin cognitif (affectant la mémoire et, selon les cas, le langage, la fonction visio-spatiale, la fonction exécutive, la praxie). La maladie a des conséquences sur la vie quotidienne du patient qui souffre d'une perte d'autonomie progressive (Williams. 2006). Une autre pathologie appelée tauopathie primaire liée à l'âge

(PART ou anciennement SD-NFT, TSPD<sup>4</sup>) se présente aussi en clinique par des pertes de mémoire mais le déclin cognitif est moins prononcé que pour la maladie d'Alzheimer et concerne une population plus âgée (Crary *et al.* 2014).

La paralysie supranucléaire progressive (PSP) est communément caractérisée en clinique par le syndrome de Richardson ((i.e. syndrome Steele–Richardson–Olszewski) associant une ophtalmoplégie supranucléaire, une paralysie pseudobulbaire, une dystonie axiale et une démence. Dans la chronologie de la pathologie, les troubles moteurs apparaissent toujours avant les symptômes cognitifs (Delacourte. 2009). Toutefois, certains patients atteints de PSP ne présentent pas le syndrome de Richardson mais développent un syndrome parkinsonien<sup>5</sup> tandis que de plus rares formes de PSP s'apparentent à une dégénérescence corticobasale ; à l'origine de diagnostics incertains.

La dégénérescence corticobasale (CBD) est caractérisée de façon prédominante par des symptômes moteurs (un syndrome parkinsonien et parfois un syndrome de la main étrangère) mais aussi des troubles cognitifs modérés (dont des troubles du langage) (Armstrong M *et al.* 2013).

La maladie aux grains argyrophiles (AGD), identifiée récemment, a un spectre clinique très hétérogène et est difficilement différentiable des stades avancés de la maladie d'Alzheimer. Les patients présentent une diminution de la mémoire à court terme, des problèmes d'élocution, d'écriture et de lecture, des troubles du comportement et sont désorientés (Delacourte and Jellinger. 2009).

La maladie de Pick (PiD) induit chez les patients un changement du comportement, principale manifestation de la pathologie, une aphasie, une perte de la mémoire et une démence. On parle ici de pathologie présénile, du fait notamment qu'elle se déclare plus précocement que la maladie d'Alzheimer (i.e avant 65 ans) (Dickson. 1998 ; Williams. 2006).

Les patients atteints du syndrome parkinsonien post-encéphalitique (PEP) peuvent présenter une bradykinésie ou une akinésie et une rigidité des membres secondairement à la contraction d'une encéphalite, souvent d'origine virale (Williams. 2006 ; Souirti *et al.* 2007). Elle a été décrite pour la première fois par le neurologue Von Economo comme une complication tardive (parfois plusieurs années après) de l'encéphalite léthargique, pathologie sévère, qui se

---

<sup>4</sup> Anciennement, les dénominations retrouvées dans la littérature pour PART pouvaient être SD-NFT pour "Senile Dementia of the NeuroFibrillary Tangle Type" (ou Démence Sénile de Type Enchevêtrements NeuroFibrillaires) ou TSPD pour « Tangle-Predominant Senile Dementia » (ou Démence Sénile avec prédominance des enchevêtrements [neurofibrillaires])

<sup>5</sup> Le syndrome parkinsonien est un syndrome moteur défini par une akinésie et au moins un des symptômes suivants : rigidité extrapyramidale, tremblements au repos et instabilité posturale (Voisin. 2015)

manifestait par une somnolence accrue et une ophtalmoplégie (Williams. 2006 ; Evidente and Gwinn. 1998 ; Rail *et al.* 1981).

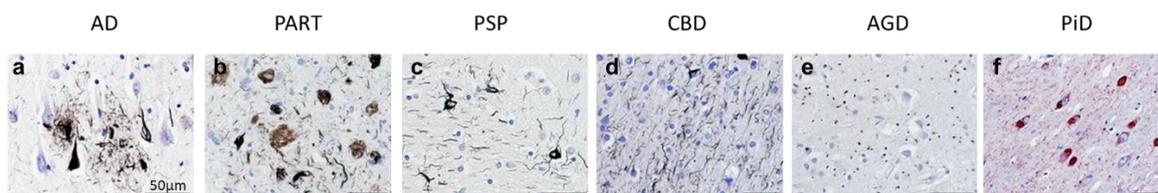
La démence fronto-temporale avec parkinsonisme associée au chromosome 17 (FTDP-17) est une tauopathie d'origine génétique survenant précocement (entre 30 et 50 ans). Elle se manifeste par des troubles cognitifs, moteurs, de comportement et des changements de personnalité. Le phénotype de la pathologie varie entre des familles porteuses de différentes mutations mais aussi au sein même d'une famille. Les manifestations motrices de la maladie rendent le diagnostic différentiel compliqué par rapport à la maladie de Parkinson et à la PSP (Williams. 2006 ; Wszolek *et al.* 2006).

Ainsi, les tauopathies sont associées à des tableaux cliniques aux symptômes partagés, surtout dans les phases précoces de la pathologie. Ces ressemblances sont à l'origine d'une incertitude de diagnostic que seul l'examen anatomo-pathologique peut venir valider.

## II.2.2 Anatomopathologie et évolution de la pathologie Tau

### II.2.2.1 Observations histologiques

Aujourd'hui encore, seul l'examen histopathologique post-mortem des régions cérébrales de patients atteints de tauopathies permet l'établissement d'un diagnostic certain. Historiquement, les dépôts pathologiques (i.e plaques séniles, dégénérescences neurofibrillaires et corps de Pick) étaient visualisés au microscope après marquage argentique, par la méthode de Gallyas publiée au début des années 70 (Uchihara. 2007). Avec l'avènement des techniques d'immunohistochimie, les anatomopathologistes disposent aujourd'hui d'un certain nombre d'anticorps permettant l'observation fine des dépôts intracellulaires de Tau. L'anticorps le plus utilisé est AT8, qui reconnaît Tau sous sa forme agrégée et phosphorylée au niveau de la sérine 202 et de la thréonine 205.



**Figure 3.** Observations histologiques de coupes hippocampales humaines de patients atteints de tauopathies. (a-e) Observations histologiques après coloration à l'argent par la méthode Gallyas-Braak de coupes hippocampales humaines de patients atteints de gauche à droite (a) AD, (b) PART, (c) PSP, (d) CBD, (e) AGD. (f) Observations histologiques par immunodétection (AT100) de coupe hippocampale humaine d'un patient atteint de la maladie de Pick. Issu de Calvaguera *et al.* 2013.

Dans le cas de la maladie d'Alzheimer, bien qu'un diagnostic ne puisse être établi à partir de l'observation globale du cerveau du patient, ce dernier présente une atrophie corticale symétrique affectant de façon prédominante les lobes médio-temporaux, ce qui a pour conséquence une apparence dilatée des ventricules latéraux. La maladie d'Alzheimer est caractérisée en histologie, d'une part, par l'agrégation d'une part du peptide A $\beta$ 40 et A $\beta$ 42 sous la forme de plaques, diffuses ou denses (davantage associée à AD) en extracellulaire dans l'isocortex. D'autre part, sont retrouvés des enchevêtrements (dits « tangles » en anglais) neuro-fibrillaires intracellulaires de la protéine Tau diffus ou focalisés en fonction du stade de la pathologie (**Figure 3.a**). Conséquemment à la mort du neurone hôte, ces enchevêtrements fibrillaires peuvent aussi être retrouvés dans le milieu extracellulaire (Serrano-Pozo *et al.* 2011). En 1991, Heiko Braak établit la chronologie de la maladie d'Alzheimer suivant 6 stades, définissant la progression des lésions intraneuronales en trois phases : précoce, intermédiaire et tardive (Braak and Braak. 1991). Durant les deux premiers stades, les dégénérescences neurofibrillaires de Tau sont essentiellement retrouvées au sein de la région transentorhinale voisine tandis qu'on observe leur progression vers les aires limbiques (notamment dans l'hippocampe) pendant les stades III et IV pour enfin atteindre le néocortex lors de la phase tardive (**Figure 4**). Cette progression spatio-temporelle de Tau agrégée a donc lieu entre régions cérébrales interconnectées, ce qui suggère un mécanisme de transmission « prion-like », de région à région via leurs conections neuronales. En 1992, Hardy et Higgins proposent l'hypothèse de la cascade amyloïde selon laquelle l'agrégation du peptide A $\beta$  serait l'agent causal de la maladie d'Alzheimer et l'agrégation de Tau, au même titre que la mort cellulaire et la démence, une conséquence directe de ce phénomène (Hardy and Higgins. 1992). Ce dogme, à l'origine de nombreuses stratégies thérapeutiques, est aujourd'hui remis en question, notamment par deux observations majeures : (1) les dépôts extracellulaires amyloïdes ne sont pas corrélés à la mort neuronale et au déclin cognitif, (2) la distribution des enchevêtrements fibrillaires de Tau ne co-localise pas avec les agrégats A $\beta$  au niveau régional, comme illustré en **Figure 4** (Ricciarelli and Fedele. 2017).

La tauopathie primaire liée à l'âge (PART) se manifeste macroscopiquement par une atrophie modérée du lobe médio-temporal. Elle est caractérisée par la présence d'enchevêtrements neuro-fibrillaires de la protéine Tau comparable à ceux observés dans le cas de la maladie d'Alzheimer. La progression de la pathologie semble suivre la même distribution spatio-temporelle que la maladie d'Alzheimer entre les stades de Braak I à III mais on n'observe pas d'évolution vers les stades ultérieurs. La différence majeure avec la maladie d'Alzheimer est qu'il n'est pas constaté d'agrégation d'A $\beta$  (**Figure 3.b**) (Kim *et al.* 2019).

L'étude macroscopique des cerveaux de patients atteints de paralysie supranucléaire progressive conduit à deux observations principales : une légère atrophie du cortex frontal et

une atrophie du mésencéphale et du noyau subthalamique. Au niveau microscopique, des enchevêtrements neuro-fibrillaires positifs pour Tau sont observés dans les ganglions de la base, le diencéphale, le tronc cérébral et la moelle épinière. De plus, dans le cadre de la PSP, il est observé des astrocytes hirsutes contenant Tau sous sa forme agrégée (**Figure 3.c**) dans le cortex moteur et le striatum. Enfin, on note également des oligodendrocytes d'aspect bobiné au sein du diencéphale et de la substance blanche cérébelleuse (Dickson *et al.* 2007).

Dans le cas de la dégénérescence cortico-basale, on observe une atrophie corticale asymétrique affectant le cortex antérieur, la région fronto-pariétale et le cortex temporal supérieur. Les ganglions de la base sont également atrophiés, notamment au niveau du noyau caudé et de la substance noire (Armstrong R. 2013). La principale caractéristique histologique de la pathologie est la présence de plaques astrocytaires, ce qui la différencie notamment de la PSP. On observe également un changement de la morphologie des neurones qui apparaissent gonflés dans le gyrus cingulaire, l'amygdale et le cortex insulaire. (**Figure 3.d**) (Parthimos and Schulpis. 2019).

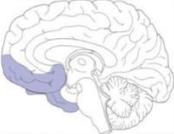
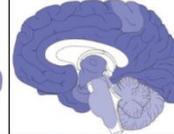
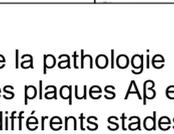
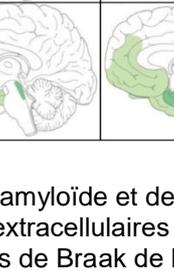
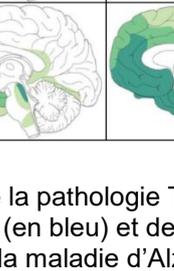
Macroscopiquement, les cerveaux des patients atteints de maladie à grains argyrophiles apparaissent inchangés ou présentent une atrophie légère. Microscopiquement, la maladie, comme son nom l'indique, présente la particulière existence de grains Tau positifs de petite taille (4-8  $\mu\text{m}$ ) dans les dendrites et leurs projections (Grinberg and Heinsen. 2009), et ce de façon abondante dans le système limbique (**Figure 3.e**). La morphologie des oligodendrocytes est également singulière ; ces derniers formant des corps bobinés (« coiled bodies »). Les astrocytes sont positifs pour Tau phosphorylé mais la protéine ne semble pas agrégée (pas de coloration avec la méthode Gallyas). Enfin, les neurones présentent des inclusions intracytoplasmiques enchevêtrées de type enchevêtrements naissants (dits « pre-tangles » en anglais) (Rodriguez and Grinberg. 2015).

Enfin, macroscopiquement, le cerveau des patients atteints de maladie de Pick révèle une atrophie fronto-temporale (Frederick. 2006 ; Kremen *et al.* 2010). Au niveau histologique, elle est très singulièrement caractérisée par la présence de corps de Pick sphériques intra-neuronaux (**Figure 3.f**), et ce essentiellement dans l'hippocampe et les couches corticales supérieures des régions frontale et temporale (Irwin *et al.* 2016).

### II.2.2.2 Evolution de la pathologique

Dans le cadre de la maladie d'Alzheimer, l'évolution des régions touchées par les dégénérescences neurofibrillaires est corrélée à l'apparition progressive des symptômes ; ce n'est pas le cas pour la pathologie A $\beta$  dont le schéma spatio-temporel est différent (**Figure 4**).

Ainsi, la classification établie par Braak décrit l'évolution de la pathologie associée à Tau du lobe médio-temporal, à l'isocortex puis au cortex primaire (Braak and Braak. 1995). La pathologie Tau suit la même évolution régionale dans le cas de la pathologie PART mais n'atteint pas les derniers stades de Braak V-VI (i.e l'isocortex et cortex primaire) (Duyckaerts *et al.* 2015 ; Cray *et al.* 2014).

Maladie d'Alzheimer	Stades I-II	Stades III-IV	Stades V-VI
Plaques A $\beta$			
Inclusions de Tau			

**Figure 4.** Schématisation de la pathologie amyloïde et de la pathologie Tau dans la maladie d'Alzheimer. Progression des plaques A $\beta$  extracellulaires (en bleu) et des inclusions intracellulaires de Tau (en vert) au cours des différents stades de Braak de la maladie d'Alzheimer. Adaptée de Goedert. 2015 et Leuzy *et al.* 2019.

Pour la PSP, les premières régions touchées sont les ganglions de la base puis sont affectés le cervelet, la moelle épinière, la partie postérieure du lobe frontal avant que la pathologie n'envahisse le cortex associatif. En comparaison, dans le cas de la Dégénérescence cortico-basale, les premières régions touchées sont les mêmes que pour la PSP mais évoluent ensuite vers le cortex moteur primaire et le lobe frontal avant d'affecter d'autres zones, notamment le tronc cérébral et le cervelet (Leuzy *et al.* 2019). Les grains argyrophiles positifs pour Tau retrouvés dans la maladie aux grains argyrophiles sont quant à eux d'abord essentiellement localisés dans le cortex entorhinal antérieur puis s'étendent à l'échelle de la région entorhinale entière, l'hippocampe, l'amygdale et l'hypothalamus pour finalement atteindre le néocortex et le tronc cérébral (Ferrer *et al.* 2008). Enfin, dans le cas de la maladie de Pick, les premières régions affectées sont les régions limbiques et le néocortex frontotemporal puis sont atteintes les structures subcorticales. La pathologie Tau touche le cortex moteur primaire et le noyau précérébelleux médullaire dans la troisième phase de la maladie pour finalement envahir le cortex visuel primaire au dernier stade (Irwin *et al.* 2016).

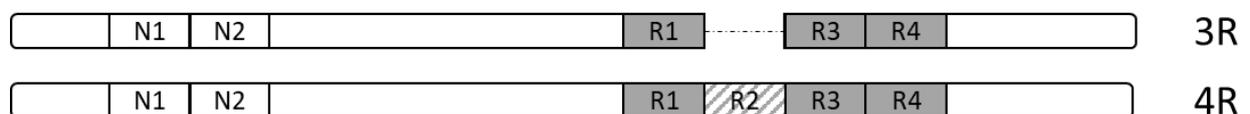
Alors que l'évolution de la pathologie Tau dans le cadre de la maladie d'Alzheimer a été exhaustivement décrite, les autres tauopathies, dont les cohortes de patients sont plus réduites, nécessiteraient d'être davantage documentées d'un point de vue anatomopathologique. Alors que les observations cliniques motivent un examen

anatomopathologique pour établir un diagnostic certain, des déterminants moléculaires et biochimiques des différentes tauopathies ont pu être identifiés et peuvent, dans une certaine mesure, servir de critères discriminants.

### II.3 Caractéristiques moléculaires et biochimiques

Les caractéristiques moléculaires et biochimiques développées ci-après sont résumées en **Table 3** (p31).

Les espèces agrégées de Tau extraites des cerveaux des patients ont été largement étudiées par microscopie électronique à transmission. Il a été montré l'existence d'assemblages protéiques de haut poids moléculaire dont la morphologie pouvait d'une part différer entre les tauopathies mais d'autre part également coexister sous différentes formes chez un même patient (e.g deux types d'agrégats dans le cas de la maladie d'Alzheimer). Ainsi, des analyses biochimiques par western-blot des espèces agrégées dans les différentes tauopathies ont été réalisées après solubilisation des corps d'inclusion en présence de détergent (usuellement le sarkosyl). Ces dernières ont révélé la présence de Tau au sein de ces agrégats sous une forme hyperphosphorylée qui se traduit par deux ou trois bandes principales comprises entre 60 et 68 kDa en fonction des pathologies et de la structure cérébrale considérée (Lee *et al.* 2001) (voir **Figure 7** et partie introductive III.I.1). Comme illustré schématiquement en **Figure 5**, la protéine Tau existe sous différentes isoformes pouvant notamment contenir 3 ou 4 séquences répétées dans la partie C-terminale de la protéine qui donnent respectivement naissance à la catégorie des isoformes 3R et des isoformes 4R.



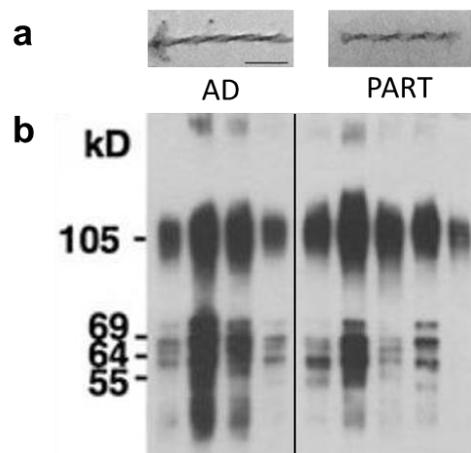
**Figure 5.** Catégories 3R et 4R des isoformes de Tau.

Les isoformes de Tau peuvent être classées dans deux catégories : les isoformes 3R (3 séquences répétées en C-terminal) et les isoformes 4R (4 séquences répétées en C-terminal).

Ainsi, après déphosphorylation, les différentes isoformes de Tau, 3R et/ou 4R, présentes dans les corps d'inclusion ont pu clairement être identifiées et quantifiées.

Dans la maladie d'Alzheimer, au sein des plaques extracellulaires, on retrouve de façon prédominante le peptide A $\beta$ 42 par rapport au peptide A $\beta$ 40. Dans le compartiment intracellulaire, les agrégats de Tau ont pour une majeure partie d'entre eux une structure dite en « Paired-Helical Filaments » ou PHFs d'environ 10nm de diamètre et une périodicité de 65 nm, constitués de 2 protofilaments. Une autre classe de fibres est également présente, en moins grand nombre, les filaments droits ou SFs (appelés « Straight Filaments » en anglais) qui semblent être composées d'un filament unique d'après les observations réalisées par MET dans les années 90, assertion plus tard démentie dans Fitzpatrick *et al.* Nature 2017. Enfin, un autre type de filaments "hybrides", dont la conformation passe de PHF à SF, a été décrit. (Serrano-Pozo *et al.* 2011 ; Crowther. 1991). On estime le ratio 4R:3R au sein des agrégats autour de 0.75 soit une presque égalité entre les deux isoformes (Togo *et al.* 2002).

Dans le cas de PART, comme illustré en **Figure 6**, les observations des agrégats par MET indiquent la présence de PHFs semblables à ceux observés pour AD (Crary *et al.* 2014) et les profils de western-blot des fractions sarkosyl-insolubles d'AD et PART sont très ressemblants. Une évaluation du ratio 4R:3R agrégé par rapport au contrôle soluble révèle une similarité de composition, légèrement plus enrichie en formes 3R pour PART (Santa-Maria *et al.* 2012).



**Figure 6.** Comparaison des déterminants biochimiques d'AD et de PART. (a) Comparaison des observations réalisées au MET sur des agrégats issus d'AD et de PART (barre d'échelle 100nm) (b) Comparaison des profils Western-Blot (anticorps HT7 reconnaissant toutes les isoformes de Tau) étudiant la composition en isoformes de Tau au sein des fractions sarkosyl-insolubles du lobe frontal de patients atteints d'AD et de PART. Issu de Santa-Maria *et al.* 2012 ; Crary *et al.* 2014.

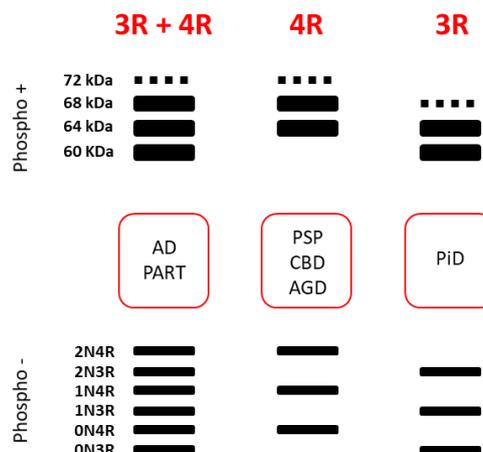
Dans la PSP, les enchevêtrements neurofibrillaires appelés « neurofibrillary tangles » (NFTs) en anglais, et les inclusions gliales sont constitués de filaments droits de Tau (diamètre de 15nm pour les NFTs) (Dickson *et al.* 2007). L'analyse par western-blot des enchevêtrements de Tau issus du pont révèle un ratio 4R:3R de 2.88 pour la PSP-RS (présentant un syndrome de Richardson) et de 1.63 pour la PSP-P (présentant un syndrome parkinsonien) ; la PSP

présente donc un enrichissement plus ou moins marqué de la forme 4R au sein des agrégats en fonction des signes cliniques considérés (Williams *et al.* 2005).

Dans la CBD, la fraction sarkosyl-insoluble du lobe frontal contient des filaments enroulés plus larges et à la périodicité plus grande que ceux retrouvés dans la maladie d'Alzheimer (largeur 13-28 nm ; périodicité 170-200 nm) (Ksiezak-Reding *et al.* 1994). La composition en isoformes de Tau des agrégats retrouvés dans la CBD est, pour certaines régions, comparable à celle de la PSP. En effet, elle est largement décrite comme étant 4R prédominante avec un ratio 4R:3R d'environ 2.5 au niveau du cortex frontal et de 3 au niveau du globus pallidus, structure appartenant aux ganglions de la base (Takanashi *et al.* 2002).

Dans l'AGD, les grains contiennent des filaments droits de Tau mesurant de 9 à 25 nm (Grinberg and Heinsen. 2009). Les analyses par western-blot des fractions sakosyl-insolubles (anticorps WKS44 ; aa 162-178) ont montré que le ratio 4R/3R était en moyenne de 2 :1 dans les premiers stades de la pathologie (Braak I) pour diminuer autour de 1 :1 voire 0.6 :1 dans les stades plus tardifs de la pathologie (Braak II, III) (Togo *et al.* 2002).

Les agrégats de Tau retrouvés chez les patients atteints de la maladie de Pick présentent, dans la substance blanche et grise, essentiellement une conformation en filaments droits (largeur 15 nm) bien que certains filaments type rubans ou encore PHF sont parfois retrouvés. D'autre part, c'est majoritairement l'isoforme 3R qui est retrouvée au sein des agrégats. On notera cependant quelques cas pour lesquels a été observée une prédominance 4R. Aucun calcul de ratio 4R:3R à partir de western-blots n'a été calculé mais David Williams décrit un ratio 4R:3R de 1:3, probablement en considérant l'assertion évoquée ci-dessus. Ces agrégats sont plus abondamment retrouvés dans la substance blanche que dans la substance grise (Zhukareva *et al.* 2002 ; Arai *et al.* 2001 ; Williams. 2006).



**Figure 7.** Schématisation du code-barre moléculaire caractéristique des différentes tauopathies. Avant déphosphorylation (haut) et après déphosphorylation (bas). Issu de Lee *et al.* 2001.

Ainsi, sur la base des ratios 4R:3R composant les agrégats, il est possible de dresser un code-barre moléculaire illustré en **Figure 7** et comprenant trois classes : (1) coexistence relativement égale des formes 3R et 4R (2) prédominance 4R (3) prédominance 3R. Cette classification est largement répandue dans la littérature et mérite toutefois d'être considérée avec précaution. En effet, d'une part, il semblerait que les ratios des isoformes agrégées puissent dépendre pour une même pathologie du stade de la maladie (Togo *et al.* 2002). De plus, la composition des assemblages pourrait également varier en fonction de la région cérébrale considérée, probablement partiellement en raison d'une expression différentielle cellulaire/tissu-spécifique des isoformes de Tau (Delacourte *et al.* 1998 ; Takanashi *et al.* 2002). Enfin, David Williams et ses collaborateurs ont montré qu'une hétérogénéité des symptômes associés à une même tauopathie pouvait être corrélée à une composition en isoformes agrégées différente (Williams *et al.* 2005). Ceci pourrait impliquer que d'autres processus moléculaires associés à Tau et à son agrégation, inconnus à ce jour, sont responsables de l'existence de sous-groupes pour une même pathologie ou que la classification des tauopathies mériterait peut-être d'être révisée. Ainsi, avec l'avancée technique actuelle (séquençage d'ARN à la cellule unique par exemple), il serait intéressant qu'une nouvelle analyse plus fine des ratios 4R:3R soit réalisée. Cette dernière pourrait ainsi considérer différents paramètres pour les différentes tauopathies : type cellulaire, région cérébrale et tableau clinique.

#### II.4 Déterminants génétiques

Les éventuels déterminants génétiques des différentes tauopathies ont fait l'objet d'intenses recherches.

Le gène *MAPT* codant pour Tau est présent sur le bras long du chromosome 17. Ce gène fait partie d'une région plus large, susceptible à un déséquilibre de liaison ayant pour conséquence les haplotypes H1 et H2 (Figure 7), H1 étant aujourd'hui considéré comme un facteur de risque dans de nombreuses maladies neurodégénératives, notamment pour la PSP et la CBD (Kalinderi *et al.* 2009 ; Togo *et al.* 2002 ; Williams *et al.* 2005 ; Williams. 2006). Il n'a cependant jamais été mis en évidence de façon certaine l'implication de l'haplotype H1 dans la maladie d'Alzheimer, PART et l'AGD (Santa-Maria *et al.* 2012 ; Togo *et al.* 2002).

La causalité de mutations du gène *MAPT* dans la maladie d'Alzheimer n'a encore jamais été établie avec certitude. De façon anecdotique, dix mutations du gène *MAPT* ont été reportées pour la PSP, la majeure partie d'entre elles étant localisées au niveau de l'exon 10 et sa région d'épissage (Im *et al.* 2015). Pour la maladie aux grains argyrophiles (AGD), une mutation dans

le gène *MAPT*, S305I (R2) présente chez une famille scandinave et une CNV 17p –Variabilité du Nombre de Copies) ont été identifiées. Alors que certains cas de démences fronto-temporales (DFT) associées à des mutations de *MAPT* ont révélé des profils cliniques extrêmement ressemblants à ceux de PiD (Williams. 2006 ; Pickering-Brown *et al.* 2000), la maladie de Pick est considérée sporadique.

Dans la maladie d'Alzheimer, des formes familiales précoces ont été associées à des mutations d'APP, précurseur du peptide A $\beta$ , ou encore des gènes codant pour Préséniline 1 et 2 (i.e PSEN1 et PSEN2), sous-unités catalytiques de la sécrétase  $\gamma$ , aussi liées à la pathologie amyloïde. De plus, l'allèle codant pour l'apolipoprotéine E4 (APOE  $\epsilon$ 4) est décrit comme un facteur de risque corrélé à l'établissement de la forme sporadique précoce de la maladie (Williams. 2006). L'analyse de la distribution allélique d'apoE n'a révélé aucune corrélation quant à la PSP (Williams *et al.* 2005).

Evoquée plus haut, la démence fronto-temporale avec parkinsonisme associée au chromosome 17 (FTDP-17) est une pathologie rare de transmission autosomique dominante survenant plus précocement que les autres tauopathies (entre 30 et 50 ans). La FTDP-17 est toujours associée à des mutations du gène *MAPT*, différentes selon les familles. Les trois mutations les plus communes, retrouvées dans 60% des cas, sont P301L, N279K et une mutation associée à l'épissage alternatif des exons 10 et 16. Toutefois, la pénétrance du phénotype associée à une mutation donnée peut être variable entre individus d'une même famille (Wszolek *et al.* 2006), suggérant que les mutations du gène *MAPT* ne sont pas seules responsables du phénotype pathologique.

Pour conclure, la plupart des tauopathies sont sporadiques (i.e maladie d'Alzheimer, PART, AGD, PiD), certaines sont associées à l'haplotype H1 (i.e PSP, CBD) et enfin la FTDP-17 a pour origine unique l'existence de mutations au sein de *MAPT*. L'effet sur la liaison aux microtubules et leur polymérisation ainsi que sur l'agrégation de Tau d'un certain nombre de mutations de *MAPT* identifiées chez des patients atteints de tauopathies sera détaillé en partie III.1. A ce jour, en dehors de la FTDP-17, nous ne savons pas si les mutations du gène *MAPT* participent activement au processus pathologique.

Des analyses génétiques à large échelle ont été réalisées ces dernières années pour tenter de mettre en évidence des facteurs de risques associés aux différentes tauopathies et participeront à mieux comprendre les formes familiales et sporadiques des tauopathies (Kouri *et al.* 2015).

Les principales caractéristiques des tauopathies les plus documentées, développées précédemment, sont récapitulées dans la **Table 3** ci-après. Bien que nous ne puissions pas être exhaustifs, d'autres tauopathies sont à référencer tels que le Syndrome de Down pour

lequel les formes de Tau 3R et 4R coexistent de façon égale ou encore la dystrophie myotonique, pathologie à prédominance 3R.

Pathologie	Prédominance Tau 3R					Tau 3R + 4R					Prédominance Tau 4R		
	PiD	AD	PART	PEP	FTDP-17	AGD	PSP	CBD					
<b>Symptômes</b>	Comportement	Mémoire	Cognition	Motricité	Motricité, Cognition		Syndrome de Richardson	Motricité					
<b>Structures touchées</b>	Hippocampe, cortex frontal et temporal	Cortex entorhinal, hippocampe, isocortex	Cortex entorhinal, hippocampe	Thalamus, ganglions de la base, tronc cérébral	Cerveau, cervelet, tronc cérébral	Hippocampe, cortex entorhinal, amygdale, hypothalamus	Ganglions de la base	Lobe pariétal, cortical					
<b>Type cellulaire</b>	Neurones Oligodendrocytes	Neurones	Neurones	Neurones + gile	Neurones +/- gile	Neurones	Neurones Astrocytes	Neurones					
<b>Histologie</b>	Corps de Pick	Dégénérescence neuro-fibrillaire, Plaques neuritiques	Dégénérescence neuro-fibrillaire	Dégénérescence neuro-fibrillaire	Dégénérescence neuro-fibrillaire	Grains fusiformes	Astrocytes hirsutes	Neurones gonflés					
<b>Morphologie agrégats</b>	PHF et SF	PHF et SF	PHF	PHF et SF	PHF	SF	PHF et SF	SF					
<b>Dépôts amyloïdes</b>	-	+	-	-	-	-	-	-					
<b>Ratio 4R:3R</b>	1:3	1:1	1:1	1:1	2:1, 1:1, 1:2	2:1	2/4:1	2:1					
<b>Déterminants génétiques</b>	Principalement sporadique	Principalement sporadique	Principalement sporadique	Sporadique	Transmission autosomique dominante	Principalement sporadique	Principalement sporadique	Principalement sporadique					
Haplotype H1/H2	-	-	-	-	-	-	H1	H1					

**Table 3. Caractéristiques de quelques tauopathies**

PiD (Maladie de Pick), AD (Maladie d'Alzheimer), PART (Primary age-associated tauopathy) PEP (Parkinsonisme post-encéphalique), FTDP-17 (Démence Fronto-Temporale et Parkinsonisme associée au chromosome 17), AGD (Démence à Grains Argyrophiles), PSP (Paralyse Supranucléaire Progressive), CBD (Dégénérescence Cortico-Basale). Table adaptée de Yamada *et al.* 2001, Williams 2006, Misra and Kalita. 2010, Jellinger. 2011, Forrest *et al.* 2019.

## II.5 Méthodes diagnostiques

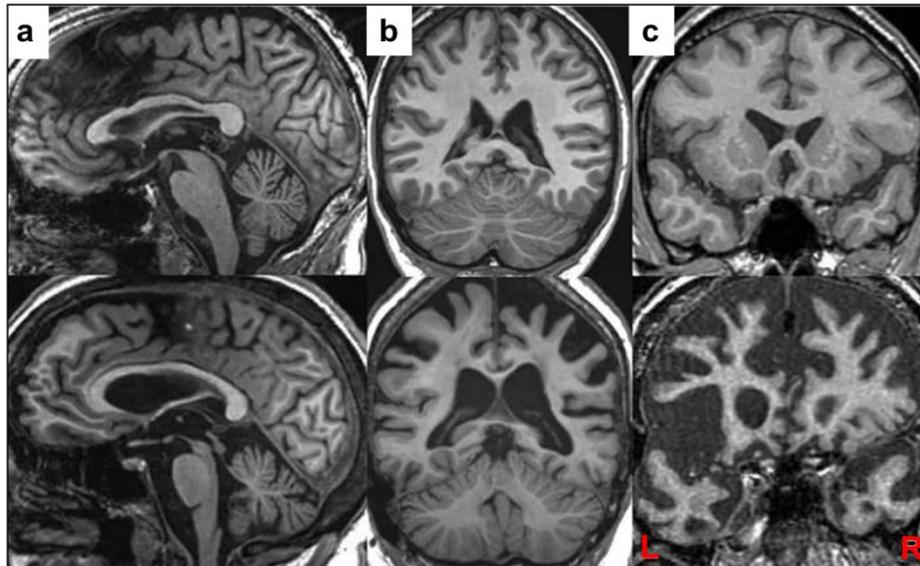
Afin de s'affranchir au mieux du chevauchement des spectres cliniques des tauopathies, le praticien dispose de plusieurs méthodes diagnostiques amenées à être associées.

### II.5.1 Examen clinique et neuropsychologique

Dans un premier temps, un examen approfondi des troubles cognitifs et moteurs est évidemment réalisé et susceptible d'être réévalué au cours des consultations ultérieures. Certaines revues destinées au corps médical mettent régulièrement à jour les critères cliniques à considérer dans le but de poser un diagnostic différentiel. (Thibodeau *et al.* 2009. ; EM Consulte). De plus, la Haute Autorité de Santé (HAS) met à disposition des recommandations de diagnostic telle que la NINCDS-ADRDA ((National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association) pour la maladie d'Alzheimer par exemple. Des tests neuropsychologiques ont été élaborés et sont largement utilisés par les praticiens ; on retiendra notamment le Mini-Mental State Examination (MMSE) établi par Folstein en 1975 ou encore le Test de 5 mots publié par Dubois en 2002. Les manifestations cliniques, décrites plus haut, permettent au neurologue, de poser un diagnostic possible de la pathologie dont la sensibilité et la spécificité restent à améliorer par d'autres examens comme l'imagerie.

### II.5.2 Imagerie

Les techniques d'imagerie sont très utilisées en clinique pour diagnostiquer les tauopathies. Un examen d'IRM (Imagerie par Résonance Magnétique) structurale d'un patient de la maladie d'Alzheimer permet par exemple de suivre l'évolution de l'atrophie cérébrale avec un agrandissement des sillons corticaux et des ventricules (Josephs. 2017). Comme évoqué plus haut, l'examen macroscopique post-mortem du cerveau peut différer en fonction de la tauopathie, et ce de façon plus évidente dans les stades avancés de la pathologie ; ceci résulte à des clichés distincts pour l'œil du praticien avisé (**Figure 8**). Cependant, le degré de robustesse et de spécificité du diagnostic par IRM est largement remis en question et servirait davantage à appuyer d'autres arguments cliniques (Whitwell *et al.* 2013).



**Figure 8.** Images par IRM du cerveau de patients atteints de tauopathies. IRM en pondération T1 suggérant une évolution d'un état normal (images du haut) vers (a) une PSP, (b) une CBD, (c) une maladie de Pick. Issu de Josephs. 2017.

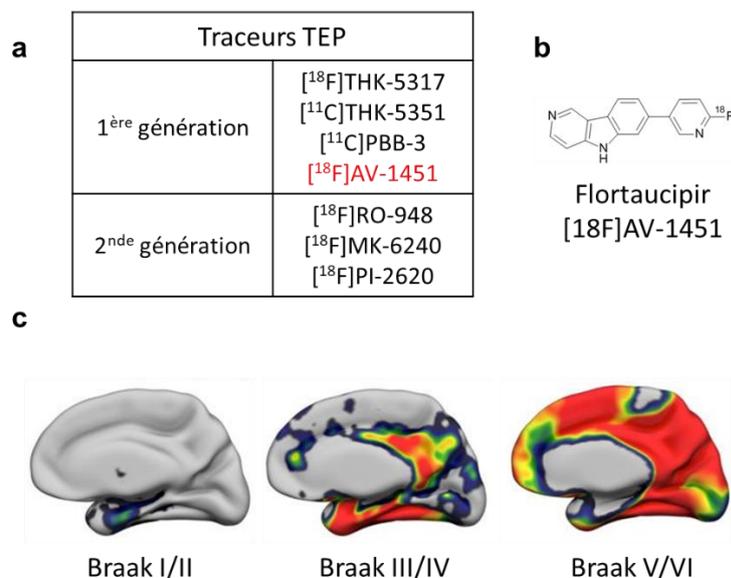
L'accent est aujourd'hui davantage porté sur l'utilisation d'un autre type d'imagerie : la Tomographie par Emission de Positons (TEP). Cette technique permet notamment de suivre les troubles métaboliques associés à la neurodégénérescence (hypométabolisme glucidique par exemple) ; on notera toutefois que l'examen SPECT de la perfusion cérébrale (ou TEMP en français), bien que moins sensible que la TEP, reste plus accessible en France à l'heure actuelle (Parthimos and Schulpis. 2019 ; Sarazin *et al.* 2012).

Ainsi, de nombreuses équipes travaillent au développement de traceurs TEP ciblant la pathologie A $\beta$  (i.e [18F]FDDNP, flobetapir<sup>6</sup>, flutemetamol<sup>7</sup>, florbetaben<sup>8</sup>) (Vlassenko *et al.* 2012), et Tau (1<sup>ère</sup> et 2<sup>nde</sup> générations) afin d'améliorer la spécificité et la sensibilité des marqueurs qui restent toutefois insuffisantes. Le traceur [18F]AV-1451, par exemple, a été extensivement étudié dans différentes tauopathies (Passamonti *et al.* 2017) et semble permettre de suivre l'évolution de la pathologie Tau au cours des différents stades de Braak chez les patients atteints de maladie d'Alzheimer (**Figure 9**) (Schöll *et al.* 2019).

<sup>6</sup> [18F]AV-45

<sup>7</sup> [18F]GE067

<sup>8</sup> [18F] BAY94-9172



**Figure 9.** Traceurs PET et suivi de la pathologie Alzheimer.

(a) Traceurs PET de la pathologie Tau (b) Structure chimique du traceur [<sup>18</sup>F]AV-1451 (c) Suivi des différents stades de Braak grâce à un examen 18F-flortaucipir (AV-1451)-TEP continu de patients atteints de la maladie d'Alzheimer. Issu de Schöll *et al.* 2019.

Cependant, [<sup>18</sup>F]AV-1451 semble présenter un défaut de spécificité et se lie à d'autres structures à feuillets  $\beta$  comme TDP-43 ou encore l' $\alpha$ -synucléine agrégées ou encore à la monoamine oxydase B (MAO-B) des ganglions de la base (Leuzy *et al.* 2019).

La structure quaternaire des structures amyloïdes et par conséquent l'accessibilité des sites de liaison du traceur influence la spécificité et la sensibilité des traceurs TEP. Ainsi, l'identification de déterminants moléculaires et conformationnels des inclusions, fonctions de la tauopathie, permettraient une amélioration du développement de ces molécules et de leur efficacité en clinique. Enfin, les installations techniques nécessaires à la synthèse et à l'utilisation de la TEP restent aujourd'hui difficiles et onéreuses à mettre en place.

### II.5.3 Analyse du Liquide Céphalo-Rachidien (LCR) – Biomarqueurs biochimiques

L'analyse du liquide céphalo-rachidien (LCR) nécessite que le patient soit soumis à une ponction lombaire, acte médical relativement invasif aujourd'hui couramment pratiqué. La pertinence du diagnostic par étude du LCR repose sur un protocole robuste de prélèvement, du stockage et de l'étude des échantillons par ELISA<sup>9</sup> ainsi que sur l'existence de biomarqueurs fiables, pathologie-spécifiques et contenus en quantité suffisante pour être

<sup>9</sup> INNOTEST ELISA kit (Fujirebio Diagnostics, Ghent, Belgium)

détectés. Les trois marqueurs les plus communément étudiés aujourd'hui sont T-tau (anticorps dirigés contre les six isoformes de Tau), p-tau (anticorps dirigé contre la T181 phosphorylée) et  $A\beta_{1-42}$  (anticorps dirigé contre  $A\beta_{1-42}$ ) (Coughlin and Irwin. 2017). Etonnement, bien que le peptide  $A\beta_{1-42}$  soit retrouvé agrégé en grande quantité au sein des plaques extracellulaires chez les patients atteints de la maladie d'Alzheimer, la quantité d' $A\beta_{1-42}$  (et le ratio  $A\beta_{1-42}/A\beta_{1-40}$ ) au sein du LCR est retrouvée diminuée par rapport aux individus contrôles ; cependant cette observation est aussi valable pour les patients atteints de démence à corps de Lewy. Les biomarqueurs t-tau et le p-tau semblent être des bons marqueurs de la maladie d'Alzheimer par rapport aux patients contrôles. En effet, il a été observé une augmentation de la concentration de la protéine Tau dans le LCR de patients atteints d'AD, dépendante du degré de sévérité de la neurodégénérescence (Lee *et al.* 2019). Pour p-T181<sup>10</sup>, bien qu'il soit observé une sensibilité très variable en fonction des études, la spécificité avoisine 80% (Vanderstichele *et al.* 2006). Une combinaison des trois marqueurs permet de gagner en spécificité mais ne semble pas palier aux problèmes associés à la sensibilité de l'analyse du LCR (Formichi *et al.* 2006). Ces dernières années, de nombreux groupes ont développé de nouveaux biomarqueurs dirigés contre d'autres résidus phosphorylés de Tau, ou encore contre des formes tronquées de Tau (Borroni *et al.* 2008 ; Molinuevo *et al.* 2018). D'autres études visaient la détection directe de peptides de Tau dans le LCR par spectrométrie de masse et leur quantification (Barthélemy *et al.* 2016). Une solution au problème de sensibilité évoqué ci-dessus pourrait être l'amplification des amorces de Tau présentes dans le LCR ; une preuve de faisabilité a d'ailleurs été réalisée pour la maladie de Pick (Saijo *et al.* 2017).

#### *II.5.4 Difficultés associées au diagnostic différentiel des tauopathies*

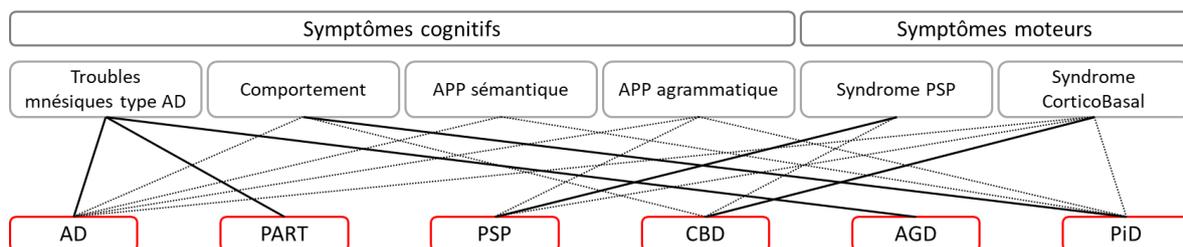
Malgré les efforts conséquents alloués au développement de méthodes diagnostiques, le diagnostic différentiel des tauopathies reste compliqué. En effet, aujourd'hui, seul l'examen anatomopathologique post-mortem permet de poser un diagnostic « certain » d'une tauopathie spécifique ; le diagnostic réalisé ante-mortem étant « possible » ou « probable ». Ceci s'explique principalement par un recouvrement des tableaux cliniques entre tauopathies et d'une hétérogénéité du spectre clinique au sein d'une même entité pathologique (**Figure 10**). Pour illustrer les difficultés rencontrées par le praticien, prenons un premier exemple, la paralysie supranucléaire progressive. En effet, elle est souvent confondue en clinique avec d'autres pathologies neurodégénératives associées à des troubles moteurs comme l'atrophie

---

<sup>10</sup> Bien qu'utilisée dans le diagnostic d'AD, la phosphorylation de la Thréonine 181 est retrouvée chez les patients atteints de la maladie d'Alzheimer, mais aussi chez les individus sains (Šimić *et al.* 2016).

multisystématisée, la démence à corps de Lewy (toutes deux des synucléinopathies) ou encore la dégénérescence corticobasale (Dickson *et al.* 2007). De façon surprenante, Gearing et collaborateurs ont publié en 1994 une étude de 13 cas diagnostiqués PSP en ante-mortem dont plus de la moitié se sont révélés être, soit atteints de la maladie d'Alzheimer, soit de tauopathies mixtes (PSP/AD). Pour tenter de contrer les erreurs de diagnostics confondant cette fois-ci la maladie de Parkinson et la PSP, pathologies toutes deux associées à des troubles moteurs, les médecins peuvent évaluer la réponse du patient au traitement à la L-Dopa, ce dernier s'avérant le plus souvent négatif chez les patients PSP contrairement aux sujets atteints de la maladie de Parkinson (Scaravilli *et al.* 2005). Il n'en reste pas moins que le diagnostic de la PSP reste pour de nombreux cas difficile à réaliser.

Le cas de la pathologie PART est quant à lui très discuté dans la communauté scientifique et illustre l'absence d'accord sur l'existence d'entités bien distinctes entre les tauopathies. En effet, cette maladie avait initialement été décrite comme une pathologie distincte de celle d'Alzheimer en raison de critères histologiques lésionnels (agrégation neurofibrillaire de Tau en absence de plaques A $\beta$ ) et de critères cliniques mal définis (démence modérée, impact moins important sur le quotidien) (Crary *et al.* 2014). Cependant, un autre courant s'accorde à dire que les apparentes différences observées entre AD et PART sont dues à un décalage temporel entre l'apparition de la pathologie A $\beta$  et celle de Tau. Un autre argument avancé est que les deux pathologies partagent une caractéristique fondamentale d'AD, la présence d'enchevêtrements neurofibrillaires dans les structures entorhinale et hippocampale (Duyckaerts *et al.* 2015). Ils concluent ainsi que PART serait un cas particulier d'AD. Ce débat renforce l'idée que les critères cliniques et histologiques sur lesquels reposent les actuels diagnostics ne suffisent pas à discriminer les tauopathies et qu'il existe un réel besoin de marqueurs robustes tauopathie-spécifiques (Coughlin and Irwin. 2017; Orr *et al.* 2017).



**Figure 10.** Illustration du recouvrement des spectres cliniques présentés par différentes tauopathies. Les caractéristiques cliniques prédominantes de chaque tauopathie sont représentées par un trait plein épais tandis que les signes cliniques secondaires retrouvés de façon non homogène en fonction des patients sont illustrés par des traits fins en pointillés. L'Aphasie Primaire Progressive (APP) est définie par l'apparition et l'aggravation de troubles du langage de nature aphasique. Cette dernière peut se présenter sous la forme sémantique (manque de mot, troubles de compréhension du mot isolé) ou encore sous la forme agrammatique/fluente (trouble des connaissances sémantiques perturbant la reconnaissance des objets et des personnes) (Moreaud. 2013). Issu de Irwin. 2016.

Cette difficulté de diagnostic robuste des différentes tauopathies participe certainement à l'échec du développement de thérapeutiques efficaces rencontré par l'industrie pharmaceutique ces dernières années. En effet, estimée à plus de 20%, l'inclusion de patients atteints de tauopathies distinctes mais aux manifestations cliniques ressemblantes à la pathologie ciblée, pourrait expliquer l'absence de réponse de certains patients et le conséquent arrêt des essais cliniques (Cummings *et al.* 2018). De plus, la communauté s'accorde à penser qu'une explication majeure à ces déconvenues serait l'inclusion des patients à des stades trop avancés de la maladie. Ainsi, il apparaît crucial d'identifier des biomarqueurs pertinents et tauopathie-spécifiques ainsi que d'approfondir les connaissances cliniques et les caractéristiques biochimiques mesurables, notamment à des stades précoces de la pathologie.

## II.6 Essais cliniques et prise en charge thérapeutique des patients

Le traitement des maladies neurodégénératives est aujourd'hui un enjeu de santé publique majeur engendrant un investissement considérable dans le développement d'agents thérapeutiques. Jusqu'à présent, les stratégies élaborées pour les tauopathies se sont essentiellement focalisées sur la maladie d'Alzheimer et la paralysie supranucléaire progressive (Medina. 2018).

### *Maladie d'Alzheimer*

La maladie d'Alzheimer étant la plus commune des maladies neurodégénératives, elle a fait l'objet de pas moins de 1778 essais cliniques de par le monde jusqu'à ce jour (clinicaltrials.gov). Le taux d'échec associé est particulièrement alarmant avec 72% d'arrêt en Phase 1, 92% en Phase 2 et 98% en Phase 3. En 2018, Cummings et collaborateurs ont dressé un bilan mis à jour des mécanismes d'action des thérapeutiques retrouvés en phase II d'essai clinique. Ils constatent que près de 40% des molécules ciblent les effets secondaires de la pathologie (cognition, symptômes psychiatriques). 57% des agents sont destinés à modifier la maladie en agissant, sur la pathologie amyloïde pour 53% d'entre eux, alors que seulement 25% de ces molécules ciblent la pathologie associée à Tau (Cummings *et al.* 2018). En 2014, le même groupe évalue que 35% de ces thérapeutiques sont des petites molécules tandis que 18% sont des immunothérapies ; cette dernière catégorie étant sur le devant de la scène pharmaceutique ces dernières années (Cummings *et al.* 2014). Les stratégies mises en place pour la pathologie amyloïde  $\beta$  sont essentiellement (1) l'inhibition de la production du peptide A $\beta$  par des inhibiteurs de la  $\beta$ -secrétase (BACE) et (2) sa dégradation par

immunothérapie active et passive. Les mécanismes ciblés pour la pathologie Tau sont principalement (1) la prévention de l'agrégation de Tau par des petites molécules inhibitrices et (2) la clairance de Tau phosphorylée par immunothérapie. Enfin, une autre cible est prévalente dans les essais cliniques : les inhibiteurs de l'activation de la microglie, modulateurs de la neuroinflammation (Graham *et al.* 2017 ; Hung and Fu. 2017).

Ainsi, seules quatre molécules sont aujourd'hui mises sur le marché ; trois ciblent la voie cholinergique et une la voie glutamatergique. En effet, au niveau cérébral chez les patients atteints de la maladie d'Alzheimer, il est observé une diminution significative de l'expression de la choline acétyltransférase corrélée avec un déclin cognitif (Bowen *et al.* 1976). Il a été également décrit une diminution des neurones cholinergiques du noyau basal de Meynert et des projections vers le cortex et l'hippocampe (Whitehouse *et al.* 1982). Compte tenu de ces deux observations, les traitements aujourd'hui disponibles pour les patients atteints d'AD sont donc de deux types (Delrieu. 2017) : des anticholinestérasiques pour les stades légers à modérés et un antagoniste des récepteurs NMDA (N-méthyl-D-aspartate) pour les stades modérés à sévères. Toutefois, en France, en raison du service médical rendu insuffisant et d'effets secondaires importants (Prescrire 2015), la Haute Autorité de Santé a publié en mai 2018 une liste actualisée de recommandations appuyant un déremboursement de ces molécules.

Plusieurs raisons peuvent expliquer cet échec du développement médicamenteux. Initialement, le principal problème résidait dans l'absence de marqueurs pharmacodynamiques de l'engagement à la cible. Cependant, quelques biomarqueurs, potentiellement perfectibles, sont aujourd'hui disponibles (LCR, TEP, IRM), notamment pour la maladie d'Alzheimer, et ont changé le cours des essais cliniques. L'utilisation de thérapies combinées adaptées à la complexité intrinsèque aux tauopathies, ciblant A $\beta$  et Tau par exemple pour la maladie d'Alzheimer, ou encore les aspects de la pathologie associant perte de fonction normale et gain de fonction toxique, sont aujourd'hui très appuyés (Khanna *et al.* 2016). Enfin, dans les phases pré-cliniques, les modèles animaux utilisés pour mimer la pathologie sont sujets à discussion.

### *Autres tauopathies*

Actuellement, d'après le VIDAL, pour la Paralyse Supranucléaire Progressive, la Dégénérescence Cortico-Basale, la maladie à Grains Argyrophiles et la maladie de Pick, il n'existe aucun médicament commercialisé ayant prouvé une efficacité.

Pour conclure, il est nécessaire de mettre au point un diagnostic différentiel robuste et précoce des différentes tauopathies, appuyé par des critères tant cliniques que biochimiques. Des biomarqueurs doivent ainsi être identifiés pour à la fois inclure des patients dont le diagnostic est certain mais également estimer l'efficacité d'un traitement en cours de développement dans le cadre des essais cliniques. Le stade d'avancement de la pathologie devrait également être examiné avec précaution ; l'efficacité d'une molécule à un stade trop avancé de la pathologie pouvant être très diminuée voire nulle. La communauté médicale appuie l'intérêt d'une prise en charge à la fois médicamenteuse et non médicamenteuse stabilisant ou diminuant le déclin cognitif et améliorant le bien-être du patient et de son entourage. Une recrudescence des travaux de recherche fondamentale sur les tauopathies et une connaissance toujours plus approfondie des mécanismes sous-jacents à la pathologie sont bien évidemment nécessaires au développement de thérapies efficaces.

### III. Tau

#### III.1 Le gène MAPT, épissage alternatif et expression des isoformes

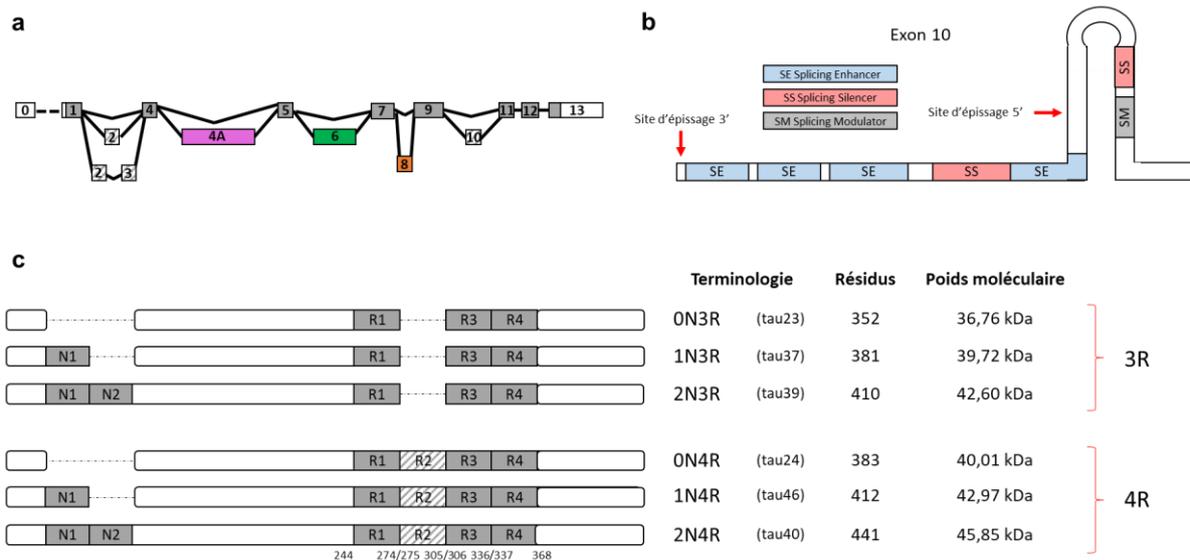
Découverte en 1975 (Weingarten *et al.* 1975), la protéine Tau appartient à la famille des MAPs (Microtubule-Associated Proteins) et est codée par un gène unique *MAPT* (134 kb) situé sur le bras long du chromosome 17 (Chr17q21). La séquence de *MAPT* est hautement conservée au sein des primates avec une homologie comprise entre 97 et 100%.

Le transcrit primaire de Tau (**Figure 11.a**) est constitué de 16 exons dans le cerveau humain dont les exons 1,4,5,7,9,11,12 et 13 sont constitutifs. L'ARNm de Tau est retrouvé à la fois dans les cellules neuronales et non neuronales du cerveau. Dans les neurones, l'unique promoteur du gène MAPT présente trois sites de liaison à des facteurs de transcription et son activité augmente avec la croissance neuronale.

Les exons 2, 3 (la présence de l'exon 3 dépendant de celle de l'exon 2) et 10 sont sujets à un épissage alternatif et sont spécifiques au cerveau humain adulte. De cet épissage alternatif, il résulte donc 6 isoformes de Tau constituées de 352 à 441 acides aminés et dont le poids moléculaire apparent sur gel varie de 45 à 65 kDa. Comme illustré en **Figure 11.c**, ces isoformes se distinguent par :

- l'absence (0N) d'un (1N) ou deux (2N) inserts (29 ou 58 aa) en N-terminal de la protéine,
- la présence de trois (3R) ou quatre (4R) séquences répétées codant pour des domaines de liaison aux microtubules en C-terminale de la protéine.

Ces six isoformes peuvent être regroupées sous deux grandes catégories : les isoformes 3R (exclusion de l'exon 10) et les isoformes 4R (inclusion de l'exon 10).



**Figure 11.** Le gène MAPT, régulation de l'épissage alternatif de l'exon 10 et isoformes de Tau. **(a)** Structure du gène MAPT : en blanc les exons non codants ; en rose, vert et orange, les exons non transcrits dans le système nerveux central humain ; en gris plein, les exons constitutifs du cerveau adulte humain ; en gris hachuré, les exons soumis à épissage alternatif. Adapté de Caillet-Boudin et al. 2015. **(b)** Eléments régulateurs de l'épissage de l'exon 10 : en bleu régulation positive (enhancers), en rose régulation négative (silencers), en gris régulation variable (modulators). Adapté de Liu et al. 2008. **(c)** Les six isoformes de Tau du cerveau humain peuvent être classées dans deux catégories : 3R (3 séquences répétées en C-ter) et 4R (4 séquences répétées en C-ter). L'ancienne terminologie est indiquée entre parenthèses à titre d'information.

Ainsi, parmi les 16 exons, les exons 4A et 8 ne sont pas transcrits dans le système nerveux central, et sont respectivement spécifiques du système périphérique d'une part, et du système nerveux bovin d'autre part. L'exon 6 est retrouvé dans la moelle épinière et le muscle squelettique. L'exon 0, quant à lui, n'est pas codant mais la taille de l'exon 1 dépend de sa présence ; cela n'affecte donc pas la séquence traduite (le 5' n'étant pas codant). La taille des exons 4A et 6 est modulée en fonction des sites d'épissage 3' (2 et 3 respectivement). Par un changement du cadre de lecture et ainsi l'introduction d'un codon stop, les versions p et d de l'exon 6 conduisent à la production d'une protéine tronquée, sans domaines de liaison aux microtubules. De plus, une comparaison des transcrits des haplotypes H1 et H2 a été réalisée (Ensembl) et a montré que l'haplotype influence l'épissage ; l'inclusion de E0 est notamment moins fréquente pour H2 (Caillet-Boudin *et al.* 2015).

L'épissage alternatif de l'exon 10, menant aux formes 3R (exclusion) et 4R (inclusion), est notamment modulé par les exons flanquants ; en effet, en amont, l'exon 9 semble promouvoir son inclusion tandis qu'en aval l'exon 11 semble jouer le rôle inverse (Andreadis. 2005). D'autres mécanismes sont impliqués dans la régulation de l'épissage de l'exon 10 (**Figure 11.b**) : des protéines riches en sérines et arginines (SR), des kinases ou encore des *cis* éléments. Parmi les *cis* éléments, on note la présence d'une structure en tige-boucle à

l'interface entre l'exon 10 et l'intron 10 contenant 4 éléments régulateurs de l'épissage (2 silenciers, 1 enhancer et 1 modulateur). Plus en amont, dans la première moitié de l'exon 10, on retrouve 3 éléments enhancers. Enfin, il a été montré que des mutations dans cette région peuvent avoir pour effet une augmentation, une diminution ou un maintien du ratio 4R:3R (Qian and Liu. 2014 ; Liu and Gong. 2008). Parmi les *trans* facteurs, on note le rôle des hnRNP (Heterogeneous Nuclear RiboNucleoProteins), complexes ARN-protéines, et des SR-Protéines, composants du spliceosome. Les SR-protéines ont la capacité de lier la séquence d'ARN par leur partie N-terminale et les autres constituants du spliceosome par leur C-terminal. Il existe plusieurs SR-protéines agissant sur l'inclusion ou l'exclusion de l'exon 10. Des mutations dans les éléments *cis*, notamment dans le cas de la FTDP-17, conduisent à une affinité de liaison réduite de ces SR-protéines et ainsi à une modulation de l'épissage de l'exon 10. Enfin, les sérines, dont sont enrichies les SR-protéines, sont sujettes à phosphorylation ce qui va, là aussi, moduler leur liaison aux éléments *cis* exoniques.

Comme nous venons de l'évoquer, on observe des mutations affectant l'épissage alternatif de l'ARNm MAPT. Une seconde catégorie de mutations (**Table 4**), faux-sens, est impliquée dans une modulation de la liaison aux microtubules (partie C-terminale de la protéine Tau) et affecte d'ailleurs aussi parfois l'épissage alternatif de l'exon 10. Ceci est en accord avec le fait que les isoformes de la protéine 4R (inclusion de l'exon 10) ont une meilleure constante d'affinité de liaison aux microtubules que les isoformes 3R (exclusion de l'exon 10) (Goode *et al.* 2000). Ces mutations peuvent se situer sur l'exon 10 (P301L, P301S), sont généralement associées à une inclusion de l'exon 10 (forme 4R prévalente) et peuvent conduire à une diminution de la polymérisation des microtubules. De plus, une agrégation de Tau concomitante est parfois observée. D'autres mutations, nombreuses, ont été décrites dans des régions communes à toutes les isoformes, affectent également la polymérisation des microtubules et sont corrélées avec une agrégation de Tau. Bien que l'information de liaison ne soit pas toujours disponible, il est postulé qu'une diminution de l'interaction aux microtubules aurait pour conséquence une augmentation des formes libres de Tau cytoplasmique favorisant ainsi leur agrégation. Enfin, une mutation particulière,  $\Delta$ K280, a été très largement décrite dans la littérature car elle a pour double effet de modifier le ratio 3R/4R (incrément des formes 3R) et de perturber la liaison aux microtubules. (Sergeant *et al.* 2005 ; Van Swieten *et al.* 2007). De façon intéressante, il a également été montré qu'une acétylation *de novo* de la lysine 280 conduisait à une augmentation de la quantité et de la neurotoxicité de Tau humaine chez la drosophile. Ceci semble montrer que la lysine 280 pourrait être impliquée dans le processus d'agrégation et la neuropathogénicité de Tau (Gorsky *et al.* 2016).



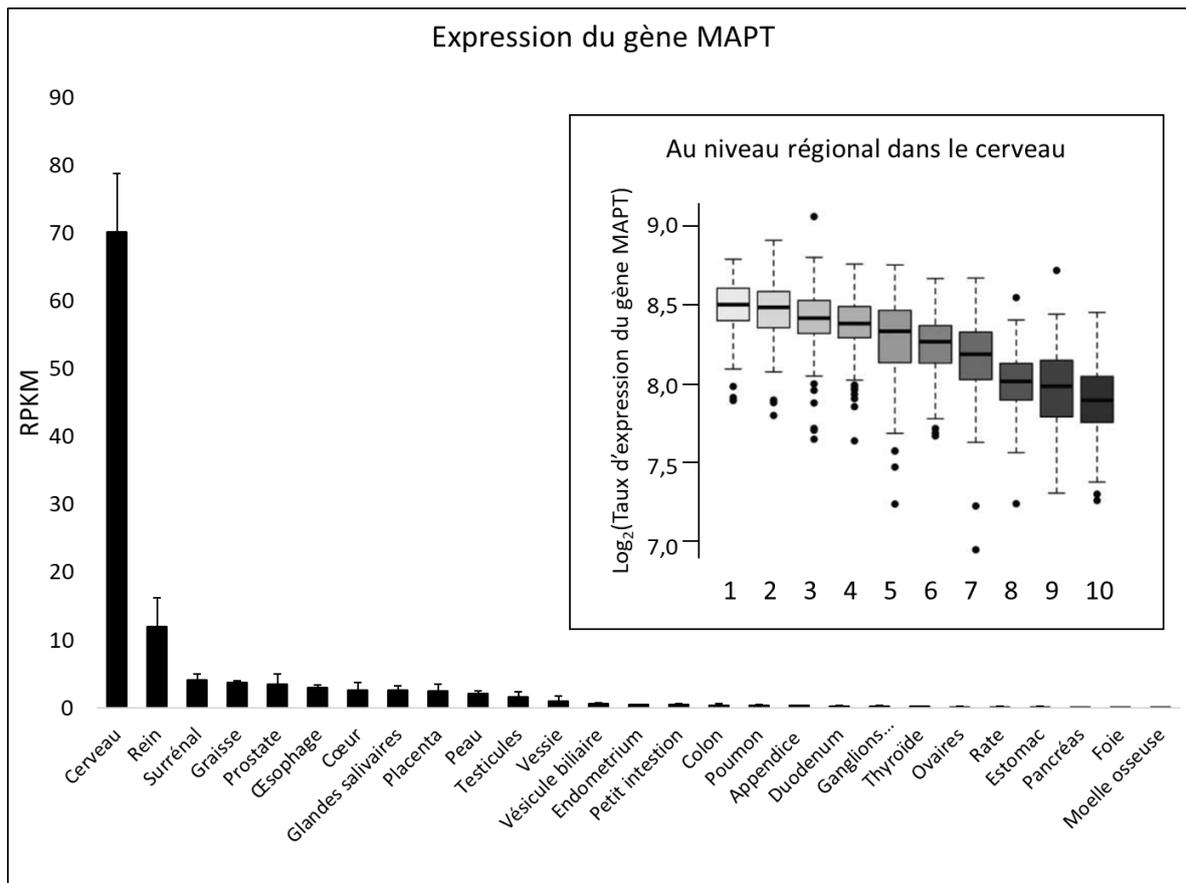
Exon	Mutations	Agrégation de Tau	Polymérisation des microtubules	Liaison aux microtubules	Inclusion de l'exon 10
1	R5H, R5L	↑ (1N/3R et 4R)	↓	ND	↔, ND
2	G55R	ND	↑	ND	ND
7	A152T	↔	↓	↓	↔
9	K257T, G272V I260V, L266V	↑ (3R) ↑ (4R)	↓	ND ou ↔	ND ou ↔
10	Δ280K	↑ (3R)	Essentiellement ↓	Essentiellement ND ou ↔ (sauf Δ280K et P301L ↓)	↓
	N279K, P301L, P301S, G303V S285R, Δ296N, N296H, K298E, P301T, S305I, S305N	↑ (4R) ND ou ↔			Essentiellement ↑ Essentiellement ↑
11	K317N	↓(3R) ↑(4R)	↓	ND	ND
	Q336H, Q336R	↑ (3R)	↑	ND	
	S320F, G335V	↑	↓	ND	
	L315R, K317M, P332S, G335S	ND ou ↔	↓ (sauf K317M et P332S ND)	ND ou ↔	
12	V337M, S352L, P364S	↑	↓ (sauf S356T ND)	ND ou ↔ (sauf V337M ↓)	ND
	K369I	↓	↓	ND	ND
	E342V, S356T, G366R,	ND ou ↔	Variable	ND	ND (sauf E342V ↑)
13	E372G, G389R (G->A), G389R (G->C), N410H	↑ (ON et 1N/3R et 4R)	↓	ND	Essentiellement ND (sauf N410H ↑)
	R406W	↔		↓	ND

**Table 4.** Liste des mutations du gène MAPT identifiées dans le cerveau humain de patients atteints de différentes tauopathies

Les flèches « ↑ » et « ↓ » indiquent une augmentation ou une diminution par rapport à Tau WT, tandis que « ↔ ou ND » signifie qu'il n'est pas observé de différence ou que l'information n'a pas été déterminée. Adaptée de Strang *et al.* 2019.

Dans un contexte normal, le profil d'expression de Tau diffère en fonction du stade développemental, du tissu considéré, du type cellulaire.

Ainsi, au stade fœtal, seul Tau 0N3R est détecté, l'expression des cinq autres isoformes est donc périnatale (Goedert and Jakes 1990 ; Hefti *et al.* 2018). Les études transcriptomiques haut-débit permettent aujourd'hui de disposer de données quantitatives de l'expression de MAPT dans l'ensemble des tissus humains où on identifie préférentiellement une expression cérébrale et de façon plus secondaire rénale. Au sein du cerveau, on observe des différences en fonction de la région considérée. En effet, les cortex frontal et temporal sont les zones cérébrales où l'expression de Tau est la plus accrue, viennent ensuite le cortex occipital, l'hippocampe et le thalamus (**Figure 12**).



**Figure 12.** Expression du gène MAPT.

Histogramme d'expression du gène MAPT en Reads Per Kilobase and per Million (RPKM). Les données présentées sont issues de l'étude « HPA RNA-seq normal tissues ». Un séquençage des ARN a été réalisé sur des tissus provenant de 95 individus humains (Illumina HiSeq2000, date de l'analyse Wed Apr 4 07:08:55 2018) (Fagerberg *et al.* 2014). En encart, expression différentielle du gène MAPT en fonction de la région cérébrale considérée : 1 Cortex frontal, 2 Cortex temporal 3 Cortex occipital 4 Hippocampe 5 Thalamus 6 Cervelet 7 Substantia Negra 8 Putamen 9 Medulla 10 Matière blanche intralobulaire (Trabzuni *et al.* 2012).

Les études transcriptomiques en cellule unique ont également permis de mettre en évidence une expression différentielle du gène *mapt* en fonction du type cellulaire considéré pour une même région cérébrale. Ainsi, dans le cortex murin, il a été montré que MAPT était deux fois plus transcrit dans les neurones que dans les oligodendrocytes. Dans les astrocytes, l'expression du gène MAPT compte pour seulement 15% des transcrits neuronaux (Zhang *et al.* 2014). Ainsi, l'expression du gène MAPT, prédominante au niveau cérébral, varie en fonction de la région et du type cellulaire considéré. Au sein du neurone, on observe une distribution asymétrique de Tau, en gradient, avec un enrichissement au niveau de la synapse au détriment du corps cellulaire (Dixit *et al.* 2008).

Chez l'Homme adulte, toutes régions cérébrales confondues, on note une prédominance d'expression des isoformes 1N (50% de Tau total), suivies des isoformes 0N (40%) ; les isoformes 2N étant de loin beaucoup moins abondantes (10%). Enfin, il n'est pas observé de

différences marquées de niveau d'expression entre les isoformes 3R et 4R (Trabzuni *et al.* 2012).

Considérant que les isoformes agrégées varient en fonction de la tauopathie étudiée, la communauté scientifique s'est questionné sur l'éventuelle expression différentielle des isoformes entre individu normal et individu atteint d'une tauopathie. En 1989, par RT-PCR, sans fractionnement tissulaire préalable, il a été montré que dans le cortex frontal, le cortex cérébelleux et le tronc cérébral, chez l'individu sain et le patient atteint de la maladie d'Alzheimer, les transcrits des deux classes d'isoformes, 3R et 4R, sont retrouvés de façon similaire. Dans le cas de la PSP, les isoformes 4R sont davantage exprimées dans le tronc cérébral par rapport aux patients contrôle et les transcrits des isoformes 3R sont appauvris dans le cortex cérébelleux, à l'origine d'un ratio de transcription 4R/3R significativement plus important (Goedert *et al.* 1989 ; Chambers *et al.* 1999). Le ratio d'ARNm 4R/3R est également retrouvé augmenté de façon significative dans le cas de la FTDP-17 alors qu'il n'est pas différent des patients contrôle pour les personnes atteintes de CBD ou de la maladie de Pick au niveau du cortex frontal (Ingelsson *et al.* 2007 ; Connell *et al.* 2005). Enfin, à l'image des deux dernières tauopathies mentionnées, il semblerait que le ratio de transcription 4R/3R soit inchangé dans le cas de la maladie aux grains argyrophiles (Unpublished Data, Cortés *et al.* 2012).

Cette différence d'expression observée dans le cas de certaines tauopathies pourrait s'expliquer par (1) des facteurs épigénétiques ou (2) l'influence des haplotypes H1/H2 et leurs sous-classes. Ainsi, des études épigénétiques du gène *MAPT* ont été réalisées et certains travaux semblent montrer une hypométhylation du gène dans le cas de certaines neuropathies telles que la PSP ou AD. Or l'hyperméthylation du gène *MAPT* conduit à une diminution de son expression et aurait donc un effet neuroprotecteur. Ainsi, le phénomène d'hypométhylation précédemment décrit participerait au processus pathologique (Caillet-Boudin *et al.* 2015). Cependant, une étude parue en 2009 et réalisée sur de nombreux patients atteints de diverses tauopathies n'a montré aucune différence significative de méthylation (Barrachina and Ferrer. 2009). Enfin, bien qu'il soit admis que l'haplotype H1c constitue un facteur de risque pour la PSP, aucune corrélation significative n'a pu être établie entre les différents haplotypes et le taux d'expression du gène *MAPT*.

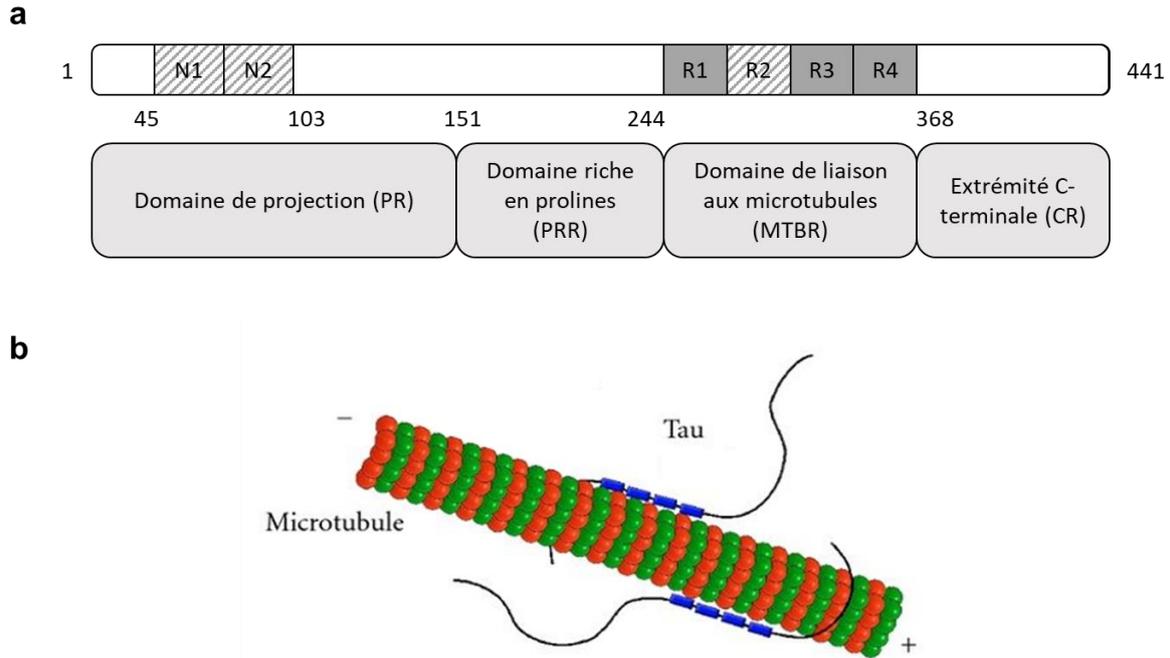
Pour conclure, grâce aux méthodes de séquençage haut-débit aujourd'hui disponibles, il a été possible d'étudier l'expression du gène *MAPT* et plus particulièrement de ses isoformes 3R et 4R. Bien que les assemblages retrouvés dans les différentes tauopathies soient isoforme-spécifiques, il n'a pas pu être montré d'implication directe de l'expression différentielle des versions de *MAPT* pour chacune des tauopathies. En effet, le ratio d'ARN messagers 4R/3R

est inchangé dans le cortex frontal par rapport aux individus contrôles pour la maladie de Pick dans laquelle les isoformes 3R sont retrouvées agrégées, ou encore les pathologies 4R dominantes telles que la dégénérescence cortico-basale et la maladie aux grains argyrophiles. Toutefois, certaines sous-populations cellulaires, où sont retrouvés de façon substantielle les agrégats, pourraient exprimer différenciellement les deux types d'isoformes. Les régions touchées, notamment au premiers stades, différant entre les tauopathies, ceci pourrait expliquer les différences de composition en isoformes des agrégats. Par exemple, les cellules granulaires du gyrus dentelé ou les cellules pyramidales de la corne d'Ammon et du subiculum (toutes situées dans l'hippocampe), où sont principalement retrouvées les corps de Pick, n'exprimeraient que les isoformes 3R. Ceci expliquerait pourquoi les patients atteints de la maladie de Pick ne présenteraient que des agrégats composés d'isoformes 3R (Goedert et al. 1989 ; Buée and Delacourte. 2006). Il serait donc intéressant d'étudier les sous-populations cellulaires des régions affectées par la pathologie Tau dans les différentes tauopathies pour déterminer si la composition différentielle 3R et/ou 4R des agrégats reflète bien une expression en isoformes de Tau spécifique au type cellulaire.

### III.2 Domaines et fonctions

Comme illustré dans la **Figure 13.a**, la protéine Tau peut être scindée en quatre domaines principaux (Heinisch and Brandt. 2016) :

- Le domaine de projection (PR)
- La région riche en prolines (PRR)
- Le domaine de liaison aux microtubules (MTBR)
- La région C-terminale (CR)



**Figure 13.** Domaines de Tau et son interaction avec les microtubules.

(a) Distribution des domaines le long de la séquence de Tau (isoforme 2N4R). (b) Représentation schématique de Tau interagissant avec un microtubule par son domaine de liaison (ici 4 séquences répétées). Issu de Kolarova *et al.* 2012.

Caractérisée par les nombreuses interactions qu'elle peut établir grâce à son domaine de projection, Tau est connue pour être associée aux microtubules par son domaine de liaison aux microtubules (**Figure 13.b**). Plus généralement, Tau (et son homologue MAP1B) joue un rôle dans la croissance axonale (Takei *et al.* 2000). Elle est impliquée dans l'établissement et le maintien de la polarité neuronale. Ainsi, nous détaillerons ci-dessous les différents domaines de la protéine et leurs fonctions.

Le domaine N-terminal de la protéine Tau est communément appelé domaine de projection. Les deux inserts, N1 et N2 (acides aminés 45-103), sont constitués de nombreux résidus acides et sont suivis d'une région basique riche en prolines (acides aminés 151-244). Ce domaine détermine l'espace entre les microtubules au sein de l'axone et pourrait ainsi contribuer à l'augmentation du diamètre axonal (Chen *et al.* 1992). De plus, c'est par ce domaine riche en prolines que la protéine Tau se lie aux filaments d'actine et de spectrine (He *et al.* 2009; Carlier *et al.* 1984), établissant ainsi une connexion entre les microtubules et les autres éléments du cytosquelette. Ce phénomène serait impliqué dans la régulation de la flexibilité microtubulaire. De plus, le domaine de projection joue un rôle dans l'interaction des microtubules avec la membrane plasmique (Brandt *et al.* 1995). Enfin, la région N-terminale serait également partie prenante dans l'activation de la voie de signalisation impliquant la phospholipase C gamma (PLC $\gamma$ ) (Hwang *et al.* 1996). En effet, le domaine riche en prolines

comprend sept motifs PXXP par lesquels Tau se lierait au domaine SH3 de PLC $\gamma$ -1, et plus secondairement de PLC $\gamma$ -2. D'autres acteurs de voies de transduction de signal à domaines SH3 tels que p85, sous-unité régulatrice de la Phosphatidylinositol (PI) 3-kinase, ou encore fyn, kinase de la voie SRC, peuvent interagir avec la région N-terminale de Tau (Reynolds *et al.* 2008). Enfin, c'est aussi par son domaine N-terminal que Tau interagit avec les moteurs moléculaires que sont la kinésine et la dinéine et en module l'activité. Ces moteurs moléculaires permettent le transport de molécules le long des microtubules et assurent donc des fonctions nécessaires à la cellule. Dans le cas de la maladie d'Alzheimer, le gradient proximo-distal de Tau au sein du neurone est inversé. Il a été proposé que cette redistribution de Tau engendrerait une inhibition du transport antérograde médié par la kinésine (Dixit *et al.* 2008).

Le domaine C-terminal, domaine de liaison aux microtubules (MTBR), comprend 3 (isoformes 3R) ou 4 (isoformes 4R) séquences répétées hautement conservées codées par les exons 9 à 12. Dès 1975, le rôle de la protéine Tau, de la famille des MAP, dans la promotion de l'assemblage des microtubules et leur stabilité a été démontré (Weingarten *et al.* 1975). C'est par la région C-terminale que Tau promeut la polymérisation et la stabilisation des microtubules et est impliquée dans le transport axonal. De plus, il a été démontré que toutes les isoformes de Tau n'ont pas la même affinité de liaison aux microtubules. En effet, le  $K_D$  des isoformes 4R est 3 fois inférieur à celui des isoformes 3R (Goode *et al.* 2000). Ainsi, les isoformes 4R sont plus efficaces que les isoformes 3R dans la promotion de la polymérisation des microtubules (Trinczek *et al.* 1995). De plus, les éventuelles modifications post-traductionnelles, notamment des phosphorylations, au sein de ce domaine rentrent en compétition avec la liaison de Tau aux microtubules et constitueraient donc un mode de régulation de cette interaction (Šimić *et al.* 2016). Ainsi, la phosphorylation de la protéine Tau réduit nettement son activité stabilisatrice des microtubules (Drechsel *et al.* 1992 ; Drubin and Kirschner. 1986). La phosphorylation du MTBR constituerait également un mode de régulation neuronale de sa distribution le long de l'axone (Dixit *et al.* 2008). De plus, certains travaux suggèrent que le MTBR, conjointement au domaine riche en prolines, est lui aussi impliqué dans la liaison de Tau à l'actine (Elie *et al.* 2015). Enfin, le MTBR participerait à la liaison de Tau avec les protéines 14-3-3, retrouvées au cœur de nombreuses maladies neurodégénératives et impliquées dans la survie cellulaire et la promotion d'agrésomes<sup>11</sup> (Foote and Zhou. 2012 ; Melková *et al.* 2019). Enfin, l'extrémité C-terminale de Tau, en aval du MTBR, est le lieu de nombreuses modifications post-traductionnelles (phosphorylations et

---

<sup>11</sup> Les agrésomes sont des dépôts cytoplasmiques de protéines mal repliées observés dans un contexte pathologique (Johnston *et al.* 1998).

truncations pour n'en nommer que deux) et pourrait être impliquée dans les processus pathologiques.

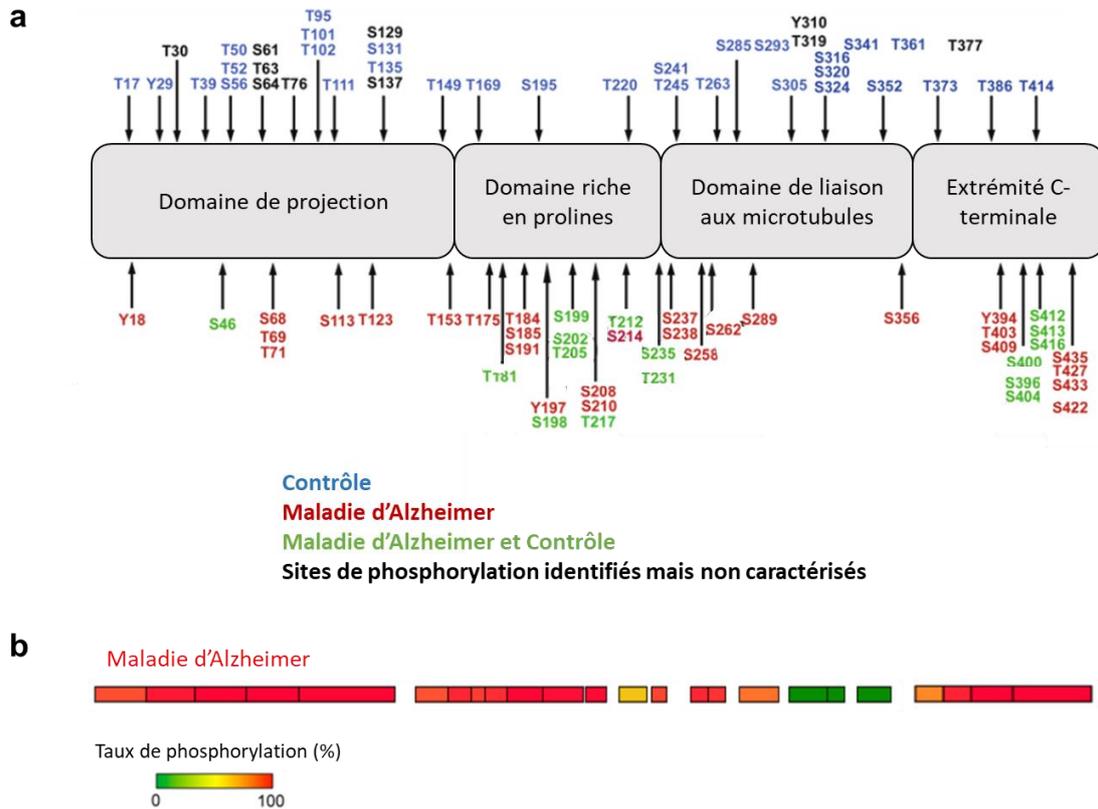
### III.3 Modifications post-traductionnelles

#### III.3.1 Addition de groupements fonctionnels

Tau, et sa forme agrégée, présentent de nombreuses modifications post-traductionnelles dont les plus communes sont la phosphorylation, l'acétylation et la O-glycosylation. Ces modifications post-traductionnelles ont un impact sur les fonctions biologiques de Tau, sa localisation et sa dégradation.

##### III.3.1.1 Phosphorylations

Alors que la protéine monomérique est, dans une moindre mesure, elle aussi sujette à phosphorylation, il est communément admis que dans la maladie d'Alzheimer les agrégats de Tau sont retrouvés hyperphosphorylés (Grundke-Iqbal *et al.* 1986). Il existe 85 sites putatifs de phosphorylation au sein de la protéine Tau entière dont 70 ont été décrits comme sujets à modification. Alors que 28 sites n'ont été retrouvés effectivement modifiés que le cas de la maladie d'Alzheimer (par rapport aux individus contrôles), certaines sont retrouvées à la fois dans le cerveau adulte normal et celui atteint de la maladie d'Alzheimer (**Figure 14.a**). Toutefois, la distribution générale des sites de phosphorylations est différente et le taux de modification peut également diverger (Šimić *et al.* 2016). Ainsi, de nombreux anticorps ont été créés pour reconnaître certains sites phosphorylés de Tau mais ne permettent pas une quantification robuste de l'état de phosphorylation. En 2016, l'étude FLEXITau, reposant sur la quantification de ces modifications post-traductionnelles (plus particulièrement des phosphorylations) sur une fraction sarkosyle de tissu cortical de patients atteints de la maladie d'Alzheimer à large échelle, a été réalisée par l'équipe de J Steen (Mair *et al.* 2016). Ainsi, en moyenne, on observe 10 phosphates par molécule de Tau au sein des agrégats issus de patients atteints de la maladie d'Alzheimer avec une distribution allant de 4 à 19 sites modifiés. Il a été noté un taux de modification accru de l'extrémité N-terminale, de la majorité de la région riche en prolines et de l'extrémité C-terminale. Au sein du domaine de liaison au microtubules, on observe un taux de modification bien moins important, en particulier au niveau des troisième et quatrième séquences répétées (R3 et R4). En amont de ce domaine, dans l'extrémité C-terminale du domaine riche en prolines (acides aminés 243-254), on observe également un taux de modification modéré. Ces observations sont illustrées en **Figure 14.b**.



**Figure 14.** Phosphorylation de Tau chez des individus sains et des patients atteints de la maladie d'Alzheimer.

(a) Sites de phosphorylation identifiés de Tau dans des cerveaux de patients contrôles (bleu), atteints de la maladie d'Alzheimer (rouge), contrôles et atteints de la maladie d'Alzheimer (vert), observés mais non caractérisés (noir). Issu de Šimić *et al.* 2016. (b) Quantification du taux de phosphorylation le long de la séquence de Tau dans des cerveaux de patients atteints de la maladie d'Alzheimer évalué par l'étude FLEXITau (Mair *et al.* 2016).

Les potentielles kinases à l'origine de cette hyperphosphorylation *in vivo* sont nombreuses. Parmi elles, certaines ont fait l'objet d'études *in vitro* et *in vivo* quant à leur activité sur Tau ; c'est par exemple le cas de la GSK3 $\beta$  (glycogen synthase kinase 3 $\beta$ ), la Cdk5 (cyclin-dependent kinase 5), la PKA (cAMP-dependent protein kinase) et la MARK (Microtubule-Affinity-Regulating Kinase) (Johnson and Stoothoff. 2004). Ces enzymes ne doivent cependant pas être seules responsables de la phosphorylation de Tau ; il est toutefois difficile d'étudier le phénomène *in vivo* sans biais, d'autant plus que cet effet peut être direct et indirect (Cavallini *et al.* 2013).

Ainsi, l'hyperphosphorylation est un marqueur de la maladie d'Alzheimer mais la question de la chronologie de la phosphorylation et de la fibrillation de Tau reste en suspens. La majorité de la communauté s'accorde à dire que la phosphorylation de Tau est à l'origine même de l'agrégation de Tau. Le schéma proposé est le suivant : la phosphorylation de Tau diminuerait la liaison de cette dernière aux microtubules, engendrant une augmentation de la

concentration des formes libres qui auraient ainsi tendance à s'agréger. Deux arguments sont toutefois en faveur d'une phosphorylation postérieure à la polymérisation :

- (1) L'addition accrue de groupements phosphates induirait un changement de charge globale de la protéine. Ceci contraindrait beaucoup l'établissement de l'agrégation en raison de l'effet de répulsion de charges associé à l'incrément de charges négatives
- (2) Il s'avère que la région peu modifiée chez les patients atteints de la maladie d'Alzheimer correspond à celle décrite comme le coeur des fibres dans la littérature. Ceci s'expliquerait davantage si le processus de phosphorylation avait lieu en aval de l'agrégation au niveau des surfaces accessibles de l'assemblage, comme le suggéraient déjà Hasegawa et ses collaborateurs au début des années 1990 (Hasegawa *et al.* 1992).

Le rapport de causalité entre l'interruption de l'interaction avec les microtubules par phosphorylation et l'agrégation doit être considéré avec précaution ; il ne s'est d'ailleurs pas toujours révélé probant lors des études de phosphorylations *in vitro* et serait site-spécifique (Schneider *et al.* 1999 ; Biernat *et al.* 1993). De façon intéressante, d'autres travaux ont montré le rôle potentiellement bénéfique de l'hyperphosphorylation de Tau sur la survie cellulaire. En effet, l'hyperphosphorylation de Tau est retrouvée conjointement à une diminution de la phosphorylation de  $\beta$ -caténine et sa conséquente translocation nucléaire, à l'origine d'une résistance à l'apoptose (Li *et al.* 2007). Ceci est en accord avec l'idée que la protéine Tau et la  $\beta$ -caténine sont toutes deux substrats de la kinase GSK3 et que la modification de l'une par la GSK3 serait compétitive de celle de l'autre (Wu *et al.* 2009).

Si certains travaux, comme FLEXITau en 2016, ont tenté des approches de phosphorylation *de novo in vitro*, une étude complète *in vitro* des modifications post-traductionnelles reste particulièrement difficile à mettre en place compte tenu du grand nombre de sites putatifs de modifications et de la grande hétérogénéité de ces dernières. Il n'en reste pas moins qu'une meilleure connaissance de ces phénomènes post-traductionnels reste essentielle et constitue un défi de premier plan pour la communauté scientifique.

### *III.3.1.2 Autres modifications*

Les lysines, au nombre de 44 au sein de Tau 2N4R, sont des résidus cibles de nombreuses modifications post-traductionnelles dont l'acétylation, l'ubiquitination, la sumoylation, la formylation et la méthylation (Zee and Garcia. 2012). Ainsi, des acétylations ont également été identifiées, notamment au niveau des lysines 174 (PRR), 274 et 280 (MTBR) au sein de cerveaux de patients atteints de la maladie d'Alzheimer, de la maladie de Pick et de la FTDP-17 (Park *et al.* 2018). *In vitro*, il a été montré que l'acétylation de Tau 2N4R monomérique au niveau de la lysine 280 promouvait son agrégation et perturbait sa fonction physiologique (i.e

inhibition de la polymérisation des microtubules) (Cohen *et al.* 2011). Des expériences similaires ont suggéré que l'acétylation au niveau de la lysine 274 conduirait aux mêmes observations (Rane *et al.* 2019). Ces acétylations pourraient moduler la phosphorylation des résidus avoisinants (et réciproquement), modulant le rôle de la protéine Tau dans la régulation de la polymérisation microtubulaire (Barbier *et al.* 2019). De plus, il semblerait que, chez les patients atteints de la maladie d'Alzheimer, l'alanine 2, résidu N-terminal de la protéine après excision de la méthionine 1, est retrouvée acétylée au sein des PHF (Hasegawa *et al.* 1992). Par ailleurs, les trois sites de poly-ubiquitinylation aujourd'hui identifiés sont les lysines 254, 311, 353 (MTBR) également sujettes à l'acétylation (Cripps *et al.* 2006). De plus, il a également été montré, sur des cellules en culture, que Tau, qui contient deux séquences consensus de SUMO ( $\Psi$ KXE<sup>12</sup>), pouvait être sumoylé au niveau de la lysine 340 (Dorval and Fraser. 2006). La pertinence biologique de cette modification et son éventuelle implication dans le processus neurodégénératif associé à Tau reste à établir. Enfin, des méthylations, au niveau des lysines 44 et 163, 174, 180 (PRR) et 254, 267 et 290 (MTBR), ont également été mises en évidence au sein de PHFs issus de cerveaux de patients atteints de la maladie d'Alzheimer. Une quantification de ces modifications semble suggérer que l'ubiquitinylation est moins prépondérante que la méthylation, notamment au niveau de la lysine 254. Bien que des travaux *in vitro* évoquent un rôle délétère de la méthylation dans le phénomène d'agrégation, une analyse fonctionnelle approfondie de l'effet de ces modifications reste à réaliser (Thomas *et al.* 2012 ; Funk *et al.* 2014).

Les glycosylations constituent l'addition d'oligosaccharides sur la chaîne latérale d'acides aminés spécifiques, asparagine pour la N-glycosylation et sérine et thréonine pour la O-glycosylation. Il existe trois sites putatifs de N-glycosylation sur Tau 2N4R (Asn 167 (PRR), 359 (MTBR), 410) (Sato *et al.* 2001) et 11 sites possibles de O-glycosylation (sérines 199,202 (PRR) et 262, 356 (MTBR) et 404, 422 (CR) ; thréonines 181, 205, 212, 214, 217 (PRR)) (Wang *et al.* 2003). La N-glycosylation semble avoir lieu sur la Tau phosphorylée, notamment au niveau des PHF de cerveaux de patients atteints de la maladie d'Alzheimer, alors que Tau n'est pas retrouvée glycosylée dans le cerveau humain normal (Wang *et al.* 1996). La O-glycosylation est plus difficile à prédire et à mettre en évidence expérimentalement car le substitut est de taille et de composition variable (Kizuka *et al.* 2017). Il a été observé une diminution du taux d'O-glycosylation de Tau chez les patients atteints d'AD par rapport aux individus contrôle (Robertson *et al.* 2004). Ainsi, alors que la N-glycosylation stabiliserait la structure des PHF, la O-glycosylation en préviendrait la formation et régulerait négativement la phosphorylation de Tau (Schedin-Weiss *et al.* 2014 ; Liu *et al.* 2004). De plus, il a été mis en évidence que les PHFs de Tau pouvaient subir des glycations, réaction qui associe

---

<sup>12</sup> Ici  $\Psi$  désigne un acide aminé hydrophobe de grande taille

également un glucide à une protéine (résidu N-terminal, lysines et arginines) mais sans l'intervention d'une enzyme. Cette réaction, dite de Maillard, génère des produits terminaux de glycation (appelés AGE pour Advanced Glycation Endproducts en anglais), retrouvés en quantité importante dans le cerveau des patients atteints de la maladie d'Alzheimer (Münch *et al.* 1997). Ainsi, des expériences ont montré que la mise en présence de Tau glyquée sur des neurones en culture génère des espèces réactives de l'oxygène à l'origine d'une dérégulation des fonctions neuronales (Avila *et al.* 2004).

Enfin, une déamidation extensive de l'asparagine 279 (MTBR) appartenant à l'hexapeptide 4R-spécifique VQIINK, central au processus d'agrégation, a été identifiée dans des fractions sarkosyl-insolubles de patients atteints de la maladie d'Alzheimer. Cette modification, à l'origine d'une faible réactivité à l'anticorps RD4 4R-spécifique largement utilisé par la communauté scientifique, ne semble pas avoir lieu dans le cadre de la PSP et de la CBD (Dan *et al.* 2013).

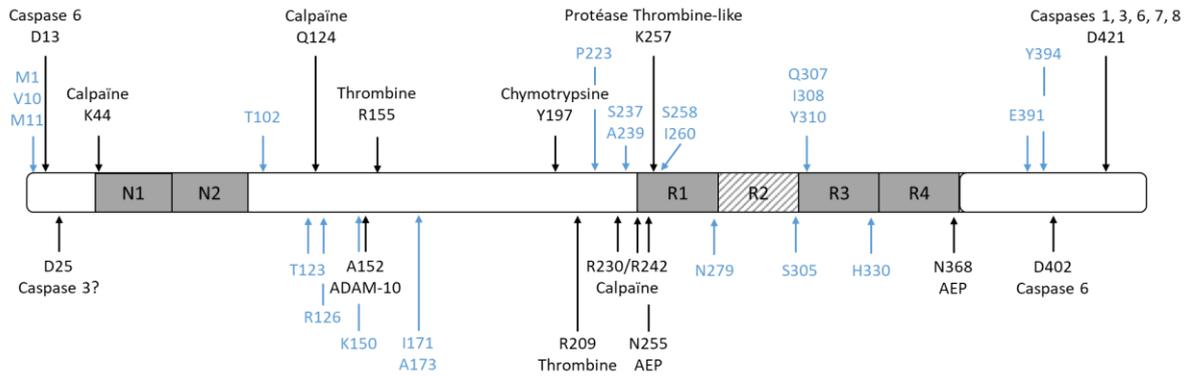
### III.3.2 Ponts disulfure

Les isoformes 3R et 4R de Tau contiennent respectivement une et deux cystéines au sein de motifs KCGS. Les isoformes 3R ne peuvent donc former que des ponts disulfures intermoléculaires tandis que les isoformes 4R peuvent aussi établir des ponts intramoléculaires. Une étude par RMN du solide de la structure des fibres formées par K19 (forme tronquée de Tau 2N3R, présentée en **Annexe 2**) a montré l'importance du pont disulfure établi entre les cystéines 322 de soit (1) chacun des deux protofilaments, soit (2) deux molécules de Tau situées sur le même protofilament. Dans des conditions réductrices, il est possible de générer des fibres de Tau K19 mais le rendement en est grandement affecté (facteur 2). A l'inverse, la formation de fibres de Tau K18 (forme tronquée de Tau 2N4R, présentée en **Annexe 2**) n'est pas réalisable en conditions oxydatives. Pour expliquer ce phénomène, on peut émettre l'hypothèse selon laquelle un pont disulfure intramoléculaire entre les cystéines 291 (R2) et 322 (R3) s'établirait en conditions oxydatives et bloquerait la polymérisation (Barghorn and Mandelkow. 2002). Une expérience complémentaire par laquelle cette cystéine est mutée en alanine montre un réarrangement du corps des fibres de K19, notamment des résidus K321-S324, en accord avec des données préliminaires obtenues sur K11 et K12 (présentées en Annexe 2) (Daebel *et al.* 2012 ; Schweers *et al.* 1995). L'implication des ponts disulfures dans le processus d'agrégation sera davantage discutée dans le paragraphe III.4.1. Enfin, l'état d'oxydo-réduction de la protéine Tau est aussi important pour la réalisation de ses fonctions normales. En effet, il a été montré que l'oxydation de Tau

par addition de peroxyde d'hydrogène ou d'ions peroxyntriques affectait la polymérisation des microtubules, phénomène réversible par addition ultérieure d'agents réducteurs (Landino *et al.* 2004). Ainsi, l'état d'oxydation des cystéines de la protéine Tau module sa fonction biologique et son agrégation de façon isoforme-dépendante, et pourrait ainsi prendre part à la pathologie.

### III.3.3 Clivages protéolytiques

La protéine Tau est la cible de multiples clivages enzymatiques (Figure 15). Il a notamment été montré que plusieurs membres de la famille des caspases (i.e caspases 3, 7 et 8) protéolysent efficacement Tau 2N4R monomérique en C-terminal de la protéine (Asp421). D'autre part, ce clivage est retrouvé au niveau des dégénérescences neurofibrillaires issues de cerveaux de patients atteints de la maladie d'Alzheimer. De plus, de façon intéressante, dans des neurones corticaux, il a été observé que Tau (dont l'état d'oligomérisation est indéterminé) était sujette à protéolyse par des caspases après induction de l'apoptose par A $\beta$ . De plus, ces formes générées, tronquées en C-terminal, s'assemblent plus rapidement que la forme entière de Tau (Gamblin *et al.* 2003). D'autres formes tronquées par la caspase 6, notamment au niveau de l'Asp13 et l'Asp402, ont été identifiées. Toutefois, la chronologie des événements de protéolyse et d'agrégation est à ce jour inconnue. Des travaux ont néanmoins suggéré que le clivage par la caspase 3 en position 421 aurait lieu en aval du processus de polymérisation de Tau 2N4R entière (Jarero-Basulto *et al.* 2013). D'autres protéases ont pour cible la protéine Tau ; il s'agit notamment des calpaïnes, activées par le Ca<sup>2+</sup> cytosolique. La calpaïne 1 génère ainsi, à partir de Tau 2N4R monomérique, *in vitro* un fragment de Tau de 17 kDa, neurotoxique pour des cultures de neurones hippocampiques et dont les extrémités N et C terminales présumées ont été assignées aux positions 43 et 229 (Park and Ferreira. 2005). D'autres troncations de Tau, dont l'enzyme impliquée n'est pas toujours connue, ont été mises en évidence depuis les années 1990 à nos jours et sont représentées en bleu dans la **Figure 15**.



**Figure 15.** Sites de clivage et protéases associées sur la séquence de Tau 2N4R. En noir, les sites de clivage dont la protéase est connue, en bleu les sites de clivage dont la protéase n'a pas encore été identifiée. Adaptée de Wang *et al.* 2010 et de Quinn *et al.* 2018.

Plus récemment, Derisbourg et collaborateurs ont développé une stratégie comprenant l'immunoprécipitation d'espèces de Tau issues de fractions corticales de cerveaux de patients atteints d'AD, leur marquage par la biotine (N-terminal et lysines) suivi d'une digestion, d'une purification par liaison avec des billes de streptavidine puis d'une analyse LC-MS/MS. En plus de tronctions dans la région C-terminale de Tau, cette approche leur a notamment permis de mettre en évidence des formes tronquées dans la partie N-terminale de Tau. Certaines de ces espèces, en se liant aux microtubules, semblent jouer un rôle dans leur stabilisation (Derisbourg *et al.* 2015). Retrouvées dans les fluides biologiques (i.e le LCR et dans une moindre mesure le sang), ces tronctions pourraient notamment constituer un biomarqueur de la pathologie (Inekci *et al.* 2015 ; Cicognola *et al.* 2019). Toutefois, nombreuses de ces tronctions sont à la fois retrouvées dans le cerveau des patients atteints de la pathologie d'Alzheimer et de personnes saines. Ainsi, pour pouvoir les utiliser comme outils diagnostiques, un examen de leur distribution et une quantification dans les différentes tauopathies seraient nécessaires.

### III.3.4 Isomérisations

Un changement de conformation de la protéine Tau, impliquant une isomérisation des prolines de la configuration *cis* vers la configuration *trans*, peut avoir lieu (Wedemeyer *et al.* 2002). Objet de nombreuses études, la protéine Pin1 (Peptidyl-prolyl *cis-trans* Isomerase NIMA-interacting 1) isomériser les motifs phospho-Ser/Thr-Pro, régulant ainsi les phénomènes de phosphorylation des protéines. Il a été démontré que Pin1 induisait la déphosphorylation de Tau par PP2A (Protéine Phosphatase 2A) en isomérisant le motif phospho-Thr231-Pro (Zhou *et al.* 2000). De façon concomitante, Pin1 est retrouvée liée aux enchevêtrements

neurofibrillaires retrouvés chez les patients atteints de la maladie d'Alzheimer et déplétée sous sa forme soluble ; ce qui n'est le cas dans le cytoplasme de neurones de patients contrôle. De plus, la promotion de la polymérisation des microtubules par la protéine Tau peut être restaurée *in vitro* par addition de Pin1. Ces observations semblent donc suggérer que Pin1 joue un rôle dans la pathologie (Lu *et al.* 1999). Enfin, une accumulation cytoplasmique granulaire de Pin1 a également été mise en évidence dans d'autres tauopathies comme la maladie de Pick et la FTDP-17 tandis que l'isomérase ne semble pas associée à la PSP (Ramakrishnan *et al.* 2003).

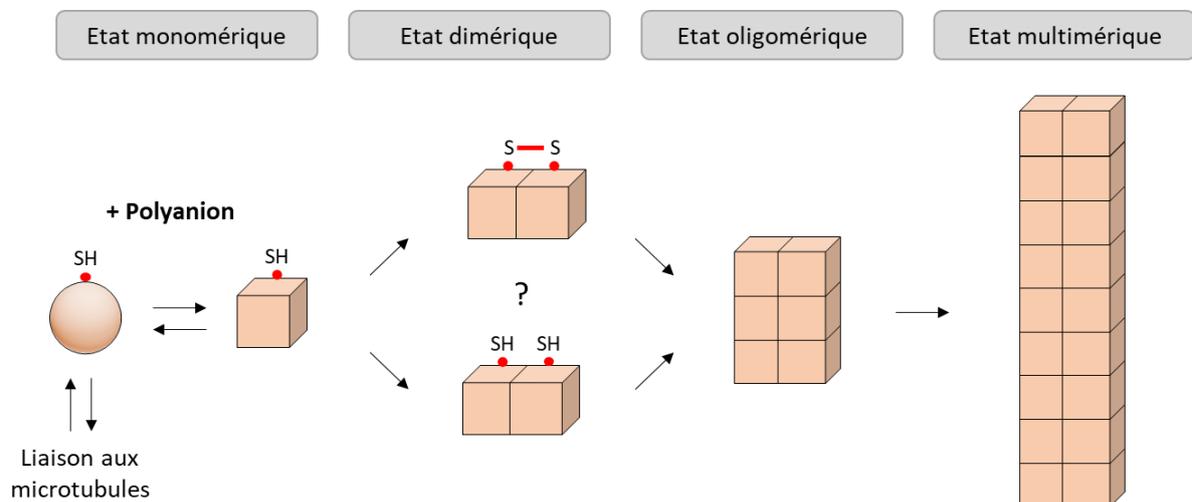
### III.4 Agrégation

#### III.4.1 Formation des agrégats

La compréhension du processus de formation des agrégats de Tau est essentielle pour appréhender les tauopathies et développer des outils thérapeutiques ou diagnostiques. La polymérisation *in vitro* des formes recombinantes de Tau permet de documenter le mécanisme d'agrégation. Cette dernière peut être suivie par différentes méthodes dont la sédimentation, la chromatographie par exclusion stérique (Flach *et al.* 2012) ou par suivi de la fixation des Thioflavines S et T (Friedhoff *et al.* 1998). Les agrégats générés peuvent ensuite être observés par microscopie électronique (Kidd. 1963). Il est à noter que Tau seule ne polymérise pas spontanément (Friedhoff *et al.* 1998b (PNAS)). Il a été montré que l'incubation de Tau 2N4R avec des glycosaminoglycanes telle que l'héparine (Goedert *et al.* 1996), de l'ARN (Kampers *et al.* 1996), de l'ADN (Hikosou *et al.* 2007) ou encore avec des acides gras libres tel que l'acide arachidonique (Wilson and Binder. 1997) permettait la formation d'assemblages de haut poids moléculaire. Il a été proposé que ces polyanions, en interagissant avec la protéine Tau monomérique déstructurée, permettent un changement de conformation (e.g structures  $\beta$  pour R2 et R3 ; hélices  $\alpha$  pour la séquence N-terminale 149-154 riche en lysines et le C-terminal du R4) favorable à l'agrégation (von Bergen *et al.* 2005; Sibille *et al.* 2006). Le rôle de l'héparine sur les changements conformationnels de Tau a notamment été étudié par FRET. La liaison de l'héparine à Tau, par des interactions électrostatiques augmenterait transitoirement lors de l'agrégation l'accessibilité au solvant des composants du MTBR initiant le processus d'agrégation, augmentant ainsi la probabilité de contact intermoléculaire entre ces régions. Ainsi, l'héparine aurait *in fine* pour conséquence une compaction de la région MTBR et de la région riche en prolines au sein des polymères. Elle ne provoquerait pas de changement observable dans la première partie du segment N-terminal du domaine de

projection (aa 17-103) tandis que des changements conformationnels sont observés dans la seconde partie (aa 103-184), consistant avec des précédents travaux prouvant une interaction entre le domaine 149-154 riche en lysines et l'héparine (Elbaum-Garfinkle and Rhoades. 2012).

Alors que certains avancent la possibilité que les fibres de Tau assemblées *in vitro* en présence d'héparine seraient différentes de celles assemblées *in vivo*, il est probable que le processus d'agrégation *in vivo* requiert un processus équivalent et impliquerait que Tau se lie à d'autres composants cellulaires dans un contexte pathologique (Chirita *et al.* 2005). Il a d'ailleurs été observé une colocalisation de dermatane sulfates et de sulfate d'héparane avec des neurodégénérescences neurofibrillaires dans l'hippocampe de patients atteints de la maladie d'Alzheimer (Snow *et al.* 1992 ; Goedert *et al.* 1996). De plus, il a été montré une accumulation des sulfates d'héparane et des sulfates de chondroïtine, deux glycosaminoglycanes (GAGs), dans l'hippocampe de patients atteints d'AD. Additionnellement, les GAGs retrouvés chez les patients AD ont une plus grande capacité de liaison avec Tau monomérique (Huynh *et al.* 2019). Ainsi, il se pourrait que les glycosaminoglycanes jouent également un rôle important dans le processus d'agrégation de Tau *in vivo*.



**Figure 16.** Schéma proposé du processus de polymérisation des fibres de Tau

Ainsi, un des mécanismes proposés pour le processus de nucléation implique la formation de dimères, notamment par établissement de ponts disulfures intermoléculaires (Friedhoff *et al.* PNAS 1998a ; Congdon *et al.* 2008). Cette nucléation serait ensuite suivie d'une oligomérisation marquant l'entrée en phase d'élongation (Patterson *et al.* 2011). La formation

de dimères par établissement de ponts disulfures supposée nécessaire au processus de nucléation est toutefois mise à mal par le succès des expériences de fibrillation en conditions réductrices (Wilson and Binder. 1995 ; Al-Hilaly *et al.* 2017). En effet, comme développé précédemment, il a été montré que la polymérisation des isoformes 3R de Tau est plus rapide en conditions oxydatives tandis que celle des isoformes 4R l'est davantage en présence d'agents réducteurs (Barghorn *et al.* 2005; Barghorn and Mandelkow. 2002). En conditions réductrices, la mutation de la cystéine 322 ne modifie pas la cinétique d'assemblage de Tau 0N3R, plus lente qu'en conditions oxydatives. En conditions réductrices, la double mutation des cystéines 291 et 322 n'affecte pas non plus la vitesse de polymérisation de Tau 2N4R (Barghorn and Mandelkow. 2002). Il faut aussi souligner que, *in vivo*, le contexte physiologique (i.e cytosol) est réducteur ; ce qui appuie le fait que les ponts disulfures joueraient peu dans le processus d'agrégation (Kuret *et al.* 2005). Il est difficile d'identifier les espèces intermédiaires tels que des noyaux de nucléation, structures instables. Des observations par MET et AFM ont mis en évidence l'existence de formes agrégées de Tau de petite taille et d'apparence sphérique appelées SNUs (Spherical Nucleation Units) (Xu *et al.* 2010). Il est postulé que ces formes agrégées constitueraient un intermédiaire structural d'agrégats riches en structures  $\beta$  qui, plus tard dans le processus de fibrillation, coexisteraient avec des agrégats de plus haut poids moléculaire telles que les fibres matures (Xu *et al.* 2010). Il n'existe pas de consensus dans la littérature concernant le nombre moyen de molécules de Tau dont ces oligomères seraient constitués : il pourrait s'agir de dimères ou de trimères d'après Lasagna-Reeves (Lasagna-Reeves *et al.* 2012a FASEB) ou d'assemblages d'environ 40 molécules de Tau d'après Maeda (Maeda *et al.* 2007) ou encore de 74 molécules d'après Xu (Xu *et al.* 2010). *In vivo*, chez les patients atteints de la maladie d'Alzheimer, ces SNUs seraient présentes dès les premiers stades de Braack, avant la formation des dégénérescences neuro-fibrillaires (Maeda *et al.* 2007 ; Lasagna-Reeves *et al.* 2012a FASEB). Les éventuelles propriétés fonctionnelles de ces espèces oligomériques, très débattues, font l'objet de nombreuses études. Ainsi, des travaux ont montré que des oligomères issus de cortex cérébral de patients atteints de la maladie d'Alzheimer pourraient agir sur la potentialisation à long terme dans l'hippocampe et perturberaient la fonction de mémoire chez la souris (Lasagna-Reeves *et al.* 2012b Sci Rep). De plus, les oligomères générés *in vitro* nuiraient à l'intégrité de membranes et auraient un effet cytotoxique *in cellulo* (Flach *et al.* 2012). Toutefois, ces données doivent être interprétées avec prudence compte tenu de l'hétérogénéité de préparation, de concentration, de taille et de structure de ces oligomères qui existent aujourd'hui dans la littérature. En effet, ces différences pourraient d'ailleurs engendrer un biais considérable quant à l'interprétation du rôle toxique de ces espèces (Glabe. 2008). La toxicité des agrégats multimériques matures, de grande taille, sera quant à elle discutée en partie III.4.3.

Les six isoformes de Tau sont capables de polymériser en présence d'acide arachidonique ou d'héparine *in vitro*. Cependant, *in vitro*, la présence ou l'absence des exons soumis à épissage alternatif (i.e exons 2, 3 et 10) module la cinétique de fibrillation. En effet, en présence de rouge thiazine, la propension à s'agréger de Tau est améliorée par les exons 2 et 10 et diminuée par l'exon 3. En présence de rouge thiazine, mais aussi d'acide arachidonique, les isoformes 4R ont une phase de latente plus courte que les isoformes 3R ; l'exon 10 influence donc le taux de nucléation (King *et al.* 2000 ; Zhong *et al.* 2012). Des études ont tenté de mettre en évidence la séquence de Tau minimale nécessaire et suffisante pour son agrégation. De façon surprenante, une forme tronquée de Tau, constituée uniquement de la séquence 317-KVTSKCGSLGNIHHKPGGG-335 (partie C-terminale de R3) est capable de polymériser en présence d'héparine. Des expériences complémentaires semblent montrer que son agrégation serait dépendante de la formation de ponts disulfure. A l'inverse, les séquences 256-VKSKIGSTENLKHQPGGG-273 (partie C-terminale du R1) et 350-VQSKIGSLDNITHVPGGG-367 (partie C-terminale du R4), seules, ne polymérisent pas dans les mêmes conditions (Pérez *et al.* 1996). De la même façon, la forme tronquée de Tau ne comprenant que la partie N-terminale de la protéine et les 7 premiers acides aminés de R1, n'est pas capable de former de fibres (Pérez *et al.* 1996). Toutefois, ceci ne présume pas que la partie N-terminale de la protéine ne module pas l'agrégation de la protéine entière. En effet, dans une autre étude, il a été montré qu'une délétion des acides aminés 2-18 réduit d'un facteur 2,4 la capacité d'assemblage de Tau en présence d'acide arachidonique (Gamblin *et al.* 2003). D'autre part, l'implication de l'extrémité C-terminale dans le processus d'assemblage a également été étudiée, en présence d'acide arachidonique. Il en résulte que la troncation de Tau en C-terminal (dès la seconde moitié du R2 au résidu C-ter, i.e 321-441) facilite la polymérisation de Tau ; la région minimale à l'origine de l'inhibition de la polymérisation par la région C-terminale serait comprise entre les acides aminés 430-441 (Abraha *et al.* 2000).

Ainsi, l'ensemble de ces résultats suggère que la région au cœur du processus de polymérisation est située au niveau du MTBR tandis que les parties N et C terminales jouent un rôle régulateur dans le mécanisme d'agrégation. Ceci suggère également que l'agrégation de modèles de Tau tronquée largement utilisés dans la littérature pourrait impliquer d'autres mécanismes à l'échelle moléculaire et que ces formes tronquées doivent être considérées avec prudence. Pour conclure, des zones d'ombre persistent autour du mécanisme d'agrégation de Tau, de l'implication de l'état d'oxydo-réduction aux différents intermédiaires conformationnels présents tout au long du processus d'assemblage.

La présence des isoformes 3R et 4R au sein des agrégats issus de cerveaux de patients atteints de la maladie d'Alzheimer et les propriétés d'agrégation propres aux différentes isoformes (prises séparément), évoquées plus haut, ont soulevé la question suivante : les formes 3R et 4R sont-elles capables de co-agréger de façon indifférenciée ou existe-t-il une « barrière de souche »<sup>13</sup> responsable de la formation d'agrégats composés d'un seul type d'isoforme de Tau?

Dans un premier temps, il a été montré par une analyse par Résonance Paramagnétique Electronique (RPE), qu'en présence d'héparine, la co-polymérisation *in vitro* des formes 0N3R et 2N4R en conditions équimolaires conduisait à la formation de fibres mixtes (3R et 4R au sein de la même entité) observables par MET (Siddiqua and Margittai. 2010). Parmi les six isoformes, Tau 0N3R présente la concentration critique à partir de laquelle Tau polymérise ( $C_c$ ) et le temps de latence les plus élevés tandis que Tau 2N4R est associée à une  $C_c$  et un temps de latence bas. Quand Tau 2N4R est incubé avec Tau 0N3R, 2N4R semble favoriser l'agrégation de 0N3R et en abaisser la valeur de  $C_c$  associée (Zhong *et al.* 2012). A l'inverse, cette étude suggère aussi que les isoformes 3R pourraient inhiber la polymérisation de Tau 4R (Adams *et al.* 2010). Plus tard, il a été suggéré l'existence d'une barrière de souche asymétrique isoforme-dépendante. En effet, alors que des amorces de Tau 3R pouvaient initier la fibrillation de Tau 3R et de Tau 4R, des amorces de 4R ne semblaient pas capables d'initier l'assemblage de Tau 3R (Dinkel *et al.* 2011). De plus, des expériences complémentaires de réactions de nucléation en chaîne ont révélé qu'une même isoforme, ici 4R, pouvait générer des fibres dont les propriétés de recrutement et de nucléation variaient en fonction de l'amorce initiale 3R et étaient transmises au cours des générations (Dinkel *et al.* 2011). Toutefois, en 2018, le même groupe contredit ces résultats en présentant que des amorces de Tau 4R auraient aussi la capacité de recruter la protéine monomérique 3R et de former ainsi des fibres (Weismiller *et al.* 2018). L'existence de souches 3R et 4R distinctes dont les déterminants structuraux seraient distincts et soumis à une barrière de souche, régulant aussi leur propagation *in vivo* restent donc à démontrer. De plus, il a été observé qu'une forme agrégée tronquée de 2N4R (comprenant uniquement les résidus 1–372) et une forme agrégée entière mutée en position P301S and P301L ne sont pas capables d'induire la polymérisation de Tau 0N3R et forment des fibres au cœur structuré divergeant. Ces résultats suggèrent que la partie C-terminale de la protéine et certaines mutations pourraient jouer un rôle déterminant dans l'établissement d'une barrière de souches structurellement distinctes (Weismiller *et al.* 2018 ; Frost *et al.* 2009 ; Aoyagi *et al.* 2007).

---

<sup>13</sup> Je choisis ici de parler de « barrière de souche » où les différentes souches correspondraient à des agrégats composés d'isoformes de Tau différentes.

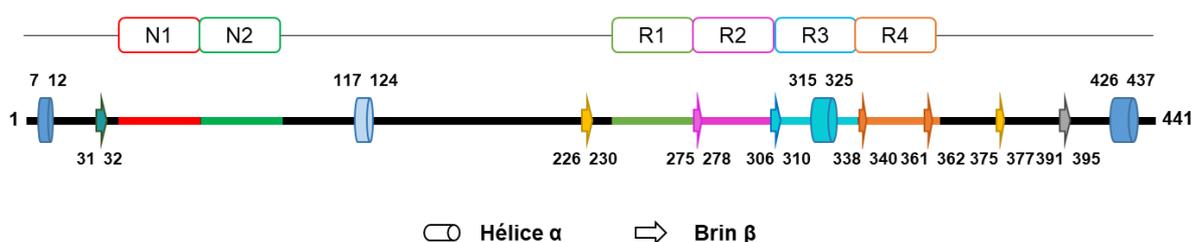
En résumé, le processus d'agrégation de Tau fait intervenir différents intermédiaires conformationnels *in vitro* et *in vivo*. L'addition de polyanions est nécessaire à la polymérisation *in vitro* ; *in vivo*, l'assemblage de la protéine Tau endogène pourrait également faire intervenir des polyanions physiologiques et/ou des partenaires protéiques cellulaires. Toutes les isoformes ne présentent pas les mêmes capacités d'agrégation suggérant l'existence de souches. Le phénomène de barrière de souche n'est à ce jour pas clairement établi et nécessiterait des travaux complémentaires, notamment pour déterminer si des fibres mixtes pourraient exister *in vivo*.

### III.4.2 Structure de Tau monomérique et de ses agrégats

Les fonctions d'une protéine (e.g rôle stabilisateur et promoteur de la formation des microtubules pour Tau) et ses propriétés (e.g capacité à polymériser pour Tau) dépendent notamment de la structure que peut prendre cette protéine. Nous avons utilisé l'annotation de séquence du plus long isoforme, i.e 2N4R, pour la suite de la discussion.

#### III.4.2.1 Prédictions de structures secondaires

Des prédictions de structures secondaires de Tau ont été établies par différentes études. En raison du degré de robustesse des algorithmes (60-70% de fiabilité), nous ne présenterons ici que les structures retrouvées de façon redondante dans la littérature, illustrées en figure 17.



**Figure 17.** Structures secondaires, hélices  $\alpha$  et brins  $\beta$ , prédites le long de la séquence de Tau 2N4R monomérique. En rouge, insert N1 ; en vert prononcé, insert N2 ; en vert pâle, R1 ; en rose, R2 ; en bleu azur, R3 ; en orange, R4. Adapté de Gamblin *et al.* 2005.

Au niveau de l'extrémité N-terminale, les séquences 7-EFEVME-12 et 31-MH-32 pourraient respectivement former une hélice  $\alpha$  et un brin  $\beta$ . En aval des régions encodées par les exons 2 et 3, la séquence peptidique 117-EAAGHVTQ-124 est prédite comme structurée en hélice  $\alpha$  et le peptide 226-VAVVR-230 en brin  $\beta$ . Au sein du MTBR, toutes les études s'accordent à

prédire que les séquences 275-VQII-278 (R2) et 306-VQIVY-310 (R3) seraient prompts à former des brins  $\beta$  amphipathiques. Egalement au sein du R3, le polypeptide 315-LSKVTSKCGSL-325 pourrait former, en conditions hydrophobes, une hélice  $\alpha$  amphipathique. Dans le R4, les séquences 338-EVK-340 et 361-TH-362 et, en C-terminal du MTBR, 375-KLT-377 et 391-EIVYK-395, constitueraient des brins  $\beta$ . Enfin, le peptide C-terminal 426-ATLADEVASLA-437 aurait tendance à se structurer en hélice  $\alpha$  de nature amphipathique (Gamblin *et al.* 2005).

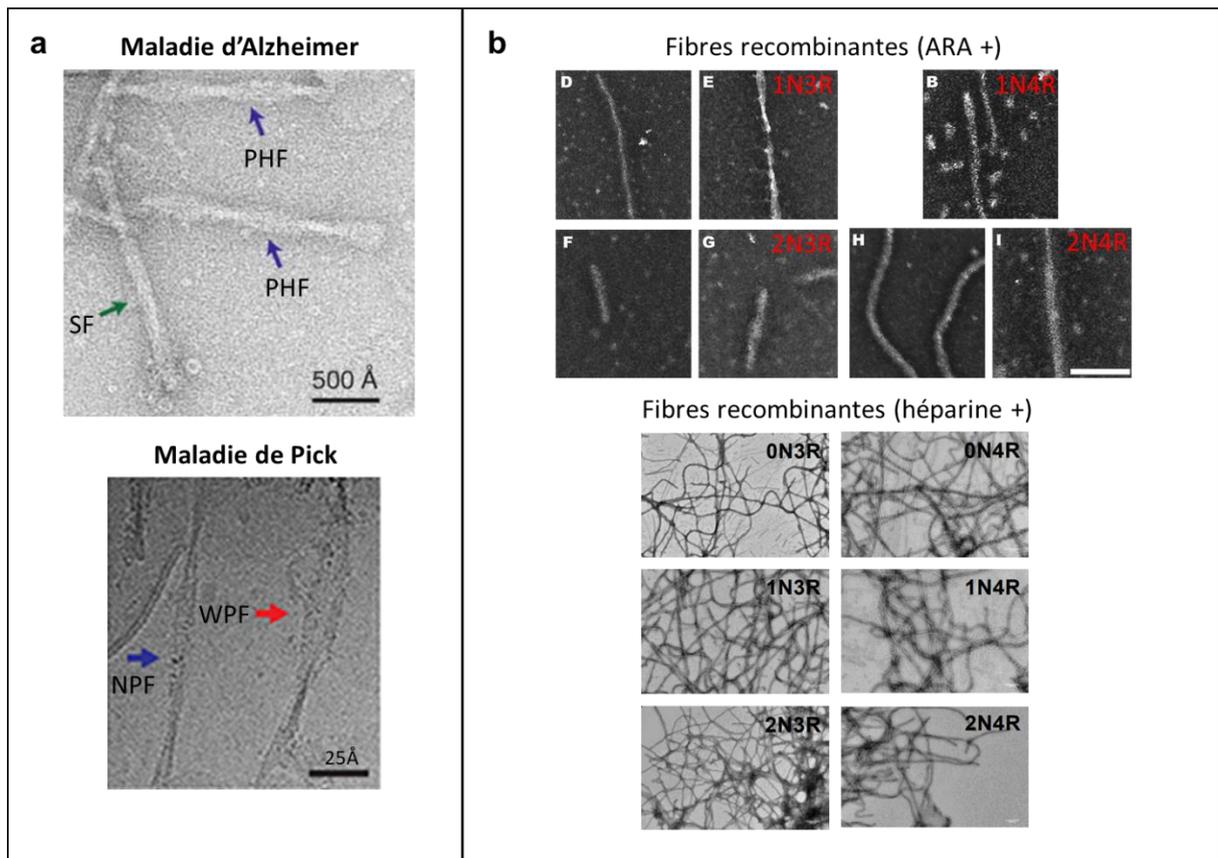
#### III.4.2.2 Etude RMN de la structure de Tau monomérique

La plus longue isoforme de Tau, 2N4R, comprend 441 résidus, dont l'arrangement est majoritairement désordonné. Tau est donc une protéine de grande taille pour des approches classiques de détermination structurale telles que la cristallographie suivie de la diffraction aux rayons X ou la RMN. De plus, les protéines dont la structure est principalement désordonnée sont caractérisées en RMN par un recouvrement de signal, à l'origine d'une incertitude d'attribution. C'est pourquoi l'essentiel des travaux ont jusqu'à très récemment porté sur des petits fragments de tau (e.g hexapeptides) ou au mieux sur des formes tronquées (e.g K18, K19). Cependant, les déterminants structuraux de la forme dite « soluble » ou plus justement monomérique de Tau 2N4R ont toutefois pu être caractérisés par RMN (Mukrasch *et al.* 2009). Bien que très majoritairement désordonnée, comme attendu, la protéine monomérique présente quelques régions pouvant transitoirement adopter des structures secondaires. Ainsi, l'étude montre que certaines régions de la protéine monomérique sont faiblement (moins de 20% en moyenne) organisées en structures  $\beta$  en amont (86-GKQAAAQ-92 (2N spécifique), 161-GQKGQA-166 et 224-KKVAVVR-230 (région riche en prolines)) et au sein du MTBR (256-VKSKIG-262 (R1), 274-KVQIINKKLDL-284 (R2), 305-SVQIVYKPVDL-315 (R3), 336-QVEVKSEKLD-345 et 351-QSKIGSL-357 (R4)). De plus, deux peptides en N-terminal et C-terminal présentent une organisation partielle (respectivement 18% et 25%) en hélices  $\alpha$  (114-LEDEAAGHVT-123 et 428-LADEVASLA-437). Enfin, ces données de RMN suggèrent que le domaine C-terminal (360–441) rentre transitoirement en contact avec la région répétée et les 40 premiers résidus N-terminaux. Pour conclure, certaines régions de la protéine Tau, bien que globalement désordonnée, peuvent transitoirement adopter une structure secondaire dont certaines ont été prédites par modélisation.

#### III.4.2.3 Structure de Tau fibrillaire

L'hypothèse prion selon laquelle la conformation d'une protéine peut définir un phénotype transmissible ou héréditaire a longtemps été débattue par la communauté et reste pour certains

discutable (Mudher *et al.* 2017 ; Polanco *et al.* 2018). En effet, les barrières énergétiques entre les différents intermédiaires conformationnels sont basses et il n'était donc pas imaginable qu'une conformation soit assez stable pour être conservée lors de la transmission aux cellules. Pour établir une haute barrière d'énergie entre les conformations, il doit exister un ensemble de liaisons non covalentes fortes qui maintiennent la protéine dans son état conformationnel pendant le processus de transmission. Ainsi, il est admis que l'état fibrillaire est maintenu par une forte densité de liaisons hydrogène. De plus, il a été observé que les amyloïdes ont des motifs steric-zipper également maintenus par de fortes liaisons de Van der Waals séquence-spécifiques. Du fait qu'il existe une grande hétérogénéité de conformations possibles des motifs steric-zipper pour une même protéine, ceci pourrait jouer un rôle important dans « l'encodage » des souches (Wiltzius *et al.* 2009). En terme de taille, de composition en acides aminés et de séquence, les peptides et protéines associés aux maladies neurodégénératives amyloïdes n'ont pas d'homologie. Il a cependant été observé une morphologie globale commune des agrégats lors des observations au MET (i.e aspect de fibres allongées, exemple des fibres de Tau en **Figure 18**). Les analyses spectrales de dichroïsme circulaire, dont l'utilisation est soumise à controverse dans le cas des fibres amyloïdes, et de FTIR ont révélé un contenu enrichi en structures  $\beta$  que la diffraction aux rayons X semble indiquer organisées en feuilletts perpendiculairement à l'axe du (des) protofilament(s) (Stefani *et al.* 2003). Le contenu en structures  $\beta$  de la protéine Tau agrégée fait aujourd'hui consensus (von Bergen *et al.* 2000).

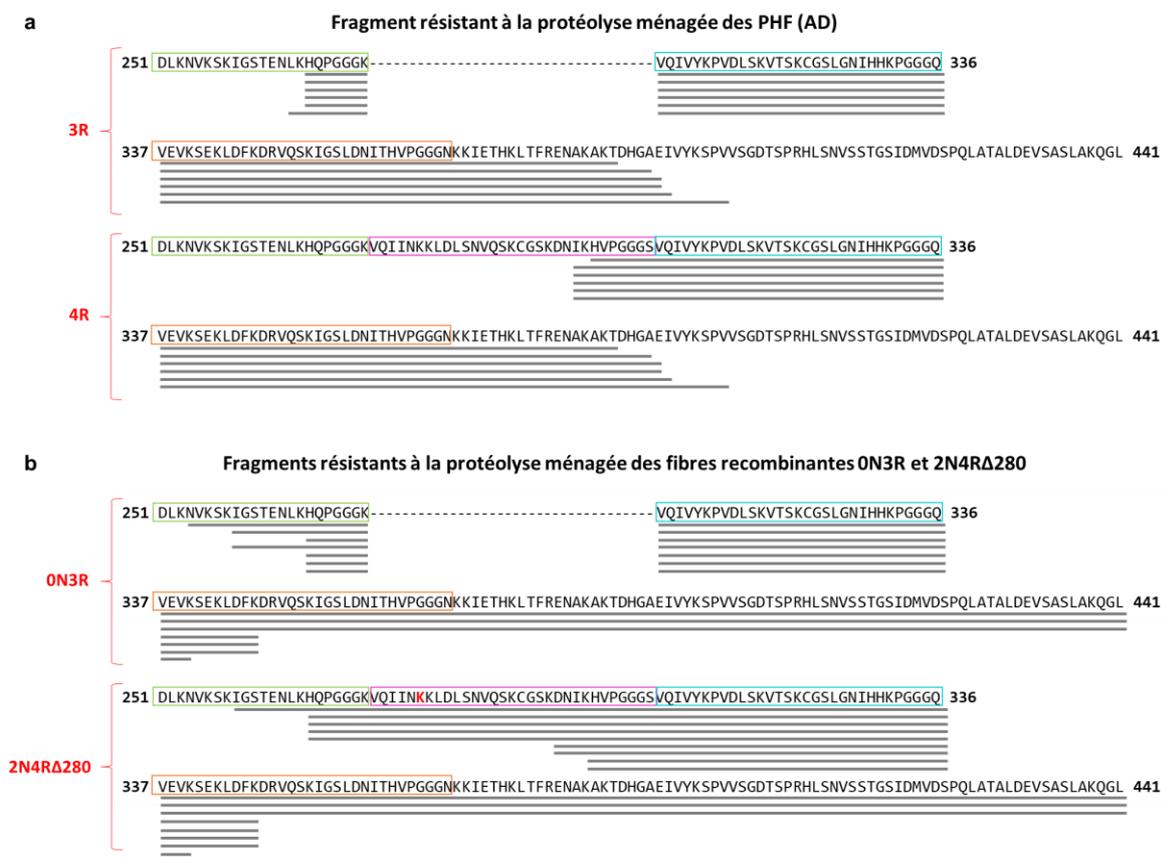


**Figure 18.** Observation par MET de fibres de Tau issues de patients et assemblées in vitro

(a) Observation par MET de fibres de Tau issues, en haut, de cerveaux de patients atteints de la maladie d'Alzheimer (PHF, Paired Helical Filaments et SF, Straight Filaments) et, en bas, de cerveaux de patients atteints de la maladie de Pick (NPF, Narrow Pick Filaments et WPF, Wide Pick Filaments). (b) Observation par MET de fibres de Tau générées *in vitro* en présence d'acide arachidonique, en haut, et d'héparine en bas. Issu de Fitzpatrick *et al.* 2017 ; Falcon *et al.* 2018; King *et al.* 2001 ; données de notre équipe (Karine Madiona, non publié).

Les premiers travaux d'étude de structure des agrégats de Tau reposent sur des expériences de protéolyse ménagée opérée par des enzymes aspécifiques comme la pronase sur des agrégats de type PHF issus de patients atteints de la maladie d'Alzheimer. Comme illustré en **Figure 19.a**, ces expériences ont montré qu'un fragment de 12 kDa (et plus minoritairement un fragment de 9.5 kDa), correspondant aux séquences 269-386 (isoforme 3R, absence de R2) et 297-396 (isoforme 4R), soit environ 90 résidus, était davantage résistant à la protéolyse. D'après ces premières données, les PHFs sont constitués des isoformes 3R et 4R et leur cœur comprend au minimum la partie C-terminale du R1 (pour 3R) et du R2 (pour 4R), R3, R4 et une séquence d'une vingtaine d'acides aminés en aval du MTBR (Wischik *et al.* 1988; Jakes *et al.* 1991 ; Novak *et al.* 1991 ; Novak *et al.* 1993). Des expériences similaires de protéolyse ménagée ont été réalisées, cette fois-ci, par la trypsine et sur des fibres obtenues *in vitro* en présence d'héparine de Tau 0N3R et d'un mutant 2N4R $\Delta$ K280 (**Figure 19.b**) (von Bergen *et al.* 2006). Ainsi, les fibres 0N3R présentent un cœur dont la partie N-terminale est relativement similaire à celle identifiée dans les PHF mais différent nettement en C-ter pour lequel on

retrouve essentiellement (1) l'ensemble de la séquence C-terminale de la protéine, (2) une troncation au niveau des résidus 340 et 347. Quant aux fibres 2N4RΔK280, elles diffèrent de la même façon dans la partie C-terminale mais aussi dans la partie N-terminale pour laquelle on retrouve une protection plus étendue, dès la fin du R1. La différence observée en C-terminale est imputée à l'action *in vivo* de la calpaïne au niveau de la séquence PEST (407-438) (Wang *et al.* 1989 ; Yang and Ksiezak-Reding. 1995). Il est à noter que la délétion de la lysine 280 a été décrite comme ayant un effet positif fort sur l'agrégation et pourrait donc impliquer d'autres mécanismes d'assemblage qui rend ici la comparaison avec d'autres fibres 4R discutable (von Bergen *et al.* 2001).



**Figure 19.** Détermination du coeur des fibres par protéolyse ménagée.

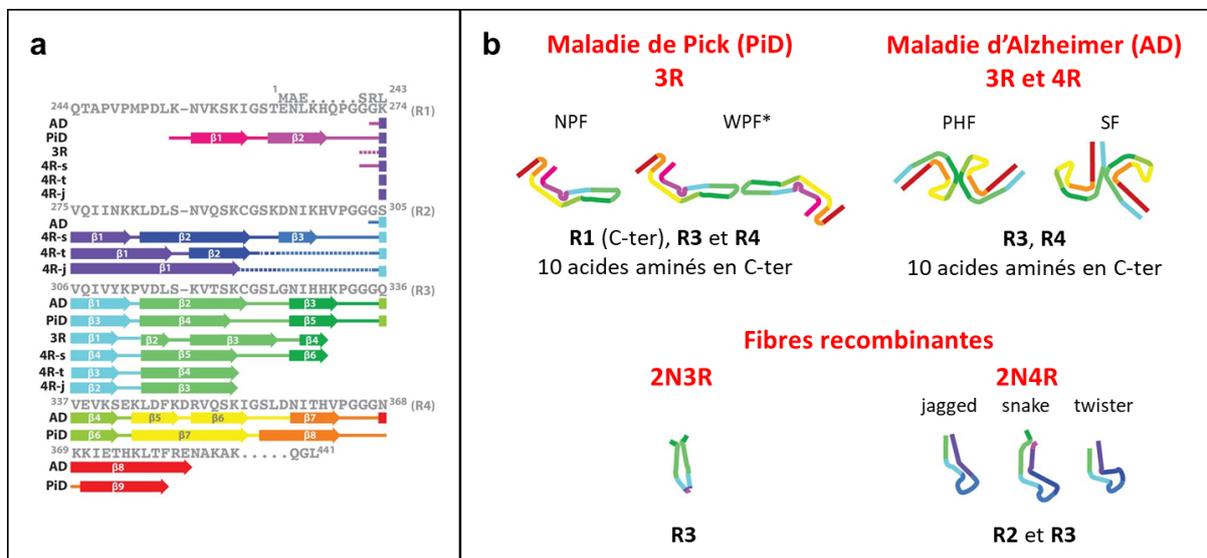
(a) Séquences correspondant au fragment de 12 kDa généré par la pronase sur des PHF issus de patients atteints de la maladie d'Alzheimer (Résidus N-terminaux identifiés par séquençage d'Edman et résidus C-terminaux putatifs associés). Issu de Novak *et al.* 1993. (b) Séquences correspondant aux fragments générés par la trypsine sur des fibres recombinantes de Tau 0N3R et Tau 2N4RΔK280 (Résidus N-terminaux identifiés par séquençage d'Edman et résidus C-terminaux déduits d'après les masses moléculaires mesurées par SELDI-MS). Issu de von Bergen *et al.* 2006.

Pour des agrégats formés à partir de 31 résidus de Tau (uniquement du domaine R3, aa 306-336, en PBS sans héparine), des analyses par échanges hydrogène-deutérium ont montré

une forte protection des résidus 308-IVYK-311 et 315-SKVTSKC-322, consistant avec la structure cristallographique de l'hexapeptide 306-VQIVYK-311 (motif tight steric-zipper) (Sawaya *et al.* 2007), tandis que les résidus 323-GSLGNIH-329 montrent un certain degré de protection (J (Stöhr *et al.* 2017). Cependant, il est à noter que les résultats de cette étude montrent que le résidu Gln307 n'est pas protégé contrairement à la structure cristallographique. Pour la forme réduite assemblée (Cys322 SH), le R3 serait structuré en 3 brins bêta espacés de 4.8Å d'après les résultats de FTIR. Les frontières des feuillets se situeraient au niveau des acides aminés V313-D314 et G326-N327. L'extrémité C terminale du R3 serait beaucoup moins structurée (Stöhr *et al.* 2017). Les régions N et C-terminales jouant un rôle dans l'agrégation, l'analogie structurale entre les formes assemblées de Tau tronquée et de Tau entière reste fragile.

Récemment, grâce à un développement technologique accru, des structures de cœurs de fibres issues de patients et générées *in vitro* ont été résolues par cryomicroscopie électronique (**Figure 20**). En 2017, la structure des agrégats de Tau purifiés du cortex temporal et frontal d'un patient atteint de la pathologie d'Alzheimer a été publiée (Fitzpatrick *et al.* 2017). Ainsi, comme attendu dans la maladie d'Alzheimer, les agrégats décrits dans cette étude sont composés des isoformes 3R et 4R de Tau ; la coexistence ou non des deux types d'isoformes au sein d'un même agrégat n'a toutefois pas été démontrée. De plus, ces travaux confirment la présence de deux structures d'agrégats distinctes observées par MET : les PHFs et les SFs (définis dans la partie introductive II.3). Les cœurs structurés de ces assemblages sont composés de deux protofilaments et de la même séquence en acides aminés, i.e V306-F378 (R3, R4 et 10 acides aminés additionnels en Cter) organisée en 8 feuillets  $\beta$  parallèles. Cependant, les structures de cryo-microscopie montrent que ces assemblages diffèrent dans leur organisation quaternaire. En effet, dans le cas des PHFs, l'interface entre les deux protofilaments est constituée des enchainements anti-parallèles des motifs 332-PGGGQ-336. La structure asymétrique des SFs engendre une hétérogénéité des séquences constituant l'interface : 321-KCGS-324 pour un premier protofilament et 313-VDLSK-317 pour le second. La résolution de la densité électronique ne permet pas l'identification d'une région de 16 acides aminés organisée en feuillet  $\beta$  environnant le cœur de ces fibres mais les auteurs suggèrent par des analyses de spectrométrie de masse complémentaires (recouvrement de séquence) qu'il s'agirait des extrémités C-terminales des répétitions 1 et 2 alternativement. De la même façon, en 2018, deux structures sont identifiées à partir d'agrégats purifiés et traités à la pronase issus du cortex temporal et frontal d'un patient atteint de la maladie de Pick (pathologie 3R prédominante) (Falcon *et al.* 2018). Un unique protofilament compose le premier type d'assemblages appelé Narrow Pick Filaments (NPFs, 93% des agrégats) tandis

que deux protofilaments sont à l'origine des Wide Pick Filaments (WPFs, 7% des agrégats). Par manque de résolution, la structure de ces derniers n'a, cependant, pas pu être décrite jusqu'à présent. Les acides aminés K254-F378 (R1, R3, R4 et 10 acides aminés additionnels en C-terminal) constituent les 9 brins  $\beta$  du corps des NPFs. Ce repliement, et la présence unique des isoformes 3R, diffère donc de celles retrouvées dans la maladie d'Alzheimer. Enfin, en 2019, une étude similaire a été réalisée sur les fibres de Tau 2N4R et Tau 2N3R assemblées *in vitro* en présence d'héparine (Zhang *et al.* 2019). Là encore, la structure du coeur de ces fibres a été élucidée après un traitement à la pronase visant à éliminer les régions désordonnées constituant le « fuzzy coat » autour du coeur structuré des fibres. Les auteurs soulignent l'hétérogénéité des fibres 2N4R avec l'identification de 4 formes distinctes appelées « snake » (45%), « twister » (30%), « jagged » (5%) et « hose » (20%). Une élucidation structurale du coeur des trois premières formes a pu être réalisée. Les fibres de Tau 2N4R assemblées en présence d'héparine sont toutes constituées d'un unique protofilament et le squelette de ces fibres est composé des acides aminés G272/K274-K321/H330 (R2 et R3). Les acides aminés 306-VQIVY-310 et 313-VDLSKVTSK-321 constituent deux brins  $\beta$  communs à l'ensemble des coeurs décrits. Dans le cas de l'assemblage *in vitro* en présence d'héparine de Tau 2N3R, on note une homogénéité des structures retrouvées (98% identiques). Le squelette de ces fibres, issues de l'association de 2 protofilaments, est constitué des acides aminés G272/K274-H330 (R3).



**Figure 20.** Etudes structurales du coeur des fibres de Tau par cryomicroscopie électronique après digestion à la pronase.

(a) Brins  $\beta$  présents le long de la séquence structurée. (b) Représentation schématique du repliement, en haut à gauche, des NPF (Narrow Pick Filaments) et des WPF (Wide Pick Filaments, non résolus) retrouvés dans le cortex de patients atteints de la maladie de Pick (Tau 3R) ; en haut à droite des PHF (Paired Helical Filaments) et des SF (Straight Filaments) retrouvés dans le cortex de patients atteints de la maladie d'Alzheimer (Tau 3R et 4R) ; en bas, des fibres recombinantes 2N3R et 2N4R (morphologie jagged, snake et twister). Issu de Fitzpatrick *et al.* 2017 ; Falcon *et al.* 2018a ; Zhang *et al.* 2019.

Ainsi, les agrégats de Tau retrouvés chez les patients atteints de la maladie d'Alzheimer et de la maladie de Pick présentent une hétérogénéité de conformation avec une prédominance de la forme PHF et NPF respectivement (Crowther. 1991; Moreno-Herrero *et al.* 2004). Ce pluralisme de formes est également retrouvé pour les fibres générées *in vitro* (Siddiqua *et al.* 2012).

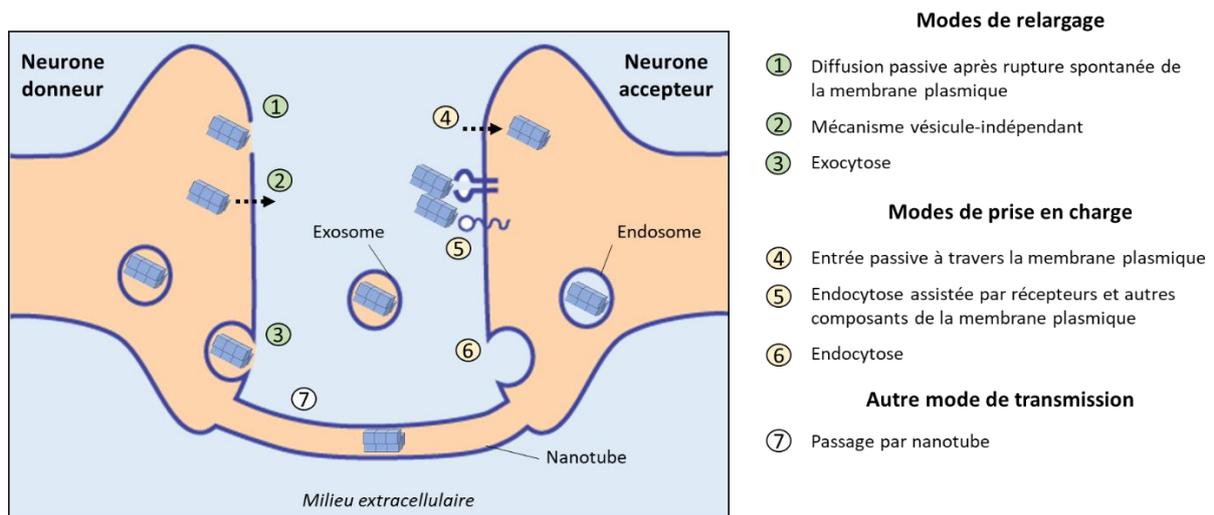
La pertinence des agrégats synthétisés en laboratoire par rapport à ceux présents dans les tauopathies humaines reste aujourd'hui discutée. Par STEM (microscope électronique en transmission à balayage), il a été établi que les fibres recombinantes présentaient une organisation en paires de filaments similaire à celles des PHF de patients atteints de la maladie d'Alzheimer mais une densité réduite (King *et al.* 2001). Les dernières études de cryomicroscopie électronique suggèrent que le cœur des fibres de Tau recombinantes diffère de celui des agrégats retrouvés chez les patients (Zhang *et al.* 2019). Toutefois, les séquences identifiées comme hautement structurées dans la pathologie et dans les assemblages formés *in vitro* sont nettement plus étendues d'après les données antérieures de protéolyse ménagée présentées plus haut. Cette divergence majeure montre qu'il est nécessaire de mettre en perspective les données de cryomicroscopie électronique et ses limitations (voir chapitre 4) et qu'un complément d'informations structurales, notamment sur les régions extérieures au cœur, est nécessaire.

L'utilisation de formes tronquées de Tau pour l'étude de son agrégation, de la structure des agrégats et de la pathologie Tau *in vivo* est aujourd'hui très débattu. En effet, plusieurs observations mettent à mal la pertinence de ces conformères. D'une part, il est avéré que les fibres de patients peuvent à la fois être composées de Tau tronquée (en N et C-terminal) (Zhou *et al.* 2018) ou de Tau entière (Fitzpatrick *et al.* 2017), la chronologie d'occurrence de la troncation n'étant pas connue. D'autre part, le cœur des fibres de type AD et de type PiD comprend une séquence C-terminale de 10 acides aminés (369-378) en aval du MTBR. D'après une étude sur le processus d'agrégation réalisée par FRET, le MTBR est davantage compacté dans la forme tronquée K16 (198-372) que dans la protéine entière, ce qui est cohérent avec de précédentes suggestions proposant que la présence des extrémités inhibe des interactions spécifiques au sein du MTBR (Elbaum-Garfinkle and Rhoades. 2012). Ces arguments mettent à mal l'utilisation massive de formes tronquées de Tau (**Annexe 2**) comme modèle pour l'étude de la structure des agrégats de Tau, de leur capacité de recrutement de la protéine native et du phénomène de propagation associé. Les fibres recombinantes générées *in vitro* permettent toutefois de disposer d'agrégats « purs », d'échantillons dont le contenu est connu, contrairement aux agrégats purifiés issus de patients. De plus, les protocoles de purification des agrégats issus de patients comprennent de multiples étapes,

susceptibles d'engendrer des changements conformationnels. Enfin, alors que certains soulignent une hétérogénéité des assemblages de fibres de Tau formées *in vitro* (Zhang *et al.* 2019), un pluralisme de formes est également observé *in vivo*. Ceci conditionne évidemment l'utilisation de méthodes peu soumises à des variations d'échantillonnage pour l'étude de la structure de ces agrégats. Ces assemblages formés *in vitro* constituent des modèles d'étude de complexité réduite, indispensables à la compréhension des mécanismes sous-jacents de l'agrégation et de la propagation.

### III.4.3 Transmission « prion-like »

Comme évoquée précédemment, une propagation de cellule à cellule des agrégats de Tau permet d'expliquer le mode d'évolution, de proche en proche, de la pathologie associée à Tau. La forme agrégée circulante n'a jusqu'à présent pas été élucidée : certains spéculent qu'il s'agirait des oligomères, d'autres de fibres matures. Différents mécanismes de propagation intercellulaire ont été proposés et sont explicités en **Figure 21**. A l'exception de la transmission par nanotubes, ces derniers nécessitent tous (1) un mode de relargage de la cellule hôte (i.e neurone donneur), (2) un mode de prise en charge par la cellule réceptrice (i.e neurone accepteur) et (3) une capacité de recrutement et d'amplification des formes monomériques de Tau par l'espèce transmise. Ces mécanismes ont suscité, et suscitent toujours, de nombreux travaux de recherche (Vasili *et al.* 2019).



**Figure 21.** Mécanismes possibles de propagation cellule à cellule des agrégats de Tau

Des travaux sur des cultures de cellules exprimant une forme tronquée de Tau (MTBR sauvage et mutants) ont montré que des agrégats provenant de cellules hôtes pouvaient se

propager à des cellules réceptrices co-cultivées et recruter la forme monomérique. Ces mêmes agrégats sont retrouvés dans le milieu extracellulaire, preuve qu'ils ont été relargués par les cellules donneuses. Par la suite, les auteurs ont montré que ces assemblages n'étaient pas encapsidés mais libres dans le milieu extracellulaire. Enfin, par microscopie à force atomique, ils ont pu observer que les espèces présentes dans le milieu extracellulaire présentaient une morphologie de type fibres (Kfoury *et al.* 2012).

Les mécanismes mis en jeu dans les processus de relargage et de prise en charge de Tau ont été activement étudiés ces dernières années. La première hypothèse proposée quant au passage du milieu intracellulaire au milieu extracellulaire de Tau agrégée consiste en une rupture de la membrane plasmique, pouvant être la conséquence de l'accumulation d'agrégats (Figure 21. Hypothèse 1). En effet, il a été montré que des enchevêtrements fantômes persistaient dans le milieu extracellulaire après la mort du neurone hôte (Bancher *et al.* 1989). Un second processus (Figure 21. Hypothèse 2), appuyé par une étude publiée en 2018, voudrait que la forme circulante de Tau (ici, des oligomères) induise une certaine perméabilité de la membrane plasmique pour ainsi être sécrétée (Merezhko *et al.* 2018). De plus, des travaux ont montré que des oligomères Tau 0N4R étaient copurifiés avec des fractions exosomales dans des cultures cellulaires et dans le liquide céphalo-rachidien de patients atteints de la maladie d'Alzheimer, essentiellement aux premiers stades de la pathologie (Saman *et al.* 2012). Ce transfert de Tau par la voie exosomale semble se produire au niveau de l'espace intersynaptique (Wang *et al.* 2017). De plus, il a été mis en évidence la présence d'oligomères de Tau dans des exosomes de souris *in vivo* (Asai *et al.* 2015). Ceci tend à valider l'hypothèse selon laquelle la transmission de cellule à cellule de Tau agrégée pourrait impliquer un transport exosomal (Figure 21. Hypothèse 3). D'autres types de vésicules, comme les ectosomes, pourraient également servir la propagation de Tau (Dujardin *et al.* 2014).

Une possible prise en charge des formes agrégées de Tau médiée par la liaison à des composants membranaires est appuyée par l'identification de clusters de fibres de Tau 1N3R et 1N4R à la membrane plasmique. Ces assemblages se lient à de nombreuses protéines dont des récepteurs transmembranaires détenant un domaine extracellulaire tels que les sous-unités GluA1 et GluA2 des récepteurs AMPA, les sous-unités GluN1 and GluN2B des récepteurs NMDA ou encore les sous-unités  $\alpha 1$ ,  $\alpha 3$  et  $\beta 1$  de la pompe Na/K-ATPase (Shrivastava *et al.* 2019). Pour être internalisées, ces formes agrégées pourraient ensuite être endocytées (Figure 21. Hypothèse 5). Il a d'ailleurs été montré que les protéoglycanes à héparane sulfate, présents au niveau de la membrane plasmique, pouvaient jouer un rôle central dans le processus d'endocytose (Figure 21. Hypothèse 5), et plus spécifiquement de macropinocytose (Figure 21. Hypothèse 6), des agrégats extracellulaires de Tau (Holmes *et*

*al.* 2013). Une récente étude rapporte que les formes monomérique et agrégée de Tau partagent une voie d'endocytose dynamine-dépendante mais n'observe une intégration par macropinocytose que pour la forme monomérique (Evans *et al.* 2018). Une autre hypothèse, encore jamais démontrée, propose que l'espèce transmise diffuse à travers la membrane de la cellule réceptrice (probablement lors d'une rupture transitoire) mais ceci n'a encore jamais été démontrée (Figure 21. Hypothèse 4).

Enfin, en 2016, il a été démontré la transmission cellule à cellule d'espèces fibrillaires de Tau 1N4R, générées au sein de notre laboratoire, par des nanotubes faisant communiquer deux cellules voisines (Figure 21. Hypothèse 7). La formation de ces nanotubes est justement induite par l'addition exogène de ces assemblages et permet à l'espèce circulante d'être transmise d'une cellule à l'autre sans nécessité de relargage ni de prise en charge (Tardivel *et al.* 2016). Ce mécanisme original de propagation par nanotubes reste à mettre en évidence *in vivo*.

Pour conclure, différents mécanismes de transmission cellule à cellule ont été proposés et démontrés en laboratoire, principalement sur des cultures cellulaires. La forme circulante, oligomérique ou fibrillaire, et la nature de Tau mise en jeu, Tau entière ou tronquée, ne font pas consensus au regard des différentes études. Les processus mis en jeu semblent pluriels et pourraient refléter une multiplicité de prise en charge de ces espèces toxiques et des voies de propagation co-existantes *in vivo*.

Le phénomène de propagation des formes agrégées de Tau a également été validé à plus large échelle *in vivo*. En effet, l'injection stéréotaxique de fractions corticales sarkosyl-insolubles issues de patients atteints de différentes tauopathies (AD, CBD et PSP) chez la souris au niveau de l'hippocampe conduit au recrutement et à l'agrégation de la protéine endogène dans les régions avoisinantes du site d'injection (Narasimhan *et al.* 2017). Ainsi, trois mois après inoculation, on observe que le développement de la pathologie Tau est plus étendu pour la PSP que pour AD et CBD. Tous les agrégats diffusent dans les régions connectées anatomiquement au site d'injection mais la CBD et la PSP montrent une étendue plus large dans les régions corticales de la pathologie Tau, suggérant des propriétés de propagation tauopathie-spécifiques (Narasimhan *et al.* 2017). Toutefois, six mois après injection, quelque soit la région d'inoculation considérée, cette différence régio-sélective entre les tauopathies n'est plus observée, en désaccord avec l'évolution humaine de ces pathologies. D'un point de vue cellulaire, alors que les agrégats de Tau sont uniquement retrouvés dans les neurones pour AD, ils sont également retrouvés dans les cellules gliales pour CBD (plaques astrocytaires) et PSP (astrocytes gonglés) ; des observations comparables à celles

réalisées dans la pathologie humaine. Au niveau biochimique, ces inclusions ont été immunodétectées par différents anticorps conformationnels spécifiques (épitopes phosphorylés ou non) et présentent des différences en fonction de l'homogénat AD, PSP et CBD administré et du type cellulaire considéré (Narasimhan *et al.* 2017). Des travaux similaires avec des homogénats totaux issus de patients atteints de AD, PART, PSP, CBD, AGD et PiD chez des souris transgéniques ALZ17<sup>14</sup> et le modèle murin sauvage ont été réalisés PiD (Clavaguera *et al.* 2013). Ainsi, toutes ces formes agrégées injectées ont eu pour conséquence un développement croissant dans le temps de la pathologie Tau, à l'exception de PiD qui semble confinée à la zone d'injection et qui, de façon surprenante, n'engendre pas le développement de corps de Pick. L'administration d'homogénats AD et PART induisent des inclusions de Tau de morphologie similaire tandis que, pour PSP et CBD, des inclusions AD-like sont retrouvées dans les neurones. Les agrégats AGD conduisent à la formation de structures agrégées ressemblant aux grains argyrophiles observés dans la pathologie humaine. Les souris inoculées avec les homogénats PSP, CBD et AGD présentent des inclusions de Tau dans les oligodendrocytes (i.e « coiled bodies ») et dans les astrocytes (astrocytes gonflés pour la PSP, plaques astrocytaires pour la CBD et une morphologie similaire à ce qui est observée dans la pathologie humaine pour AGD). Enfin, aucune inclusion gliale n'est relevée pour AD, PART et PiD (Clavaguera *et al.* 2013).

En résumé, les fractions sarkosyl-insolubles contenant des agrégats de Tau issues de patients atteints de tauopathie sont capables d'induire un phénotype pathologique chez l'animal. Ces agrégats sont capables de recruter la protéine endogène et d'en induire l'agrégation, à l'image des autres protéines prion-like. Les inclusions résultantes diffèrent entre les tauopathies et leurs caractéristiques cellulaire-spécifiques et biochimiques semblent conservées, suggérant un phénomène de souche.

Comme développé précédemment, à partir d'une même protéine (e.g PrP,  $\alpha$ -synucléine) peuvent être générées différentes souches, aux propriétés conformationnelles distinctes et à l'origine de phénotypes différents.

Plusieurs équipes ont donc cherché à étudier la potentielle existence de souches de Tau agrégée. Les travaux portant sur la structure du cœur des fibres présentes dans les différentes tauopathies semblent confirmer un phénomène de souches (Taniguchi-Watanabe *et al.* 2016;

---

<sup>14</sup> Le modèle transgénique de souris ALZ17 comprend l'expression de Tau humaine 2N4R sauvage sous le contrôle du promoteur neural murin Thy-1.2. Le transgène est exprimé 5 fois plus que la Tau endogène de souris (Probst *et al.* 2000)

Mudher *et al.* 2017). Différentes souches, dont les caractéristiques biochimiques ont été élucidées par SDD-AGE et protéolyse ménagée, ont pu être générées et stabilisées dans des cellules HEK293 en culture. Ces souches présentent des propriétés de nucléation hétéroclites et sont à l'origine de profils pathologiques différents chez la souris P301S. Ces divergences permettent de discriminer les souches notamment en fonction de la pathologie astrocytaire, de la régio-sélectivité et de la vitesse de propagation au sein du réseau neuronal (Sanders *et al.* 2014 ; Kaufman *et al.* 2016).

Les particularités conformationnelles de différentes souches pourraient donc jouer un rôle majeur dans l'encodage de la tauopathie (Falcon *et al.* 2015). Ces différents épitopes structuraux pourraient conduire au développement de précieux outils moléculaires pour le diagnostic différentiel précoce des tauopathies et pour l'innovation thérapeutique associée. L'identification de ces déterminants conformationnels, nécessite la mise en œuvre d'approches structurales, adaptées aux assemblages de haut poids moléculaire.

## IV. Approches structurales applicables à l'étude d'assemblages de haut poids moléculaire insolubles

### IV.1 Méthodes de caractérisation structurales classiques

#### IV.1.1 Diffraction aux rayons X

La diffraction aux rayons X, technique de référence en biologie structurale, a permis très tôt d'établir une description commune aux fibres amyloïdes. Ce profil est caractéristique qu'elles soient non orientées (halo à 4,3 Å et un cercle large moins intense à 9,8 Å) ou orientées (motif cross- $\beta$ <sup>15</sup>) (Eanes and Glenner. 1968). Ainsi, les fibres amyloïdes partagent des similarités de réflexion méridionale mais divergent quant aux réflexions équatoriales (espacement des feuillettes  $\beta$ ) (Sunde *et al.* 1997).

La caractérisation structurale des agrégats par diffraction aux rayons X a ainsi été réalisée, entre autres, sur les formes agrégées *in vitro* de A $\beta$ (1-40) (Malinchik *et al.* 1998) et des souches d' $\alpha$ -synucléine (Serpell *et al.* 2000 ; Bousset *et al.* 2013). Cependant, contrairement aux peptides de Tau agrégée, l'étude des fibres de Tau 2N4R entière n'a pas aboutie (Giannetti *et al.* 2000), probablement en raison de la trop grande taille de ces assemblages et du manteau déstructuré observable par MET.

La cristallisation suivie de la diffraction aux rayons X est aujourd'hui la méthode à l'origine du plus grand nombre de structures protéiques dans la Protein Data Bank. Cette technique permet la détermination à haute résolution de la conformation des protéines à l'échelle de l'atome avec une résolution moyenne de 2.17 Å en 2016 (Grimes *et al.* 2018). Comme son nom l'indique, cette technique nécessite une cristallisation de l'objet protéique et, malgré l'automatisation du protocole de cristallisation et la pluralité des tampons disponibles et testés, cette étape reste limitante. De plus, compte-tenu de l'effet de l'environnement (pH, sels, etc) sur la conformation des fibres, la cristallisation pourrait ne pas refléter la structure des fibres dans les conditions décrites. Bien que la taille des objets étudiés ait largement augmenté ces dernières années (taille moyenne de 100 kDa en 2016), cette méthode n'est toujours pas suffisamment résolutive pour les assemblages de protéines considérés dans le cadre de nos travaux. En effet, à ce jour, aucune structure d'agrégats de Tau entière à haute résolution n'est référencée, seules des structures d'hexapeptides de Tau agrégés sont disponibles (PDB

---

<sup>15</sup> Feuillettes  $\beta$  anti-parallèles disposés dans l'axe z de la fibre et des brins  $\beta$  de façon transverse à l'axe de la fibre

Accession Numbers : 3OVL, 3Q9G, 2ON9, 4NP8) (Landau *et al.* 2011 ; Liu *et al.* 2011 ; Sawaya *et al.* 2007 ; Wiltzius *et al.* 2009). De la même façon, seule la structure de petits peptides assemblés d' $\alpha$ -synucléine a été élucidée (PDB Accession Numbers : 4RIK, 4R0U, 4R0W) (Rodriguez *et al.* 2015 ; Li *et al.* 2014). Enfin, la PrP et le peptide amyloïde- $\beta$  agrégés n'ont également pas été décrits structuralement par cristallographie suivie d'une diffraction aux rayons X.

#### *IV.1.2 Spectroscopie infrarouge à transformée de Fourier (FTIR)*

Le FTIR (Fourier Transform Infrared spectroscopy) est une technique spectroscopique d'absorption vibrationnelle permettant de déterminer le contenu en structures secondaires d'une protéine. Ici, un détecteur mesure l'absorption de l'échantillon<sup>16</sup> d'un faisceau de lumière passé par un interféromètre (modulation des fréquences). On obtient alors un interférogramme qui sera converti en spectre infrarouge par transformée de Fourier. Ainsi, au niveau de la liaison peptidique, les vibrations de la liaison C=O conduisent à la bande Amide I sur le spectre tandis que les vibrations de la liaison N-H conduisent à la bande Amide II. Ces deux liaisons étant impliquées dans des liaisons hydrogène établissant les structures secondaires, la localisation de ces bandes sur le spectre est modulée par le contenu en structures secondaires. L'intensité de l'absorption à une longueur d'onde donnée est proportionnelle à la quantité de structures secondaires données (feuillets  $\beta$ , hélices  $\alpha$ , coudes, ou à l'inverse état déstructuré) (Gallagher. 2009 ; Hering *et al.* 2002).

Cette technique a permis de mettre en évidence une caractéristique majeure des agrégats amyloïdes : les feuillets  $\beta$  (Fink. 1998). De façon intéressante, il a aussi été mis en évidence par FTIR un riche contenu en hélices  $\alpha$  au sein des PHFs purifiés à partir de cerveaux de patients atteints de la maladie d'Alzheimer (Sadqi *et al.* 2002). Cette méthode a mis en lumière des différences conformationnelles entre des souches d' $\alpha$ -synucléine et entre des souches de PrP murine (Bousset *et al.* 2013 ; El Moustaine *et al.* 2008). Elle a même permis d'identifier des variations de structures  $\beta$  en fonction du type d'agrégats observé (i.g agrégats amorphes et fibrillaires) (Shivu *et al.* 2013). Il existe toutefois un problème majeur intrinsèque à cette technique : l'eau absorbant, on observe une saturation du signal pour les échantillons en solution aqueuse (Grdadolnik. 2002). Pour finir, cette technique est informative quant au contenu de structures secondaires mais ne renseigne pas sur la séquence primaire concernée.

---

<sup>16</sup> On note aujourd'hui une large utilisation du mode ATR (Attenuated Total Reflection) en FTIR.

### IV.1.3 Résonance Magnétique Nucléaire (RMN)

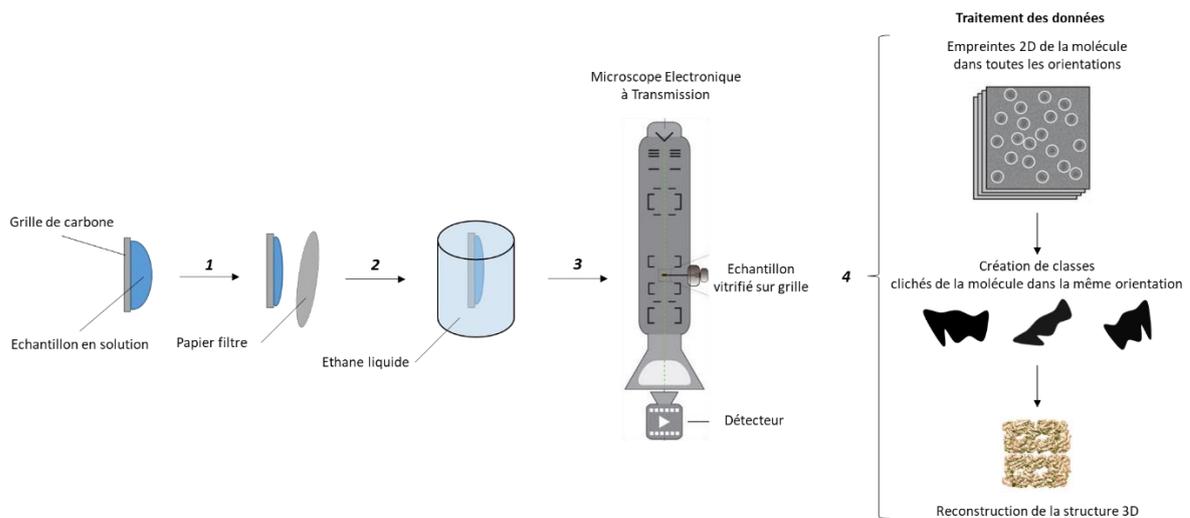
La RMN est une technique reposant sur la résonance magnétique des noyaux atomiques placés dans un champ magnétique et excités par un rayonnement radiofréquence. La fréquence de résonance va notamment refléter l'environnement proche de l'atome et en conditionner son déplacement chimique (RMN du proton).

La RMN du liquide est limitée par la taille des objets d'étude ; la haute concentration en protéines qu'elle nécessite rend compliquée l'étude du phénomène d'agrégation « naturelle » en temps réel. La facilitation de l'interprétation des spectres issus d'analyses RMN multidimensionnelles et le développement de la RMN à l'état solide permet aujourd'hui de résoudre des structures d'assemblages amyloïdes (Karamanos *et al.* 2015). En effet, plusieurs équipes ont notamment contribué à résoudre la structure de fibres de peptides A $\beta$ (1-40) (Bertini *et al.* 2011), d'A $\beta$ (1-42) (Xiao *et al.* 2015) et d' $\alpha$ -synucléine (Figure 24) par RMN du solide (Gath *et al.* 2014 ; Tuttle *et al.* 2016). A ce jour, la PDB ne recense aucune structure d'assemblages de Tau non tronquée élucidée dans sa totalité par RMN (Sillen *et al.* 2005). De nombreux polypeptides de Tau assemblés ont été étudiés par RMN et les structures de K18 et K19 agrégées, représentant un tiers de la protéine Tau entière, ont toutefois été résolues par RMN du solide (Andronesi *et al.* 2008 ; Daebel *et al.* 2012). Ainsi, la RMN du solide permet de mettre en évidence le contenu en structures secondaires, le long de la séquence protéique (contrairement à la diffraction aux rayons X et au FTIR) mais ne permet pas de disposer d'informations d'accessibilité de surface ou au contraire de protection des résidus au solvant.

### IV.2 Cryo-Microscopie

La cryo-microscopie électronique a connu une avancée technique fulgurante ces dernières années et a même conduit Richard Henderson, Jacques Dubochet et Joachim Franck à l'obtention du Prix Nobel de chimie en 2017 (« for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution », The Nobel Foundation). Le principe de la cryo-microscopie électronique repose sur (1) l'absorption en couche mince de l'échantillon aqueux sur une grille de carbone, (2) sa vitrification rapide dans une solution d'éthane liquide, (3) l'acquisition de clichés par microscopie électronique à transmission et (4)

le traitement des images 2D et la reconstruction de la structure 3D par informatique (**Figure 22**).



**Figure 22.** Principe de l'étude structurale par cryo-microscopie

Le développement, d'une part de la préparation des échantillons en amont de la microscopie, et, d'autre part de la détection et de la reconstruction 3D en aval, a permis à la cryo-microscopie électronique, de par sa haute résolution, de devenir une méthode de référence pour l'étude structurale des protéines mais aussi des assemblages macromoléculaires comme les agrégats protéiques. Ainsi, en 2008, la structure de fibres de peptide A $\beta$ 1-40 a mis en évidence une organisation en paires de feuillets  $\beta$  mais la résolution de 8 Å ne permettait pas d'identifier la séquence primaire (Sachse *et al.* 2008). Plus récemment, deux études réalisées cette fois sur des fibres de peptide A $\beta$ 1-42 ont atteint une résolution de 5 puis de 3 Å (Schmidt *et al.* 2015 ; Gremer *et al.* 2017), suffisante pour décrire l'organisation de la séquence primaire au sein de la structure. Des travaux similaires ont permis de décrire certaines souches d' $\alpha$ -synucléine fibrillaire à une résolution de respectivement 3.4, 3.7 et 3.4 Å et ont mis en évidence l'existence d'une hétérogénéité des espèces pour une même souche, ajoutant un degré de complexité à la diversité structurale (Guerrero-Ferreira *et al.* 2018 ; Li *et al.* 2018 ; Guerrero-Ferreira *et al.* 2019). Enfin, comme développé précédemment, les groupes de Michel Goedert et de Sjors Scheres à Cambridge sont à l'origine de plusieurs publications sur l'élucidation structurale par cryo-microscopie électronique d'assemblages de Tau issus de patients (Goedert *et al.* 2019 ; Falcon *et al.* 2019) et de fibres générées *in vitro* (Zhang *et al.* 2019). En dehors du coût très élevé de la technologie, la cryo-microscopie électronique présente certaines limites. En effet, une première contrainte de la technique est le très grand nombre de clichés à acquérir nécessaire à l'obtention d'une structure 3D de bonne résolution. Alors

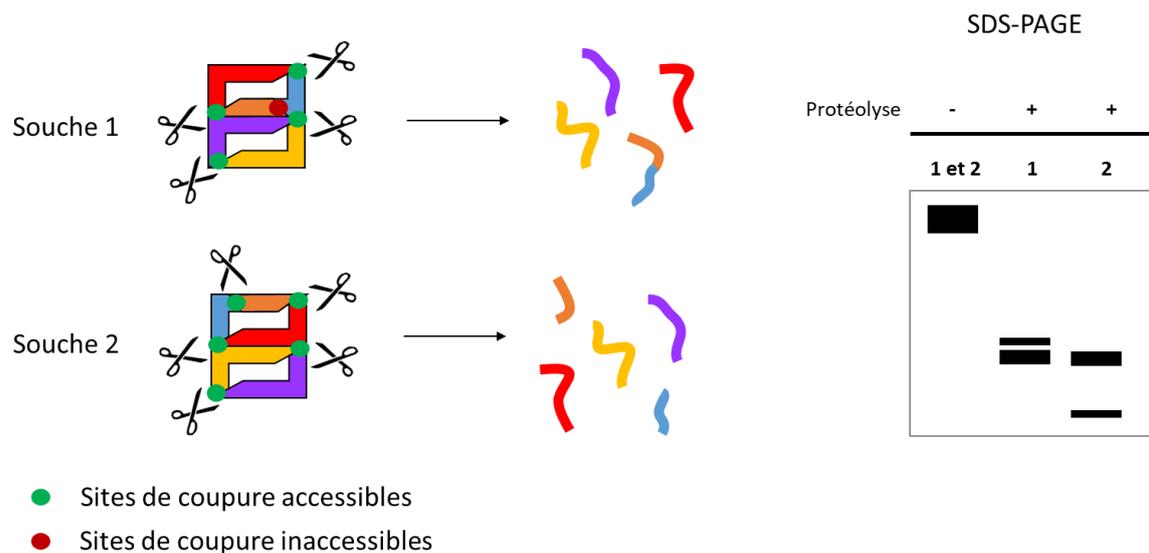
que les structures secondaires comme les feuillets  $\beta$  et les hélices  $\alpha$  sont visibles dès 5 Å, une résolution inférieure à 3,5 Å est nécessaire à l'attribution des chaînes latérales des acides aminés. Certaines grandes figures du domaine pointent un certain « optimisme » quant à l'attribution de cartes de densité et le lissage des données pour atteindre une résolution inférieure à 3 Å à l'origine de structures erronées (Baker. 2018). De plus, l'utilisation de logiciels de reconstruction 3D différents (i.e algorithmes distincts sur lesquels ils reposent) conduit à l'élucidation de structures distinctes pour un même échantillon (Editorial Nature Methods. 2018). Aussi, bien que moins contrainte d'un point de vue dynamique que la cristallographie associée à la diffraction aux rayons X (congélation des molécules dans toutes les orientations ; certains soulignent toutefois l'existence d'orientations privilégiées lors de l'absorption sur papier filtre), la cryo-microscopie électronique nécessite une certaine rigidité de structure permettant une résolution plus fine (Lyumkis. 2019). Ainsi, des états conformationnels moins peuplés d'un point de vue stoechiométrique pourraient échapper à l'œil de l'expérimentateur. Cet effet d'échantillonnage est abrogé pour les méthodes d'étude telles que les marquages covalents et non covalents analysés par spectrométrie de masse où les valeurs observées reflètent une moyenne d'évènements. Enfin, la cryo-microscopie électronique sur les agrégats issus de patients nécessite une succession d'étapes de purification qui pourraient moduler la structure initiale des agrégats et sélectionner certaines souches (extraction sarkosyl, gradient de sucrose). Pour conclure, la cryo-microscopie électronique permet de mettre en évidence le cœur ultra-structuré d'assemblages de haut poids moléculaire mais ne délivre aucune information sur les surfaces accessibles de ces agrégats.

### IV.3 Protéolyse ménagée

La détermination de régions structurées et de surfaces exposées au sein d'une protéine par protéolyse ménagée repose sur l'hypothèse selon laquelle le repliement d'une protéine peut induire (1) une inaccessibilité ou (2) une exposition accrue d'un site de clivage à l'enzyme dans des conditions limitantes (temps de digestion et ratio  $[E]/[S]$  contrôlés). Alors qu'une protéine déstructurée sera en théorie sujette à protéolyse au niveau de tous les sites de clivage possibles de l'enzyme ; après repliement, certains sites contenus dans des régions structurées, hélices  $\alpha$  par exemple, seront moins sensibles à la digestion que d'autres situés dans des régions restant déstructurées ou dans des boucles par exemple (Fontana *et al.* 2004). Les produits de protéolyse ménagée peuvent être observés par SDS-PAGE et identifiés selon différentes méthodes, qui seront développées plus loin. Ainsi, la protéolyse ménagée

peut permettre de suivre le repliement différentiel d'une même protéine en fonction de son environnement ou de la fixation d'un ligand (Zettlmeissl *et al.* 1983 ; Fontana *et al.* 1997).

Elle peut également participer à la compréhension des mécanismes d'agrégation, des différents conformères formés au cours du processus et de leurs propriétés d'accessibilité résidu-spécifiques distinctes (Polverino de Laureto *et al.* 2003). Enfin, elle peut permettre d'identifier des signatures conformationnelles en comparant la forme monomérique et fibrillaire d'une protéine prion-like (Bousset *et al.* 2004) mais également entre souches d'agrégats distincts (**Figure 23**) (Wemheuer *et al.* 2017).



**Figure 23.** Schématisation du principe de l'étude structurale de souches par protéolyse ménagée suivie par SDS-PAGE

La protéolyse ménagée a été utilisée à la fois sur les assemblages formés *in vitro* (Bousset *et al.* 2013) mais aussi sur des agrégats extraits de cerveaux de patients (Von Bergen *et al.* 2006). Elle peut reposer sur l'utilisation d'enzymes aspécifiques (e.g Pronase, Protéinase K) ou des enzymes spécifiques (e.g Trypsine) ; l'identification des fragments protéolytiques formés étant plus difficile dans le premier cas. L'optimisation des conditions (cinétique et ratio [E]/[S]) pour lesquelles sont mises en évidence des différences conformationnelles peut s'avérer longue.

La protéolyse ménagée a ainsi permis de caractériser un certain nombre d'agrégats, dont le peptide A $\beta$ (1-40) agrégé (Kheterpal *et al.* 2001), et a même participé à la meilleure compréhension du rôle des extensions polyQ sur l'agrégation des protéines associées (Scarff *et al.* 2015).

## IV.4 Méthodes associant la spectrométrie de masse

Dans cette partie, je présenterai, de manière non exhaustive, quelques approches associant la spectrométrie de masse et ayant déjà été utilisées pour l'étude structurale d'assemblages de haut poids moléculaire. En annexe 3, sont explicités le principe de la spectrométrie de masse et les différents instruments que j'ai utilisé au cours de mon travail de thèse.

### IV.4.1 Marquage non covalent : Echanges Hydrogène-Deutérium

L'élucidation des surfaces accessibles de protéines par échanges hydrogène-deutérium a été développée dès les années 1950 par Linderstrøm-Lang (Englander *et al.* 1997). Cette méthode repose sur les propriétés d'échange des atomes hydrogène de la protéine avec le solvant deutéré en présence (solvant en eau deutérée D<sub>2</sub>O). Les atomes d'hydrogène contenus dans les chaînes latérales au sein des groupements –OH, –SH, –NH<sub>2</sub>, –COOH et –CONH<sub>2</sub> échangent très rapidement avec le solvant, à une vitesse qui n'est pas mesurable par l'expérimentateur. A l'inverse, le taux d'échange des hydrogènes de la liaison peptidique est suffisamment lent pour être observé en temps réel et reflète l'environnement proche de chaque acide aminé et son organisation dans la structure tridimensionnelle de la protéine (Tsiatsiani *et al.* 2017). Cette technique permet, après digestion, d'identifier des peptides et de mesurer le nombre de deutérium incorporés au niveau des liaisons peptidiques. Ainsi, un hydrogène présent dans une région dite « protégée » (i.e structurée, enfouie, en contact avec un autre facteur) sera moins rapidement échangé. Cette technique permet ainsi d'observer des changements conformationnels dynamiques au cours du temps. Le principe de cette technique comprend les étapes suivantes : (1) Cinétique d'incubation de la protéine avec du D<sub>2</sub>O, (2) Arrêt de la réaction par diminution du pH ( $\approx 2.5$ ) et de la température ( $\approx 0^\circ\text{C}$ ), (3) Mesure de la masse entière de la protéine pour mesurer le taux d'échanges global, (4) Digestion à basse température et à pH acide (4) Analyse par spectrométrie de masse (incrément de masse) précédée d'une séparation des peptides par chromatographie liquide ou analyse par RMN.

Cette technique présente toutefois des difficultés techniques de mise en place. En effet, selon les conditions analytiques, le rétro-échange peut être parfois trop rapide. De plus, la digestion devant se faire à pH acide et à froid, la plupart des protéases ne peuvent être utilisées ; la pepsine restant l'enzyme de référence ou d'autres enzymes de même type ; la pepsine n'étant pas très spécifique (Tsiatsiani *et al.* 2017). La durée de la digestion et de l'éventuelle séparation chromatographique ultérieure doit être minimale pour éviter le rétro-échange. Ainsi, afin de réduire le temps de l'expérience, une automatisation du protocole, en chaîne, a été

développée (Hamuro *et al.* 2003). L'aspécificité de la pepsine, à l'origine d'une grande hétérogénéité de la population peptidique, conduit à de nombreuses espèces isobares qui compliquent considérablement l'analyse mais augmentent le recouvrement de séquence et la résolution de la localisation des échanges. Le mode de fragmentation le plus communément utilisé en spectrométrie de masse est le CID (dissociation induite par collision avec un gaz) mais n'est pas adapté pour localiser l'incorporation des deutérium sur une séquence peptidique. En effet, le CID induit un phénomène de transfert du deutérium aux liaisons peptidiques voisines (i.e « hydrogen scrambling ») au sein du peptide. Ainsi, associé à des paramétrages particuliers, le mode ETD (dissociation par transfert d'électrons) peut être utilisé pour l'étude des échanges hydrogène-deutérium mais la résolution à l'échelle de l'acide aminé reste techniquement difficile à obtenir (Wei *et al.* 2014).

Dès le début des années 2000, cette méthode a été utilisée pour l'étude structurale des agrégats (Kheterpal *et al.* 2006), notamment du peptide A $\beta$  (Wang *et al.* 2003), de la protéine prion HET-s du champignon filamenteux *Podospora anserina* (Nazabal *et al.* 2003) ou encore de la protéine prion de levure Ure2p (Redeker *et al.* 2007). Plus récemment, des études sur des fibres générées *in vitro* de Tau par échanges hydrogène-deutérium ont fait l'objet de publications (Ramachandran and Udgaonkar. 2013 ; Huang *et al.* 2018). Toutefois, la taille de ces assemblages multimériques, d'autant plus pour les polymères de protéines entières, complique la mise en place de cette approche (en particulier par analyse RMN). Enfin, une étape de dissociation des agrégats après l'arrêt de la réaction, est nécessaire à la mesure des échanges hydrogène deutérium par MS (Kheterpal *et al.* 2000). Or, il est difficile de réaliser une telle dissociation, à froid, dans un temps limité. A ces complications, s'ajoutent celles évoquées précédemment, intrinsèques à la technique.

#### IV.4.2 Marquage covalent

Comme discuté précédemment, l'étude structurale des assemblages par échanges hydrogène-deutérium reste techniquement difficile à mettre en place. Le caractère labile du marquage, à l'origine du phénomène de back-exchange, peut toutefois être contourné en utilisant des réactions de marquage covalent comme l'hydroxylation radicalaire (Hydroxyl Radical Footprinting (HRF)). Les radicaux hydroxyles ( $\bullet$ OH) peuvent être générés par différentes méthodes comme la chimie de Fenton, la radiolyse de l'eau par rayonnement synchrotron (Baud *et al.* 2017) ou encore par oxydation photochimique (FPOP pour Fast Photochemical Oxydation of proteins) (Cornwell *et al.* 2018). Ces radicaux hydroxyles peuvent réagir avec les chaînes latérales d'acides aminés. Le principe de la FPOP, par exemple,

repose sur les étapes suivantes : (1) Cinétique de marquage de la protéine par des radicaux hydroxyles  $\bullet\text{OH}$ , (2) Arrêt de la réaction par addition d'une solution de catalase et de méthionine amide, (3) Digestion enzymatique (4) Analyse des incréments de masse par spectrométrie de masse ou par RMN (Cornwell *et al.* 2018).

Ici, la durée de l'expérience, digestion et chromatographie comprises, n'a pas à être réduite et les conditions de digestion ne sont pas restrictives (i.e le choix de la protéase n'est pas contraint par le pH), ce qui peut permettre notamment d'obtenir un meilleur recouvrement de séquence. De plus, le mode de fragmentation CID peut être utilisé. La chimie radicalaire va cependant cibler en premier lieu essentiellement les acides aminés hydrophobes et ne sera donc pas particulièrement informative sur les régions hydrophiles de la protéine. De plus, la réactivité des chaînes latérales et l'effet de l'environnement proximal sur le potentiel de marquage d'un résidu donné est toujours peu connu. Enfin, l'oxydation peut induire un changement de conformation de la protéine ; la réaction de marquage par FPOP et son arrêt doivent donc absolument être contrôlés (Chea and Jones. 2018). Récemment, cette méthode a été appliquée à l'étude structurale d'agrégats protéiques. Elle a ainsi permis de mettre en évidence les régions et les acides aminés exposés pour des assemblages de peptides A $\beta$ 1-42 (Li *et al.* 2016), de prion PrP de souris (Li *et al.* 2018b) ou de superoxyde dismutase (Sheng *et al.* 2019). Cette méthode permet de mettre en évidence les résidus exposés au solvant.

D'autres types de marquage covalent permettent d'étudier l'accessibilité d'acides aminés au solvant. Le marquage par la NHS-Biotine permet ainsi le marquage préférentiel des lysines et du résidu N-terminal accessibles au sein de protéines monomériques (Gabant *et al.* 2008, Armengaud *et al.* 2005). Il a aussi permis de mettre en évidence l'impact sur l'accessibilité au solvant des résidus de l'hétérodimérisation des sous-unité PB1 et PA de l'ARN polymérase du virus Influenza (Dreger *et al.* 2009). Ce marquage n'a toutefois jamais été utilisé pour élucider les surfaces accessibles au sein d'agrégats protéiques. Un autre type de marquage, celui des acides aminés carboxylés (Glu, Asp, acide aminé C-terminal), est quant à lui réalisable grâce à la chimie GEE-EDC (Glycine Ethyl Ester - 1-Ethyl-3-(3-Diméthylaminopropyl)-Carbodiimide) (Zhang *et al.* 2012). Il permet de disposer d'informations d'accessibilité de résidus au solvant mais est techniquement difficile à mettre au point. Là encore, à ce jour, une telle approche n'a jamais été publiée dans le cadre d'une étude structurale d'assemblages insolubles de haut poids moléculaire. Ces deux derniers types de marquage pourraient être intéressants compte tenu de leur caractère non labile, et de la nécessité d'intégrer une étape de dissociation des agrégats en conditions dénaturantes en amont de la digestion et de l'analyse par spectrométrie de masse. Toutefois, contrairement aux échanges hydrogène-deutérium, ces marquages covalents consistent en l'addition d'un groupement de taille plus conséquente que celle induite par l'échange H/D et nécessiteraient un contrôle fin de la réaction (i.e quantité de marqueur et

temps de réaction) pour s'assurer qu'ils n'induisent pas de changements de structure des agrégats..

#### IV.4.3 Pontage covalent

Le principe du pontage covalent, aussi appelé réticulation chimique, consiste en la formation d'un pont covalent entre deux acides aminés proches ou en interaction non covalente grâce à un agent de pontage a minima bi-fonctionnel. Cet agent de pontage est constitué d'un bras défini par une longueur donnée qui contraint l'établissement d'un pontage entre deux acides aminés réactifs situés à une proximité adéquate à la longueur du bras (Leitner *et al.* 2010). Il peut être homobifonctionnel (liaison covalente entre deux groupements réactifs identiques) ou hétérobifonctionnel ou trifonctionnel (liaison covalente entre deux ou trois groupements réactifs distincts) (Kluger and Alagic.2004 ; Petrotchenko and Borchers. 2010). Cette approche est particulièrement employée pour mettre en évidence les régions d'interaction entre différentes sous-unités de complexes hétéromériques (Rappsilber. 2011) ou entre deux protéines partenaires (Redeker *et al.* 2012 ; Nury *et al.* 2015). Ainsi, par réticulation chimique, trois types de peptides peuvent être obtenus (Leitner *et al.* 2010 ; Paramelle *et al.* 2013) :

- Un pontage inter-peptide (type 2) traduisant une interaction intraprotéine ou interprotéine
- Un pontage intra-peptide ou pontage cyclique (type 1) traduisant une proximité intraprotéine
- Un pontage dead-end (type 0) traduisant une accessibilité des résidus au solvant

Ainsi, le pontage covalent associé à la spectrométrie de masse, de par les nombreuses informations que cette méthode peut délivrer, tant sur les interactions intra et inter protéiques que sur l'accessibilité de résidus au solvant, constitue une approche pertinente pour l'étude structurale de complexes protéiques.

L'étude structurale de complexes protéiques multimériques par pontage covalent reste difficile. En effet, d'une part, les pontages en conditions natives restreignent le nombre de marqueurs utilisables. D'autre part, ces approches nécessitent une longue mise au point du protocole expérimental, en particulier sur la nature et quantité d'agent pontant (Jones *et al.* 2019). De plus, concernant les assemblages homomériques, au premier abord, l'expérimentateur ne pourra pas faire la différence entre un pontage intramoléculaire et un pontage intermoléculaire. Toutefois, il est à noter qu'il existe aujourd'hui différentes stratégies pour répondre à cette question tels que l'expression et l'oligomérisation de protéines lourdes et légères en amont du

pontage ou encore des calculs de score différentiels (Lima *et al.* 2018 ; Gaber *et al.* 2019). D'autre part, l'expérimentateur peut rencontrer des difficultés lors de l'étape de digestion enzymatique des agrégats insolubles pontés. Ceci aurait pour conséquence une diminution du rendement protéolytique et un faible nombre de peptides pontés. Ainsi, cette approche nécessiterait une étape supplémentaire d'enrichissement des peptides pontés (Rey *et al.* 2018). Ce type d'approche a été employé sur des assemblages oligomériques d' $\alpha$ -synucléine (Ruesink *et al.* 2019), de peptide amyloïde  $\beta$  (Moore *et al.* 2009) ou de Tau (Patterson *et al.* 2011 ; Kim *et al.* 2015) pour la stabilisation de ces espèces et l'étude ultérieure de l'état d'oligomérisation, des propriétés physico-chimiques de ces espèces et de leur effet sur l'homéostasie cellulaire. Cependant, dans l'ensemble de ces études, les acides aminés impliqués dans l'établissement de ponts covalents n'ont pas été spécifiquement décrits. Le pontage covalent de monomères de Tau 2N4R a révélé des interactions intramoléculaires au sein des séquences R1 et R2 du MTBR (Mirbaha *et al.* 2018). Une expérience de pontage chimique sur des formes de Tau tronquées monomériques a aussi permis de mettre en évidence des interactions entre acides aminés de la partie N-terminale du MTBR, ceux de la partie C-terminale mais aussi entre des résidus N et C-terminaux du MTBR (Chen *et al.* 2019). Ainsi, les nombreux verrous expérimentaux liés à la mise en place de ce type d'approche sur des assemblages insolubles expliquent en partie qu'à ce jour aucune étude structurale d'assemblages insolubles de haut poids moléculaire (i.e fibres matures) par réticulation chimique n'ait encore été publiée.

Pour conclure, malgré la nature multimérique et insoluble des polymères protéiques, il existe différentes méthodes d'étude permettant d'obtenir des informations structurales sur les agrégats protéiques et leurs conformères, qu'ils soient produits *in vitro* ou *in vivo*. Chacune de ces approches nécessite bien sûr une optimisation antérieure à l'analyse de données, adaptée à la protéine étudiée et son environnement. Elles présentent toutes des avantages et des limitations qu'il est essentiel de mettre en perspective afin de ne pas sur-interpréter/mésinterpréter les résultats. Enfin, l'association de plusieurs méthodes de caractérisation des assemblages de haut poids moléculaire permet d'obtenir des informations robustes et complètes de structure. Récemment, par exemple, des travaux conséquents portant sur des monomères et des oligomères de la protéine PrP combinaient plusieurs approches de protéomique telles que la protéolyse ménagée, le marquage de surface covalent et non covalent et la réticulation qui ont permis l'identification robuste de déterminants structuraux (Serpa *et al.* 2013).

## II. Objectifs de la thèse

L'objectif principal de cette thèse, financée par la Fondation Vaincre Alzheimer, était de caractériser structurellement des polymorphes fibrillaires de Tau pour identifier des déterminants moléculaires de surfaces qui distinguent ces polymorphes de Tau dans le but de pouvoir proposer la conception de nouveaux outils moléculaires de diagnostic pour discriminer les tauopathies

Pour ce faire, nous avons purifié et assemblé Tau 1N3R et 1N4R recombinante humaine, correspondant aux deux isoformes les plus représentées dans le cerveau adulte, en fibres *in vitro*. Deux approches de protéomique structurale ont alors été développées dans le but d'identifier des signatures conformationnelles isoforme-spécifiques propres à chacun des deux types d'assemblage.

La première stratégie repose sur l'utilisation combinée de différentes protéases dans des expériences de protéolyses ménagées des fibres 1N3R et 1N4R. Ces « codes-barres » protéolytiques isoforme-spécifiques obtenus, identifiés par une combinaison de méthodes, mettent en lumière des résidus accessibles ou protégés spécifiques de la conformation des assemblages 1N3R et 1N4R. Une protéolyse avancée et la quantification des peptides générés ont permis de décrire de façon plus fine ces différences d'accessibilité à l'enzyme le long de la séquence.

La seconde stratégie développée consiste en un marquage covalent des lysines accessibles au solvant des fibres et équivalents monomériques par la Sulfo-NHS-Biotine. Nous avons ainsi optimisé les conditions de marquage, l'identification des résidus marqués et la quantification relative des peptides associés par spectrométrie de masse afin de dresser une cartographie des lysines accessibles et protégées.

Nos résultats révèlent que l'extrémité N-terminale est accessible pour les deux types de fibres mais contient des éléments structurés résistants à la protéolyse. A l'inverse, la région C-terminale est accessible pour Tau 1N4R et protégée pour Tau 1N3R, marquant une claire différence entre ces agrégats. Le marquage de surface des assemblages de Tau 1N3R et Tau 1N4R, à la résolution de l'acide aminé, a également mis en évidence des différences dans la région C-terminale, en aval des domaines de liaison au microtubules.

Ainsi, par deux approches originales, nous avons pu identifier des signatures conformationnelles isoforme-spécifiques des assemblages de Tau 1N3R et Tau 1N4R générés *in vitro*. A l'avenir, il serait intéressant d'appliquer ces mêmes méthodes sur des dépôts intracellulaires de Tau issus de cerveaux de patients atteints de différentes tauopathies. Les épitopes conformationnels spécifiques que nous avons identifiés pourront servir à terme pour concevoir de nouveaux outils moléculaires permettant de discriminer les différentes tauopathies.

### III. Résultats

#### III.1 Mise en place de méthodes dédiées à l'analyse des surfaces accessibles des fibres de Tau

##### A. Identification de surfaces accessibles et de régions protégées par protéolyse ménagée

Mon travail de thèse visait à mettre en évidence des déterminants structuraux des fibres de Tau 1N3R et de Tau 1N4R. La première approche reposait sur l'utilisation de protéolyses ménagées médiées par des enzymes spécifiques. Deux types d'informations structurales importantes peuvent être obtenues selon le temps de digestion. Les extrémités N et C-terminales des fragments générés dans les premiers temps de protéolyse correspondent à des sites accessibles au sein des fibres. A l'inverse, à des temps plus longs de protéolyse ménagée, les fragments résistants correspondent à des régions protégées. Après optimisation des conditions de digestion, l'identification de ces fragments protéolytiques a fait appel à un large spectre de méthodes qui seront explicitées plus en avant.

##### A.1 Présentation des enzymes disponibles et de leurs sites de clivage

Nous avons choisi d'utiliser des enzymes spécifiques qui présentent des sites de coupure théoriques répartis tout au long de la séquence de Tau. Ainsi, les sites de coupure potentiels de la trypsine, la chymotrypsine, l'ArgC, la LysC, la LysN, la GluC, l'AspN sur les séquences de Tau 1N3R et de Tau 1N4R sont présentés en **Figure 24**.

## Tau 1N3R

GluC/AspN (49)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIVYKPVLDL SKVTSKCGSLGNIHHPGGGQVEVKSEKLDKDRVQSKIGSLDNI 301 THVPGGKNKIIETHKLTFRENAKAKTDHGAEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV 361 DSPQLATLADEVASLAKQGL </pre>
Chymotrypsine (8)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIVYKPVLDL SKVTSKCGSLGNIHHPGGGQVEVKSEKLDKDRVQSKIGSLDNI 301 THVPGGKNKIIETHKLTFRENAKAKTDHGAEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV 361 DSPQLATLADEVASLAKQGL </pre>
Trypsine (52)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIVYKPVLDL SKVTSKCGSLGNIHHPGGGQVEVKSEKLDKDRVQSKIGSLDNI 301 THVPGGKNKIIETHKLTFRENAKAKTDHGAEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV 361 DSPQLATLADEVASLAKQGL </pre>
LysC/LysN (38)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIVYKPVLDL SKVTSKCGSLGNIHHPGGGQVEVKSEKLDKDRVQSKIGSLDNI 301 THVPGGKNKIIETHKLTFRENAKAKTDHGAEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV 361 DSPQLATLADEVASLAKQGL </pre>
ArgC (8)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIVYKPVLDL SKVTSKCGSLGNIHHPGGGQVEVKSEKLDKDRVQSKIGSLDNI 301 THVPGGKNKIIETHKLTFRENAKAKTDHGAEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV 361 DSPQLATLADEVASLAKQGL </pre>

## Tau 1N4R

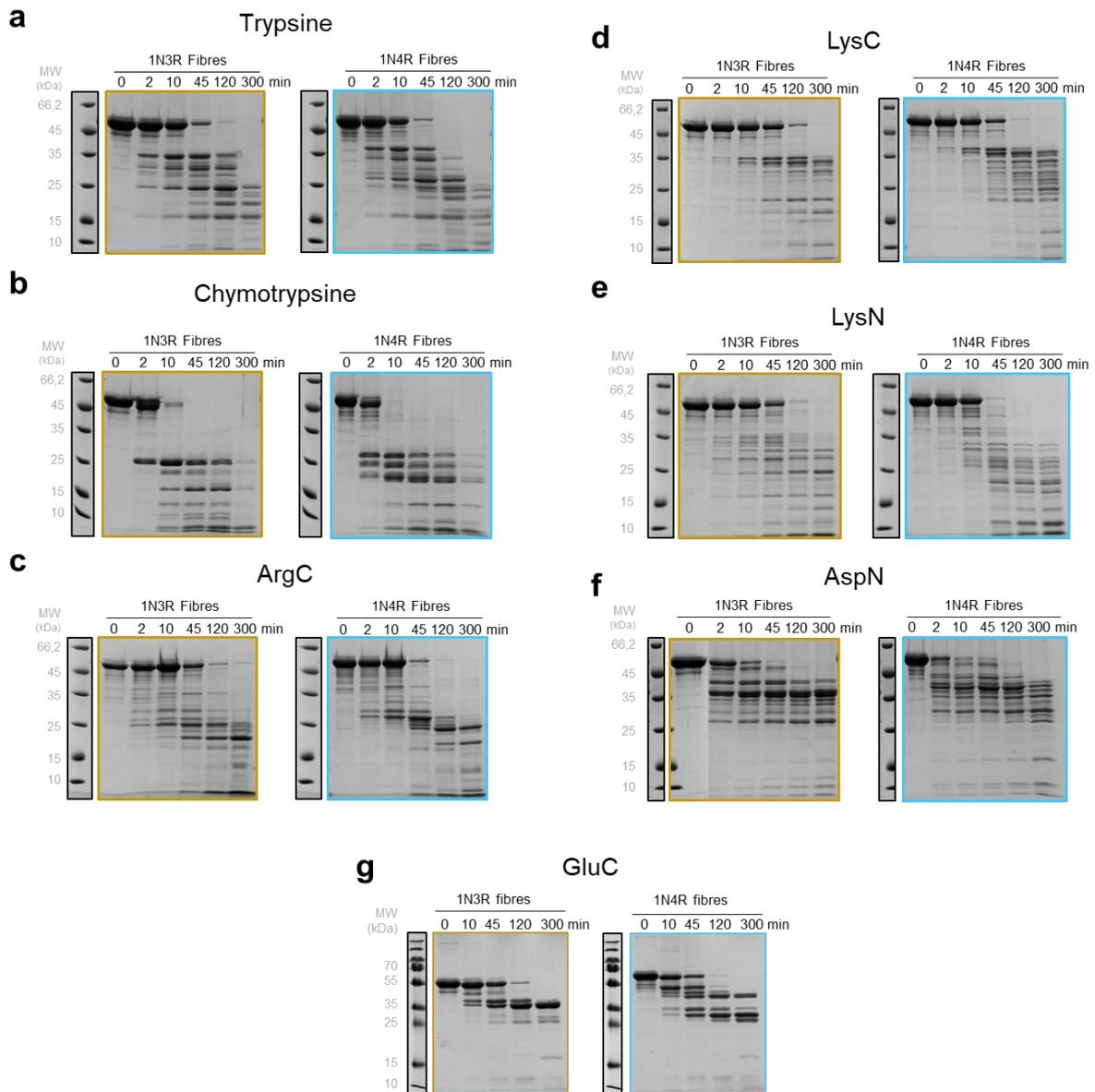
GluC/AspN (51)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIINKKLDL SNVQSKCGSKDNIKHVPGGGSVQIVYKPVLDL SKVTSKCGSLGNIH 301 HKPGGGQVEVKSEKLDKDRVQSKIGSLDNI THVPGGKNKIIETHKLTFRENAKAKTDHG 361 AEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV DSPQLATLADEVASLAKQGL </pre>
Chymotrypsine (8)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIINKKLDL SNVQSKCGSKDNIKHVPGGGSVQIVYKPVLDL SKVTSKCGSLGNIH 301 HKPGGGQVEVKSEKLDKDRVQSKIGSLDNI THVPGGKNKIIETHKLTFRENAKAKTDHG 361 AEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV DSPQLATLADEVASLAKQGL </pre>
Trypsine (57)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIINKKLDL SNVQSKCGSKDNIKHVPGGGSVQIVYKPVLDL SKVTSKCGSLGNIH 301 HKPGGGQVEVKSEKLDKDRVQSKIGSLDNI THVPGGKNKIIETHKLTFRENAKAKTDHG 361 AEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV DSPQLATLADEVASLAKQGL </pre>
LysC/LysN (43)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIINKKLDL SNVQSKCGSKDNIKHVPGGGSVQIVYKPVLDL SKVTSKCGSLGNIH 301 HKPGGGQVEVKSEKLDKDRVQSKIGSLDNI THVPGGKNKIIETHKLTFRENAKAKTDHG 361 AEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV DSPQLATLADEVASLAKQGL </pre>
ArgC (8)	<pre> 1 MAEPRQEFVEMEDHAGTYGLDRKDQGGYTMHQDECDTADAGLKEsplQTPTEdGSEEPG 61 SETSDAKSTPTAEAEAGIGDTPSLEDEAAGHVTQARMVSKDGTGSDDKKAKGADGKT 121 KIATPRGAAPPQKQGANATRIPAKTPPAKTPPSSGEPKSGDRSGYSSPGSPGTPGSR 181 SRTPSLPTPPTREPKKVAVVRTPPKSPSSAKSRLQAPVPMFDLKNVKSIGSTENLKHQ 241 PGGGKVQIINKKLDL SNVQSKCGSKDNIKHVPGGGSVQIVYKPVLDL SKVTSKCGSLGNIH 301 HKPGGGQVEVKSEKLDKDRVQSKIGSLDNI THVPGGKNKIIETHKLTFRENAKAKTDHG 361 AEIVYKSPVVGSDTSPRHLSNVSSSTGSDIMV DSPQLATLADEVASLAKQGL </pre>

**Figure 24.** Présentation des enzymes disponibles et de leurs sites de clivage

La chymotrypsine (coupure préférentiel en C-terminal des acides aminés aromatiques) et l'ArgC (coupure en C-terminal des arginines) se distinguent des autres enzymes par leur nombre réduit de résidus cibles (huit pour les deux isoformes). Ces enzymes devraient donc être à l'origine d'un nombre contraint de fragments protéolytiques. Les résidus acides, ciblés par la GluC et l'AspN, sont nombreux au sein de la séquence de Tau. Les lysines, substrats de la trypsine, la LysC, et la LysN sont également multiples.

## A.2 Choix des conditions de protéolyse ménagée

L'optimisation des rapports massiques (Enzyme/Tau) (ici désignés (E/S)p/p) nous a conduit à l'obtention de profils protéolytiques présentés en **Figure 25** pour l'ensemble des enzymes listées en **Figure 24**.



**Figure 25.** Cinétique des protéolyses ménagées suivie par SDS-PAGE de fibres 1N3R et fibres 1N4R par (a) la Trypsine (ratio 1/7200), (b) la Chymotrypsine (ratio 1/550), (c) l'ArgC (ratio 1/480), (d) la LysC (ratio 1/4800), (e) la LysN (ratio 1/4800), (f) l'AspN (ratio 1/120) et (g) la GluC (ratio 1/150). Les ratios sont ici exprimés en rapport massique (Enzyme/Tau).

Les profils protéolytiques des fibres de Tau 1N3R et 1N4R sont distincts quelle que soit l'enzyme utilisée, suggérant des différences structurales.

Les profils protéolytiques obtenus à partir d'une digestion à la Trypsine (**Figure 25.a**), de l'ArgC (**Figure 25.c**), de la LysC (**Figure 25.d**) et de la LysN (**Figure 25.e**) comprennent de nombreuses espèces sur gel et rendent l'analyse différentielle très compliquée. Cette dernière ne permettra pas de mettre en évidence quelques fragments protéolytiques caractéristiques, marquant une différence claire entre les isoformes 1N3R et 1N4R. Comme on peut le constater en **Figure 25.c**, contrairement à nos attentes, l'ArgC est à l'origine de nombreuses espèces sur gel à un ratio massique (E/S) de 1/500 et ce dès 2 minutes de protéolyse. Une analyse de peptides générés par protéolyse à l'ArgC à des temps plus longs a permis de mettre en évidence une activité « aspécifique » de l'enzyme en C-terminal des lysines, en accord avec des observations publiées (Golghalyani et al. J. Proteome Res. 2017). L'AspN génère également de nombreux fragments à partir des fibres de Tau 1N4R après 5 heures de protéolyse (**Figure 25.f**). Au ratio (E/S)<sub>p/p</sub> 1/120, avec l'AspN, on observe une disparition de la forme entière à 120 minutes de protéolyse et une bande majoritaire d'environ 35 kDa après 300 minutes, identique à celle observée pour la protéolyse ménagée à la GluC (**Figure 25.g**).

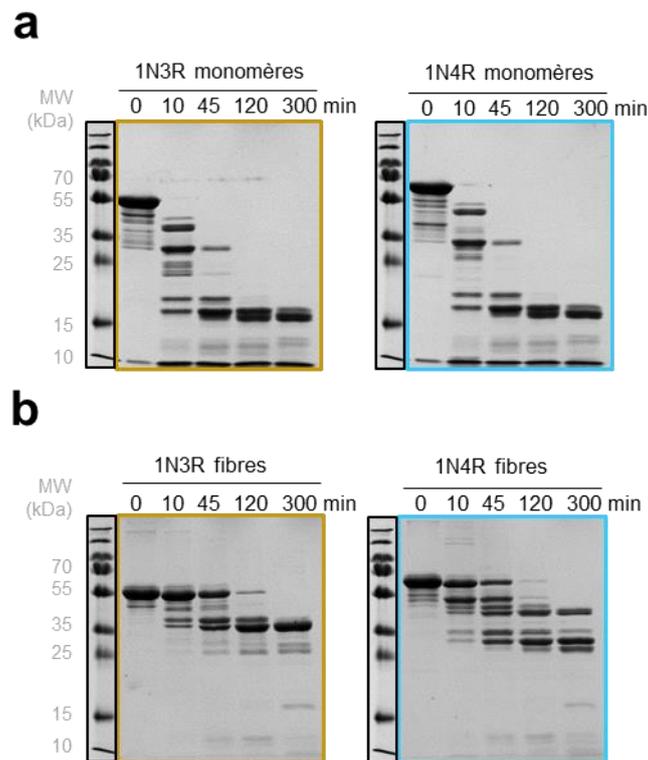
Les premiers temps de protéolyse ménagée par la chymotrypsine (**Figure 25.b**), et en particulier à 2 minutes, semblent révéler des différences majeures entre les deux isoformes. Le contenu en résidus aromatiques au sein de la séquence de Tau étant limité et identique entre 1N3R et 1N4R, ceci est donc directement imputable à un arrangement conformationnel distinct.

La protéolyse ménagée par la GluC (**Figure 25.g**), dont les sites potentiels de coupure sont nombreux au sein de la séquence de Tau, s'est révélée intéressante. En effet, de façon surprenante, même après 5 heures de digestion, on ne constate que la présence de bandes de forte intensité et de haut poids moléculaire (supérieur à 25 kDa). Ceci semble indiquer l'existence d'une large région résistante à la digestion enzymatique au sein des fibres. De plus, les profils sur gel de la protéolyse des fibres 1N3R et des fibres 1N4R sont nettement différents, suggérant une conformation distincte isoforme-spécifique des assemblages.

Ainsi, les protéolyses ménagées par la GluC et par la Chymotrypsine ont été sélectionnées pour mettre en évidence des signatures conformationnelles isoforme-spécifiques.

Comme expliqué ci-dessus, les différences observées avec la chymotrypsine ne peuvent résulter de différences de séquence entre 1N3R et 1N4R. Ce n'est pas le cas de la GluC. En effet, on compte deux sites de coupure additionnels dus à la présence de la deuxième répétition dans l'isoforme 1N4R par rapport à 1N3R. Nous avons donc cherché à vérifier que ces différences majeures de prise en charge par la GluC étaient bien causées par l'arrangement des molécules de Tau au sein des fibres et non par ces divergences de séquence.

Ainsi, une protéolyse ménagée par la GluC au même ratio massique (i.e (E/S)<sub>p/p</sub> = 1/150) a été réalisée en parallèle sur les fibres et les formes monomériques de Tau 1N3R et 1N4R. Les profils protéolytiques observés après SDS-PAGE sont présentés en **Figure 26.a**, monomères et **b**, fibres).



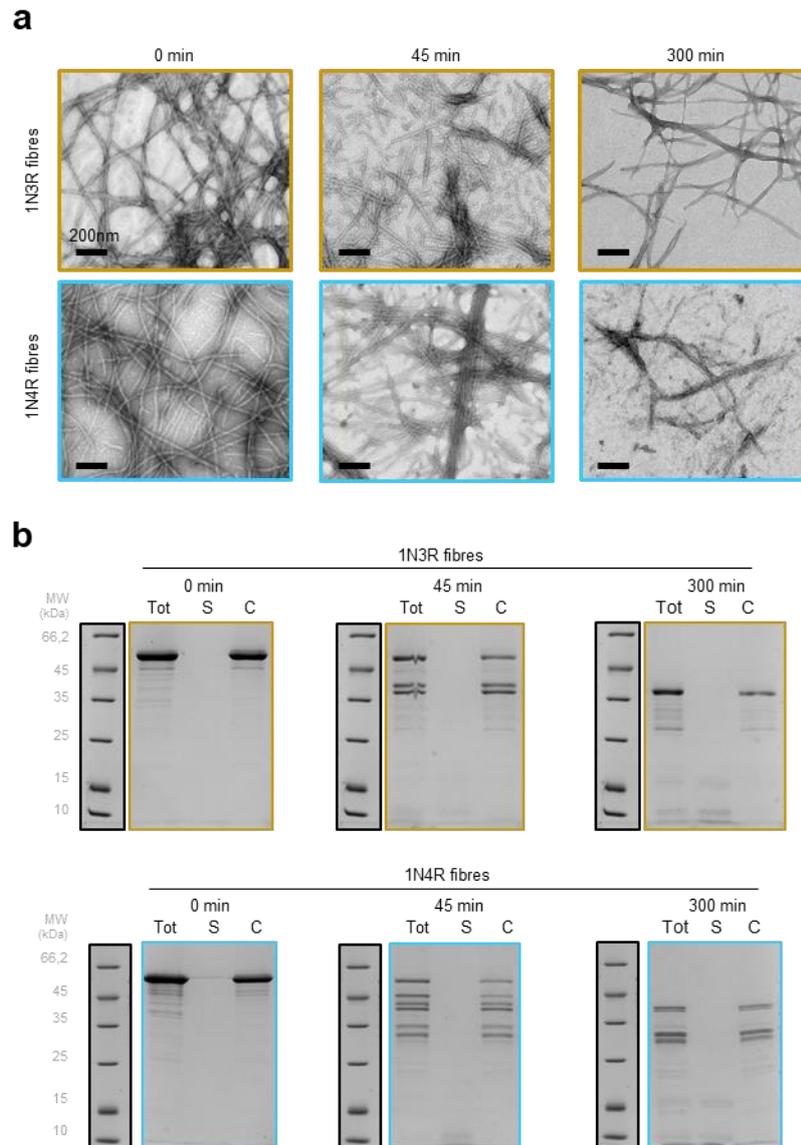
**Figure 26.** Cinétique des protéolyses ménagées suivie par SDS-PAGE des monomères et des fibres de Tau 1N3R et 1N4R.

(a) Cinétique des protéolyses ménagées par la GluC suivie par SDS-PAGE de monomères 1N3R et 1N4R. (b) Cinétique des protéolyses ménagées par la GluC suivie par SDS-PAGE de fibres 1N3R et 1N4R

On peut constater que les profils de protéolyse ménagée par la GluC des monomères de Tau 1N3R et de Tau 1N4R sont très semblables. La forme entière de la protéine disparaît entre 0 et 10 minutes de protéolyse. La seule distinction entre Tau 1N3R et 1N4R monomérique observable sur gel est la présence à T 10 minutes d'un doublet de bandes d'environ 23 kDa plus intense pour 1N3R que pour 1N4R, espèces protéiques qui ne sont plus observables à 45 minutes de protéolyse. Ainsi, les différences majeures de profils protéolytiques GluC entre les fibres 1N3R et 1N4R sont bien conformationnelles spécifiques et non dues directement à la séquence en acides aminés.

Concernant les fibres, nous avons exploré la nature, soluble ou agrégée, des fragments protéolytiques. Pour cela, nous avons suivi la protéolyse ménagée GluC des fibres de Tau

1N3R et 1N4R par d'autres techniques : la microscopie électronique à transmission et la sédimentation. Les résultats sont présentés en **Figure 27**.



**Figure 27.** Les fragments protéolytiques de haut poids moléculaire générés par la GluC sont fibrillaires.

**(a)** Clichés de MET en coloration négative du mélange réactionnel de la protéolyse des fibres de Tau 1N3R (haut) et de Tau 1N4R (bas) par la GluC à 0, 45, et 300 minutes (de gauche à droite).

**(b)** Sédimentation des produits de protéolyse ménagée des fibres 1N3R (haut) et des fibres 1N4R (bas) par la GluC à 10, 45 et 300 minutes suivie par SDS-PAGE. Tot = Fraction totale, S = Surnageant, C = Culot.

Ainsi, par MET, nous pouvons constater qu'après cinq heures de protéolyse ménagée à la GluC, des fibres sont toujours présentes dans le mélange réactionnel et ce, tant pour Tau 1N3R, que pour Tau 1N4R. De plus, l'analyse du culot par SDS-PAGE après sédimentation des fibres protéolysées témoigne de la nature fibrillaire des principaux fragments protéolytiques. L'addition d'AEBSF (4-(2-aminoethyl)benzenesulfonyl fluoride

hydrochloride)<sup>17</sup>, à la concentration finale élevée de 10mM, ne parvient pas à arrêter complètement l'activité enzymatique de la GluC, ce qui explique la diminution d'intensité des bandes après sédimentation. Toutefois, l'ensemble de ces expériences montre que les fragments protéolytiques résultant de la protéolyse ménagée à la GluC sont de nature fibrillaire. Ceci démontre que les fibres présentent une conformation résistante à la protéolyse.

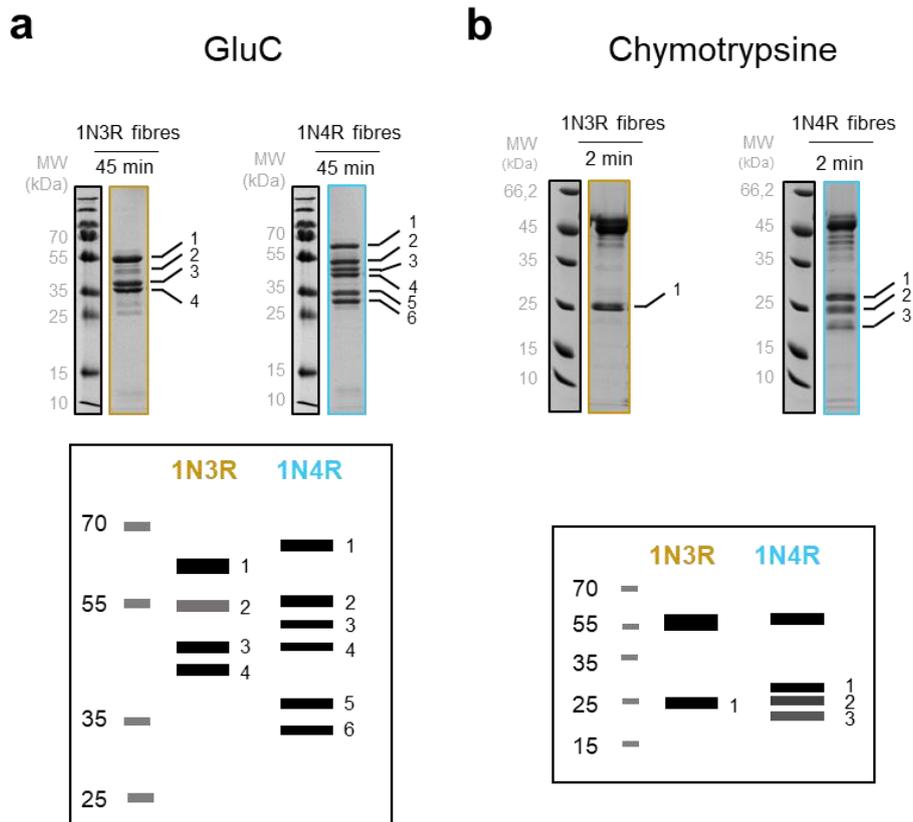
Pour conclure, un large spectre d'enzymes a été utilisé pour mettre en lumière des signatures conformationnelles isoforme-spécifiques. Parmi elles, les protéolyses ménagées par la GluC et par la Chymotrypsine ont révélé des différences majeures entre les fibres 1N3R et 1N4R. Ces fragments protéolytiques fibrillaires ont par la suite été identifiés grâce à une combinaison d'approches.

### A.3 Identification des fragments protéolytiques de haut poids moléculaire

Les conditions de protéolyse ménagée que nous avons optimisées ont permis de mettre en évidence des fragments protéolytiques GluC (**Figure 25.g**) et chymotryptiques (**Figure 25.b**) spécifiques des assemblages de Tau 1N3R et Tau 1N4R. Ces profils protéolytiques spécifiques de Tau 1N3R et 1N4R constituent des code-barres moléculaires (**Figure 28**). La séquence de ces fragments protéolytiques de haut poids moléculaire séparés sur gel SDS-PAGE a été identifiée par une combinaison de méthodes dont les apports et les limitations seront discutés dans cette partie.

---

<sup>17</sup> Inhibiteur irréversible de protéases à sérine



**Figure 28.** Codes-barres moléculaires des fibres de Tau 1N3R et 1N4R protéolysées.

(a) Identification du code barre moléculaire conformationnel-spécifique des fibres de Tau 1N3R et 1N4R à 45 minutes de protéolyse ménagée par la GluC. (b) Identification du code barre moléculaire conformationnel-spécifique des fibres de Tau 1N3R et 1N4R à 2 minutes de protéolyse ménagée par la Chymotrypsine.

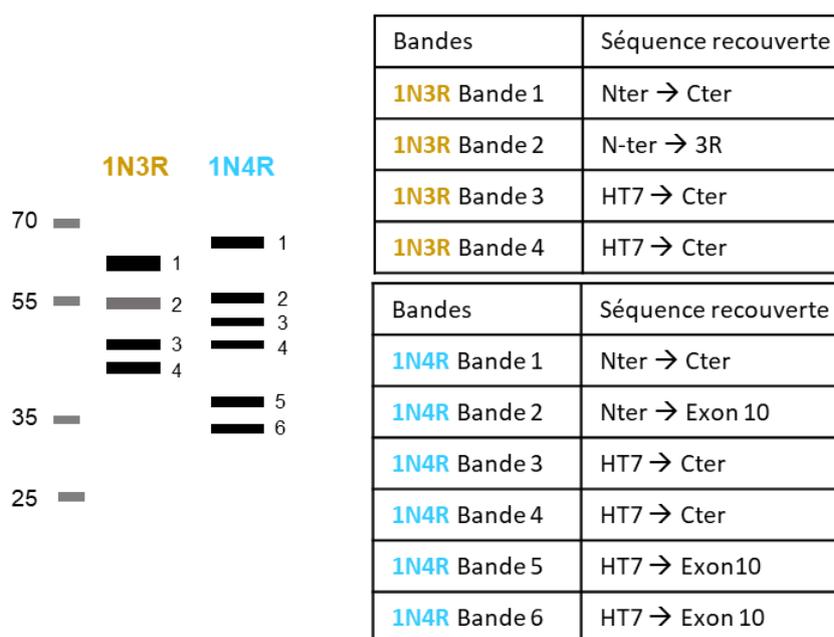
### A.3.1. Analyse par immunodétection (Western-Blots) (crédit : Karine Madiona)

En première intention, l'analyse par western-blot est une technique de choix pour l'identification de fragments protéiques issus d'expériences de protéolyse ménagée (Mazzoni *et al.* 2009). Cette technique est basée sur immunodétection après transfert de protéines séparées sur SDS-PAGE. Pour ce faire, nous avons utilisé les anticorps présentés en **Table 5**.

Anticorps	Séquence	Isoforme reconnue	Références
Nter	1-19	1N3R et 1N4R	1
HT7	131-135	1N3R et 1N4R	2
Tau1	160-178	1N3R et 1N4R	3
Tau5	182-202	1N3R et 1N4R	4
3R	239-249	1N3R	5
Exon10	246-262	1N4R	6
Cter	398-412	1N3R et 1N4R	1

**Table 5.** Présentation des différents anticorps utilisés au cours de notre étude  
 Références : (1) Equipe Alzheimer et Tauopathies dirigée par Luc Buée, INSERM, Lille, France, (2) MN1000 Thermo Fisher, Waltham, Etats-Unis, (4) AHB0042 Thermo Fisher, Waltham, Etats-Unis (5) 05-803 Merck Millipore, Darmstadt, Germany (6) 05-804 Merck Millipore, Darmstadt, Germany

Ainsi, par cette première approche, nous avons pu déterminer des bornes régionales des fragments protéolytiques GluC (**Figure 29**).



**Figure 29.** Identification des fragments issus de la protéolyse ménagée à la GluC des fibres de Tau 1N3R et de Tau 1N4R par Western Blot. Les deux anticorps les plus éloignés sur la séquence, définissant les bornes minimales des fragments protéolytiques.

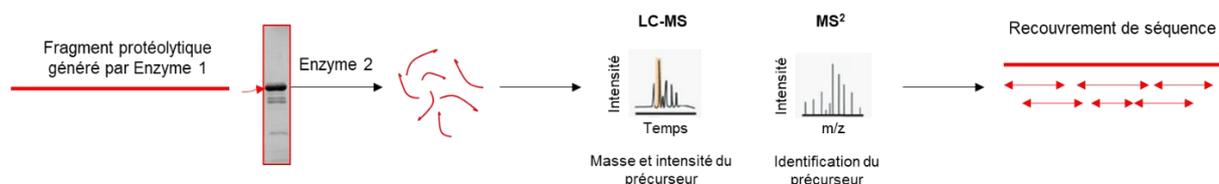
L'absence de reconnaissance des bandes 3 et 4 pour 1N3R et 3, 4 5 et 6 pour 1N4R par l'anticorps NTer polyclonal rend compte de l'accessibilité de la partie N-terminale de la protéine agrégée pour les deux isoformes. Les bandes intenses 2 et 6 spécifiques de 1N4R (la bande 2 est retrouvée à l'état de traces pour 1N3R) illustrent quant à elles une nette différence

d'accessibilité de la partie C-terminale au sein des fibres. En effet, la région C-terminale de Tau 1N4R agrégée semble accessible à l'enzyme tandis qu'elle semble protégée pour les fibres 1N3R.

Ainsi, cette première approche d'identification des fragments protéolytiques signatures issus de la digestion par la GluC est informative sur la région à l'origine des différences sur gel mais ne permet en aucun cas l'attribution d'une séquence particulière aux espèces protéiques en présence. Pour compléter nos résultats, nous avons donc eu recours à d'autres méthodes.

### A.3.2 Analyse du recouvrement de séquence des fragments par une approche protéomique Bottom-up

En spectrométrie de masse, l'approche la plus courante d'identification d'une protéine et de la caractérisation de son recouvrement de séquence, dite « Bottom-up », repose sur la digestion totale en gel d'une bande considérée. Une fois extraite, la population protéolytique générée à partir de la bande protéique sera analysée par séparation chromatographique couplée à la spectrométrie de masse de haute résolution, LC-MS/MS. L'analyse des données de masse et de fragmentation nous permettra, par traitement informatique, d'identifier les peptides protéolytiques et ainsi de reconstruire un recouvrement de séquence de la bande d'intérêt (**Figure 30**).

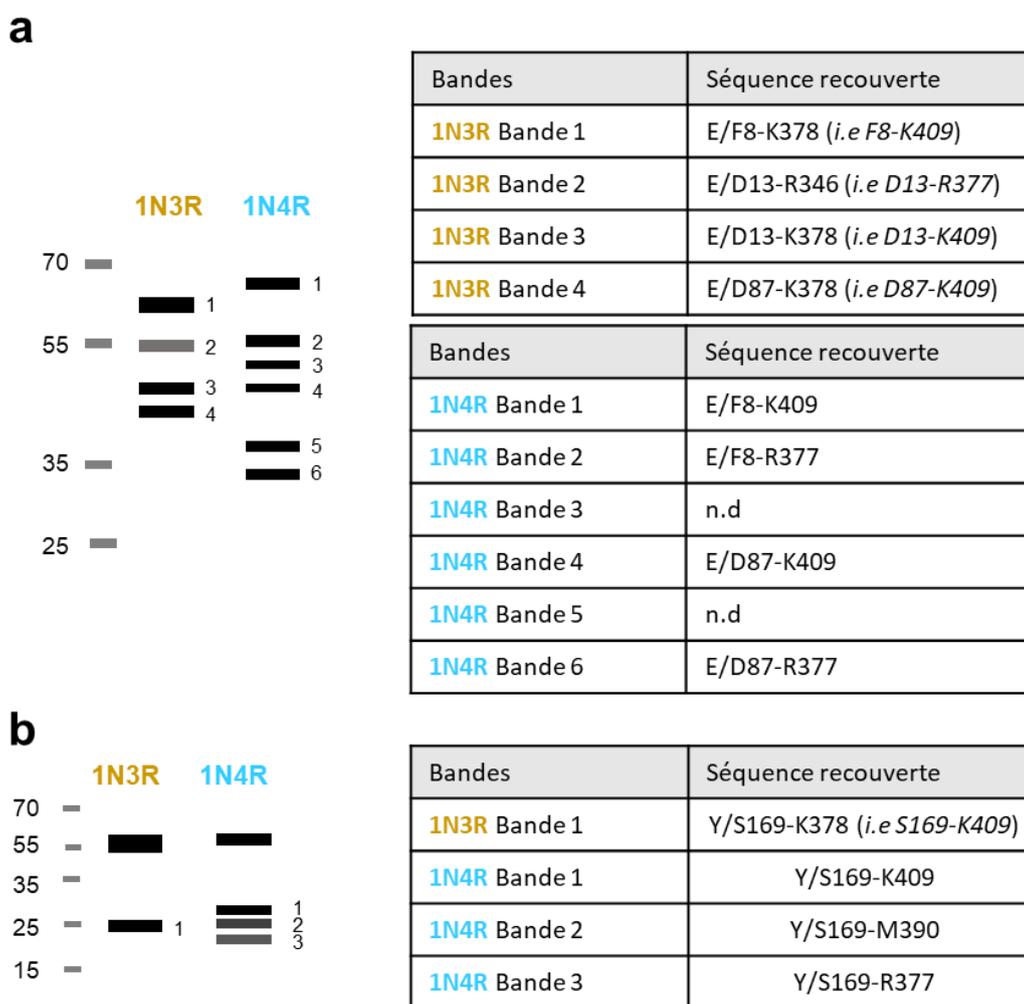


**Figure 30.** Stratégie d'identification des fragments protéolytiques par spectrométrie de masse et approche Bottom-up.

Nous avons donc tenté d'identifier les fragments protéolytiques par cette approche.

Dans le but d'obtenir le meilleur recouvrement de séquence des fragments protéolytiques possible, nous avons évalué le recouvrement de séquence obtenu sur la protéine entière après digestion en gel par différentes protéases et analyse LC-MS/MS. Dans le cas présent, il serait nécessaire d'utiliser une enzyme différente de celle utilisée pour l'expérience de protéolyse ménagée.

La trypsine permet d'obtenir le meilleur recouvrement de séquence à la fois pour Tau 1N3R (94%) et pour Tau 1N4R (81%) comme présentés en **Annexe 4**. Cette enzyme a donc été retenue pour la protéolyse systématique des fragments issus des protéolyses ménagées par la GluC et par la Chymotrypsine. Les résultats de cette approche menée sur l'analyse des fragments de protéolyses ménagées constituant les codes-barres moléculaires sont décrits en **Figure 31**.



**Figure 31.** Recouvrement de séquence des fragments générés par protéolyse ménagée de Tau 1N3R et 1N4R par la GluC (a) et par la Chymotrypsine (b) après 45 minutes et 2 minutes de digestion respectivement. Ce recouvrement de séquence a été établi après digestion tryptique et recherches des peptides semi-tryptiques.

Le premier point sur lequel nous nous devons d'insister est le fait que le recouvrement de séquence obtenu sur la forme entière de Tau par la digestion tryptique n'est pas de 100% (**Annexe 4**) du fait de l'absence du peptide N-terminal pour les deux isoformes et du peptide C-terminal seulement pour 1N4R. En effet, d'une part, le peptide N-terminal 2-AEPR-5 (Mth 471,24) est de trop petite taille pour être détecté sous son état bichargé dans les conditions

analytiques utilisées et semble être élué sous sa forme monochargé (MH+ 472,25) en tout début de gradient. D'autre part, pour le long peptide tryptique C-terminal 347-HLSNVSTGSIDMVDSPQLATLADEVASLAK-378 pour 1N3R (i.e 378-409 pour 1N4R) (Mth 3258.59 sous sa forme oxydée) on observe des différences de rendement de digestion et/ou d'extraction du gel. Ce constat constitue une première limite à cette approche.

Certaines incohérences résultent de l'analyse des recouvrements de séquence des fragments protéolytiques. En effet, des bandes de plus haut poids moléculaire ont un recouvrement de séquence plus restreint que des bandes de plus bas poids moléculaires (i.e Protéolyse ménagée GluC 1N3R Bande 2 par rapport à 1N3R Bande 3). Ceci pourrait s'expliquer par des différences d'abondance des fragments protéolytiques observés sur gel ou par l'existence au sein d'une même bande de plusieurs fragments protéolytiques dont les séquences conjointes recouvrent une plus large région que celle attendue d'après le poids moléculaire apparent observé sur gel. Cependant, au vu des données de western-blot, cette seconde hypothèse semble peu probable. De plus, il a déjà été montré que la diffusion des bandes et la capacité d'ionisation différentielle des peptides (certains peptides contaminants même à l'état de traces peuvent parfaitement ressortir de l'analyse de part leur bonne efficacité d'ionisation) pouvaient conduire à des erreurs d'interprétation de recouvrement de séquence. De plus, toutes les extrémités N-terminales identifiées pour les fragments chymotryptiques semblent identiques (i.e S169).

Dans l'état, nos données ne nous permettent donc pas de déterminer avec confiance la séquence constituant les différents fragments protéolytiques. Pour ce faire, nous avons donc décidé d'identifier de manière ciblée les résidus N- et C-terminaux des différents fragments.

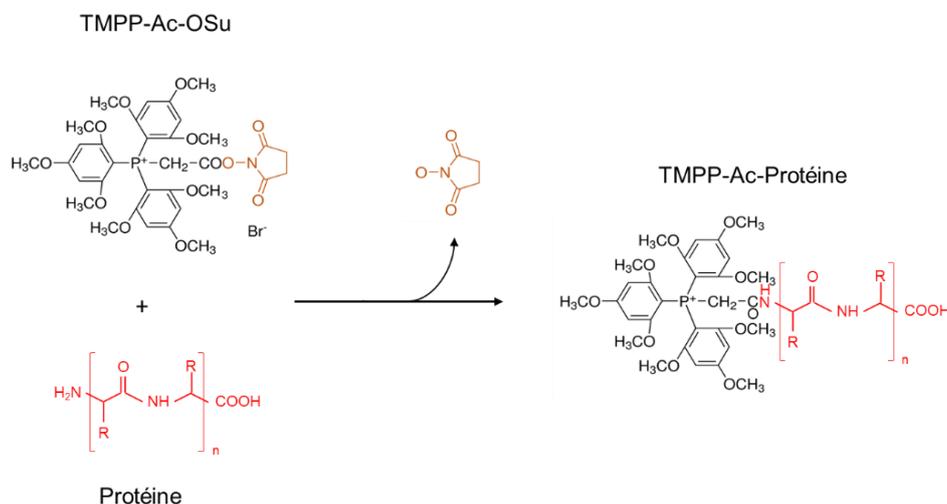
### *A.3.3 Identification du résidu N-terminal des fragments protéolytiques*

a. Par dérivatisation du N-terminal au TMPP et analyse par spectrométrie de masse

#### *a.1 Mise au point du marquage covalent spécifique du N-terminal par le TMPP*

Le marquage du résidu N-terminal par le TMPP-Ac-OSu ((N-succinimidyl)oxycarbonylmethyl)-tris(2,4,6-triméthoxyphényl)-phosphonium bromide, que nous appellerons ici TMPP), un dérivé de phosphine, repose sur une substitution nucléophile engageant une liaison covalente entre le groupement amine du N-terminus et le TMPP-Ac à un pH 8,2 (**Figure 32**) (Gallien et

al. 2009). En effet, le pKa de ce groupement est en moyenne de 8<sup>18</sup>, une valeur inférieure à celui de l'amine de la lysine (i.e 10.5). Ainsi, en se plaçant à un pH réactionnel de 8.2, l'expérimentateur favorise le marquage de l'amine terminale à défaut de celui des chaînes latérales des lysines.



**Figure 32.** Réaction de dérivatisation du groupement amine primaire spécifique du résidu N-terminal de la protéine par le TMPP à pH 8,2.

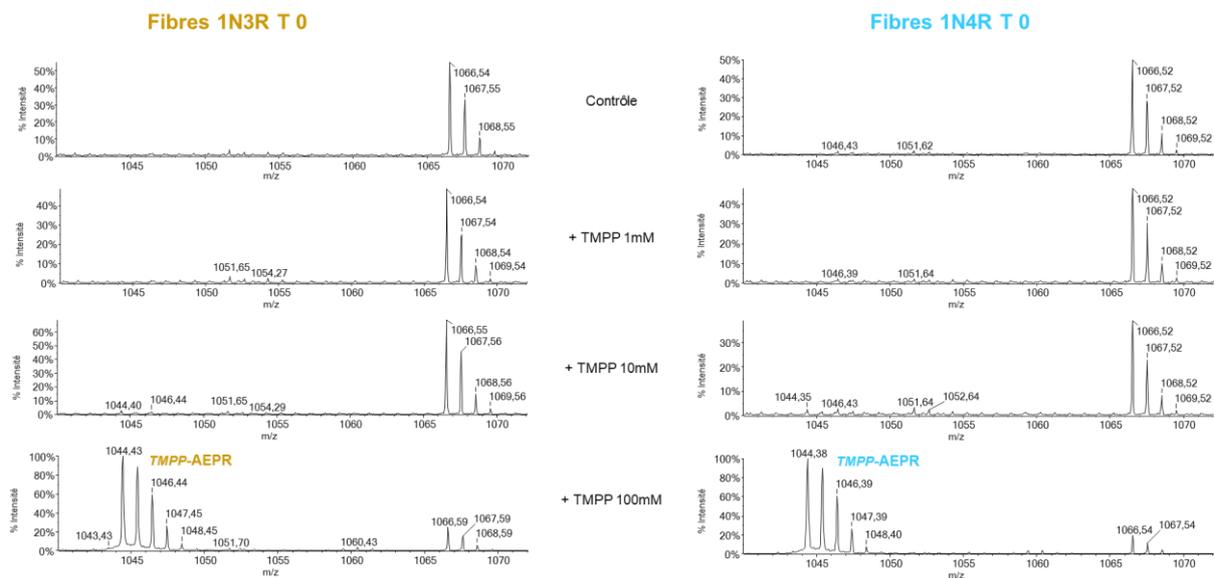
Plusieurs équipes ont travaillé à la mise au point de cette méthode de dérivatisation pour l'identification du N-terminal de protéines (Gallien *et al.* 2009 ; Bertaccini *et al.* 2013 ; Ayoub *et al.* 2015). Ce réactif permet d'améliorer l'efficacité d'ionisation du peptide N-terminal modifié grâce à la charge permanente portée par la phosphine. Il permet également d'augmenter le degré d'hydrophobicité des peptides marqués, à l'origine d'un retard d'élution et d'une séparation chromatographique par rapport au reste de la population peptidique, facilitant leur identification. Enfin, la dérivatisation est associée à un incrément de masse de +572.18 Da, distinguant le peptide marqué de son homologue non modifié.

Le principe de la technique repose sur l'excision d'une bande d'intérêt d'un gel SDS-PAGE (pourcentage d'acrylamide 12%) qui sera dans un premier temps lavée avant que les cystéines ne soient réduites puis alkylées. Ensuite, l'échantillon sera incubé pendant une heure dans la solution de marquage (TMPP repris dans un mélange d'eau et d'acétonitrile 80/20 %) additionnée de tampon Tris-HCl à pH vérifié 8,2 en excès. La réaction sera ensuite arrêtée par addition d'hydroxylamine en large excès et incubation d'une demi-heure. Afin d'éviter au mieux les réactions aspécifiques du TMPP avec d'autres acides aminés, le gel est soumis à quatre cycles de lavage en aval de la réaction de marquage, avant d'être soumis à digestion

<sup>18</sup> Attention, il faut tout de même considérer que cette valeur peut varier de 6.8 à 9.1 (Grimsley *et al.* 2009)

enzymatique. Les conditions de marquage et d'arrêt de la réaction reprennent celles publiées par Daniel Ayoub (Ayoub *et al.* 2012) et ont été optimisées pour la protéine Tau. Le mélange peptidique résultant est ensuite analysé par spectrométrie de masse.

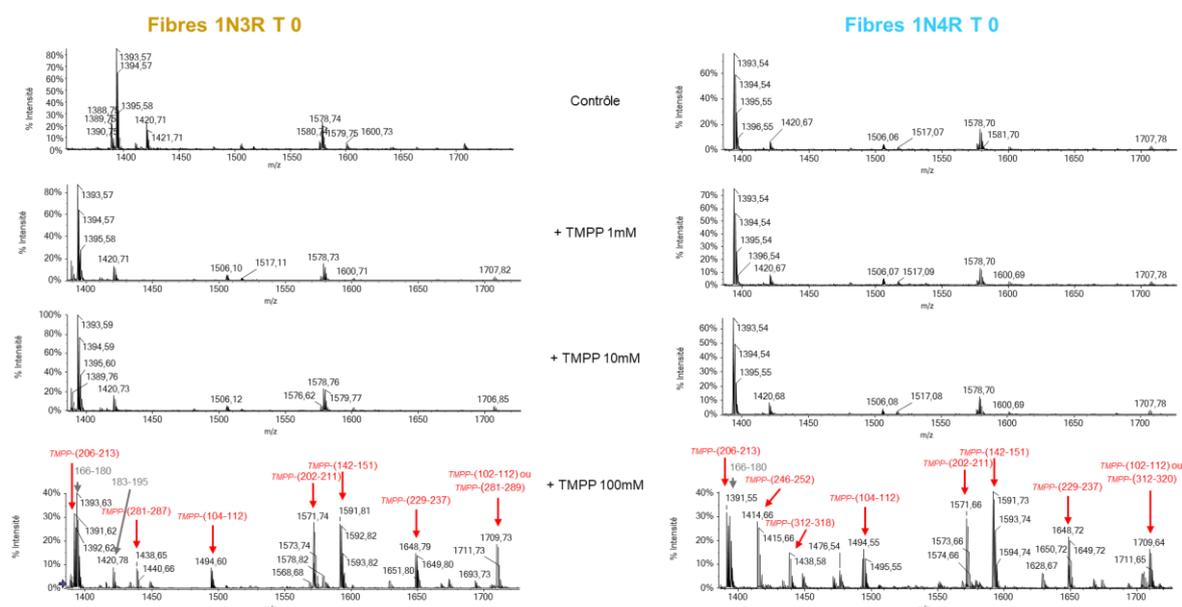
Un premier marquage au TMPP en gel des formes entières de Tau a été réalisé en gamme de TMPP (1m M, 10 mM, 100 mM). Une analyse rapide par MALDI-MS, présentée en figure 10, a permis de mettre en évidence que seule la plus haute concentration (soit 100 mM) donnait lieu à un marquage efficace du résidu N-terminal de la protéine (**Figure 33**). Ce peptide marqué au TMPP est, comme attendu, le plus intense du spectre, rendant son identification simple à réaliser.



**Figure 33.** Dérivatisation du résidu N-terminal : preuve de faisabilité.

Le peptide N-terminal tryptique modifié est l'ion le plus intense du spectre MALDI en présence d'une forte concentration de TMPP (100 mM). Spectres MALDI-MS acquis en mode réflectron (Matrice HCCA reprise à 10 mg/mL en ACN 50%, TFA 0,1%).

Nous avons toutefois noté l'apparition conjointe de nombreux peptides marqués à la même concentration de TMPP (**Figure 34**).

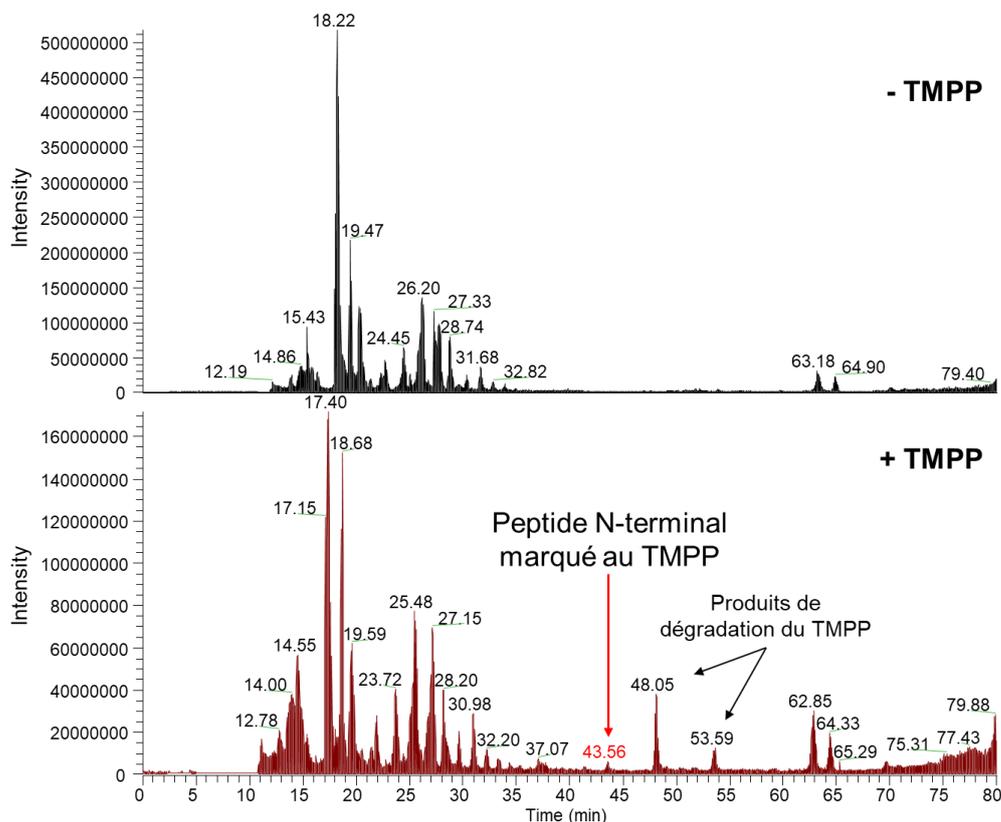


**Figure 34.** Aspécificité du marquage par le TMPP.

La dérivatisation du résidu N-terminal est accompagné de nombreux marquages aspécifiques sur des peptides contenant des lysines, des sérines, des thréonines et des tyrosines, observés à une forte concentration de TMPP (100 mM). Spectres MALDI-MS acquis en mode réflectron (Matrice HCCA reprise à 10 mg/mL en ACN 50%, TFA 0,1%).

L'intensité de ces espèces atteint jusqu'à 40% de celle du peptide N-terminal dérivatisé en analyse MALDI-TOF. Le peptide N-terminal marqué au TMPP restant majoritaire, il semblait possible de se dédouaner de cette réaction secondaire et de poursuivre l'analyse par LC-MS/MS du marquage des fragments protéolytiques marqués au TMPP pour identifier les peptides marqués sur leur acide aminé N-terminal. Ces marquages aspécifiques ont toutefois été implémentés dans les paramètres de recherche automatique des analyses LC-MS/MS.

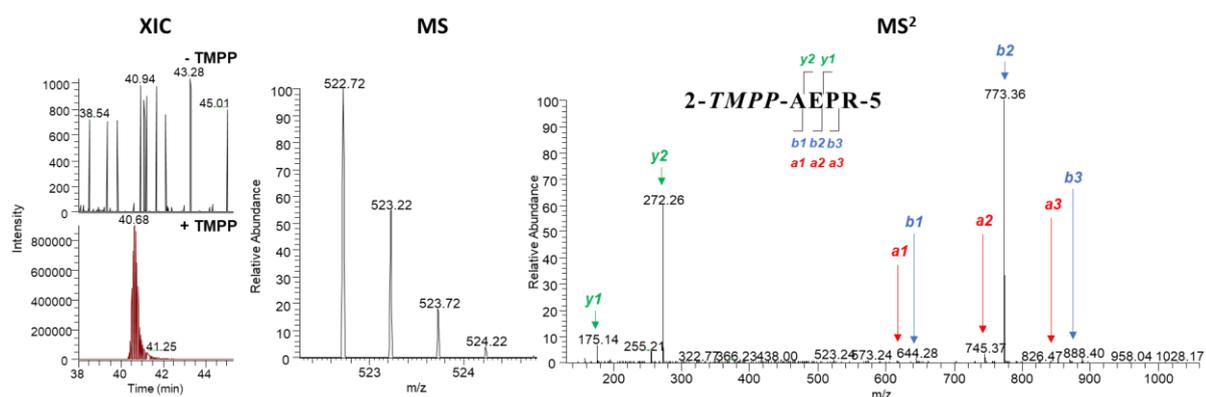
Ces mêmes échantillons ont donc ensuite été analysés par nanoLC-MS/MS (LTQ-Orbitrap), avec un gradient d'acétonitrile (de à 5 à 60% d'acétonitrile, TFA 0,1%) de 70 minutes (**Figure 35**) pour *in fine* pouvoir identifier les peptides N-terminaux des fragments de protéolyse ménagée observés sur gel.



**Figure 35.** Chromatogramme (intensité d'ions) de Tau 1N3R entière digérée en gel par la trypsine après marquage par le TMPP Sans (haut) ou avec (bas) dérivatisation du résidu N-terminal obtenu lors de la séparation sur une colonne C18 (15 cm).

Comme décrit précédemment, l'addition de l'étiquette TMPP engendre un retard d'élution en sortie de colonne C18 (Gallien *et al.* 2009). On observe donc une séparation des peptides non marqués, en début de chromatogramme (RT 0-30 minutes), et des peptides modifiés, en fin de chromatogramme ; ce qui facilite leur identification. Deux produits de dégradation caractéristiques du TMPP sont aussi observés sur le chromatogramme en **Figure 35**. Ces derniers,  $MH^+$  694.26 (RT  $\approx$  48 min) et 591.20 (RT  $\approx$  53.5 min), correspondraient respectivement à un produit de réaction du TMPP avec le Tris (TMPP-Ac-NH-C(CH<sub>2</sub>OH)<sub>3</sub>) et à une forme acide issue de l'hydrolyse du TMPP-Ac-OSu (TMPP-Ac-OH). Ces espèces avaient déjà été décrites auparavant (Ayoub. 2012), et persistent malgré les cycles de lavage supplémentaires.

Bien que des traces (intensité de l'ordre de  $10^3$ ) du peptide N-terminal marqué (i.e 2-TMPP-AEPR-5) soient observables à une concentration de 10 mM de TMPP, la concentration maximale testée (i.e 100 mM) permet de détecter un pic évident sur le chromatogramme pour le peptide N-terminal de la protéine entière marquée au TMPP. Le XIC, le spectre MS et le spectre MS<sup>2</sup> qui lui sont associés sont présentés en **Figure 36**.



**Figure 36.** XIC, spectres MS et MS<sup>2</sup> du peptide N-terminal de la protéine Tau entière marqué par le TMPP.

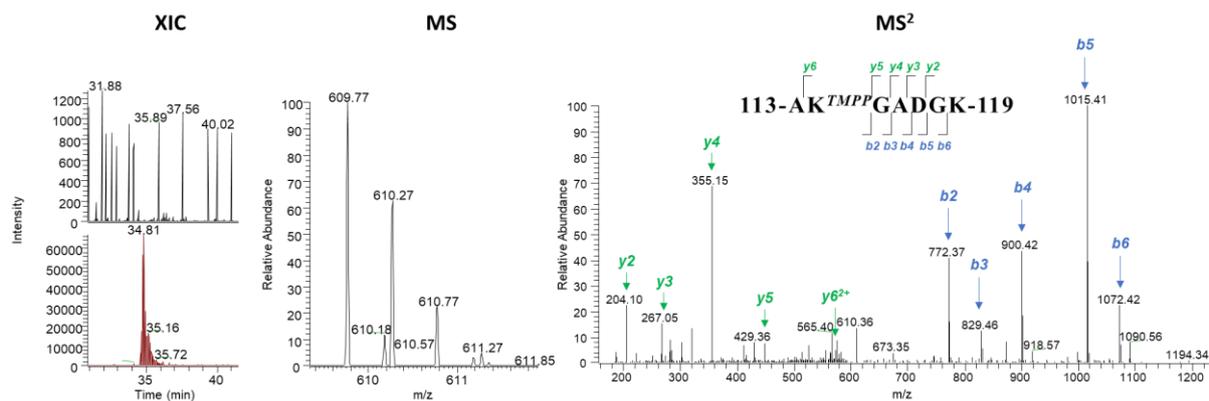
A gauche, spectres XIC du précurseur MH2+ 522.72 du peptide semi-tryptique TMPP-AEPR (Mth 1043.42) pour l'échantillon Tau 1N3R non marqué (haut) et marqué par le TMPP (bas). (Fenêtre d'extraction de m/z de 200 mmu). Au milieu, spectre MS. A droite, spectre MS<sup>2</sup> du peptide N-terminal 2-AEPR-5 marqué en position 1 par le TMPP.

Le spectre de fragmentation est en accord avec la fragmentation d'un peptide portant la charge fixe du TMPP en N-terminal : singulièrement enrichi en ions fragments a et b, comme relaté auparavant (Ayoub. 2012). Cette particularité de fragmentation conduit à l'obtention de spectre enrichi dans la partie droite (hautes masses) et est souvent à l'origine de mauvais scores d'identification. C'est d'ailleurs une des raisons pour laquelle le peptide N-terminal, de petite taille (4 acides aminés), présente une identification avec une confiance moyenne<sup>19</sup> avec une recherche classique avec le moteur de recherche SEQUEST et qu'il n'est pas identifié lorsqu'on choisit le filtre de Confiance Elevée<sup>20</sup>. Cette observation a donc été prise en compte dans la suite de nos analyses.

Au seuil de Confiance Moyenne sont toutefois identifiés un certain nombre d'autres peptides marqués par le TMPP sur des acides aminés de la séquence interne. Parmi eux, le peptide 113-AK<sup>TMPP</sup>GAGGK-119 est identifié avec une confiance élevée et une intensité de  $6,7 \cdot 10^4$  en XIC. A titre d'exemple, nous présentons brièvement en **Figure 37** le XIC et les spectres MS et MS<sup>2</sup> de ce peptide marqué de façon aspécifique sur une lysine interne.

<sup>19</sup> Probabilité seuil de faux positifs fixée à 0,0484

<sup>20</sup> Probabilité seuil de faux positifs fixée à 0,0084



**Figure 37.** XIC, spectres MS et MS<sup>2</sup> d'un peptide interne de la protéine Tau entière marqué par le TMPP. De gauche à droite, XIC du précurseur MH2+ 609.77 du peptide tryptique 113-AK<sup>TMPP</sup>GADGK-119 dans l'échantillon Tau 1N3R exposé (bas) ou non (haut) au TMPP (fenêtre d'extraction 200 mmu) ; spectre MS ; spectre MS<sup>2</sup> annoté.

L'analyse des peptides après dérivatisation des fragments au TMPP a donc été appliquée aux fragments protéolytiques.

### *a.2 Identification de l'extrémité N-terminale des fragments protéolytiques GluC*

L'identification du résidu N-terminal par marquage au TMPP s'est avérée laborieuse en raison du marquage aspécifique évoqué plus haut. En effet, de nombreux peptides modifiés au niveau des lysines, ou plus ponctuellement sur les sérines, thréonines et tyrosines ont introduit une hétérogénéité considérable des peptides présents et, par conséquence, identifiés par analyse automatique avec les moteurs de recherche classique (SEQUEST, Massai). En effet, on note en moyenne 3 à 5 fois plus de peptides aspécifiques identifiés par rapport aux peptides N-terminaux GluC-spécifiques, bien que généralement moins intenses que le peptide marqué sur son N-terminal. Sur la protéine entière utilisée comme preuve de concept de la méthode, cette hétérogénéité est atténuée par la présence d'un unique résidu N-terminal, très abondant, ce qui lève l'ambiguïté sur son identification.

Pour pallier au problème lié aux peptides aspécifiques, nous avons utilisé plusieurs filtres :

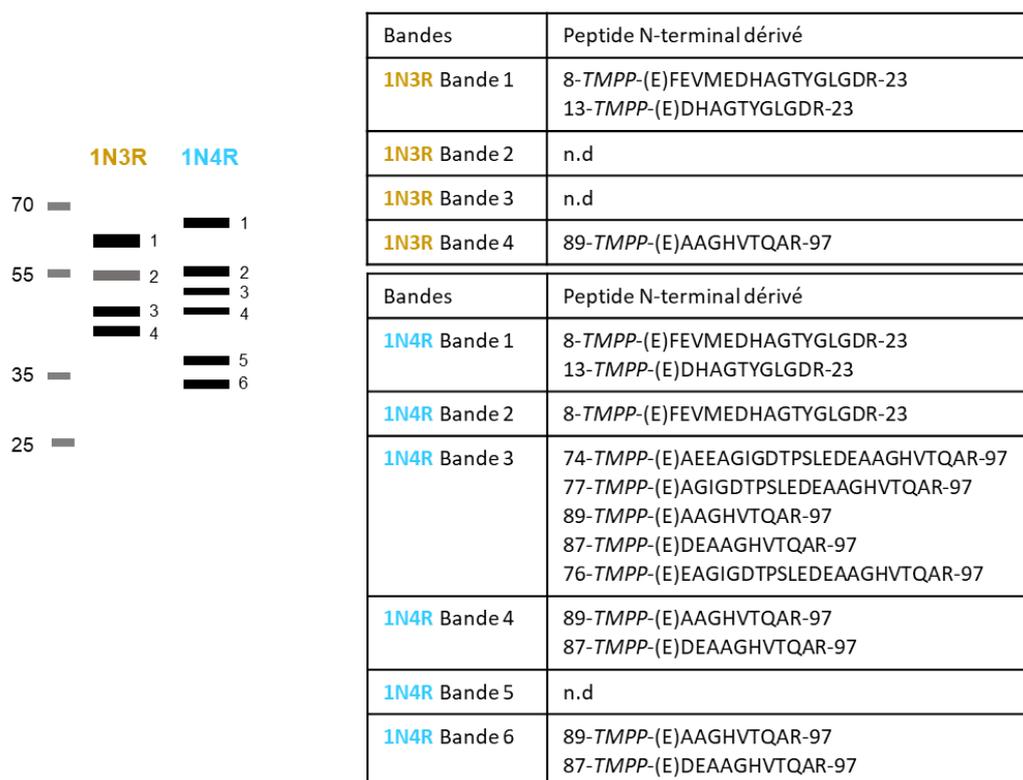
- (1) la concordance de l'extrémité N-terminale avec une coupure par la GluC (l'acide aminé qui précède le peptide marqué doit être un glutamate ou un aspartate),
- (2) un critère de spectral count<sup>21</sup> minimum (fixé à 3),
- (3) une vérification manuelle du spectre MS<sup>2</sup>,

<sup>21</sup> Nombre de spectres MS<sup>2</sup> identifiés pour un même peptide

(4) une intensité significative du précurseur dans l'analyse MS ( $10^3$ ).

Enfin, deux peptides ont été identifiés à partir de la bande correspondant à la protéine entière : 55-*TMPP*- GSEEPGSETSDAK-67 et 105-*TMPP*- GTGSDDKK-112. Ces deux contaminants, ont été exclus de la liste des peptides pertinents.

De cette manière nous avons pu identifier l'extrémité N-terminale de certains fragments protéolytiques GluC comme présenté dans la **Figure 38**.



**Figure 38.** Identification des extrémités N-terminales des fragments protéolytiques GluC des fibres de Tau 1N3R et 1N4R par marquage au TMPP et analyse LC-MS/MS  
Code-barre moléculaire à T 45 minutes de protéolyse ménagée par la GluC des fibres de Tau 1N3R et 1N4R à gauche. Identification du peptide N-terminal par marquage au TMPP suivi d'une analyse LC-MS/MS du code barre moléculaire à droite (n.d, non déterminé). Les spectres MS<sup>2</sup> sont présentés en Annexe 5.

Ainsi, pour le fragment protéolytique 1 de 1N3R, le résidu N-terminal identifié (12 spectres MS<sup>2</sup>), particulièrement intense dans les temps courts de protéolyse, correspond à la phénylalanine 8 ; un second résidu N-terminal, l'aspartate 13 a également été mis en évidence (6 MS<sup>2</sup>). Ce fragment traduit une grande accessibilité à l'enzyme au niveau des premiers acides aminés de la protéine. Pour le fragment 4 des fibres 1N3R, espèce majoritaire dans les temps avancés de la digestion, le peptide marqué 89-97 est très abondant (30 MS<sup>2</sup>), traduisant une accessibilité de cette région et une protection en aval de cette région au niveau de la

surface des fibres. Enfin, le marquage par le TMPP du résidu N-terminal n'a livré aucune information concernant les bandes 2 et 3.

Pour les fibres 1N4R, le profil protéolytique correspondant au code-barre moléculaire GluC est plus complexe avec un total de 6 bandes caractéristiques. Pour les fragments protéolytiques 1 et 2, intenses à des temps courts de digestion, le résidu 8 est identifié comme N-terminal. Là aussi, pour le fragment 1, le résidu 13 est également identifié comme extrémité N-terminale (6 MS<sup>2</sup>). Suivant les critères stringents définis plus haut, nous avons pu identifier jusqu'à 5 résidus N-terminaux différents pour la bande 3; rangés selon un nombre d'identifications décroissant A74 (37 MS<sup>2</sup>), A77 (29 MS<sup>2</sup>), A89 (12 MS<sup>2</sup>), D87 (10 MS<sup>2</sup>) et E76 (3 MS<sup>2</sup>). Les extrémités N-terminales des fragments 2 et 3 du code-barre moléculaire 1N3R n'ayant pu être élucidées, nous ne pouvons tirer de conclusion comparative concernant la région 74 à 77, si ce n'est qu'elle semble en partie accessible à la GluC. Les fragments protéolytiques 4 et 6, majoritaires aux temps longs de protéolyse, présentent deux extrémités N-terminales communes, l'aspartate 87 et l'alanine 89 (peptides 87-97 et 89-97 : 7 et 6 MS<sup>2</sup> dans la bande 4, 27 et 50 MS<sup>2</sup> dans la bande 6). Ainsi, les fibres 1N3R et 1N4R présentent une accessibilité au niveau de la région 87-89 et une protection de la région en aval.

Après 5 heures de protéolyse ménagée par la GluC, on note la présence majoritaire d'une bande unique pour 1N3R et de deux bandes de poids moléculaire distincts pour 1N4R partageant les mêmes extrémités N-terminales. Ceci suggère une différence d'accessibilité à l'enzyme entre les fibres 1N3R et 1N4R dans la partie C-terminale de la molécule Tau.

### *a.3. Identification de l'extrémité N-terminale des fragments protéolytiques chymotryptiques*

A l'image de la stratégie de dérivation des fragments protéolytiques GluC, un marquage du résidu N-terminal par le TMPP des fragments protéolytiques chymotryptiques a été réalisé. Le temps de protéolyse étant court (2 minutes), nous pouvons envisager que les sites privilégiés par la chymotrypsine sont hautement spécifiques (coupure en C-terminal des résidus F, Y et W). Les résultats de la dérivation au TMPP sont présentés en **Figure 39**.



**Figure 39.** Identification des extrémités N-terminales des fragments protéolytiques chymotryptiques des fibres de Tau 1N3R et 1N4R par marquage au TMPP et analyse LC-MS/MS  
Code-barre moléculaire à T 2 minutes de protéolyse ménagée par la chymotrypsine des fibres de Tau 1N3R et 1N4R à gauche. Identification du peptide N-terminal par marquage au TMPP suivi d'une analyse LC-MS/MS du code barre moléculaire à droite.

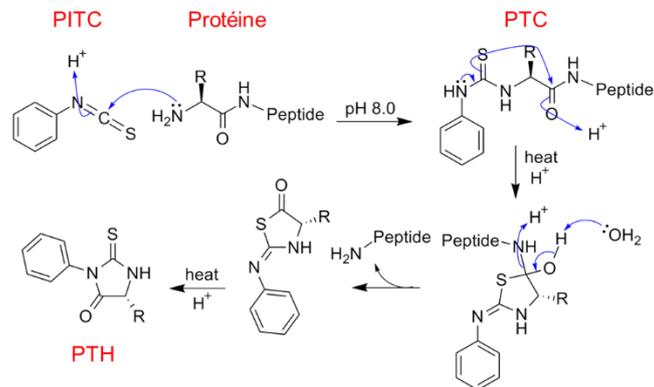
L'identification du résidu N-terminal des fragments chymotryptiques par la dérivation au TMPP suivie par spectrométrie de masse confirme l'extrémité N-terminale identifiée par analyse LC-MS du recouvrement de séquence. Ainsi, tous les fragments protéolytiques partagent le même résidu N-terminal, la sérine 169 (spectre MS<sup>2</sup> en Annexe 5) ; le site de coupure diffère donc bien en C-terminal de la protéine.

La stratégie de marquage du résidu N-terminal par le TMPP et son identification par analyse nanoLC-MS/MS a mis en évidence, d'une part l'existence de marquage aspécifique de chaînes latérales de plusieurs acides aminés. Ce phénomène complique grandement l'analyse. D'autre part, l'efficacité du TMPP s'est révélée être séquence dépendante. En effet, le peptide N-terminal de la protéine entière (AEPR) dérivatisé au TMPP efficacement et a été identifié sans ambiguïté. Nous noterons toutefois que cette bande était très intense et qu'il s'agissait de l'unique résidu N-terminal de l'échantillon. Toutefois, l'identification des fragments protéolytiques GluC reste incomplète. Cependant peu ou pas décrite dans les articles utilisant cette approche, la dérivation préférentielle dépendante de la séquence de l'extrémité N-terminale rejoint les conclusions du groupe de travail de l'ABRF (Association of Biomolecular Resource Facilities) (ARBD. 2014).

#### b. Par séquençage d'Edman

La dérivation du N-terminal par le TMPP s'est avérée plus délicate et compliquée à mettre en œuvre que prévu et a conduit à une identification incomplète des extrémités N-terminales des fragments protéolytiques GluC. Nous avons donc confirmé nos résultats par séquençage d'Edman (Speicher *et al.* 1997), dont le mécanisme est présenté en **Figure 40**. Cette technique présente l'avantage de modifier chimiquement et spécifiquement le groupe α-amino terminal

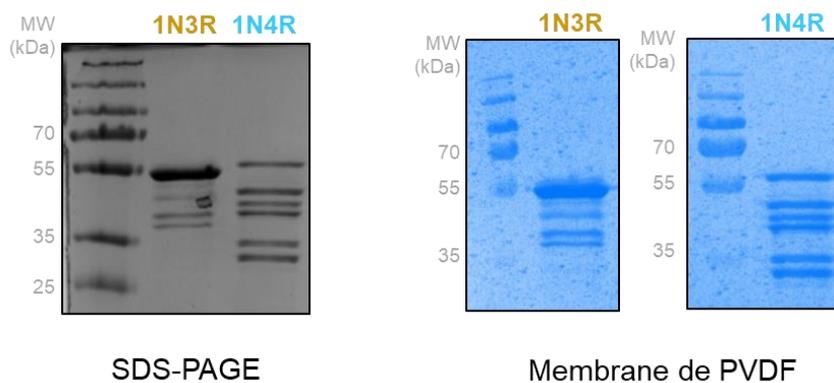
d'un polypeptide. Après cyclisation et hydrolyse en conditions acides, l'acide aminé N-terminal modifié est identifié par HPLC.



**Figure 40.** Mécanisme de la dégradation d'Edman  
PITC (isothiocyanate de phényle), PTC (phénylthiocarbamide) et PTH (phénylthiohydantoïne)

Ce mécanisme de clivage de l'acide aminé N-terminal est répété pendant plusieurs cycles jusqu'à obtenir séquentiellement la composition en acides aminés à partir de l'extrémité N-terminale.

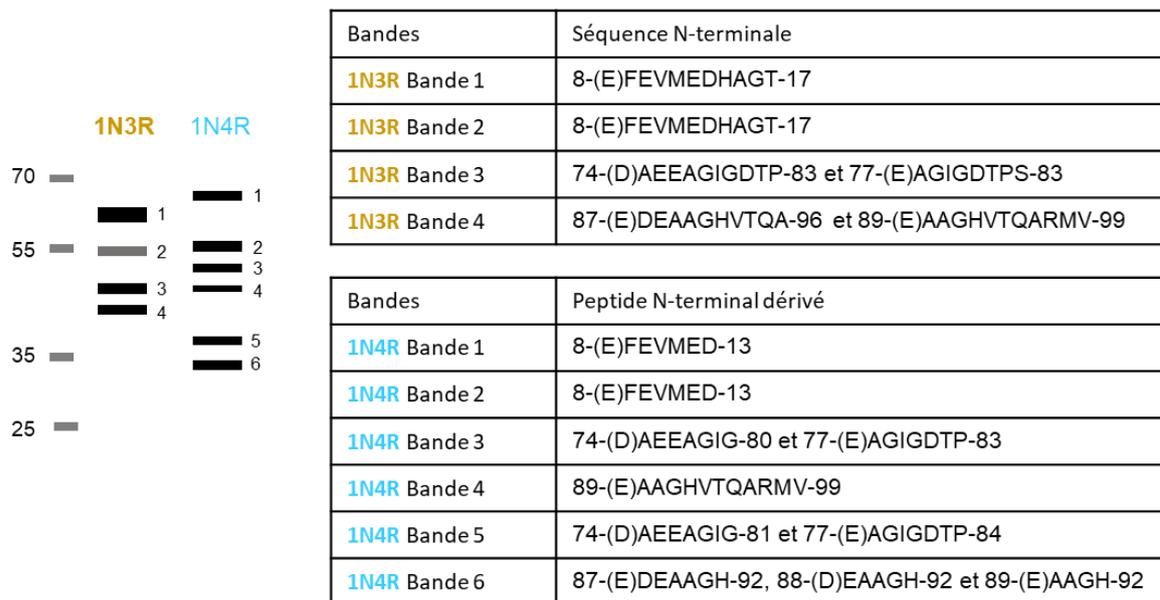
Cette méthode de séquençage s'applique à l'étude de fragments protéolytiques en général et avait précédemment été utilisée lors d'expériences de protéolyse ménagée de fibres de Tau (von Bergen *et al.* 2006). Pour ce faire, nous avons électro-transféré les fragments protéolytiques GluC séparés par SDS-PAGE vers une membrane de Polyfluorure de vinylidène (PVDF) ensuite colorée au bleu de Coomassie pour visualiser les bandes protéiques d'intérêt. Ces bandes protéiques ont été découpées pour être séquencées. Ainsi, nous avons transféré efficacement les fragments protéiques des fibres de Tau 1N3R et 1N4R à T 45 minutes de protéolyse ménagée par la GluC (**Figure 41**).



**Figure 41.** Electrotransfert des fragments protéolytiques GluC des fibres 1N3R et 1N4R générés après 45 minutes de digestion ménagée.

Les séquenceurs d'Edman étant devenus très rares, le séquençage a été effectué sur la plateforme de Protéomique de l'IMM (Institut de Microbiologie de Méditerranée) qui dispose d'un séquenceur Shimadzu PPSQ 31B. Après lavage des membranes par un mélange éthanol/eau (90/10), dix cycles de dégradation ont été effectués.

Les résultats du séquençage sont présentés en **Figure 42**.



**Figure 42.** Identification des extrémités N-terminales des fragments protéolytiques GluC des fibres de Tau 1N3R et 1N4R par séquençage d'Edman  
Code-barre moléculaire à T 45 minutes de protéolyse ménagée par la GluC des fibres de Tau 1N3R et 1N4R. Identification des peptides N-terminaux par séquençage d'Edman des fragments observés.

Ainsi, le séquençage d'Edman nous a permis de confirmer et de compléter l'identification du N-terminal par marquage au TMPP analysé par spectrométrie de masse.

Pour compléter l'identification des fragments protéolytiques, nous avons cherché à déduire le résidu C-terminal des fragments observés.

#### A.3.4 Déduction du résidu C-terminal des fragments protéolytiques

##### a. Fragments protéolytiques GluC : d'après des mesures de masse entière

Par western-blot, nous avons pu établir que les trois bandes les plus intenses sur gel à 45 minutes correspondaient à des fragments protéolytiques de Tau 1N3R dont l'extrémité C-terminale est intacte. Pour 1N4R, ce n'est pas le cas des fragments protéolytiques des bandes

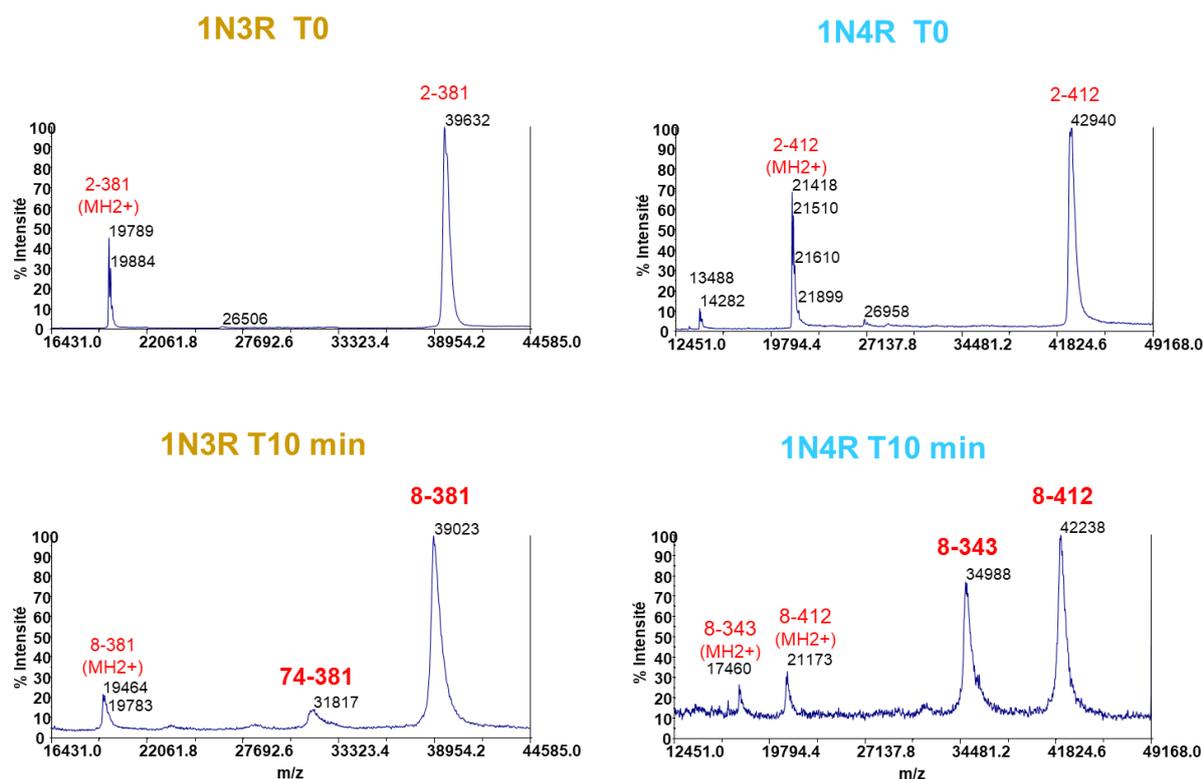
2, 5 et 6 pour lesquelles l'extrémité C-terminale de la protéine n'est pas immunodétectée. Nous avons donc tenté de déduire le C-terminal de ces fragments. Les sites de coupure potentiels de la GluC (en C-terminal des aspartates et glutamates) sont très nombreux. Après protéolyse totale par la trypsine, aucun peptide spécifique C-terminal correspondant à une coupure GluC (i.e se terminant par un aspartate ou un glutamate) n'a pu être mis en évidence. Concernant les peptides tryptiques, nous avons tout de même noté les observations suivantes, en accord avec les résultats de western-blots :

- le peptide 324-IGSLDNITHVPGGGNK-339 est détecté dans l'ensemble des bandes 1N3R et 1N4R

- le peptide 357-TDHGAEIVYK-366 n'est pas détecté dans les bandes 1N4R 2 et 6

Ceci suggère que la coupure en C-terminal des fragments protéolytiques 1N4R contenus dans les bandes 2 et 6 se situe entre l'acide aminé 340 et 366. Dans cette région, on note la présence de trois résidus glutamate 343, 351 et 362. Afin de vérifier et préciser ces sites potentiels, nous avons donc mesuré la masse entière des fragments protéolytiques issus de la protéolyse ménagée par la GluC à 10 minutes par MALDI-MS.

Nous avons opté pour une mesure de masse par MALDI-MS, davantage tolérante que l'ESI-MS à l'analyse de protéines initialement agrégées. Cette analyse peut être précédée d'une séparation rapide par chromatographie à polarité de phase inversée. Contrairement à l'approche Bottom-Up, l'analyse MS n'est ici pas associée à une fragmentation, l'identification *de facto* des fragments protéolytiques n'étant donc pas complètement vérifiée. Toutefois, comme discuté plus haut, le résidu N-terminal des fragments protéolytiques GluC étant désormais connu et les sites potentiels restreints, l'information de masse entière nous permettra de définir le résidu C-terminal probable des ces espèces. Dans le but d'améliorer la qualité des spectres et du signal pour les fragments d'intérêt, et du fait du meilleur potentiel d'ionisation des espèces peptidiques de petite taille qui gênent l'ionisation des fragments d'intérêt de haut poids moléculaire, nous avons inclus une étape de pré-fractionnement de l'échantillon sur zip-tip C18 en amont de l'analyse MS. Les espèces de petite taille ont ainsi été éluées à 10% d'acétonitrile en TFA et l'ensemble des fragments de plus grande taille, nos espèces protéiques d'intérêt, à 70% d'acétonitrile en TFA 0.1%. Les spectres obtenus sont présentés en **Figure 43**.



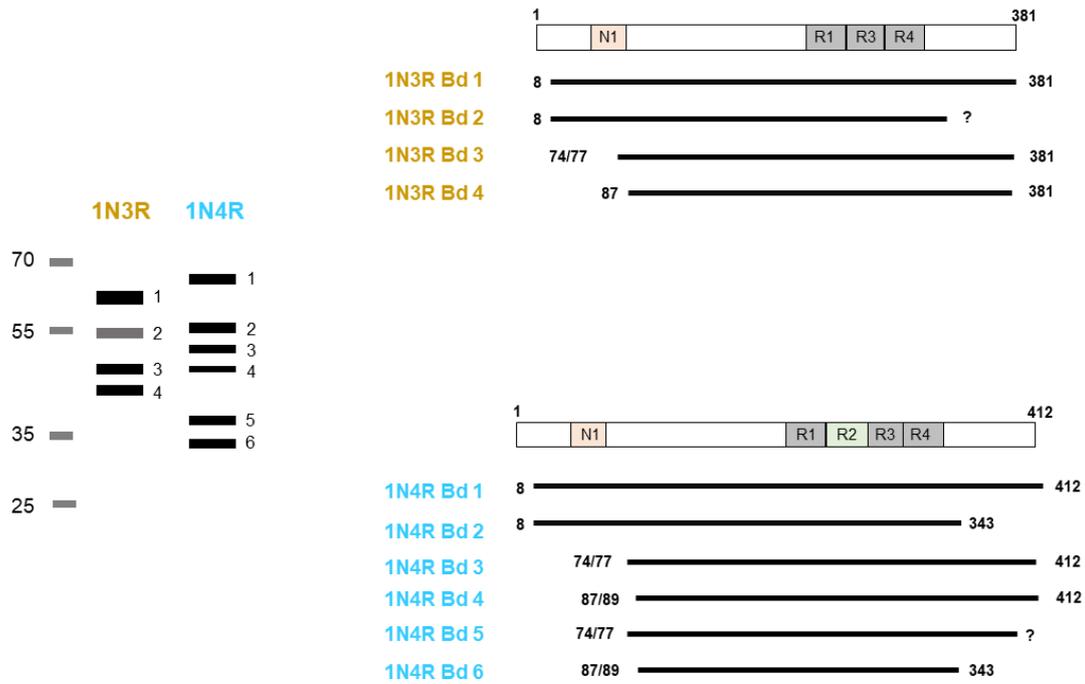
**Figure 43.** Spectres MALDI-MS des produits de protéolyse ménagée des fibres 1N3R et 1N4R par la GluC. Fragments à T10 minutes de protéolyse co-cristallisés en acide sinapinique (lissage gaussien de 5 points) et séquences associées possibles en adéquation avec le résidu N-terminal identifié ( $|\Delta m| < 150\text{ppm}$ ).

Après 10 minutes de protéolyse ménagée des fibres 1N3R, on constate la présence majoritaire d'une espèce à un m/z de 39023 soit une différence de 608 Da par rapport à la protéine entière (T0). Au vu de l'identification de l'extrémité N-terminale, nous pouvons supputer qu'il pourrait s'agir de la séquence N-terminale (2-AEPRQE-7). Une autre espèce, peu intense, est retrouvée avec un m/z d'environ 31817 et pourrait correspondre à la séquence 74-381. Ainsi, en accord avec les données de western-blot, nous confirmons que l'extrémité N-terminale est accessible au sein des fibres de Tau 1N3R tandis que la région C-terminale semble protégée après assemblage.

De la même façon, après 10 minutes de protéolyse des fibres 1N4R, on observe une espèce majoritaire à m/z 42238 qui correspondrait également à une prise en charge par la GluC de l'extrémité N-terminale et résulterait par le fragment 8-412. Une seconde espèce, abondante ici, de m/z 34988 correspondrait à un fragment de Tau 1N4R pris en charge au niveau des extrémités N- et C-terminales : la séquence 8-343. En accord avec une coupure dans la région 340-360, l'analyse en masse entière nous permet d'affirmer que la GluC cible le glutamate 343 en C-terminal de la protéine. Ainsi, comme les fibres 1N3R, l'extrémité N-terminale est accessible au sein des fibres 1N4R. En accord avec les données de western-blot, la région C-

terminale est tronquée par la GluC, témoignage d'une accessibilité de cette région au sein des fibres 1N4R.

Pour conclure, le code barre moléculaire des fibres de Tau 1N3R et 1N4R généré par la GluC est représenté en **Figure 44**.



**Figure 44.** Extrémités N et C-terminales des fragments protéolytiques générés après 45 minutes de digestion par la GluC.

#### b. Fragments chymotryptiques : d'après le recouvrement de séquence

Les fragments chymotryptiques partageant tous la même extrémité N-terminale et au vu du faible nombre de sites de coupure candidats (3), nous avons pu identifier des peptides dont l'extrémité C-terminale résulte d'une coupure par la chymotrypsine par l'approche Bottom-Up des fragments protéolytiques chymotryptiques. Comme discuté plus haut, en raison du phénomène de diffusion de bande, l'analyse du recouvrement de séquence ne s'était pas avérée suffisamment informative quand à l'identification des fragments protéolytiques par analyse du recouvrement de séquence total. Cependant, au vu du faible nombre de sites potentiels en accord avec une coupure spécifique et le poids apparent sur gel, nous avons dresser une liste de candidats relativement restreinte : F349, Y365 et L412 (pas de clivage).

Pour la bande 1 issue des échantillons 1N3R et 1N4R, nous n'avons pas identifié de peptides dont le résidu C-terminal était aromatique. De plus, deux peptides semi-tryptiques en C-terminal de la protéine 400-ADEVASLAK-409 et 399-LADEVSASLAK-409 ont été retrouvés

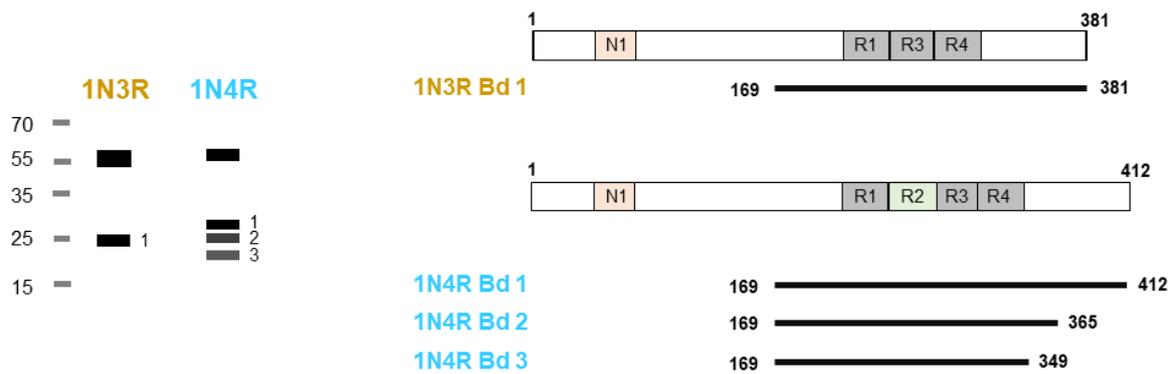
de façon prépondérante dans ces bandes par rapport aux bandes 2 et 3 1N4R. En accord avec les western-blot et le marquage du N-terminal par le TMPP, nous pouvons affirmer que la bande 1 1N3R correspond au fragment protéolytique 169-389 (Mth 22387,44 Da) et la bande 1 1N4R au fragment protéolytique 169-412 (Mth 25633.17 Da).

Pour la bande 2 de l'échantillon 1N4R, les peptides 355-AKTDHGAEIVY-365 et 357-TDHGAEIVY-365 sont clairement identifiés et relativement abondants d'après la recherche SEQUEST alors qu'ils n'ont pas pu être identifiés à partir des autres bandes. Manuellement, pour la bande 2, l'intensité est de  $2,46 \cdot 10^6$  u.a pour le précurseur deux fois chargé du peptide 357-365 ( $MH2+ = 502.74$ ) alors qu'il est présent à l'état de traces (100 fois moins) dans les bandes 1 et 3 de l'échantillon 1N4R probablement en raison de la diffusion des bandes. De la même façon, le précurseur deux fois chargé du peptide 355-365 ( $MH2+ = 602.30$ ) présente une intensité de  $1,23 \cdot 10^6$  u.a alors qu'il est 100 fois moins intense dans les bandes 1 et 3 de l'échantillon 1N4R. La bande 2 de l'échantillon 1N4R correspondrait donc au fragment protéolytique 169-365 dont la masse théorique est 20896.92 Da.

Pour la bande 3 de l'échantillon 1N4R, la recherche SEQUEST identifie le peptide 342-IETHKLTF-349 alors qu'il n'est pas identifié dans les autres bandes. Son intensité atteint  $4,26 \cdot 10^4$  u.a alors qu'il n'est pas détecté dans les autres bandes. La bande 3 de l'échantillon 1N4R correspondrait donc au fragment protéolytique 169-349 dont la masse théorique est 19112,96 Da.

Ainsi, la protéolyse ménagée des fibres de Tau 1N3R et Tau 1N4R par la chymotrypsine révèle des différences d'accessibilité dans la partie C-terminale de la protéine au premier temps de digestion, en particulier de la tyrosine 365 mais aussi de la phénylalanine 349. A l'inverse, ceci semble également montrer que la tyrosine respectivement en position 250 pour 1N3R ou 281 pour 1N4R, et la phénylalanine respectivement en position 286 pour 1N3R ou 317 pour 1N4R, ne sont pas accessibles à l'enzyme à la surface des fibres 1N3R.

Pour conclure, le code barre moléculaire des fragments protéolytiques des fibres de Tau 1N3R et 1N4R généré par la Chymotrypsine après 2 minutes de réaction est représenté en **Figure 45**.



**Figure 45.** Extrémités N et C-terminales des fragments protéolytiques générés après 2 minutes de digestion par la chymotrypsine.

Pour conclure, l'identification des fragments protéolytiques issus des protéolyses ménagées, en particulier celle des espèces générées par la GluC illustre bien les difficultés que peut poser l'identification sans *a priori* des résidus N et C-terminaux de fragments protéolytiques, même dans le cas d'échantillons peu complexes (i.e ici une protéine purifiée). En effet, nous avons dû associer différentes méthodes pour identifier, de façon certaine, les bornes des fragments issus de protéolyses ménagées et ainsi mettre en évidence des déterminants structuraux propres aux fibres de Tau 1N3R et 1N4R qui sont discutés dans l'article présenté en Chapitre IV.

#### A.4 Analyse de peptides protéolytiques issus d'une protéolyse avancée

Nous avons effectué une digestion avancée des formes fibrillaires et monomériques de Tau 1N3R et de Tau 1N4R en conditions natives. Pour ce faire, les échantillons protéiques ont été digérés par la GluC pendant 15h au ratio [Enzyme/Tau] de 1/50. A l'issue de cette protéolyse, aucun fragment protéolytique de haut poids moléculaire apparent n'est observable sur gel SDS-PAGE. Nous avons donc analysé par LC-MS/MS la population peptidique générée. De façon intéressante, grâce au moteur de recherche MASCOT, nous avons pu constater que le recouvrement de séquence pour les formes monomériques et fibrillaire était drastiquement différent, comme illustré dans la **Table 6**.

	Forme monomérique	Forme fibrillaire
<b>Tau 1N3R</b>	75%	48%
<b>Tau 1N4R</b>	74%	62%

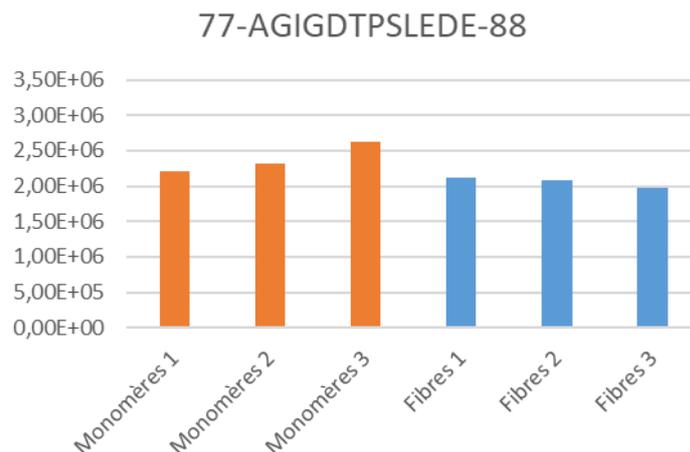
**Table 6.** Pourcentage de recouvrement de séquence de Tau après digestion par la GluC en solution. Identifié par le moteur de recherche MASCOT à partir de données de LC-MS/MS issues de l'analyse de digestion de fibres et de monomères de Tau 1N3R et 1N4R en conditions natives par la GluC.

Cette recherche nous a permis de dresser une liste de peptides à partir des données obtenues pour les monomères peu ou pas structurés, présentant une bonne accessibilité à l'enzyme. Cette liste de peptides identifiés ainsi que leur m/z, leur état de charge, leur temps de rétention, ont ensuite été utilisés pour réaliser une quantification relative « label free » de ces peptides entre les formes monomériques et les formes fibrillaires. Ainsi, leur intensité (« Extracted Ion Chromatogram » ou XIC) a été extraite pour chacune des expériences de digestion avancée.

En raison d'une différence de recouvrement de séquence entre les formes monomérique et fibrillaire, nous ne pouvions normaliser les valeurs d'intensité des peptides d'intérêt en fonction du coup d'ions total. Nous avons donc recherché un peptide interne qui puisse servir de peptide de référence notamment pour réduire au maximum l'impact d'une différence éventuelle de quantité de protéine Tau entre les échantillons (Barthélemy *et al.* 2016). Le peptide candidat devait présenter quatre caractéristiques :

- (1) se situer dans une région jusqu'à présent jamais décrite comme protégée au sein des fibres (d'après les données présentes dans la littérature et celles mises en évidence par notre étude de protéolyse ménagée présentée plus haut),
- (2) possède une bonne capacité d'ionisation (peptide intense),
- (3) présente peu ou pas de coupures incomplètes,
- (4) ne présente pas de variations importantes entre les échantillons.

Ainsi, il apparaissait judicieux de choisir un peptide situé dans la partie N-terminale de la protéine. Retrouvé de façon homogène dans les échantillons, le peptide 77-AGIGDTPSLEDE-88 est plus intense et homogène que le peptide 77-AGIDTPSLE-86 correspondant à une digestion complète. L'intensité du peptide utilisé pour la normalisation dans différents échantillons est présentée en **Figure 46**.



**Figure 46.** Intensité du peptide 77-88 choisi pour la normalisation  
 Histogramme de l'intensité brute du peptide de référence (XIC) correspondant à la séquence 77-88 pour la normalisation des intensités des peptides dans les échantillons de Tau 1N4R monomérique et fibrillaire.

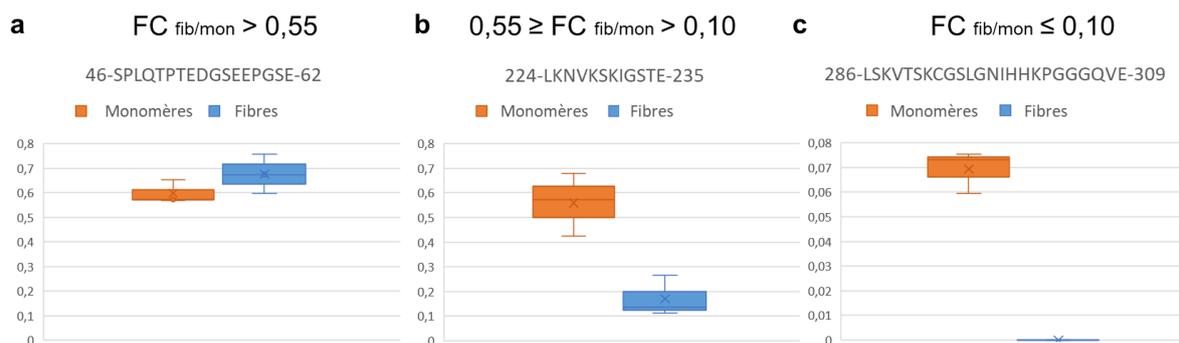
Ainsi, nous avons normalisé l'ensemble des valeurs d'intensité brute des différents peptides d'après la relation mathématique 1.

$$(1) \quad \frac{\text{Intensité du peptide d'intérêt}}{\text{Intensité du peptide standard}}$$

Nous avons ensuite pu comparer l'intensité d'un peptide donné entre les formes monomérique et fibrillaire d'une isoforme de Tau donnée en calculant un « fold-change »,  $FC_{\text{fib/mon}}$  entre l'intensité d'un peptide d'intérêt pour la forme fibrillaire et l'intensité du même peptide pour la forme monomérique d'une même isoforme (équation 2) pour mettre en évidence des différences de régions protégées entre les formes monomérique et fibrillaire pour chaque isoforme de Tau, 1N3R et 1N4R.

$$(2) \quad FC_{\frac{\text{fib}}{\text{mon}}} = \frac{\text{Moyenne (Intensité du peptide d'intérêt normalisée)}_{\text{Tau fibrillaire}}}{\text{Moyenne (Intensité du peptide d'intérêt normalisée)}_{\text{Tau monomérique}}}$$

Afin d'être stringent et de révéler des signatures conformationnelles isoforme-spécifiques majeures qui permettent de mettre en évidence le cœur protégé des agrégats de Tau 1N3R et 1N4R, nous avons fixé une limite de fold-change suffisamment élevée ou basse soit  $FC_{\text{fib/mon}} > 1,8$  et  $FC_{\text{fib/mon}} < 0,55$ . Un second seuil de fold-change a été considéré : celui pour lequel la différence d'intensités normalisées entre formes fibrillaire et monomérique est supérieure à 10 fois ( $FC_{\text{fib/mon}} \leq 0,1$ ). Ce second seuil reflète une diminution drastique de l'accessibilité à l'enzyme après assemblage. Ces catégories de peptides sont illustrées en **Figure 47**.



**Figure 47.** Box-plots représentant les trois catégories de peptides associés à différents « fold-changes » entre formes fibrillaire et monomérique de Tau 1N4R. (a)  $F_c > 0,55$  (Fibrés  $\approx$  Monomères), (b)  $0,1 < F_c < 0,55$  (Fibrés < Monomères) et (c)  $F_c \leq 0,1$  (Fibrés  $\ll$  Monomères).  $n=3$

Enfin, nous avons appliqué un test statistique non paramétrique de type Kruskal-Wallis pour déterminer la significativité de la différence.

Cette méthode optimisée de cartographie des régions protégées a été appliquée à l'analyse des fibres de Tau 1N3R et 1N4R. Ces résultats sont présentés dans le Chapitre III.2 (Article).

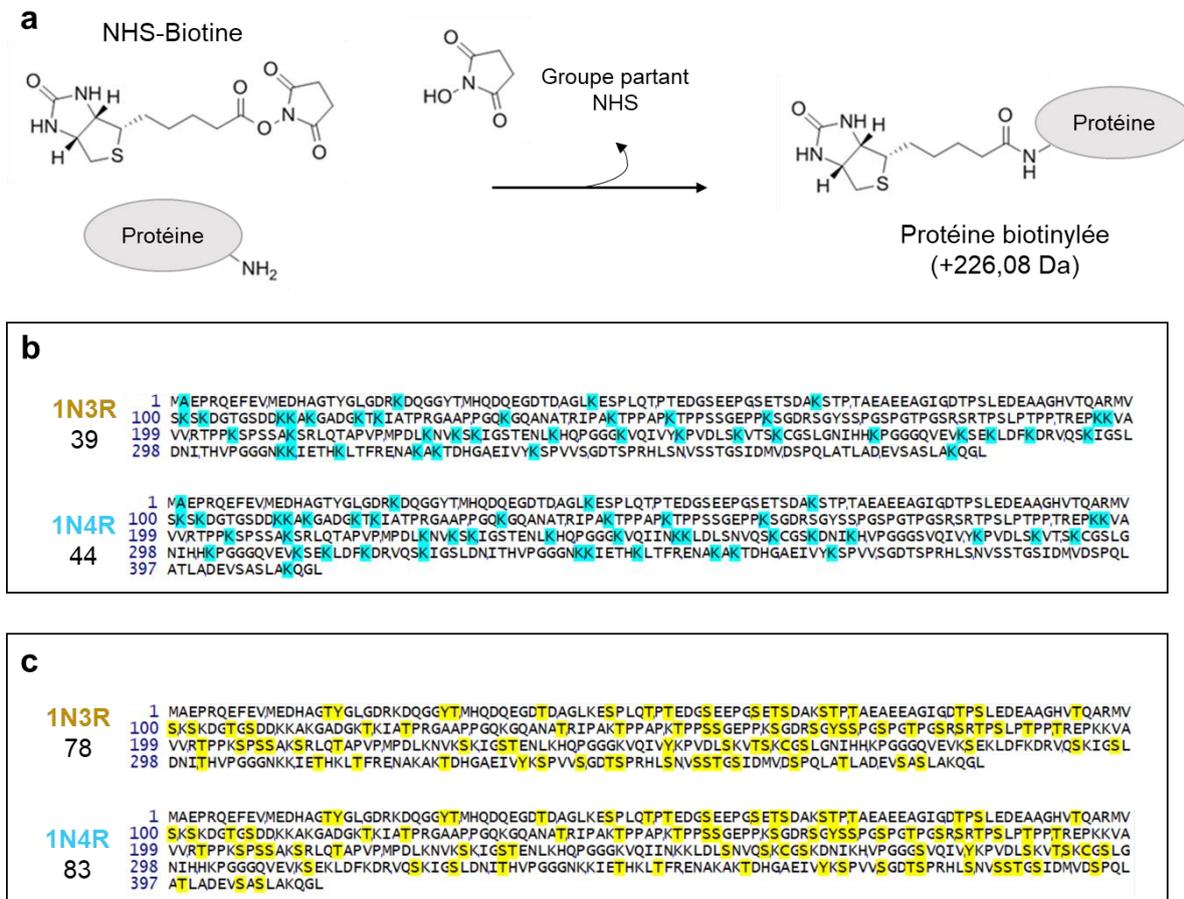
## B. Marquage covalent des surfaces accessibles par la NHS-Biotine

### B.1 Optimisation des conditions de marquage

#### *B.1.1 Choix du marquage covalent par la NHS-Biotine*

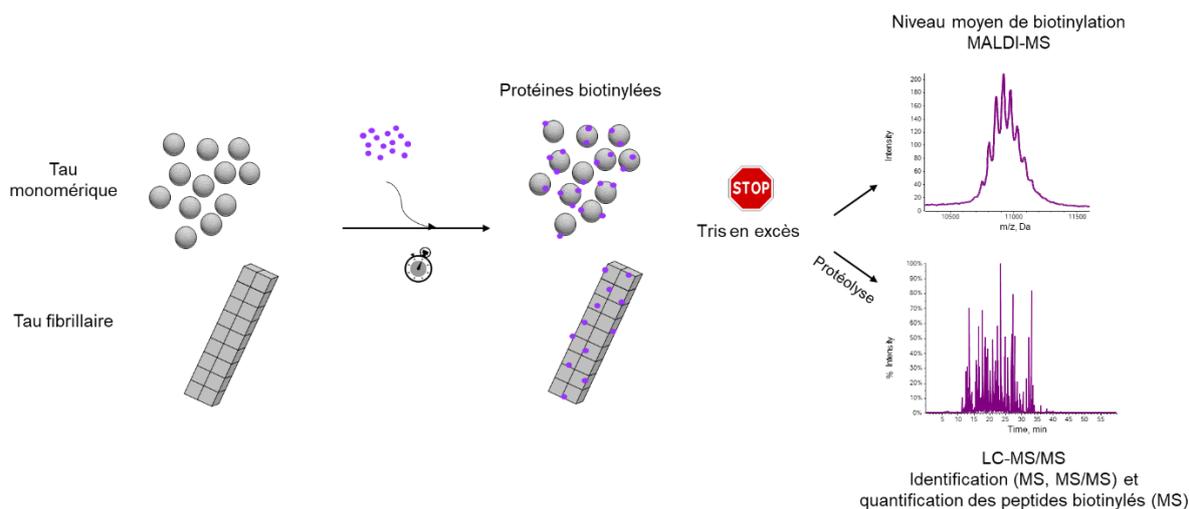
Comme discuté dans la partie introductive IV.4, il existe différents types de marquage de surfaces accessibles au solvant combiné à l'analyse des produits marqués par spectrométrie de masse dont le plus connu est la prise d'empreintes par échanges hydrogène-deutérium. Avec une technologie adaptée et un savoir-faire certain, cette technique peut être très puissante pour déterminer l'accessibilité de surface de systèmes protéiques solubles et les changements conformationnels (Canul-Tec *et al.* 2017). Récemment, cette méthode appliquée à Tau a été publiée en 2018 par Huang et ses collaborateurs (Huang *et al.* 2018). Dans cette étude, décrivant l'accessibilité de surface « d'agrégats solubles » de Tau 2N4R, deux courts peptides recouvrant la séquence 308-315 (R3) présentent une diminution de la vitesse d'échange hydrogène/deutérium, suggérant une protection de cette région au cours du processus de fibrillation. Toutefois, on peut questionner la nature de ces espèces de Tau et l'analogie qui peut être faite avec la structure des fibres matures. Du fait du phénomène de « back-exchange » évoqué plus haut dans la partie introductive IV.3.1, de la difficulté de combiner les contraintes de la méthode et des problèmes de solubilité des fibres de Tau, nous avons choisi de mettre en place un marquage de surface covalent. De plus, la protéine Tau présente un enrichissement en lysines par rapport à l'ensemble du protéome humain, à savoir 38 lysines pour 1N3R et 43 lysines pour 1N4R (Brüne *et al.* 2018). Enfin, la répartition des lysines le long de la séquence de Tau, notamment au sein de la répétition R2 pour 1N4R, est homogène (Figure 1.b). Ainsi, la dérivatisation des amines primaires par la chimie NHS-ester nous a semblé adaptée à l'étude de l'accessibilité des résidus après agrégation. Le marquage par la NHS-Biotine avait d'ailleurs permis de mettre en évidence des changements d'accessibilité de surfaces de résidus lors de changements conformationnels (Azim-Zadeh *et al.* 2007 ; Dreger *et al.* 2009). Concernant la réactivité de Tau à cette molécule, le marquage par la NHS-biotine de Tau monomérique a été utilisé précédemment comme outil moléculaire pour des études fonctionnelles, sans perturber son rôle dans la polymérisation des microtubules au ratio molaire (K/NHS-Biotine) de 6/1 (Makrides *et al.* 2003). Le nombre de biotine par molécule de Tau n'avait toutefois pas été déterminé. Le mécanisme de dérivatisation de protéines par la NHS-Biotine, préférentiellement au niveau des amines primaires (chaines latérales des lysines et extrémité N-terminale) et secondairement des hydroxyles (chaines latérales des tyrosines, théronines, sérines et cystéines) est explicité en

**Figure 48.a** et la répartition le long de la séquence de Tau 1N3R et 1N4R des résidus cibles présentée en **Figure 48.b et c**.



**Figure 48.** Principe du marquage par la NHS-Biotine et sites putatifs le long de la séquence de 1N3R et 1N4R. **(a)** Principe du marquage par la NHS-biotine. **(b)** Sites hautement réactifs le long de la séquence de Tau 1N3R et de Tau 1N4R : Acide aminé N-terminal et lysines. **(c)** Sites secondairement réactifs le long de la séquence de Tau 1N3R et 1N4R : cystéines, sérines, thréonines et tyrosines.

Le plan d'expérience élaboré est illustré en **Figure 49**. Brièvement, les protéines monomériques et fibrillaires sont incubées en présence de la NHS-Biotine. Les ratios molaires (K/NHS-Biotine) et les temps d'incubation ont été optimisés. La réaction de marquage est arrêtée par addition de Tris en excès (compétiteur). Le niveau moyen de biotinylation a ensuite été déterminé sur la protéine entière par MALDI-MS. Ensuite, les échantillons sont digérés, en conditions dénaturantes (en gel SDS-PAGE) d'une part et en conditions natives (en solution) d'autre part. La population des peptides protéolytiques générés est analysée par nano-LC-MS/MS. Les peptides biotinylés ont ensuite été quantifiés, normalisés et comparés un à un dans les différents échantillons.



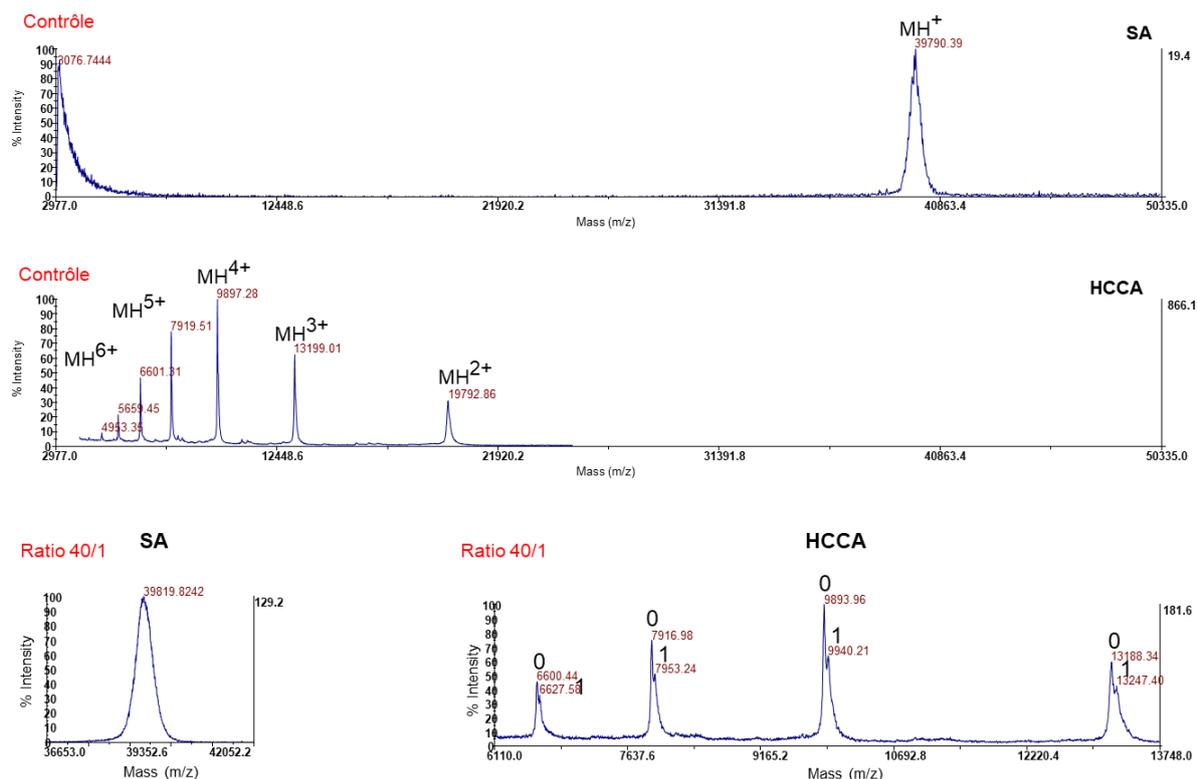
**Figure 49.** Schématisation de la stratégie de marquage de surface de Tau 1N3R et 1N4R monomérique et fibrillaire et son analyse par spectrométrie de masse

### B.1.2 Mesure du niveau moyen de marquage par MALDI-MS

Afin d'évaluer le niveau moyen de marquage des fibres et des monomères de Tau 1N3R et 1N4R, nous avons déterminé la masse entière de nos échantillons marqués par MALDI-MS.

Dans un premier temps, nous avons tenté d'observer la différence de masse entre la protéine Contrôle et la protéine biotinylée au plus fort ratio, condition pour laquelle on s'attend à observer un plus grand degré d'hétérogénéité en utilisant l'acide sinapinique comme matrice. Comme illustré en **Figure 50**, l'utilisation de l'acide sinapinique (SA) comme matrice conduit à la formation d'espèces monochargées essentiellement. La résolution et l'intensité du signal pour les échantillons co-cristallisés en acide sinapinique étaient de mauvaise qualité. Ce type de dépôt ne pouvait donc pas être utilisé pour déterminer le nombre de biotine par molécule de Tau. Moins fréquente pour l'étude de hauts poids moléculaires, la matrice Acide  $\alpha$ -cyano-4-hydroxycinnamique (HCCA) est originellement destinée à l'étude d'espèces de taille plus restreintes (i.e peptides) (Frecklington. 2007). Cependant, cette dernière peut également être utilisée pour générer des espèces multichargées de grosses molécules qui permettent de mesurer la masse après déconvolution, améliorant la résolution, la sensibilité et la précision de mesure de masse de protéines de taille relativement importante (Signor and Erba. 2013). Comme présenté en **Figure 50**, cette méthode a amélioré la qualité des spectres MALDI-MS, nous permettant de mesurer le nombre moyen de biotine ajoutée.

## Fibres de Tau 1N3R



**Figure 50.** Influence de la matrice sur la qualité de l'analyse de masse de molécules de haut poids moléculaire. En haut, spectre MALDI-MS Tau 1N3R non marqué (échantillon fibrillaire) en utilisant le SA ou l'HCCA, tout deux repris en ACN 50%, TFA 0.1%. En bas, évaluation du nombre de biotine ajoutée par molécule de Tau au ratio [K/NHS-Biotine] de 40/1 par MALDI-MS (lissage gaussien 5 points).

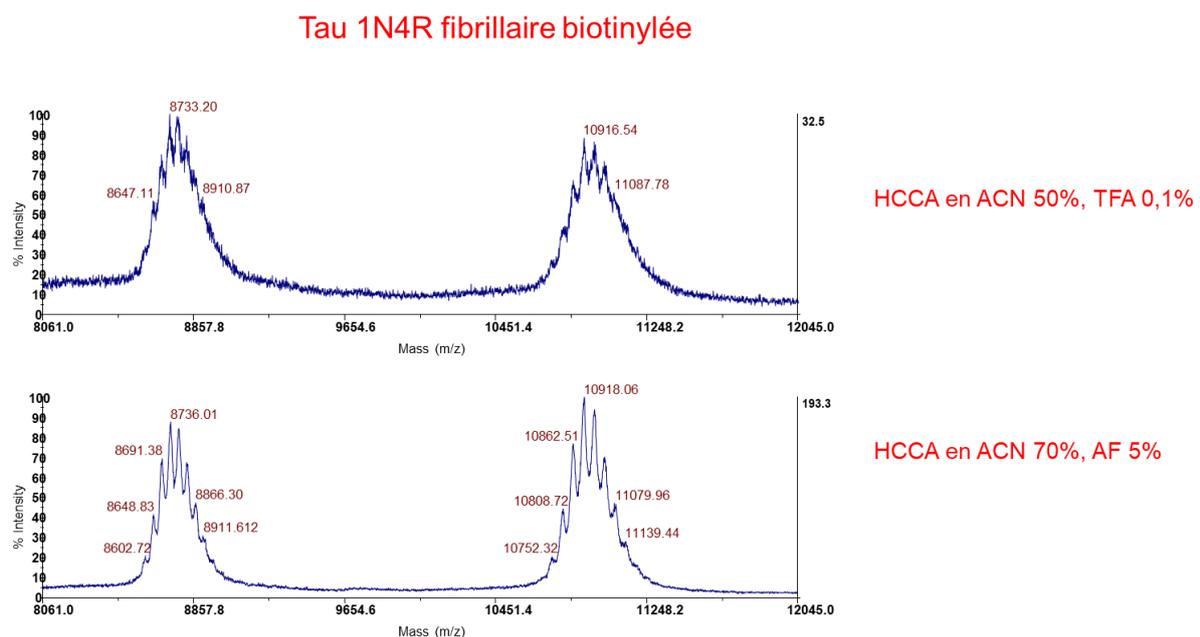
L'analyse des espèces multi-biotinylées très hétérogènes, observées en particulier pour les formes monomériques de Tau, a nécessité une mise au point de préparation d'échantillon pour augmenter la sensibilité et la résolution du signal MS. Nous avons donc fait varier les paramètres suivants : quantité de protéine déposée, type de dépôt (goutte sèche ou couche mince), tampon de solubilisation de la matrice, sonication.

La concentration protéique de Tau à déposer pour un signal MS optimal a été déterminé à partir d'une gamme de concentration de la protéine marquée à un fort ratio molaire (K/NHS-Biotine). Nous avons ainsi observé que les concentrations initiales de 4 et 8  $\mu\text{M}$  semblent appropriées pour l'analyse du niveau moyen de biotinylation.

La dispersion des fibres par rapide sonication des échantillons, notamment fibrillaires, permet d'augmenter la qualité spectrale ; elle a donc systématiquement été utilisée pour la détermination du nombre de molécules de biotine par molécule de Tau. Plusieurs matrices et types de dépôt ont été testés, en particulier le dépôt sur couche mince qui permet d'augmenter le signal MS de protéines de haut poids moléculaire analysées par MALDI-MS (Cadene and

Chait, 2000, Signor and Erba, 2013). Bien que nous ayons pu constater que la qualité spectrale peut être améliorée lorsque le dépôt est effectué sur couche mince, nous avons observé un manque de reproductibilité d'un dépôt à l'autre.

Le tampon de solubilisation de la matrice peut également jouer sur l'efficacité d'ionisation et donc sur la résolution spectrale. Une augmentation du pourcentage d'acétonitrile, notamment, favorise la solubilisation de la matrice. Ainsi, nous avons pu observer que la qualité spectrale était largement améliorée lorsque l'HCCA était solubilisée en ACN 70% et AF 5% à la concentration finale de 20 mg/mL par rapport à l'ACN 50%, TFA 0,1% à la concentration de 10 mg/mL, tampon de solubilisation classiquement utilisé pour l'analyse des protéines (**Figure 51**).



**Figure 51.** Influence du tampon de solubilisation de matrice HCCA sur la qualité spectrale. Spectres MALDI –MS d'un échantillon de Tau 1N4R marqué au ratio molaire (K/NHS-Biotine) 2,5/1 cocristallisé en HCCA (ACN 50%, TFA 0,1%, 10 mg/mL ou ACN 70%, AF 5%, 20 mg/mL).

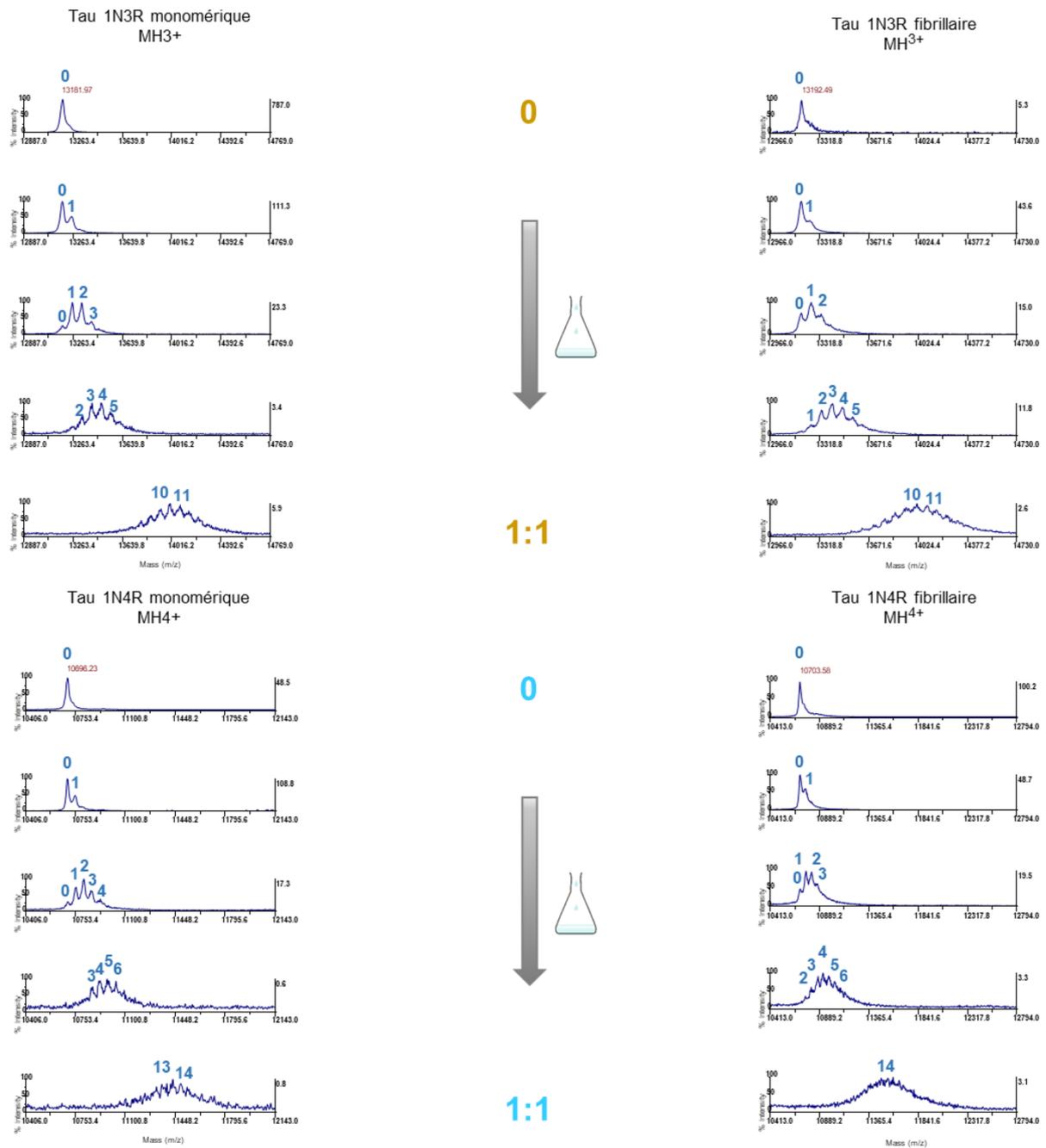
Pour conclure, l'étude du niveau de biotinylation de la protéine Tau entière monomérique et fibrillaire par MALDI-MS, pour l'optimisation des conditions de marquage, a donc été réalisée après sonication des échantillons fibrillaires et co-cristallisation de l'échantillon à la concentration comprise entre 4 et 8  $\mu$ M en HCCA solubilisé en ACN 70%, AF 5% à la concentration finale de 20 mg/mL. Par ailleurs, ce protocole optimisé a récemment été utilisé pour déterminer le niveau de marquage de fibres de Tau par la Sulfo-NHS-SS-Biotine utilisée comme hameçon pour purifier les partenaires membranaires de Tau par des approches de « pull-down » (Shrivastava *et al.* 2019).

### *B.1.3 Détermination des conditions de marquage optimales*

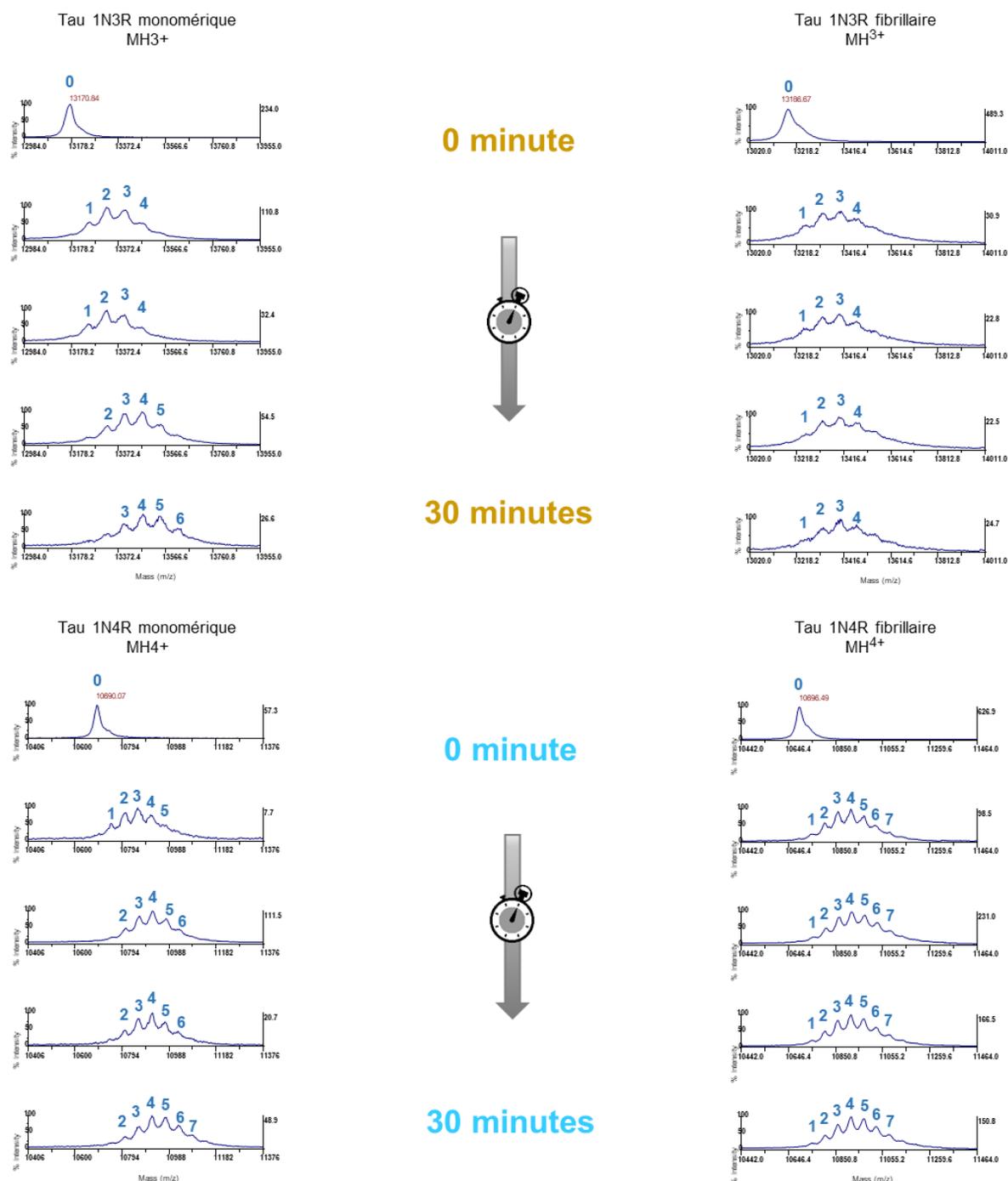
Nous avons déterminé le ratio molaire (K/NHS-Biotine) et le temps d'incubation optimaux pour un marquage de Tau 1N3R et 1N4R monomérique et fibrillaire permettant d'explorer l'accessibilité au solvant des acides aminés modifiés. Pour cela, nous avons réalisé, conjointement, une cinétique et une gamme de ratios molaires (K/NHS-Biotine) de marquage.

Dans un premier temps, nous avons analysé l'efficacité de biotinylation en fonction du ratio [K/NHS-Biotine], celui-ci évoluant de 40/1 (NHS-Biotine limitante) à 1/1 (marqueur et substrat en quantités équivalentes). Nous avons choisi de ne présenter ici que les données issues du temps d'incubation de 10 minutes (**Figure 52**). Assez logiquement, nous pouvons constater qu'à un temps d'incubation donné, le nombre moyen de molécules de NHS-Biotine par molécule de Tau augmente avec la quantité de NHS-Biotine ajoutée. A 10 minutes, le nombre moyen de biotine pour l'ensemble des ratios molaires semble inchangé entre les formes monomérique et fibrillaire pour une même isoforme. Au plus fort ratio, on peut constater que l'isoforme 1N4R est plus marquée que 1N3R.

Au vu de la gamme de ratio [K/NHS-Biotine] présentée en **Figure 52**, nous avons choisi de présenter la cinétique de 20 secondes à 30 minutes d'incubation au ratio molaire (K/NHS-Biotine) de 2,5/1 en **Figure 53**. On peut constater que le marquage de Tau 1N3R et 1N4R, sous leur forme fibrillaire, semble atteint dès 20 secondes d'incubation et est inchangé au cours du temps (respectivement 3 et 4 molécules de biotine par molécule de Tau). De façon fine, concernant la protéine monomérique, on peut aussi noter que le nombre moyen de biotines ajoutées majoritaire semble légèrement varier au cours du temps passant de 2 à 4 pour Tau 1N3R et 3 à 4-5 pour Tau 1N4R (**Figure 53**). Ceci semble suggérer que le marquage arrive à saturation pour les fibres, conformères présentant des régions structurées, moins accessibles au solvant, tandis que la saturation des sites cibles de la forme monomérique, déstructurée, n'est pas atteint jusqu'à 30 minutes d'incubation.



**Figure 52.** Mesure de l'efficacité de biotinylation de Tau 1N3R et 1N4R monomérique et fibrillaire en fonction d'une gamme de ratios molaires (K/NHS-Biotine). Tau 1N3R monomérique et fibrillaire (haut, gauche et droite respectivement) et Tau 1N4R monomérique et fibrillaire (bas, gauche et droite respectivement). Gamme de ratios molaires [K/NHS-Biotine] de 0, 40/1, 10/1, 2,5/1 et 1/1 par MALDI-MS. Ici, le temps d'incubation a été fixé à 10 minutes.



**Figure 53.** Mesure de l'efficacité de biotinylation de Tau 1N3R et 1N4R monomérique et fibrillaire en fonction d'une cinétique de marquage. Tau 1N3R monomérique et fibrillaire (haut, gauche et droite respectivement) et Tau 1N4R monomérique et fibrillaire (bas, gauche et droite respectivement). Cinétique de marquage (temps d'incubation) de 0, 20 sec, 1 min, 10 min et 30 min par MALDI-MS. Ici, le ratio molaire (K/NHS-Biotine) a été fixé à 2,5/1.

Comme évoqué plus haut, nous avons réalisé les mesures à tous les ratios [K/NHS-Biotine] et pour différents temps de réaction. Nous avons noté que pour des ratios molaires (K/NHS-Biotine) plus élevés 40/1 et 10/1 (correspondants à une quantité de biotine limitante), ce

phénomène de légère évolution du nombre moyen de biotine par molécule de Tau est aussi observé (données non montrées). Enfin, à un ratio [K/NHS-Biotine] de 1/1, cet effet est plus marqué. En effet, on observe un nombre moyen de biotines par molécule de Tau de 10 biotines pour les fibres 1N3R n'évoluant pas dans le temps alors que ce niveau de biotinylation augmente de 7 à 13 en 30 minutes pour la forme monomérique 1N3R. Il en est de même pour 1N4R ; en effet, d'une part, on compte un nombre fixe dans le temps de 14-15 biotines par molécule de Tau pour les fibres de Tau 1N4R. D'autre part, le niveau de biotinylation augmente de 11 à 17 de 20 secondes à 30 minutes de temps d'incubation en présence de marqueur pour l'espèce monomérique de Tau 1N4R (données non montrées).

Notre objectif étant de quantifier d'éventuelles différences d'accessibilité site-spécifiques entre les formes monomérique et fibrillaire pour chacune des deux isoformes de Tau, nous avons choisi un ratio molaire (K/NHS-Biotine) et un temps de réaction pour lesquels le nombre moyen de biotine par molécule de Tau soit limité. De cette façon, nous voulions éviter d'induire des éventuels changements de conformation, mais que ce niveau de biotinylation soit suffisant pour que les peptides biotinylés soient identifiables et quantifiables. Par conséquent, les conditions retenues pour notre analyse des surfaces accessibles des fibres de Tau 1N3R et 1N4R présentée dans le Chapitre III.2 (article) correspondent à un temps de réaction de 10 minutes et un ratio molaire (K/NHS-Biotine) de 2,5/1.

## B.2 Choix des conditions de digestion des protéines biotinylées

Dans le cadre de notre approche de marquage covalent, la digestion de Tau 1N3R et 1N4R monomérique et fibrillaire a été réalisée dans deux conditions :

- (1) Conditions dénaturantes, en gel.
- (2) Conditions natives, en solution

Ainsi, les produits de la digestion en conditions dénaturantes dépendent de l'accessibilité de surface des résidus au solvant tandis que les produits de la digestion en conditions natives dépendent de deux paramètres, i.e l'accessibilité à l'enzyme et l'accessibilité de surface au solvant.

L'enzyme utilisée pour ces digestions a été choisie de manière à obtenir le meilleur recouvrement de séquence possible. Malgré sa grande efficacité, nous avons décidé de ne pas utiliser la trypsine dans le cadre de nos travaux. En effet, les résidus sujets à la dérivatisation par la NHS-Biotine étant principalement les lysines (et le résidu N-terminal), le marquage d'une lysine empêchera la trypsine de couper et introduirait un biais dans l'interprétation de la quantification des peptides marqués (Mendoza and Vachet. 2009).

Pour ne pas introduire de biais au niveau de la digestion, nous avons utilisé la GluC pour la digestion en solution et en gel. Pour la digestion en conditions natives, on observe un recouvrement de séquence variable entre les formes monomériques et fibrillaires de Tau 1N3R et 1N4R (en raison des différences d'accessibilité à l'enzyme) tandis qu'on restaure un recouvrement de séquence d'environ 80% pour l'ensemble des échantillons en conditions dénaturantes (Chapitre III.2 Article). Pour les digestions en gel, nous avons constaté que les peptides marqués de certaines régions d'intérêt majeur, notamment localisés dans la seconde répétition 4R-spécifique, n'étaient pas détectés. Ceci pourrait notamment être imputé à une moins bonne activité de la GluC dans le gel par rapport à la protéolyse en solution et/ou à une extraction moins efficace de ces peptides. Pour remédier à cela, nous avons complété le recouvrement de séquence par une digestion en gel par l'AspN.

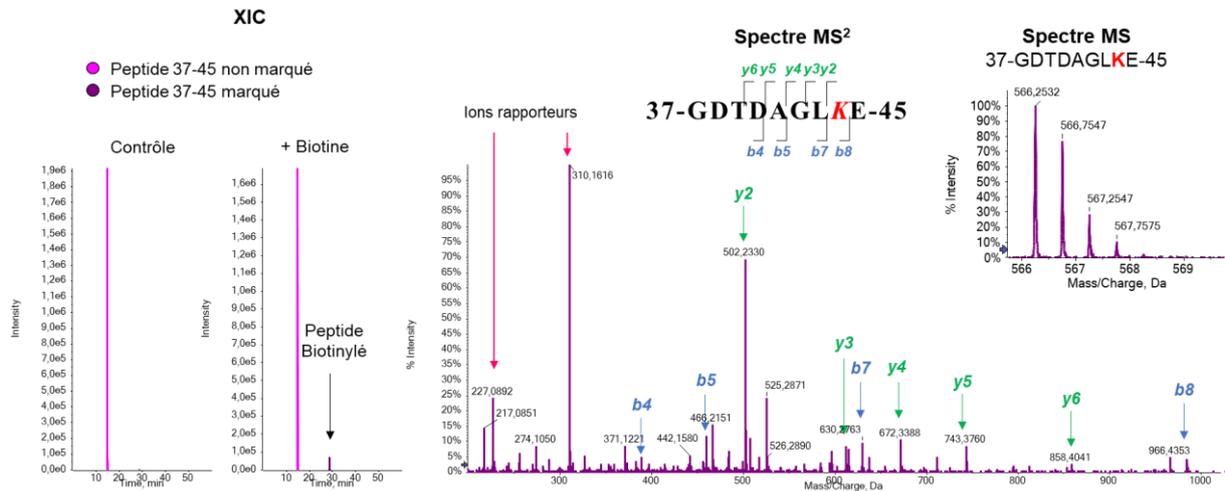
De façon intéressante, nous avons pu observer une homologie des tendances entre les peptides marqués et non marqués issus respectivement de la digestion native par la GluC des échantillons marqués et celle des échantillons contrôle. Ceci renforce l'idée que le marquage par la NHS-Biotine n'engendre pas de changement conformationnel majeur au niveau de la structure des fibres (cf. Article Figure 2 et 4).

### B.3 Identification et quantification des peptides biotinylés

#### B.3.1 Identification des peptides biotinylés

Dans un premier temps, tous les peptides fragmentés sont identifiés grâce au moteur de recherche MASCOT avec une tolérance de masse possible de 40 ppm pour les précurseurs et de 0,05 Da pour les ions fragments. Nous avons recherché les modifications suivantes : biotinylation (+226,08 Da) du résidu N-terminal et des cystéines, sérines, lysines et thréonines. Cette recherche nous a permis de dresser une liste de peptides non marqués et biotinylés à partir des données obtenues pour les monomères. En effet, le monomère étant *a priori* moins ou peu structuré, il présente davantage de lysines accessibles potentiellement biotinylés. Cette liste de peptides aussi identifiés avec un score MASCOT supérieur à 13 a permis non seulement de déterminer le recouvrement de séquence de Tau mais aussi de dresser une liste de peptides biotinylés à quantifier en intégrant leur m/z, leur état de charge, leur temps de rétention de manière à extraire leur intensité (XIC) pour chacune des expériences de marquage. Pour certains peptides, cette liste a été complétée à partir de l'analyse des échantillons fibrillaires. Ultérieurement, lors de la quantification, ces spectres MS<sup>2</sup> et le temps de rétention associé nous servent de contrôle pour la bonne identification du peptide extrait. Pour ce faire, l'avantage de l'étiquette Biotine a été utilisé. En effet, cette étiquette génère des

ions rapporteurs signatures dans le spectre MS<sup>2</sup> (m/z 227,08 et 310,16 ion immonium de la lysine biotinyllée) (Dreger *et al.* 2009, Udeshi *et al.* 2017, Rey *et al.* 2018). Ces ions fragments signatures permettent de rapidement vérifier la biotinylation des peptides marqués identifiés, notamment pour ceux présentant un score relativement faible. La **Figure 54** illustre l'identification d'un peptide biotinyllé.



**Figure 54.** XIC, spectres MS et MS<sup>2</sup> d'un peptide biotinyllé. De gauche à droite, XIC du peptide 37-45 marqué et non marqué au sein d'un échantillon de Tau 1N4R Monomérique Contrôle et biotinyllé ; Spectres MS<sup>2</sup> et MS du peptide 37-45 marqué.

A partir de cette liste complète de peptides biotinyllés, nous avons procédé à la quantification des peptides marqués.

### B.3.2 Quantification des peptides biotinyllés

Au cours de nos travaux, nous avons procédé à une quantification relative sans marquage, dite « label free » des peptides biotinyllés entre les différents types d'échantillon : Tau 1N3R monomérique et fibrillaire et Tau 1N4R monomérique et fibrillaire. La quantification relative « label free » repose sur soit le XIC (comparaison de l'intensité des signaux MS d'un même peptide entre différents échantillons, ici fibrillaire et monomérique) ou le « Spectral Counting » (comparaison du nombre d'évènements de fragmentation d'un même peptide entre différents échantillons). La quantification relative par XIC étant plus précise, en particulier pour des peptides peu abondants, nous avons donc extrait l'intensité du courant ionique pour chaque peptide de notre liste et intégré l'aire sous le pic pour chaque peptide. Cette extraction a été réalisée avec l'outil XIC Manager de PeakView, avec une attention particulière pour les nombreux peptides isobares. La seconde approche n'a pas été adoptée en raison d'un certain

nombre de limitations. D'une part, bien qu'il ait été montré une corrélation entre le nombre d'évènements de fragmentation et l'abondance de la protéine, la linéarité, compte tenu de la gamme dynamique (notamment pour les espèces saturantes), reste questionnable (Gao *et al.* 2005 ; Bantscheff *et al.* 2012) ; ce qui n'est pas le cas du XIC (Silva *et al.* 2005).

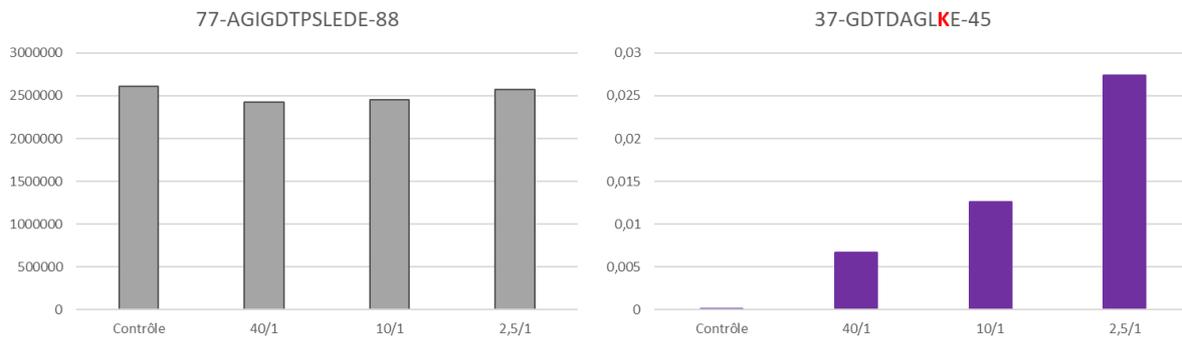
Nous avons ensuite procédé à la normalisation de l'intensité des peptides biotinylés, notamment pour réduire au maximum l'impact d'une différence éventuelle de quantité de protéine Tau entre les échantillons. Deux principales méthodes de normalisation se présentaient à nous : une normalisation par rapport à la valeur totale de coups d'ions de l'acquisition (équation 3) ou une normalisation par rapport à un peptide interne (Barthélemy *et al.* 2016) remplissant certaines caractéristiques (équation 4). La première méthode est utilisée dans de nombreuses études protéomiques, notamment de mélanges protéiques complexes. Dans le cas de nos travaux, le mélange peptidique est relativement simple (une protéine unique) mais sa population peptidique peut potentiellement différer significativement d'un échantillon à l'autre. En effet, alors que l'échantillon Contrôle ne comprend que des peptides non marqués, les autres échantillons comptent de nombreux peptides biotinylés. L'addition d'une biotine à un peptide peut en changer sa capacité d'ionisation et, par conséquent, influence la valeur de coups d'ions totale alors que la quantité de matériel n'en est pas à l'origine. Ceci conduirait donc à un biais dans la normalisation. La seconde méthode consiste à identifier un ou des peptide(s) interne(s) à la protéine qui remplisse(nt) les critères suivants : (1) ne contient pas de lysine, (2) possède une bonne capacité d'ionisation (peptide intense), (3) se situe dans une région qui n'a jusqu'à ce jour pas été décrite comme hautement structurée dans la littérature (nécessaire pour la digestion en conditions natives), (4) ne peut être sujet qu'à peu ou pas de coupure incomplète et (5) dont l'intensité varie peu d'un échantillon à l'autre. Étonnement, au vu de la répartition des lysines le long de la séquence, peu de peptides candidats sont disponibles. Un d'entre eux nous permet toutefois de normaliser nos données ; il s'agit du peptide 77-AGIGDTPSLEDE-88 situé dans la partie N-terminale de la protéine.

$$(3) \quad \frac{\text{Intensité du peptide biotinylé}}{\sum(\text{Intensités de tous les peptides})}$$

$$(4) \quad \frac{\text{Intensité du peptide biotinylé}}{\text{Intensité du peptide standard}}$$

De cette manière, nous avons normalisé l'ensemble des valeurs d'intensités extraites des peptides biotinylés et pu comparer l'intensité d'un peptide donné entre les formes monomérique et fibrillaire d'une isoforme de Tau donné. Un rapide examen de l'intensité des peptides biotinylés issus d'échantillons incubés à différents ratios molaires (K/NHS-Biotine)

pendant un laps de temps donné nous a permis de constater que l'intensité des peptides biotylés augmente avec la concentration de NHS-Biotine, comme illustré en **Figure 55**.



**Figure 55.** Intensité du peptide standard 77-88 dans différents échantillons et intensité normalisée d'un peptide biotylé en fonction des différents ratios [K/NHS-Biotine]. Intensité du peptide standard dans les échantillons monomériques 1N4R non marqué et marqués aux différents ratios molaires [K/NHS-Biotine]. Augmentation de l'intensité du peptide marqué 37-GDTDAGLKE-45 (Mth 1130,49 ; MH2+ 566,25) normalisée par rapport au peptide interne de référence pour l'échantillon Tau 1N4R monomérique marqué pendant 10 minutes.

Cette augmentation, fonction du ratio [K/NHS-Biotine], de l'intensité normalisée des peptides marquées suggère bien que l'augmentation de l'intensité est corrélée au degré de modification et donc à l'accessibilité de surface des résidus au sein des conformères.

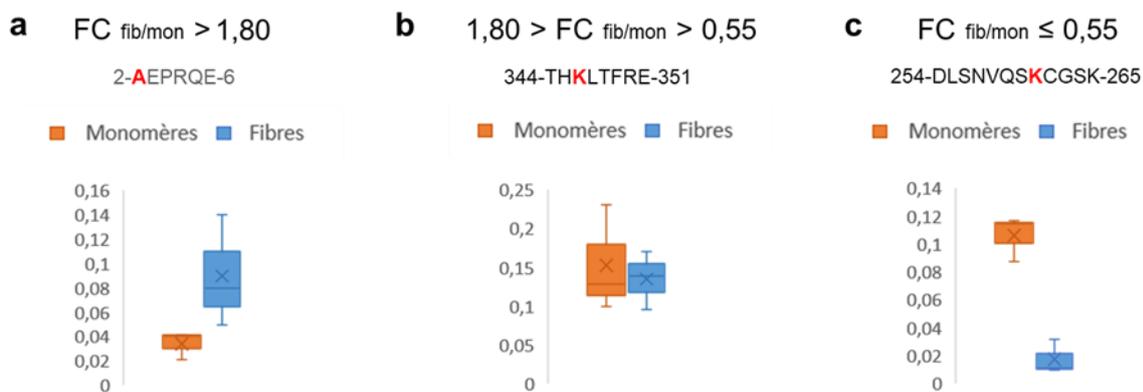
Deux approches nous étaient offertes pour effectuer la comparaison des données normalisées de quantification des échantillons monomériques et fibrillaires. La première est notamment utilisée pour l'étude différentielle de l'occurrence de modifications post-traductionnelles (acétylation et méthylation) au niveau d'histones (O'Connor *et al.* 2014) ou encore pour l'étude de l'accessibilité de surface par hydroxylation radicalaire (Gau *et al.* 2013). Il s'agit du calcul du taux de modification régi par l'équation 5. Cependant, ce type de calcul présente quelques limitations. En effet, certaines modifications peuvent conduire à un changement de l'efficacité d'ionisation des peptides marquée et ainsi influencer l'abondance mesurée par rapport au peptide correspondant non marqué. Ce phénomène peut d'ailleurs être exacerbé par la chromatographie liquide en amont : les peptides marqués et non marqués sont élués à des temps de rétention différents et donc dans des compositions de solvant et en présence de peptides différents ; pouvant ainsi modifier davantage l'efficacité d'ionisation (Mendoza and Vachet. 2009). Pour cette raison, et en adéquation avec la question biologique posée, nous avons préféré opter pour un calcul de « fold-change » de l'intensité d'un peptide marqué pour la forme fibrillaire et l'intensité du même peptide marqué pour la forme monomérique d'une même isoforme (équation 6). Comme explicité dans la partie introductive, la protéine Tau monomérique étant très majoritairement destructurée, cette relation mathématique nous

permet de mettre en évidence des différences d'accessibilité conformationnelles entre les formes monomérique et fibrillaire pour chaque isoforme de Tau, 1N3R et 1N4R.

$$(5) \quad \text{Taux de biotinylation} = \frac{\text{Intensité du peptide biotinylé}}{\sum(\text{Intensité du peptide biotinylé}, \text{Intensité du peptide non biotinylé})}$$

$$(6) \quad FC_{\frac{fib}{mon}} = \frac{\text{Moyenne (Intensité du peptide biotinylé normalisée)}_{\text{Tau fibrillaire}}}{\text{Moyenne (Intensité du peptide biotinylé normalisée)}_{\text{Tau monomérique}}}$$

Afin d'être stringent et de révéler des signatures conformationnelles isoforme-spécifiques majeures qui permettent de distinguer de façon robuste les agrégats de Tau 1N3R et 1N4R, nous avons fixé une limite de fold-change suffisamment élevée ou basse soit  $FC_{\text{fib/mon}} > 1,8$  et  $FC_{\text{fib/mon}} < 0,55$  (**Figure 56**). Un second seuil de fold-change a été considéré : celui pour lequel la différence d'intensités normalisées entre formes fibrillaire et monomérique est supérieure ou égale à 10 fois. Ce second seuil reflète une diminution drastique de l'accessibilité de surface des résidus concernés après assemblage.



**Figure 56.** Box-plots représentant les trois catégories de peptides biotinylés associés à différents « fold-changes » entre formes fibrillaire et monomérique de Tau 1N4R. (a)  $Fc > 1,8$  (Fibrilles > Monomères), (b)  $1,8 > Fc > 0,55$  (Fibrilles  $\approx$  Monomères) et (c)  $Fc \leq 0,55$  (Fibrilles < Monomères).  $n=3$

Enfin, nous avons appliqué un test statistique non paramétrique de type Kruskal-Wallis pour déterminer la significativité de la différence. Le choix de ce test a été conditionné par le nombre de réplicats (trois expériences d'assemblage et de marquage indépendantes), qui ne nous permettait pas de considérer un autre test. Le test de Kruskal-Wallis n'est pas dépendant des hypothèses de normalité et d'homoscédasticité, contrairement à l'ANOVA, et est basé sur un classement de rang. Ce test, bien que moins puissant que d'autres tests statistiques nécessitant de nombreux réplicats, nous a semblé le plus adapté à notre étude et la plus pertinent.

Cette méthode optimisée de cartographie des surfaces accessibles et protégées a été appliquée à l'analyse des surfaces des fibres de Tau 1N3R et 1N4R. Ces résultats font l'objet du Chapitre III.2 Article.

### **III.2 Application aux fibres de Tau (article)**

Les optimisations des méthodes décrites précédemment ont été appliquées à l'analyse des fibres recombinantes 1N3R et 1N4R.

L'ensemble de ce travail est présenté dans l'article qui suit.



# **Protein surface accessibilities reveal isoform-specific signatures of 1N3R and 1N4R Tau assemblies**

Emilie Caroux, Karine Madiona, Luc Bousset, Virginie Redeker\* and Ronald Melki\*

\* *Corresponding authors: [virginie.redeker@cnr.fr](mailto:virginie.redeker@cnr.fr), [ronald.melki@cnr.fr](mailto:ronald.melki@cnr.fr)*

Institut François Jacob, Molecular Imaging Center (MIRCen), Commissariat à l'Energie  
Atomique et aux Energies Alternatives (CEA), CNRS UMR 9199, Laboratory of  
Neurodegenerative Diseases, Université Paris-Sud, Université Paris-Saclay, Fontenay-aux-  
Roses, France.

## **Abstract**

Intracellular deposits of Tau protein aggregates are the common hallmark of tauopathies, a range of neurodegenerative diseases including Alzheimer's disease. Levels of tau with three (3R) or four (4R) microtubule binding repeats are found similar in the normal adult brain, whereas they differ in neuropathological intracellular Tau inclusions, according to the type of tauopathy. Conformation of the rigid core of Tau filaments was shown to be different in Alzheimer's (with equivalent 3R/4R Tau ratios) and Pick's brains (with 3R Tau), as reported by cryo-electron microscopy. Here, we used proteomic approaches to determine common and different structural molecular determinants of recombinant *in vitro* assembled 1N3R and 1N4R human Tau fibrils. Using limited proteolysis we identify the proteolytic fragments composing the molecular "bar-code" for each type of fibril. Optimization of a chemical covalent surface modification approach using NHS-Biotin combined with mass spectrometry, we mapped the regions on the surface of 1N3R and 1N4R Tau fibrils that are either accessible or protected. We show that our recombinant Tau fibrils are different from those used for the recent high-resolution structure. We observed that, while the N-terminal part of the protein remains accessible within 1N3R and 1N4R fibrils, the C-terminal region is protected within 1N3R yet solvent accessible for 1N4R assemblies. The proteomic approach we implemented might now be applied to Tau aggregates from 3R-tau and 4R- tau pathologies. Furthermore our results pave the way to the design of new molecular tools targeting the C-terminal domain to discriminate 4R from 3R Tau aggregates.

## **Keywords:**

Tau fibrils; tauopathies; protein conformation; structural proteomics; mass spectrometry; limited proteolysis; chemical surface modification

Intracellular deposits of Tau protein aggregates are the common hallmark of tauopathies, a range of neurodegenerative diseases including Alzheimer's disease (Lee *et al.* 2001; De Silva *et al.* 2003; Kovacs. 2015). The microtubule-associated Tau protein regulates assembly and stabilization of the neuronal microtubule network (Weingarten *et al.* 1975). In the normal adult human brain, six tau isoforms can be generated by alternative mRNA splicing of the single MAPT tau gene (Goedert *et al.* 1989; Liu and Gong, 2008). Exons 2 and 3 encode two N-terminal domains, involved in modulation of Tau interaction with membranes and tyrosine kinases, and define 0N, 1N or 2N tau isoforms that differ by the absence or presence of one or two of these exons. Presence or absence of exon 10, encoding the second of four microtubule-binding repeat domain (R2), located in the C-terminal domain of Tau, define tau isoforms containing all four (4R) or only three (3R) microtubule-binding repeat domains. Similar levels of 3R and 4R isoforms are expressed in the normal adult human brain, with 1N3R and 1N4R Tau isoforms being the most represented (Trabzuni 2012). However, in intracellular tau inclusions of distinct neuronal subpopulations, composition in 3R or 4R Tau aggregates differ according to the type of tauopathy (Buée and Delacourte. 1999; Lee and Leugers 2012; Sergeant *et al.* 2005; Goedert and Spillantini. 2017). As a consequence, tauopathies can be divided into at least three major neuropathological categories characterized by either predominant 4R-tau (progressive supranuclear palsy PSP; Corticobasal degeneration, CBD; Argyrophilic grain disease, AGD), or predominant 3R-tau (Pick's disease, PiD), or mixed 3R-/4R-tau (Alzheimer's Disease, AD; Tangle Disease (TD) (Lee *et al.* 2001) pathologies. Additionally, not only Tau isoform composition, but also cellular and regional distribution, and morphology of Tau inclusions vary according to the tauopathy (Buée and Delacourte. 1999; Williams. 2006 ; Park *et al.* 2016; Lee and Leugers 2012; Sergeant *et al.* 2005). Additionally, it has been shown that distinct tau strains propagate distinct fibrillar conformations *in vitro* and *in vivo*, inducing distinct tauopathies (Sanders *et al.* 2014; Kaufman *et al.* 2016; Guo *et al.* 2016; Narasimhan *et al.* 2017; Takeda *et al.* 2015). Thus, progressive spread of Tau inclusions from cell to cell, in a "prion-like" way, might explain the clinical progression of tauopathies.

All together, this suggests that distinct Tau strains might propagate distinct tauopathy-specific fibril conformations. Thus, in order to distinguish tauopathies and design tools to inhibit their progression, there is a need to identify structural molecular determinants of 3R and 4R tau aggregates, in particular those exposed at the surface of distinct Tau fibrils.

Limited proteolysis of AD and PiD patients-derived aggregates have revealed an accessible fuzzy N-terminal domain and a protease resistant core composed of the microtubule-binding domains, starting in the first R1 repeat, and reaching the C-terminus of Tau, either to the 20 first aminoacids of the C-terminal flanking region using pronase (Wischnik *et al.* 1988; Novak *et al.* 1993) or to the C-terminus using trypsin (Von Bergen *et al.* 2006).

Recently, high resolution cryo-electron microscopy (cryo-EM) structures of the protease resistant core of distinct Tau filaments purified from the brain of individuals with either Alzheimer's disease (Falcon *et al.* 2018b; Fitzpatrick *et al.* 2017) or Pick's disease (Falcon *et al.* 2018) have shown that have a different fold. For AD purified Tau filaments, the protease resistant core correspond to the R3 and R4 repeats together with ten additional amino acids at the beginning of the C-terminal flanking region, whereas for Pick's purified Tau filaments, the core is longer and includes the 20 C-terminal amino-acids of R1, in addition to the R3 and R4 repeats and the ten residues located after the R4 repeat (Falcon *et al.* 2018 ; Fitzpatrick *et al.* 2017). Furthermore, filaments of AD patients are immuno-gold decorated by antibodies of the R1 and R2 repeats before but no more after pronase treatment, suggesting an accessibility of this region (Falcon 2018b). Finally, cryo- and immuno-EM of Tau fibrils from several sporadic and inherited Alzheimer's diseases show that there is no significant variation in the structure of tau fibrils between AD individuals (Falcon *et al.* 2018b). But cryo and immuno-EM of fibrils assembled from recombinant full-length human 2N3R and 2N4R tau in presence of heparin, used in an increasing number of *in vitro* and *in vivo* studies post-injection, administration of or exposure to Tau fibrils (Références), have reported that they fold differently, with shorter pronase-resistant cores composed only of R2-R3 or R3 only, compared to AD or PiD aggregates respectively (Zhang *et al.* 2019). As cryo-EM describes ultra-structured non-

dynamic regions (Lyumkis. 2019), there is still a need to elucidate solvent accessible Tau surfaces.

Here, we used two complementary proteomic approaches to determine structural molecular determinants of full-length 1N3R and 1N4R Tau fibrils assembled *in vitro* in presence of heparin. Using limited proteolysis with GluC we identified a 1N3R and 1N4R specific proteolytic profile or molecular "bar-coding". With chemical covalent surface modification using NHS-biotin combined with mass spectrometry, we mapped the regions on the surface of 1N3R and 1N4R Tau fibrils that are either solvent-accessible or protected, and provide new structural insights on Tau fibrils made of full-length 1N3R or 1N4R. Our results highlight that our recombinant Tau fibrils present a different fold than the previous reported structure of recombinant 2N3R and 2N4R fibrils and more common hallmarks with Tau aggregates from AD or PiD brain patients. Furthermore, we observe interesting similarities and differences in the surface exposed regions of 1N3R and 1N4R Tau fibrils. First, we show for both 1N3R and 1N4R fibrils, that the N-terminal part of Tau remains globally accessible, whereas the C-terminal region presents a large protected region from the first (R1) to the middle of the fourth (R4) microtubule binding domain. Second, we identified some conformation differences between 1N3R and 1N4R aggregates, with the major one located in the C-terminal tail, including the last part of R4 and the C-terminal flanking region: this region composed of 22 amino acids is globally protected in 1N3R fibrils and accessible in 1N4R fibrils. Our results illustrate how structural proteomic approach can compare solvent-accessible and protected surfaces of distinct Tau fibrils, and opens the way towards the design of new molecular tools which could discriminate 3R from 4R Tau aggregates or even neutralize them.

## Results

### **Recombinant 1N3R and 1N4R hTau proteins are assembled into fibrils in in presence of heparin**

To date, most studies performed on Tau filaments investigate the structure of the protease resistant domains, but only very little is known about the surfaces of Tau filaments. As the amino-acid residues exposed at the surfaces of Tau fibrils might allow to distinguish distinct Tau fibrils, they also control the interactions they establish with numerous cellular components including proteins, phospholipids, membranes, extracellular matrices. In order to map the solvent accessible surfaces of 3R and 4R Tau filaments, in complement to the high-resolution structures of the protease resistant cores recently reported (Fitzpatrick *et al.* 2017; Falcon *et al.* 2018; Zhang *et al.* 2019) we set up complementary structural proteomics approach using recombinant 1N3R and 1N4R Tau proteins.

We have expressed and purified full-length 1N3R and 1N4R hTau, (Barghorn *et al.* 2005; Tardivel *et al.* 2016) corresponding to the most abundant Tau isoforms expressed in adult brain, representing respectively 26 and 23 % of Tau proteins in the adult human cortex (Trabzuni *et al.* 2012). We further assembled them into fibrils in presence of heparin, as previously described (Tardivel *et al.* 2016). Quality of assembly of monomeric 1N3R and 1N4R Tau into fibrils was controlled by ThioflavinT binding, sedimentation of the fibrils produced after five days of assembly (not shown) and transmission electron microscopy (TEM) (Figure 1 a and b). After five days of assembly, both 1NR3 and 1N4R Tau fibrils were collected by sedimentation, and concentration of Tau proteins were adjusted to 40 $\mu$ M (equivalent monomer concentration) for both monomeric and fibrillary Tau proteins, as described in material and methods. Using transmission electron microscopy (TEM), no significant difference in morphology between 1N3R and 1N4R Tau fibrils was visible at this low resolution (Figure 1 a, b).

We further analyzed soluble monomeric and insoluble fibrillar 1N3R and 1N4R tau samples in order to document the changes in solvent accessibility upon assembly into fibrils, using two

complementary structural proteomics methods using in particular limited proteolysis and chemical residue labeling combined with mass spectrometry.

### **Limited proteolysis of 1N3R and 1N4R Tau fibrils identifies accessible surfaces**

Limited proteolysis provides information on the regions of Tau that are exposed to the solvent and accessible to a proteolytic enzyme. First cleavage sites of the protein, using short incubation times of the protein with low concentrations of a proteolytic enzyme, are restricted to the surfaces of the proteins that are the most accessible to the active site of the protease, whereas the overall quaternary and tertiary structure are maintained. Proteolysis has been commonly used on amyloid aggregates to identify isoform-specific differences of the protected domains for distinct Tau assemblies (Von Bergen *et al.* 2006; Hasegawa *et al.* 2014). Most of the proteases generally used have low specificities as pronase or proteinase K, few are specific as trypsin or V8 GluC (Hasegawa *et al.* 2014; von Bergen *et al.* 2006; Zhang *et al.* 2019; Fitzpatrick *et al.* 2017). To date the first times of limited proteolysis have not been explored in detail in order to identify protease-accessible surfaces of Tau filaments.

We have used seven specific proteolytic enzymes (Trypsin, Chymotrypsin, GluC, AspN, LysC, LysN, ArgC) with theoretical cleavage sites distributed along the amino acid sequence of tau (data not shown). After separation of the long proteolytic products by SDS-PAGE, we have selected endoprotease GluC for detailed characterization of the cleavage sites using an enzyme-to-Tau ratio of 1:150 (w/w) and analysis of the proteolytic fragments after 10, 45, 120 and 300 min digestion, for both monomeric and fibrillary 1N3R and 1N4R Tau (Figure 1 c-f). No major difference in the proteolytic fragments generated from monomeric 1N3R and 1N4R Tau monomers was observed (Figure 1 c,d), confirming that the solvent accessibility is similar for 1N3R and 1N4R monomers with complete degradation of intact full-length Tau as soon as 10 min of proteolysis. Interestingly, the proteolytic profiles produced from 1N3R or 1N4R Tau fibrils are different (Figure 1 d,e). After 45 min of digestion with GluC, we observe a clear and distinct proteolytic fragment pattern for 1N3R and 1N4R fibrils, highlighting a specific proteolytic 3R- and 4R- molecular code bar (Figure 1 g,h) reflecting differences in conformation

and surface accessibilities of GluC proteolytic sites and not in primary sequences. Three abundant proteolytic fragments with apparent molecular weights between 35 and 55 kDa are observed for 1N3R Tau fibrils, whereas six products with apparent molecular weights between 25 and 60 kDa are observed for 1N4R Tau fibrils. Among the other proteolytic enzyme tested, interesting differences in the proteolytic profiles between Tau 1N3R and 1N4R fibrils were observed with Chymotrypsin in the first times of digestion, but the number of bands generated was low, due to the presence of only eight Chymotrypsin cleavage sites compared to twenty four GluC cleavage sites the sequence of Tau (data not shown). Interestingly, the proteolytic fragments with molecular weights above 20 kDa, observed for both GluC and Chymotryptic proteolysis also were observed after sedimentation of the so-digested Tau fibrils (data not shown). This indicates that the fibrillar conformation has been preserved after limited proteolysis.

We further identified the N- and C-termini of the distinct proteolytic fragments composing the GluC proteolytic code bar, as they correspond to the glutamate residues of Tau that are accessible at the surface of the fibrils using (Figure 1 i and j). The exact N-terminus of each proteolytic product observed at 45 min was identified by both Edman sequencing and the N-TOP (N-Terminal Oriented Proteomics) strategy based on chemical derivatization of the N-terminus of the polypeptides fragments isolated on SDS-PAGE band, with TMPP (N-succinimidylloxycarbonylmethyl)- tris(2,4,6-trimethoxyphenyl)- phosphonium bromide) before classical in-gel digestion and mass spectrometric identification (Bertaccini 2013). The data summarized in figure 1 i and j, highlight that at initial GluC digestion times (10 and 45 min), fragments with identical N-termini at residues F8, A74, A77, D87 and D89 are generated from cleavages of 1N3R and 1N4R fibrils at residues E7 (protein bands 1 and 2 for both 1N3R and 1N4R), E73 and E76 (protein bands 3 for 1N3R, and 3 and 5 for 1N4R), and E86 and E88 (protein bands 4 for 1N3R, and 4 and 6 for 1N4R).

This suggests that 1N3R and 1N4R tau fibrils share similar accessible epitopes in the N-terminal portion, with an accessible portion around Glu-8 and folds in the structure that protect residues E9, E12, E36, E45, E53, E57, E58, E62 from very high accessibility to the enzyme

and rapid proteolysis. Interestingly, E73, E76, E86 and E88 represent a stretch of at least sixteen residues that are accessible at the surface of both 1N3R and 1N4R Tau fibrils. Immunocharacterization by western-blot analysis of the fragments was assessed using antibodies specific to the N-terminus (N-ter), the C-terminus (C-ter), internal sequences from the first half part of the protein (HT7 and Tau5) and the sequence specific to 3R and 4R Tau sequences (3R and exon10 antibodies, respectively). Immunocharacterization confirmed the N-terminal identifications and revealed differences in enzyme accessibility in the C-terminal part of the proteolytic fragments generated from 1N3R and 1N4R Tau fibrils.

For Tau 1N3R fibrils, all abundant GluC fragments contain the epitope of the Cter antibody, corresponding to the fifteen C-terminal residues (bands 1, 3 and 4), whereas for 1N4R Tau fibrils three out of six abundant proteolytic fragments at 45 min lack the Cter epitope (bands 2, 5 and 6). This indicates that the C-terminus is more accessible in 1N4R than in 1N3R fibrils.

We further deduced the C-terminal residue of the proteolytic fragments, by analyzing the sequence coverage after in-gel tryptic digestion for each fragment band of interest and measured the masses of the fragments present in the GluC reaction mixtures by MALDI MS. For 1N3R fibrils, we detected two preponderant C-termini corresponding to E371 and the C-termini at L381, whereas for 1N4R we detected three preponderant C-termini corresponding to E402, the C-termini at L412 and E343. Similar observation was made with chymotrypsin cutting at F349 for 1N4R but not 1N3R fibrils.

These limited proteolysis data strongly suggest that 1N3R and 1N4R Tau fibrils present isoform-specific conformational differences in the C-terminal region of Tau, with a highly protease accessible surface for 1N4R Tau fibrils, around residue E343 and F349, whereas this region appears protected in Tau 1N3R fibrils

### **Advanced proteolysis identifies protected fibrillar Tau surfaces.**

We characterized more deeply the solvent protected surfaces of 1N3R and 1N4R Tau fibrils, using an advanced GluC proteolysis performed in-solution, during 15h with an enzyme-to-protein ratio of 1/50 (w/w). Advanced proteolysis was confirmed by the absence of any

proteolytic fragment at molecular weights higher than 15 kDa, visible by SDS-PAGE. In order to determine the changes in enzyme accessibility upon assembly into fibrils, we digested both Tau monomers and fibrils, and focused on the identification and quantification of short GluC proteolytic peptides, containing less than 45 amino acids, by nanoLC-MSMS analysis. First, proteolytic peptides were identified by tandem mass spectrometry and a database search using the 1N3R and 1N4R Tau sequences. We obtained a sequence coverage of 75% and 74% for monomeric 1N3R and 1N4R Tau respectively, and only 48% and 62% for fibrillar 1N3R and 1N4R Tau respectively (Fig.S1.). This qualitative analysis suggests that 1N3R fibrils exhibit a higher number of solvent protected GluC cleavage sites than 1N4R fibrils.

We used this list of identified GluC proteolytic peptides to further quantified each peptide by integrating the XIC intensity (Extracted Ion Current) of identical GluC peptides ion peaks in the ion chromatograms, matched across samples and replicates, using their MS2 spectrum, m/z and retention time. To visualize in detail the variation in enzymatic cleavage accessibility, we performed experiments in triplicate, and normalized the MS intensity of each GluC peptide to the total amount of Tau using the GluC peptide [77-88] as a reference of the total level of Tau. The rationale of this normalization was the use of a peptide that was highly accessible in both 1N3R and 1N4R Tau fibrils, as indicated by our limited proteolysis experiments.

Changes in enzyme accessibility of Tau surfaces upon assembly of monomers into fibrils was measured by the calculation of the fold changes ( $F_c$ ) of peptide ion intensities as a ratio between the normalized MS intensity of a GluC proteolytic peptide generated from fibrillary Tau and the normalized MS intensity of the equivalent peptide generated from monomeric Tau ( $F_{c\ F/M\ peptide} = Intensity_{Fib\ peptide} / Intensity_{Mono\ peptide}$ ). We determined two categories of peptides significantly less produced from fibrils than from monomers: the first category of protected region is at least 1.8 times less accessible within the fibrils ( $0.1 < F_{c\ F/M} < 0.55$ , in blue), the second is at least 10 times less produced or not even detected after proteolysis of fibrils ( $F_{c\ F/M} \leq 0.1$ , in red). Peptide intensities and the corresponding  $F_{c\ F/M}$  fold changes are given in Tables S1 and S2. Significant protection to enzyme accessibility upon assembly into fibrils are represented colored along the 1N3R and 1N4R amino acid sequences, while grey peptides

show regions accessible to the enzyme upon assembly (Figure 2). Changes measured for each individual peptide are presented in boxplots in Figure S7a and S7b. For sake of the comparison Figure S7 presents data following the amino acid sequence numbering of the 1N4R sequence. In the N-terminal region, from the N-terminus to residue E8, and between residues E36 and E88, no significant difference in enzyme accessibility upon assembly of both 1N3R and 1N4R Tau. Interestingly, the region from D13 to E36 appears protected from enzyme digestion for both 1N3R and 1N4R Tau, as indicated by the  $F_{c\ F/M}$  of the proteolytic peptides [13-25] and [13-36] ( $F_{c\ F/M\ 1N3R\ [13-25]} = 0.45$ ,  $F_{c\ F/M\ 1N3R\ [13-36]} = 0.36$ ;  $F_{c\ F/M\ 1N4R\ [13-25]} = 0.23$ ,  $F_{c\ F/M\ 1N4R\ [13-36]} = 0.18$ ). These results are in accordance with the high accessibility observed by limited proteolysis at E8, E73 and E76, and E86 and E88. Further in the N-terminal region, we identified a sequence with a remarkable protection to enzyme accessibility in 1N3R fibrils, significantly less pronounced for 1N4R fibrils, as highlighted by peptide [89-104] ( $F_{c\ F/M\ 1N3R\ [89-104]} = 0.02$ ;  $F_{c\ F/M\ 1N4R\ [89-104]} = 0.62$ ). As limited proteolysis showed a similar accessibility of residues E86 and E88, this difference suggests that the region around residue E104 may be more protected in 1N3R than in 1N4R fibrils.

The second part of the protein, starting at the end of the Proline rich region (from E193) exhibit a high level of solvent-protection to enzyme accessibility in both 1N3R and 1N4R fibrils, significantly more marked for 1N3R than for 1N4R Tau fibrils, both in intensity of the protection and in size of the protected region. In 1N3R Tau fibrils, a high enzyme-protection is observed: first from E193 to E320 including the second part of the Proline rich region, the R1, R3 and R4 repeats from Microtubule Binding Region (MTBR) and twelve additional amino acids after the R4 repeat; second from T343 to E371 located in the C-terminal flanking region. In 1N4R Tau fibrils part of this region is protected but the highly protected part is limited: first to the region ranging from N236 to D329, including the second moiety of the R1 repeat, the entire R2 and R3 repeats, and the first moiety of the R4 repeat; second to the region from T374 to E402. Very interestingly, according to limited proteolysis experiments, we observe that the region from N352 to D373 remains accessible to GluC upon assembly of 1N4R into fibrils, and that the

region from T344-E351 is only moderately protected, and more importantly 10.25 fold more accessible in 1N4R than in 1N3R fibrils ( $F_{c_{F/M\ 1N3R\ [313-321]}} = 0.04$ ;  $F_{c_{F/M\ 1N4R\ [344-351]}} = 0.41$ ). Finally, the C-terminus, corresponding to the ten C-terminal residues is moderately protected in 1N3R fibrils, but finally not significantly different from 1N4R Tau fibrils ( $F_{c_{F/M\ 1N3R\ [372-381]}} = 0.60$ ;  $F_{c_{F/M\ 1N4R\ [403-412]}} = 0.84$ ).

Advanced proteolysis confirm limited proteolysis observations and bring new information about the protease-accessibilities and the differences between 1N3R and 1N4R fibrils located in the C-terminal flanking region.

### **Chemical surface labeling identifies accessible and protected lysine residues within 1N3R and 1N4R Tau fibrils**

We further mapped the accessible surfaces of Tau fibrils at the amino acid level using covalent NHS-biotin modification of lysine residues, as used previously to assess surface accessibility of amino acid residues of proteins, alone or in complex, when conventional structural methods were unsuccessful (Carven, 2005 ; Gabant 2008 ; Dreger *et al.* 2009). The advantage of using a covalent chemical modification to explore solvent accessible regions of Tau fibrils comes first from the size of the probe, as NHS-biotin is a small water-soluble reagent compared to the size of proteolytic enzymes, and second from the stability of the labeling much more compatible with the sample processing required for insoluble aggregates than hydrogen-deuterium exchanges.

In order to determine the changes in NHS-biotin accessibility of lysine residues upon assembly of Tau monomers into fibrils, we first optimized the labeling conditions, in particular the time of reaction and the molar ratio of lysine residues (K) to the chemical modifier used (NHS-Biotin). Reaction durations of 1 to 30 min and K/NHS-Biotin molar ratios of 40/1 to 1/1 were tested (not shown) in order to select conditions providing not only a limited level of Tau modification in order to preserve the structure of Tau fibrils and monomers, but also a similar level of biotinylation for both monomeric and fibrillar Tau and both isoforms. We selected a 10 min

reaction with a K/NHS-Biotin molar ratio of 2.5/1. In these conditions, we observed the binding of an average total number of biotin molecules ranging from one to seven biotin per Tau molecule, with a maximum number of 3 to 4 and 4 to 5 biotins per Tau for 1N3R and 1N4R respectively (Fig S7).

Then, we performed two types of proteolysis of both monomeric and fibrillary 1N3R and 1N4R Tau, in order to identify and quantify the biotinylated peptides by tandem mass spectrometry. First we conducted an in-gel digestion after SDS-PAGE migration of the entire biotinylated protein, using GluC and AspN proteases. This digestion, performed under denaturing conditions, abolished the enzyme accessibility constraint in order to assess only the level of biotinylation (Fig3.a.b). The second proteolysis was performed in solution under non denaturing conditions using GluC, in order to combine both the level of surface biotinylation and enzyme accessibility. In both cases we identified and quantified biotinylated proteolytic peptides by nanoLC-MSMS analysis.

As expected, after in-gel digestion, we obtained a high sequence coverage upon both GluC and AspN digestions, ranging from 78 to 92%, for both biotinylated monomeric and fibrillar 1N3R and 1N4R Tau and non-biotinylated control samples (Fig.S2 to S5): Overall 18 lysine residues are covered by both GluC and AspN peptides. As GluC and AspN proteolytic peptides present similar changes in biotinylation upon Tau assembly, GluC peptides were more intense because of a higher specificity of GluC compared to AspN. For this reason, we present only AspN peptides that cover lysine residues non covered by GluC digestion. NHS-Biotine has been reported to present a high reactivity to primary amines (Lysine residues and N-terminus of proteins), and a low reactivity to hydroxyl groups of Threonine, Tyrosine, Serine and sulfhydryl groups of Cysteine (Ref). According to this, we identified only minimal amounts of peptides modified on the less reactive residues. For this reason we further only considered biotinylation on N-terminus and lysine residues. All together, we identified and quantified 28 residues (27 Lysines and the N-terminus) out of 39 for 1N3R and 32 residues (31 lysines and the N-terminus) out of 44 for 1N4R. Biotinylated peptides were quantified as we did and previously described for peptides generated by in-solution proteolysis (Fig 2), except that we

normalized AspN biotinylated peptides using an AspN specific reference peptide [392-400] (or 1N3R [361-369]) located in the C-terminal tail of Tau, without any potential biotinylation, as it is the case for the GluC reference peptide [77-88]. After identification of the peptides and biotinylation sites, changes in lysine biotinylation were calculated as a fold change ( $F_c$ ) of a specific biotinylated peptide intensities as described previously ( $F_c = \frac{Intensity_{Fib\ Biotinylated\ Peptide}}{Intensity_{Mono\ Biotinylated\ Peptide}}$ ), as reported in Table S3 and S4. Changes for each individual biotinylated peptide and site are represented by boxplots in Figure S9.

Figure 3 represents the four categories of biotinylated peptides we observed. Two correspond to solvent accessible regions at the surface of Tau fibrils: in grey, peptides biotinylated in a similar extend in both monomeric and fibrillary Tau ( $0.55 < F_c \leq 1.8$ ); in green, peptides more extensively biotinylated in Tau fibrils than in monomers ( $F_c > 1.8$ ) suggesting an exposure at the surface of fibrils of residues possibly involved in intermolecular interactions in the monomeric form of Tau. Two categories highlight regions that become protected from biotinylation upon assembly into fibrils: in blue, regions moderately protected ( $0.1 < F_c < 0.55$ ) suggesting their involvement in new intermolecular interactions and structures upon assembly; in red, regions highly protected or even not accessible to solvent at all ( $0.1 < F_c$ ).

First we observe that the twelve reactive residues of the N-terminal part of Tau (A2, K24, K44, K67, K101, K103, K111, K112, K114, K119, K121, K134) were all accessible to the solvent within the fibrils, at least at a similar extent as in the monomers. The N-terminus (A2) and K67 are highly exposed in both 1N3R and 1N4R Tau fibrils residue ( $F_{C_{F/M\ 1N3R\ (A2)}} = 3.44$  and  $F_{C_{F/M\ 1N4R\ (A2)}} = 2.63$ ;  $F_{C_{F/M\ 1N3R\ (K67)}} = 2.38$  and  $F_{C_{F/M\ 1N4R\ (K67)}} = 2.24$ ), in accordance with what we observed in the proteolysis experiments. Most of these lysine residues, except K44, are located in regions not previously covered by the proteolysis experiments. The accessibility of K44 is in accordance with advanced proteolysis data we obtained for the GluC peptide [37-45] for both 1N3R and 1N4R Tau (Figure 2). Interestingly, advanced proteolysis highlighted a protection to enzyme-accessibility in the region covered by GluC peptides [13-25] and [13-36] for both isoforms, and confirmed by the cleavage at E7 by limited proteolysis. Thus the

accessibility of K24 to the NHS-biotin, a small modification reagent compared to the size of GluC enzyme, suggests that the protection to the solvent concerns a limited region in the N-terminal region between position 8 and to 36, located before the N1 repeat is sufficient to hinder enzyme accessibility. Biotinylation of residues K101 to K134 give solvent accessible information not covered by any GluC proteolytic peptide (Figure 2). No proteolytic peptide was recovered between residue N138 and E193, neither GluC nor AspN digestion, leaving a large part of the [123-215] proline rich region of Tau, without any information about solvent accessibility. This was not due to solvent protection, but rather to the absence of adequate cleavage sites. The only information about lysine residues accessibilities we obtained are located in the C-terminal part of this domain, with K195, K196, K205 and K211 significantly protected from solvent accessibility upon assembly of both 1N3R and 1N4R Tau into fibrils, as indicated by the fold changes  $F_{C_{F/M}}$  comprised between 0.16 and 0.28, and in accordance with what we observed by proteolysis (Figure 2).

For the Microtubule Binding Region composed of three or four Microtubule Binding Repeats for 1N3R (from T216 to N308) and 1N4R (from T216 to N339) respectively, we obtained quantitative information of surface modification changes upon assembly into fibrils for 9 out of 13 and 13 out of 18 Lysine residues for 1N3R and 1N4R respectively.

From the level of biotinylation we measured for the lysine residues located in the first repeat R1 (from Q215 to K245), we observe a similar accessibility for K225 in both 1N3R and 1N4R monomers and fibrils ( $F_{C_{F/M} \text{ 1N3R (K225)}} = 0.69$ ;  $F_{C_{F/M} \text{ 1N3R (K225)}} = 0.64$ ). For K228, protection was moderate and comparable for both 1N3R and 1N4R fibrils ( $F_{C_{F/M} \text{ 1N3R (K228)}} = 0.38$ ;  $F_{C_{F/M} \text{ 1N3R (K228)}} = 0.44$ ). Differences in biotinylation were observed in the second moiety of R1: K230 and K238 are protected in 1N3R fibrils whereas they are accessible in 1N4R fibrils ( $F_{C_{F/M} \text{ 1N3R (K230)}} = 0.12$  and  $F_{C_{F/M} \text{ 1N4R (K230)}} = 0.69$ ;  $F_{C_{F/M} \text{ 1N3R (K238)}} = 0.16$  and  $F_{C_{F/M} \text{ 1N4R (K238)}} = 0.94$ ). The last K245 residue of R1, just before the 246-VQIINK-252 motif corresponding to the start of the R2 domain, was not identified for 1N3R fibrils, and was not even modified in 1N4R Tau fibrils, indicating its extremely high protection. In the R2 repeat, specific to 1N4R Tau, all of the lysine

residues presented a protection toward biotinylation ( $F_{c\ F/M}$  from 0.15 to 0.46), at a similar extent to the protection observed for the last part of the proline-rich region. Regarding the R3 repeat, the comparison between 1N3R and 1N4R was made difficult by the fact that only the first lysine (K282) out of the four lysines of the R3 repeat was identified for 1N4R, whereas three were identified for 1N3R Tau (Figure 3). Because of a distinct sequence coverage of 1N4R due to the R2 repeat, we got only limited information for 1N4R. However, our results clearly show that the R3 repeat is highly protected within 1N3R Tau fibrils ( $F_{c\ F/M\ 1N3R}$  for K251, K257, K261 range from 0.07 to 0.1), and relatively accessible to NHS-biotin in 1N4R Tau fibrils ( $F_{c\ F/M\ K282} = 0.52$ , equivalent to K251 in the 1N3R sequence), whereas this region was identified as protected to enzyme-accessibility in fibrils for both isoforms. It is worth to note that this lysine residue is within the VQIVYK hexapeptide, corresponding to sequences [246-251] and [277-282] in 1N3R and 1N4R respectively (Ref). These results illustrate the complementary information and resolution of the two proteomic approaches in mapping the solvent accessibility surfaces of Tau fibrils.

Finally, in the R4 domain, two identical residues are identified and quantified for both 1N3R and 1N4R Tau (K283 and K293 for 1N3R and K314 and K324 for 1N4R), together with the two residues coming just after the C-terminus of R4 (K309/K310 and K340/341 for 1N3R and 1N4R respectively). All these identified lysines are accessible in 1N4R fibrils whereas they are all protected in 1N3R fibrils ( $F_{c\ F/M\ 1N3R}$  from 0.02 to 0.14 and  $F_{c\ F/M\ 1N4R}$  from 0.88 to 0.99) as shown in Figure 3, Tables S3, S4, and Figure S9)

This difference in lysine accessibility spread in the C-terminal flanking region in which we covered all the seven lysines present in the sequence for both 1N3R and 1N4R Tau. Except the most C-terminal Lysine residue (K378 for 1N3R and K409 for 1N4R) biotinylated at the same level in both monomers and fibrils for both isoforms ( $F_{c\ F/M\ 1N3R\ (K378)} = 0.71$ ;  $F_{c\ F/M\ 1N4R\ (K409)} = 1.25$ ), the other positions were all protected in 1N3R Tau fibrils whereas they were all solvent accessible in 1N4R Tau ( $F_{c\ F/M\ 1N3R\ (K309/310)} = 0.03$ ;  $F_{c\ F/M\ 1N4R\ (K340/341)} = 0.99$ ;  $F_{c\ F/M\ 1N3R\ (K315,323,325,335)} = 0.30$  to 0.39;  $F_{c\ F/M\ 1N3R\ (K346,354, 356, 366)} = 0.62$  to 0.99). This is in accordance with

our previous proteolysis data showing an enzyme-accessibility from N352 to D373 in 1N4R but not 1N3R fibrils. Again, we observe a higher resolution in surface accessibilities information with surface modification than with enzyme accessibility in some regions, but are still complementary in regions lacking lysine residues, or in which lysine are not recovered after complete proteolysis in denaturing conditions.

All together our data highlight the accessibility of the N-terminal domain, a solvent protection from the proline-rich region to the first part of the R2 domain in both isoforms, and a major difference in the C-terminal domain starting at the beginning of the R4 repeat up to the at least the first moiety of the C-terminal identified as protected in 1N3R fibrils and accessible in 1N4R fibrils..

### **Chemical surface labeling combined to protease accessibility confirm the major solvent accessible region within 1N3R and 1N4R Tau fibrils**

Finally, the same biotinylated Tau protein samples were digested 15h in-solution with GluC, as previously done with non-labeled Tau (Figure 2), in parallel of the in-gel digestion (Figure 3). We performed this additional in-solution digestion, in order to combine the information of both surface labeling with a small probe (NHS-biotin) and enzyme accessibility with a large probe (GluC), and further highlight and confirm the significant differences in solvent accessible surfaces upon assembly into fibrils we observed (Figure 2 and 3). In-solution proteolysis of biotinylated Tau fibrils results in an important difference in the sequence coverage between monomeric and fibrillary Tau, as previously observed with non biotinylated Tau (Fig.2): 73% of the sequence was covered for both monomeric 1N3R and 1N4R Tau, whereas only 45% and 62% of the sequence was covered for 1N3R and 1N4R Tau fibrils, respectively (Fig.S6.).

Changes in proteolytic biotinylated peptides upon assembly of 1N3R and 1N4R Tau into fibril, quantified as previously described, are presented in Figure 4. Values of intensities and fold changes together with the corresponding boxplots representations are shown in Tables S5,S6 and Figure S10. The data followed more or less the data obtained by advanced proteolysis (Figure 3): a good accessibility of the N-terminal domain and a protection of the C-terminal

domain for both 1N3R and 1N4R Tau fibrils, which is much more pronounced and spread much further in the C-terminal flanking regions for 1N3R than for 1N4R fibrils. The only difference can be observed in the N-terminal region, with the absence of a significant protection for the biotinylated [13-36] peptide consecutive to assembly into fibrils, observed for proteolysis of non biotinylated samples (Figure 2), but not with modification of accessible residues only (Figure 4). At the end, as the sequence coverage is quite low, this approach is not the most informative compared to the two either proteolysis experiments alone, or residue biotinylation alone, however it reinforces the significance of the differences in accessibility between 1N3R and 1N4R fibrils, highlighted previously.

## Discussion

In order to obtain information on the solvent accessible surfaces of Tau upon assembly into fibrils, in complement to the recent high-resolution structures obtained by cryo-electron microscopy, we implemented and explored a combination of proteomic methods using 1N3R and 1N4R Tau fibrils assembled *in vitro* in presence of heparin using recombinant full-length human Tau protein. Indeed these solvent exposed regions at the surfaces of Tau fibrils are essential not only to address the organization of the entire Tau protein, and not only the rigid core of the fibrils involved in polymerization, but mainly because these surfaces might play a key role in regulating interaction with partner proteins, in the way fibril surfaces might be neutralized to inhibit their toxic function and/or propagation from cell-to-cell. Additionally, identification of differences in the regions exposed at the surface of fibrils between 3R and 4R fibrils might pave the way for the design of molecular tools that could discriminate both types of aggregate.

Using limited proteolysis of 1N3R and 1N4R fibrils with GluC, we observed a different proteolytic profile for 1N3R and 1N4R that represents a molecular “bar-code” specific for each type of fibril, as it has been reported previously using Proteinase K on different  $\alpha$ -Synuclein strains (Bousset *et al.* 2013; Griboaud *et al.* 2019). Characterization of the fragments of this bar-code revealed two hubs of enzyme accessibility in the N-terminal domain located in the ten first amino acids and at the end of the N1 projection domain around position 74 to 88 for both 3R and 4R fibrils and highlighted a different enzyme accessibility in the C-terminal flanking region accessible in 1N4R fibrils but not 1N3R fibrils. In order to identify protected region and compare our recombinant fibrils assembled in presence of heparin to previous published data, we further quantified proteolytic peptides generated by advanced proteolysis using mass spectrometry. Finally we implemented a surface modification approach using biotinylation of accessible amines (lysine and N-terminus) with the small modifier NHS-biotin, to gain in resolution in the mapping of solvent accessible surfaces of 1N3R and 1N4R fibrils.

The results obtained by these two approaches are summarized on the map of surfaces of 1N3R and 1N4R fibrils that become either protected or remain accessible upon assembly into

fibrils presented in Figure 5. This map highlight the complementary of both approaches. All together these data highlight the accessibility in the N-terminal region, with an enzyme-protection observed at the end of the N1 projection domain more marked for 1N3R than 1N4R fibrils. Regions that appear protected to complete enzyme digestion upon assembly into fibrils illustrate that the C-terminal domain is structured from the proline region domain to the end of the fourth microtubule binding repeat R4, and even twelve amino acid further for 1N4R (up to E351) and almost to the C-terminus for 1N3R fibrils. This protection map highlights a remarkable difference between 1N3R and 1N4R fibrils, the first exhibiting a high enzyme-protection, whereas the second are accessible to protease at least between E351 and D373. Interestingly the lysine accessibility map confirms the higher solvent accessibility of the C-terminal flanking region of 1N4R fibrils compared to 1N3R fibrils. This results show for the first time that proteolysis experiments, although indicative of conformation changes, should be taken with caution if integrated in the building of structural models. It is worth to mention that in order to cleave a peptide bond a protease needs to bind to several amino acids, more than ten as reported (Hubbard. 1998), in a flexible region capable to adapt to the catalytic site of the enzyme. Our data demonstrate that this difference in solvent accessibility observed in the C-terminal domain might be targeted either with large probes, as enzymes or Fab antibodies, capable of interacting with the 22 amino-acids sequence [352-373], or with small chemical probes interacting in a larger C-terminal region spreading from the R4 repeat to the C-terminus of the Tau.

Several previous studies have used limited proteolysis to document the protease resistant core of Tau filaments (Wishik *et al.* 1988; Jakes *et al.* 1991, von Bergen *et al.* 2006, Hasegawa *et al.* 2014; Taniguchi-Watanabe *et al.* 2016) and high-resolution cryo-EM to resolve the structure of the pronase resistant core using either PHFs. In order to compared the solvent/protection map we obtained and our major conclusions with the data reported for either recombinant Tau fibrils or filaments recovered from Alzheimer's or Pick's cases (Figure 6), we used the 2N4R sequence numbering in this figure and in the following text..

We observed that the N-terminal domain is accessible from the N-terminus to the solvent up to the Proline region. The N-terminus and the N1 projection domain are accessible, with an exception in the [13-36] and [118-134] regions in which a hindrance to GluC accessibility is observed for both 1N3R and 1N4R fibrils, more pronounced for the second region within 1N3R fibrils.

Interestingly this accessible projection domain establish interactions with several partners as for example cytoskeletal components, plasma membrane, kinases, and is involved in normal Tau functions (Sergeant *et al.* 2005). It is worth to note that upon assembly into fibrils this domain still is available at least partly for interactions. Upon assembly into fibrils immunostaining experiments before pronase treatment have described an accessibility of part of the N-terminal domain we describe here, and others have reported to constitute the so-called “fuzzy coat” (Ksiezak-Reding and Yen. 1991; Zhang *et al.* 2019). Our data on recombinant Tau fibrils conclude the same than what has been observed for both recombinant Tau filaments (Zhang *et al.* 2019) and tauopathies cases (Fitzpatrick *et al.* 2017; Falcon *et al.* 2018a,b).

As structures determined by cryo-EM were obtained after pronase treatment, no information outside the ultra-structured core was obtained (Fitzpatrick *et al.* 2017). It has been proposed that the polypeptide 7-EFEVME-12 might not only regulate interactions and participate to Tau aggregation but also interact with the microtubule binding domain as suggested by the high-resolution cryoEM structures (Fitzpatrick *et al.* 2017) or be detected by the conformational antibody Alz50 (Carmel *et al.* 1996). As 117-EAAGHVTQ-124 has been predicted to form an  $\alpha$ -helix (Gamblin. 2005), this might explain our limited proteolysis data.

Finally, we observe a solvent protection at the end of the proline rich region up to the R1 repeat, and both with the protease and NHS-biotin.

The Microtubule Binding Region (MBTR), reported to be involved in tubulin binding when Tau is unfolded, is tightly involved in Tau aggregation, as it assembles at his own and part of it constitute the rigid core of Tau fibrils, both for Tau fibrils formed in vitro or purified from AD or Pick's cases (Fitzpatrick *et al.* 2017; Falcon *et al.* 2018a,b; Zhang *et al.* 2019). In the current study, we observed a high enzyme protection in the MTBR for both 1N3R and 1N4R, including three HXPGGG repeat located in R1, R3 and R4 for 1N3R fibrils and three HXPGGG repeat located in R1, R2 and R3 for 1N4R fibrils. The last part of the R4 repeat was only moderately protected to enzyme accessibility. This was confirmed by the accessibility to NHS-biotin of all the lysine residues of R4 from 1N4R Tau fibrils. In the R3 domain, we obtained limited information but at least we found the hexapeptide 306-VQIVYK-311 highly enzyme protected, but NHS-biotine protection of the last lysine of this hexapeptide was significantly observed for 1N3R fibrils, and only moderately for 1N4R fibrils. Recently, a structural characterization using hydrogen-deuterium exchanges on 2N4R assembly (Huang *et al.* 2018) reported one unique protection upon assembly into fibrils within this polypeptide 308-315. We further compared our data with the numerous cryo-EM structures recently published for the highly structured and non-dynamic ultra-cores of tau filaments from AD and PiD cases (Fitzpatrick *et al.* 2017; Falcon *et al.* 2018a,b) and recombinant 2N3R and 2N4R tau fibrils (Zhang *et al.* 2019), that are resistant to pronase.

Comparison of the structures and fold identified by cryo-EM studies (Goedert 2019) of distinct Tau filaments conformers from both AD and PiD cases report that the ultra-structured core is composed of amino acids G273 to E380 for 3R-tau and G304 to E380 for 4R-tau tau filaments from AD, and K254 to F378 for 3R-tau tau filaments from PiD (Figure 6). However, cryo-EM structures of recombinant 2N3R and 2N4R fibrils (Zhang *et al.* 2019) identified different conformers and different ultra-structured cores composed of amino-acids G272 to H330 or G274 to K321 for 2N4R fibrils and a packing of G272/274 to H330 for 2N3R fibrils. Clearly our in vitro assembled recombinant fibrils are different: we identified a much larger highly protected region for 1N3R fibrils from P223 to E380 by proteolysis, and from K275-to K370 by surface biotinylation, much closer to what has been reported for 3R-tau filaments from AD or PiD cases

than the structure observed for recombinant 2N3R fibrils. However for 1N4R fibrils, we observed a smaller highly protected region than what we observed for our 1N3R fibrils: the highly enzyme protected region ranges from N265 to L357, whereas the surface protected from biotinylation we identified is reduced, it covers the sequence 268-HQPGGGKVQIINK-280 located at the junction between the R1 and R2 repeats. Very interestingly we observed that the hexapeptide 275-VQIVYK-280, important in the assembly of Tau monomers into fibrils (Lippens and Gigant 2019 + Ref), is protected from GluC cleavage and surface biotonylation, for both 1N3R and 1N4R, but much more pronounced for 1N3R Tau fibrils . This is in accordance with previous data that have reported that VQIVYK inhibitors inhibit assembly of 3R Tau in presence of heparin, but not of 4R tau, and that VQIINK inhibitors blocked heparin-induced assembly of 4R tau (Seidler *et al.* 2018). Although the regions we covered differs slightly from the one reported for AD and PiD tau filaments, mainly because we did not use the same enzyme, we observe an extended protection for 1N3R Tau fibrils compared to 1N4R tau fibrils, similarly to what has been observed for AD and PiD tau filaments. However our data differ significantly in the repeats of the MBTR involved in the protected regions we identified protected (end of R1 to at least end of R4 and even further for 1N3R and end of R1 to R3 for 1N4R) compared to what has been reported from the cryo-EM structures. It is worth to note that densities observed by cryo-EM surrounding the ultra-structured core were not enough resolved and might correspond to repeats R1/R2. Interestingly, previous studies on pronase-treated PHF from AD brains have identified the minimal pronase resistant core from N266/K268 to I392 (Fig6). We also observed a protease resistance in the R2 repeat 4R-specific. Similarly to cryo-EM structures of PiD (3R) and AD (3R/4R) tau filaments fold, the R1 polypeptide K254-K274 has been found highly structured only for PiD. In our study, we observed digestion resistance within both 1N3R and 1N4R fibrils, although the protection measured as a fold change between fibrillar and monomeric forms is much higher for 1N3R than for AN4R fibrils ( $FC_{f/m\ 1N3R}(253-265)=0.09$ ;  $FC_{f/m\ 1N4R}(253-265)=0.39$ ).

Finally, our data highlight a major difference in the surface accessibility of the C-terminal flanking region between 1N3R and 1N4R fibrils. Of particular interest is the GluC-accessible region from N381 to D402 and the biotiny-accessible region from K343 to K395 in 1N4R fibrils that are protected in 1N3R Tau fibrils. Very interestingly, in late stages of Pick's disease, the E391 cleavage was found drastically reduced compared to first stages of the pathology (less aggregated state) and to AD (Mondragón-Rodríguez 2008).

In conclusion, our data demonstrate the utility of GluC limited proteolysis to identify 3R or 4R proteolytic molecular code-bars 3R- or 4R- Tau specific. Also, our results show that the proteomic methods we implemented for recombinant Tau fibrils should be used to further characterize Tau filaments from AD or PiD cases, in complement to high-resolution structures. Our results highlight in particular a difference in solvent accessibility of the R4 repeat and the first part of the C-terminal flanking region, between 1N3R and 1N4R tau fibrils, with a particular interest in the sequence 381- NAKAKTDHGAEIVYKSPVVS GD-402 we identified accessible in 1N3R fibrils but protected in 1N4R fibrils. Further analysis of this region on tau filaments from AD, PSP and PiD with similar proteomic approaches, complemented with targeted mass spectrometric quantification of the proteolytic peptides that define this sequence will confirm its 4R- solvent accessible specificity. If so, this region might pave the way to the design of new molecular tools targeting specifically this 22 amino-acids sequence, accessible in 4R- but not 3R- tau filaments.

## Methods

### Expression and Purification of human Tau 1N3R and Tau 1N4R

Full-length human Tau1N3R or Tau 1N4R cDNA was cloned in the pET14b vector. hTau (Tau1N3R or Tau 1N4R) was expressed in *E. coli* BL21 DE3 CodonPlus cells (Stratagene). Cells were grown in LB medium to an optical density at 600 nm of 0.8 absorbance units. hTau expression was induced with 0.5 mM IPTG for 3 h. The cells were then harvested by centrifugation (4000 g, 10 min). The bacterial pellets were resuspended in lysis buffer (20 mM MES pH 6.8, 500 mM NaCl, 1 mM EGTA, 0.2 mM MgCl<sub>2</sub>, 5 mM di-thiothreitol, 1 mM PMSF + 1 tablet of Complete (Roche)) per liter and lysed by sonication. Cell extracts were clarified by centrifugation at 14,000 g, 30 min. The lysate was heated to 80 °C for 20 min and centrifuged at 14,000 g for 30 min. The supernatant was dialyzed against 100 volumes of buffer A (20 mM MES pH 6.8, 50 mM NaCl, 1 mM EDTA, 1 mM MgCl<sub>2</sub>, 2 mM DTT, 0.1 mM PMSF) at 4 °C. The dialyzed protein mixture was loaded on an SP Sepharose column (60 ml bed volume). Proteins were separated with a linear gradient of 0 to 100% buffer B (20 mM MES pH 6.8, 1 M NaCl, 1 mM EGTA, 1 mM MgCl<sub>2</sub>, 2 mM DTT, 0.1 mM PMSF). Fractions were analyzed via SDS-PAGE stained with Coomassie blue. Fractions containing hTau were pooled and dialyzed against 100 volumes of PBS buffer containing 1 mM DTT. The hTau concentration was determined spectrophotometrically using an extinction coefficient at 280 nm of 7450 M<sup>-1</sup>.cm<sup>-1</sup> for Tau 1N3R and 7575 M<sup>-1</sup>.cm<sup>-1</sup> for Tau 1N4R.

### Aggregation of protein into amyloid fibrils

Fibrillation of human Tau 1N3R and 1N4R was achieved at 40 μM in the presence of 10 μM heparin by shaking 0.5 ml solution aliquots at 37 °C in an Eppendorf Thermomixer set at 600 rpm for 5 days. Fibrils were spun for 30 min at 20 °C and 50,000 rpm. The amount of fibrillar material was estimated by subtraction of the soluble fraction remaining after centrifugation from the initial concentration. The pelleted material was resuspended in PBS 1X DTT 1mM at an equivalent monomeric hTau1N4R concentration of 40 μM.

## **Transmission Electron microscopy**

The morphology of hTau1N4R and Sup35NM assemblies was assessed using 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).

## **Limited Proteolysis monitored by SDS-PAGE**

Soluble and polymerized proteins were digested with the endoprotease GluC (Roche) at a E/S (w/w) ratio of 1/150, at 37 °C under agitation (300 rpm). The proteolysis reaction was stopped for at 0, 2, 10, 45, 120 and 300 min by addition of AEBSF final concentration 10mM and mixed with Laemmli buffer 3 X (vol:vol). After 7 minutes of heating, 5 $\mu$ L of sample was dropped on a 12% Tris-Glycine and migrated during 45 minutes (200V). After Coomassie Blue staining, gels were imaged using ChemiDoc™ MP(BioRad) and analysed with ImageLab version 5.2.1.

## **Identification of the proteolytic fragments**

### *Western-Blots*

Limited proteolysis samples were denaturated in Laemmli buffer for 10 min at 95 °C. The samples were analyzed on 12 % Tris-glycine SDS-PAGE and transfered on nitrocellulose membranes (Cat #10600002 Protran, Amersham). Membranes were stained with Ponceau Red before immunodetection with Nter and Cter rabbit polyclonal antibodies diluted at 1/10000 (Gift from Luc Buée, INSERM, Lille, France), HT7 mouse monoclonal antibody diluted at 1/4000 (Cat #MN1000, Thermo Fisher, Waltham, United States), Tau5 mouse monoclonal antibody diluted at 1/20000 (Cat #AHB0042 Thermo Fisher, Waltham, United States), anti-3R Tau (3-repeat isoform RD3) mouse monoclonal antibody clone 8E6/C11 diluted at 1/10000 (Cat #05-803 Merck Millipore, Billerica, United States), anti-4R Tau (4-repeat isoform RD4) mouse monoclonal antibody, clone 1E1/A6 diluted at 1/1000 (Cat #05-804 Merck Millipore,

Billerica, United States). Westerns were developed with Pierce™ ECL Western Blotting Substrate (Cat #32106) and imaged on a Chemidoc MP Imaging System (Biorad).

*N-TOP: TMPP-N terminal labelling and nano-LC MSMS identification of N-terminal peptides*

7µL (20µM, ≈ 5µg) of sample were dropped on a 12% Tris-Glycine and migrated during 45 minutes (200V) followed by Coomassie staining. Bands were excised and washed 3 times using acetonitrile and 25 mM NH<sub>4</sub>HCO<sub>3</sub>. After reduction (10 mM DTT) and alkylation (55 mM iodoacetamide) of cysteines, N-terminal residues were labelled using 1.5µL of a 100mM TMPP solution (N-succinimidylloxycarbonylmethyl)-tris(2,4,6-trimethoxyphenyl)-phosphonium bromide) in Tris HCl 100 mM pH 8.2 for 1 hour under agitation (600rpm) at room temperature. Reaction was stopped by addition of 5 µL of 16M hydroxylamine (30 minutes of incubation) and bands were re-washed 3 times using acetonitrile and 25 mM NH<sub>4</sub>HCO<sub>3</sub>. Trypsin (Promega Gold; Promega, Madison, WI) digestion was performed overnight at 37°C under agitation in 50 mM NH<sub>4</sub>HCO<sub>3</sub>. The volume of the trypsin solution (10 ng/µl) was adapted to the size of the protein band excised from the gel (e.g 25 µL of a 4 ng/µL solution, (E/S)<sub>w/w</sub> = 1/50). Peptides were extracted by incubation during 1 h at room temperature after addition of 1.2 volumes of 60% CH<sub>3</sub>CN, 0.1% HCOOH followed by incubation in 2 volumes of CH<sub>3</sub>CN. The extracted peptides were vacuum dried and stored at -20 °C until MS analysis.

Tryptic peptides were resuspended in 10µL 0.1% TFA and 2µL were analysed by nanoLC-MS/MS using an EASY-nLC II high performance liquid chromatography (HPLC) system (Proxeon, ThermoScientific, Waltham, MA) coupled to the nanoelectrospray ion source of an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). Peptide separation was performed on a reversed-phase C18 nano HPLC column (100 µm inner diameter, 5-µm C18 particles, 15-cm length, NTCC-360/100-5) from Nikkyo Technos (Nikkyo Technos Co., Ltd., Tokyo, Japan). The peptides were loaded at a pressure-dependent flow rate corresponding to a maximum pressure of 200 bars and eluted at a flow rate of 300 nl/min using a gradient of 5 to 60% solvent B for 70 min, followed by a washing step at 100% solvent B. Solvent A was 0.1% formic acid in water, and solvent B was 0.1% formic acid in 100% acetonitrile. NanoLC-

MS/MS experiments were conducted in the data-dependent acquisition mode. The mass of the precursors was measured with a high resolution (60,000 full weight at half maximum) in the Orbitrap. The 20 most intense ions, above an intensity threshold of 2000 counts, were selected for CID fragmentation and analysis in the LTQ.

NanoLC-MS/MS data were processed automatically using the Thermo Proteome Discoverer software (version 1.3) and the Sequest search engine with the following chemical modifications: methionine oxidation, cysteine carbamidomethylation, and TMPP labelling of N-ter peptides and lysines. MS<sup>2</sup> spectrum of resulting N-terminus were checked.

### *Edman Sequencing*

The proteolytic fragments were transferred onto 0.2 µm PVDF membranes (Applied Biosystems), using) using a Trans-Blot Turbo Transfer System (Biorad) during 10min (25V 1,3A) and the following transfer buffer: 50 mM Tris, 50 mM Boric Acid, 0.1% SDS, 10% Ethanol, pH 8.3. Membranes were stained with Coomassie blue.

Protein Sequencing. Peptide sequence data for full-length and fibrillar GluC proteolytic fragments were obtained by automated Edman degradation using a sequencer (PPSQ-1 Shimadzu).

PVDF membranes were washed three times in Ethanol / water (90/10), dried and loaded in the cartridge of the Edman sequencer Shimadzu PPSQ 31B. Ten Edman degradation cycles were performed.

### **Chemical Surface Labeling of monomeric and fibrillar Tau**

#### *Buffer exchange on Sephadex column to PBS 1X DTT 1mM of monomeric Tau*

Monomeric aliquots were quickly defrosted and buffer exchange to PBS 1X DTT 1mM was performed following manufacturer's instructions on PD10 column (GE Healthcare). Concentration of each eluted aliquots and final concentration of the pool of aliquots of interest were determined by spectrophotometry (wavelength 280nm and 320nm).

#### *Determination of labelling parameters based on kinetics and molar ratios range*

Monomeric and fibrillar Tau 1N3R and 1N4R were labelled with Sulfosuccinimidyl biotin (Ref.B5161, Sigma) at at [Lysines/NHS-Biotine] ratio of 2.5/1 during 10 min at room temperature and reaction was stopped by addition of Tris pH7.5 at final molar concentration of 40-50mM. Later, fibrillar Tau went through ultracentrifugation (30 min, 50.000rpm, 20°C) and supernatant was replaced by the equivalent of initial volume of 50mM Tris pH7.5)

#### *Determination of the level of biotinylation of monomeric and fibrillary Tau by MALDI-MS*

Samples were sonicated 10 sec and diluted to 4-8  $\mu$ M in H<sub>2</sub>O milliQ and. 0.8 $\mu$ L of sample was mixed with 0.8 $\mu$ L of matrix HCCA at 20mg/ml in acetonitrile 70%, Formic Acid 5%) and spotted on a silver plate x396. MALDI-MS measurements were performed with a MALDI-TOF-TOF 4800 (ABSciex) mass spectrometer in the linear mode, after external calibration using synuclein. T2D spectra were Gaussian smoothed on PeakView® Software 1.2. AB SCIEX.

#### **In-solution digestion under native conditions**

Soluble and polymerized proteins, biotinylated or not, were digested for 15 hours at 37°C, under agitation 300rpm with endoprotease GluC (Roche) at a E/S (w/w) ratio of 1/50 using a final Tau concentration of 10 $\mu$ M. Digestion was stopped by flash freezing in liquid nitrogen and storage at -80°C for further analysis. Samples were diluted to a final concentration of 4 $\mu$ M in 0.4% TFA before nanoLC-MSMS analysis. Three replicate experiments including the assembly into fibrils, digestion and nanoLC-MSMS were performed.

#### **In-gel digestion under denaturing conditions**

SDS-PAGE of 4 $\mu$ g (100pmol) of control and biotinylated protein was performed as described above (3 replicates). Bands were excised and washed 3 times using acetonitrile and 25 mM NH<sub>4</sub>HCO<sub>3</sub>. After reduction (10 mM DTT) and alkylation (55 mM Iodoacetamide) of cysteines, GluC and AspN were added to perform digestions at a 1/20 Enzyme-to-Substrate ratio during 15h at 37°C under agitation (300 rpm). Proteolytic peptides were extracted by addition of 1.2

volumes of 60% CH<sub>3</sub>CN, 0.1% HCOOH and incubation during 1 h at room temperature. A second extraction was performed after recovery of the first peptide extract by addition of 2 volumes of pure CH<sub>3</sub>CN. The peptides from the two extractions were pooled, and further vacuum dried to be stored at -20 °C until MS analysis. GluC and AspN peptides were resuspended in 0.1% TFA, just before nanoLC-MSMS analysis.

### **nanoLC-MSMS analysis of proteolytic peptides**

3 replicates of 5µL each (corresponding to about 20pmol of initial protein) were analysed separately by nanoLC-MS/MS with the Triple-TOF 4600 mass spectrometer (ABSciex) coupled to the nanoRSLC ultra-performance liquid chromatography (UPLC) system (Thermo Scientific) equipped with a trap column (?) and an analytical column (C18, 50cm). Peptides were loaded at 6 µL/min with 0.05% TFA in 5% acetonitrile, and peptide separation was performed at a flow rate of 300 nL/min with a 5–35% solvent B gradient in 40 min. Solvent A was 0.1% formic acid in water, and solvent B was 0.1% formic acid in 100% acetonitrile. NanoLC-MS/MS experiments were conducted in a data-dependent acquisition method by selecting the 20 most intense precursor ions, with charge states from 1 to 5 and above an intensity threshold of 1000, for CID fragmentation. NanoLC-MS/MS raw data were processed using the MS Data Converter software (AB Sciex) for generating mgf data files.

Database searching for protein identifications was performed using the Mascot search engine (Matrix Science, London, UK; version 2.4.1) against the SwissProt database in which we added the 1N3R and 1N4R Tau sequences, including a decoy database search. Mascot was searched with a fragment ion mass tolerance of 0.05 Da and a parent ion tolerance of 40 ppm. Oxidation of methionine; Biotinylation of N-ter, Lys and Cys, Ser, Thr, Tyr were specified in Mascot as variable modifications and carbamidomethylation of Cys as constant modification for in-gel digestion.

### **nanoLC-MS/MS data analysis**

Using the function XIC-manager (PeakView® Software 1.2, AB SCIEX), the extracted-ion chromatogram (XIC) was recovered from the entire data set for each peptide presenting at least 2 identifications in the MASCOT search. Using the same function, the extracted-ion chromatogram (XIC) was recovered from the entire data set for each identified biotinylated peptides among monomeric and fibrillar samples.

The base peak intensity within a mass tolerance window of 0.1 Da around the more intense mass-to-charge ratio was plotted in a time window of 0.3-0.6 min in the analysis. The accordance between MASCOT identification and MS2 spectrum associated with the extracted intensity on XIC Manager was verified. The intensity values were normalized using peptide [77-88] as reference for normalization of the peptide intensity as performed previously by Barthélemy (Barthelemy *et al.* 2016). This peptide contains no lysine and covers an accessible region as observed by limited proteolysis. The average normalized intensity of the three replicates are represented in boxplots (Figures S7, S9, S10). Fold change ( $F_{c\ F/M}$ ) between fibrillar and monomeric Tau was calculated as a ratio of the three replicates of the averaged normalized peptide intensity for the fibrils to the averaged normalized peptide intensity for the monomers. Standard deviation was calculated using the normalized intensities. Significance of the  $F_c$  was assessed by a Kruskal-Wallis Test ( $p$ -value < 0.050).

## **Acknowledgements**

We acknowledge the Microscopy platforms and the proteomic platform SICaPS from the I2BC (CNRS UMR9198, Gif-sur-Yvette, France). We thank David Cornu and Laila Sago for assistance in mass spectrometric analysis, Pascal Mansuelle for the Edman sequencing service (IMM, Marseille, France) and Luc Buée for the gift of anti-Tau antibodies (INSERM, Lille, France). We are grateful to the Vaincre Alzheimer Foundation for its financial support (EC, PhD grant). We acknowledge the support of the CNRS (for KM, LB, RM) and the INSERM (for VR).

## References

- Adams, S. J., de Ture, M. A., McBride, M., Dickson, D. W., & Petrucelli, L. (2010). Three repeat isoforms of tau inhibit assembly of four repeat tau filaments. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0010810>
- Barghorn, S., & Mandelkow, E. (2002). Toward a unified scheme for the aggregation of tau into Alzheimer paired helical filaments. *Biochemistry*. <https://doi.org/10.1021/bi026469j>
- Barghorn, S., Biernat, J., & Mandelkow, E. (2005). Purification of recombinant tau protein and preparation of alzheimer-paired helical filaments in vitro. In *Methods in Molecular Biology*. <https://doi.org/10.1385/1-59259-874-9:035>
- Barthélemy, N. R., Gabelle, A., Hirtz, C., Fenaille, F., Sergeant, N., Schraen-Maschke, S., ... Lehmann, S. (2016). Differential mass spectrometry profiles of tau protein in the cerebrospinal fluid of patients with Alzheimer's disease, progressive supranuclear palsy, and dementia with lewy bodies. *Journal of Alzheimer's Disease*, 51(4), 1033–1043. <https://doi.org/10.3233/JAD-150962>
- Bertaccini, D., Vaca, S., Carapito, C., Arsène-Ploetze, F., Van Dorsselaer, A., & Schaeffer-Reiss, C. (2013). An improved stable isotope N-terminal labeling approach with light/heavy TMPP to automate proteogenomics data validation: DN-TOP. *Journal of Proteome Research*, 12(6), 3063–3070. <https://doi.org/10.1021/pr4002993>
- Biancalana, M., & Koide, S. (2010). Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochimica et Biophysica Acta - Proteins and Proteomics*. <https://doi.org/10.1016/j.bbapap.2010.04.001>
- 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*. <https://doi.org/10.1038/ncomms3575>
- Buée, L., & Delacourte, A. (1999). Comparative Biochemistry of Tau in Progressive Supranuclear Palsy, Corticobasal Degeneration, FTDP-17 and Pick's Disease. *Brain Pathology*, 9, 681–693. <https://doi.org/10.1111/j.1750-3639.1999.tb00550.x>
- Carmel, G., Mager, E. M., Binder, L. I., & Kuret, J. (1996). The structural basis of monoclonal antibody Alz50's selectivity for Alzheimer's disease pathology. *Journal of Biological Chemistry*. <https://doi.org/10.1074/jbc.271.51.32789>
- Carven, G. J., & Stern, L. J. (2005). Probing the ligand-induced conformational change in HLA-DR1 by selective chemical modification and mass spectrometric mapping. *Biochemistry*. <https://doi.org/10.1021/bi050972p>
- De Silva, R., Lashley, T., Gibb, G., Hanger, D., Hope, A., Reid, A., ... Lees, A. (2003). Pathological inclusion bodies in tauopathies contain distinct complements of tau with three or four microtubule-binding repeat domains as demonstrated by new specific monoclonal antibodies. *Neuropathology and Applied Neurobiology*. <https://doi.org/10.1046/j.1365-2990.2003.00463.x>
- Dreger, M., Bo, W. L., Brownlee, G. G., & Deng, T. (2009). A quantitative strategy to detect changes in accessibility of protein regions to chemical modification on heterodimerization. *Protein Science*, 18(7), 1448–1458. <https://doi.org/10.1002/pro.159>
- Falcon, B., Zhang, W., Murzin, A. G., Murshudov, G., Garringer, H. J., Vidal, R., ... Goedert, M. (2018, September 6). Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature*, Vol. 561, pp. 137–140. <https://doi.org/10.1038/s41586-018-0454-y>

- Falcon, B., Zhang, W., Schweighauser, M., Murzin, A. G., Vidal, R., Garringer, H. J., ... Goedert, M. (2018). Tau filaments from multiple cases of sporadic and inherited Alzheimer's disease adopt a common fold. *Acta Neuropathologica*. <https://doi.org/10.1007/s00401-018-1914-z>
- Fitzpatrick, A. W. P., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., ... Scheres, S. H. W. (2017). Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature*, *547*(7662), 185–190. <https://doi.org/10.1038/nature23002>
- Gabant, G., Augier, J., & Armengaud, J. (2008). Assessment of solvent residues accessibility using three Sulfo-NHS-biotin reagents in parallel: Application to footprint changes of a methyltransferase upon binding its substrate. *Journal of Mass Spectrometry*. <https://doi.org/10.1002/jms.1328>
- Gamblin, T. C. (2005). Potential structure/function relationships of predicted secondary structural elements of tau. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, Vol. 1739, pp. 140–149. <https://doi.org/10.1016/j.bbadis.2004.08.013>
- Goedert, M., Spillantini, M. G., Potier, M. C., Ulrich, J., & Crowther, R. A. (1989). Cloning and sequencing of the cDNA encoding an isoform of microtubule-associated protein tau containing four tandem repeats: differential expression of tau protein mRNAs in human brain. *The EMBO Journal*, *8*(2), 393–399. <https://doi.org/10.1002/j.1460-2075.1989.tb03390.x>
- Goedert, Michel, Falcon, B., Zhang, W., Ghetti, B., & Scheres, S. H. W. (2018). Distinct Conformers of Assembled Tau in Alzheimer's and Pick's Diseases. *Cold Spring Harbor Symposia on Quantitative Biology*. <https://doi.org/10.1101/sqb.2018.83.037580>
- Goedert, Michel, & Spillantini, M. G. (2017). Propagation of Tau aggregates Tim Bliss. *Molecular Brain*, *10*(1), 1–9. <https://doi.org/10.1186/s13041-017-0298-7>
- Gribaudo, S., Tixador, P., Bousset, L., Fenyi, A., Lino, P., Melki, R., ... Perrier, A. L. (2019). Propagation of  $\alpha$ -Synuclein Strains within Human Reconstructed Neuronal Network. *Stem Cell Reports*, *12*(2), 230–244. <https://doi.org/10.1016/j.stemcr.2018.12.007>
- Guo, J. L., Narasimhan, S., Changolkar, L., He, Z., Stieber, A., Zhang, B., ... Lee, V. M. Y. (2016). Unique pathological tau conformers from alzheimer's brains transmit tau pathology in nontransgenic mice. *Journal of Experimental Medicine*. <https://doi.org/10.1084/jem.20160833>
- Hasegawa, M., Watanabe, S., Kondo, H., Akiyama, H., Mann, D. M. A., Saito, Y., & Murayama, S. (2014). 3R and 4R tau isoforms in paired helical filaments in Alzheimer's disease. *Acta Neuropathologica*. <https://doi.org/10.1007/s00401-013-1191-9>
- Huang, R. Y. C., Iacob, R. E., Sankaranarayanan, S., Yang, L., Ahljianian, M., Tao, L., ... Chen, G. (2018). Probing Conformational Dynamics of Tau Protein by Hydrogen/Deuterium Exchange Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, *29*(1), 174–182. <https://doi.org/10.1007/s13361-017-1815-8>
- Hubbard, S. J. (1998). The structural aspects of limited proteolysis of native proteins. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*. [https://doi.org/10.1016/S0167-4838\(97\)00175-1](https://doi.org/10.1016/S0167-4838(97)00175-1)
- Jakes, R., Novak, M., Davison, M., & Wischik, C. M. (1991). Identification of 3- and 4-repeat tau isoforms within the PHF in Alzheimer's disease. *The EMBO Journal*, *10*(10), 2725–2729. <https://doi.org/10.1002/j.1460-2075.1991.tb07820.x>

- Kaufman, S. K., Sanders, D. W., Thomas, T. L., Ruchinskas, A. J., Vaquer-Alicea, J., Sharma, A. M., ... Diamond, M. I. (2016). Tau Prion Strains Dictate Patterns of Cell Pathology, Progression Rate, and Regional Vulnerability In Vivo. *Neuron*, 92(4), 796–812. <https://doi.org/10.1016/j.neuron.2016.09.055>
- Kovacs, G. G. (2015). Invited review: Neuropathology of tauopathies: Principles and practice. *Neuropathology and Applied Neurobiology*, 41(1), 3–23. <https://doi.org/10.1111/nan.12208>
- Ksiezak-Reding, H., & Yen, S. H. (1991). Structural stability of paired helical filaments requires microtubule-binding domains of Tau: A model for self-association. *Neuron*, 6(5), 717–728. [https://doi.org/10.1016/0896-6273\(91\)90169-Z](https://doi.org/10.1016/0896-6273(91)90169-Z)
- Lee, G., & Leugers, C. J. (2012). *Tau and Tauopathies*. <https://doi.org/10.1016/B978-0-12-385883-2.00004-7>
- Lee, V. M. Y. (2011). Pathological Tau Aggregation. *Nature Communication*, 2(252), 1–18. <https://doi.org/10.1038/ncomms1255>.The
- Lippens, G., & Gigant, B. (2019). Elucidating Tau function and dysfunction in the era of cryo-EM. *Journal of Biological Chemistry*, 294(24), 9316–9325. <https://doi.org/10.1074/jbc.REV119.008031>
- Liu, F., & Gong, C. X. (2008). Tau exon 10 alternative splicing and tauopathies. *Molecular Neurodegeneration*, 3(1), 1–10. <https://doi.org/10.1186/1750-1326-3-8>
- Lyumkis, D. (2019). Challenges and opportunities in cryo-EM single-particle analysis. *Journal of Biological Chemistry*, 294(13), 5181–5197. <https://doi.org/10.1074/jbc.REV118.005602>
- Mondragón-Rodríguez, S., Mena, R., Binder, L. I., Smith, M. A., Perry, G., & García-Sierra, F. (2008). Conformational changes and cleavage of tau in Pick bodies parallel the early processing of tau found in Alzheimer pathology. *Neuropathology and Applied Neurobiology*, 34(1), 62–75. <https://doi.org/10.1111/j.1365-2990.2007.00853.x>
- Narasimhan, S., Guo, J. L., Changolkar, L., Stieber, A., McBride, J. D., Silva, L. V., ... Lee, V. M. Y. (2017). Pathological tau strains from human brains recapitulate the diversity of tauopathies in nontransgenic mouse brain. *Journal of Neuroscience*, 37(47), 11406–11423. <https://doi.org/10.1523/JNEUROSCI.1230-17.2017>
- Novak, M., Kabat, J., & Wischik, C. M. (1993). Molecular characterization of the minimal protease resistant tau unit of the Alzheimer's disease paired helical filament. *The EMBO Journal*, 12(1), 365–370. <https://doi.org/10.1002/j.1460-2075.1993.tb05665.x>
- Park, S. A., Ahn, S. II, & Gallo, J. M. (2016). Tau mis-splicing in the pathogenesis of neurodegenerative disorders. *BMB Reports*. <https://doi.org/10.5483/BMBRep.2016.49.8.084>
- Pérez, M., Valpuesta, J. M., Medina, M., Montejo de Garcini, E., & Avila, J. (1996). Polymerization of tau into filaments in the presence of heparin: the minimal sequence required for tau-tau interaction. *Journal of Neurochemistry*, 67(3), 1183–1190. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8752125>
- Probst, A., Tolnay, M., Langui, D., Goedert, M., & Spillantini, M. G. (1996). Pick's disease: Hyperphosphorylated tau protein segregates to the somatoaxonal compartment. *Acta Neuropathologica*. <https://doi.org/10.1007/s004010050565>

- Saman, S., Kim, W. H., Raya, M., Visnick, Y., Miro, S., Saman, S., ... Hall, G. F. (2012). Exosome-associated tau is secreted in tauopathy models and is selectively phosphorylated in cerebrospinal fluid in early Alzheimer disease. *Journal of Biological Chemistry*, 287(6), 3842–3849. <https://doi.org/10.1074/jbc.M111.277061>
- Sanders, D. W., Kaufman, S. K., DeVos, S. L., Sharma, A. M., Mirbaha, H., Li, A., ... Diamond, M. I. (2014). Distinct tau prion strains propagate in cells and mice and define different tauopathies. *Neuron*, 82(6), 1271–1288. <https://doi.org/10.1016/j.neuron.2014.04.047>
- Seidler, P. M., Boyer, D. R., Rodriguez, J. A., Sawaya, M. R., Cascio, D., Murray, K., ... Eisenberg, D. S. (2018). Structure-based inhibitors of tau aggregation. *Nature Chemistry*. <https://doi.org/10.1038/NCHEM.2889>
- Sergeant, N., Delacourte, A., & Buée, L. (2005, January 3). Tau protein as a differential biomarker of tauopathies. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, Vol. 1739, pp. 179–197. <https://doi.org/10.1016/j.bbadis.2004.06.020>
- Takeda, S., Wegmann, S., Cho, H., Devos, S. L., Commins, C., Roe, A. D., ... Hyman, B. T. (2015). Neuronal uptake and propagation of a rare phosphorylated high-molecular-weight tau derived from Alzheimer's disease brain. *Nature Communications*, 6. <https://doi.org/10.1038/ncomms9490>
- Taniguchi-Watanabe, S., Arai, T., Kametani, F., Nonaka, T., Masuda-Suzukake, M., Tarutani, A., ... Hasegawa, M. (2016). Biochemical classification of tauopathies by immunoblot, protein sequence and mass spectrometric analyses of sarkosyl-insoluble and trypsin-resistant tau. *Acta Neuropathologica*, 131(2), 267–280. <https://doi.org/10.1007/s00401-015-1503-3>
- Tardivel, M., Bégard, S., Bousset, L., Dujardin, S., Coens, A., Melki, R., ... Colin, M. (2016). Tunneling nanotube (TNT)-mediated neuron-to neuron transfer of pathological Tau protein assemblies. *Acta Neuropathologica Communications*, 4(1), 117. <https://doi.org/10.1186/s40478-016-0386-4>
- Trabzuni, D., Wray, S., Vandrovcova, J., Ramasamy, A., Walker, R., Smith, C., ... Ryten, M. (2012). MAPT expression and splicing is differentially regulated by brain region: Relation to genotype and implication for tauopathies. *Human Molecular Genetics*, 21(18), 4094–4103. <https://doi.org/10.1093/hmg/dds238>
- Von Bergen, M., Barghorn, S., Müller, S. A., Pickhardt, M., Biernat, J., Mandelkow, E. M., ... Mandelkow, E. (2006). The core of tau-paired helical filaments studied by scanning transmission electron microscopy and limited proteolysis. *Biochemistry*, 45(20), 6446–6457. <https://doi.org/10.1021/bi052530j>
- Williams, D. R. (2006). Tauopathies: Classification and clinical update on neurodegenerative diseases associated with microtubule-associated protein tau. *Internal Medicine Journal*, Vol. 36, pp. 652–660. <https://doi.org/10.1111/j.1445-5994.2006.01153.x>
- Wischnik, C. M., Novak, M., Thogersen, H. C., Edwards, P. C., Runswick, M. J., Jakes, R., ... Klug, A. (1988). Isolation of a fragment of tau derived from the core of the paired helical filament of Alzheimer disease. *Proceedings of the National Academy of Sciences*, 85(12), 4506–4510. <https://doi.org/10.1073/pnas.85.12.4506>
- Xue, C., Lin, T. Y., Chang, D., & Guo, Z. (2017). Thioflavin T as an amyloid dye: fibril quantification, optimal concentration and effect on aggregation. *Royal Society Open Science*, 4(1), 160696. <https://doi.org/10.1098/rsos.160696>

Zhang, W., Falcon, B., Murzin, A. G., Fan, J., Crowther, R. A., Goedert, M., & Scheres, S. H. (2019). Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *ELife*, 8. <https://doi.org/10.7554/elife.43584>

## Figure Legends

### Figure 1. Limited proteolysis of 1N3R and 1N4R Tau monomers and fibrils using endoproteinase GluC.

Negatively-stained transmission electron microscopy (TEM) of 1N3R (a) and 1N4R (b) Tau fibrils. Proteolytic degradation profiles of 1N3R and 1N4R Tau (40 $\mu$ M monomer concentration) monomers (c and d respectively) and fibrils (e and f respectively) with GluC were monitored over time by coomassie stained SDS-PAGE. Time courses of proteolysis (min) and molecular weight markers (MW, kDa) are indicated. Most intense proteolytic fragments products generated at 45 min proteolysis constituting the molecular code-bar of 1N3R (bands 1 to 4) and 1N4R (Bands 1 to 6) Tau fibrils were further identified (g, h). Identification of proteolysis products and the entire protein at time zero (band 0) for 1N3R and 1N4R are summarized in tables i and j respectively. Identifications were performed by immune-detection after western-blotting, and N-terminal sequencing using both Edman sequencing after transfer in a PVDF membrane and specific N-terminal TMPP-labelling followed by proteolysis and mass spectrometry identification of the N-terminally modified peptide. Epitopes of antibodies Nter, HT7, Tau5, 3RD, 4RD and Cter used for immunodetection are indicated on sequence under the tables.

### Figure 2. Changes in enzyme accessibility upon assembly of monomeric 1N3R and 1N4R Tau into fibrils.

After in-solution digestion of monomeric and fibrillar Tau 1N3R (a) and 1N4R (b) with GluC, resulting proteolytic peptides were identified and quantified by nanoLC-MSMS. Changes in enzyme accessibility were measured by calculating the fold changes ( $F_c$ ) between fibrillary and monomeric Tau, for both 1N3R (a) and 1N4R (b) isoforms for each GluC proteolytic peptide ( $F_c = \text{Intensity}_{\text{Fib peptide 1}} / \text{Intensity}_{\text{Mono peptide 1}}$ ). Only peptides with significant changes in enzyme accessibility upon Tau assembly into fibrils are represented: peptides with  $0.1 \leq F_{cF/MF/M} \leq 0.55$  (in blue) are at least 1.8 times less accessible within the fibrils; peptides with

$F_{c\ F/M} < 0.1$  (in red) are at least 10 times less intense or not even detected after proteolysis of fibrils. Peptides with no significant changes are grey ( $0.55 < F_{c\ F/M} < 1.8$ ). Cleavage sites of the quantified peptides are in bold. The amino acid sequences covered by the identified proteolytic peptides are in black, while the non-identified sequences are in grey. The complete sequence coverage is detailed in supplementary information.

### **Figure 3. Changes in lysine modification upon assembly of monomeric 1N3R and 1N4R Tau into fibrils**

Changes in accessibility of reactive amino acid residues (Lysine and N-terminal residue) to NHS-Biotin upon assembly of 1N3R (a) and 1N4R (b) Tau into fibrils followed by proteolysis with GluC and AspN using a complete in-gel digestion of denatured Tau samples. Resulting biotinylated peptides, together with their site(s) of biotinylation(s), were identified and quantified by nanoLC-MSMS. Changes in lysine accessibility was measured by calculating the fold changes of biotinylated peptide intensity between fibrillar and monomeric 1N3R (a) and 1N4R (b) Tau: ( $F_{c\ F/M\ Biotinylated\ Peptide\ at\ K\ position} = \frac{Intensity_{Fib\ Biotinylated\ Peptide}}{Intensity_{Mono\ Biotinylated\ Peptide}}$ ). Four categories of biotinylated peptides are shown. Two correspond to solvent accessible lysine at the surface of Tau fibrils: in grey, peptides biotinylated in a similar extend in both monomeric and fibrillary Tau ( $0.55 < F_c \leq 1.8$ ); in green, peptides more extensively biotinylated in Tau fibrils than in monomers ( $F_c > 1.8$ ). Two categories highlight regions that become protected from biotinylation upon assembly into fibrils: in blue, regions moderately protected ( $0.1 < F_c < 0.55$ ); in red, regions highly protected or even not accessible to solvent at all ( $0.1 < F_c$ ). Colored square label the position of the biotinylated lysine according to the previous color code.

### **Figure 4. Changes in enzyme accessibility combined to lysine biotinylation upon assembly of monomeric 1N3R and 1N4R Tau into fibrils**

Changes in enzyme accessibility combined to changes in lysine and N-terminus modification were obtained for both 1N3R (a) and 1N4R (b) Tau, after in-solution digestion of entire

biotinylated Tau and comparison of the intensities of biotinylated GluC peptides generated from Tau fibrils or monomers. Resulting biotinylated peptides, together with their site(s) of biotinylation(s) (colored squares), were identified and quantified by nanoLC-MSMS. The four same peptides categories as those described in legend of figure 3 are represented.

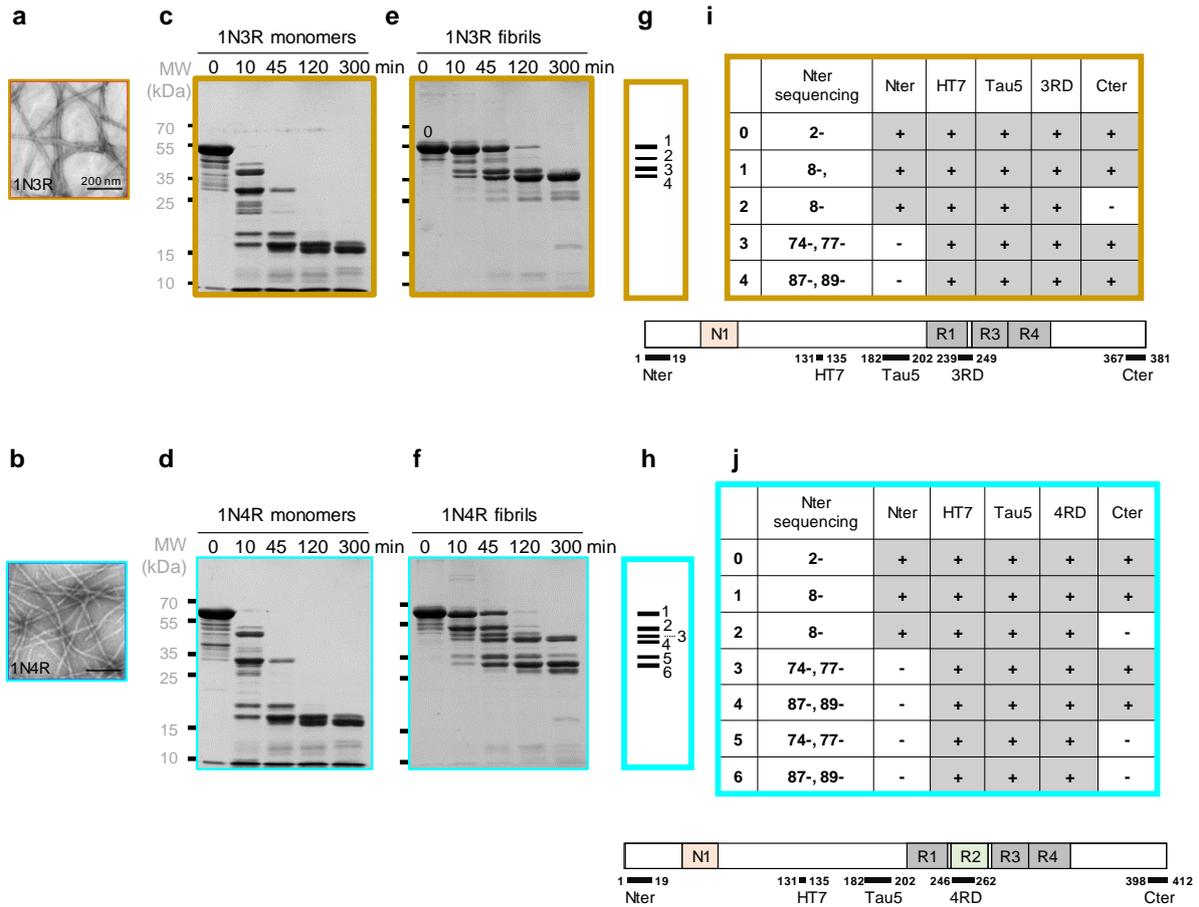
**Figure 5. Summary map of 1N3R and 1N4R Tau surface accessibility and protection changes upon assembly of monomers into fibrils.**

Graphic representation along the amino acid sequence of Tau of the changes in solvent-accessibility upon assembly of 1N3R (a) and 1N4R (b) Tau into fibrils. Grey histograms correspond to enzyme accessibility only presented in Figure 2 and black lines correspond to lysine and N-terminus biotinylation only presented in Figure 3. For easy comparison of the differences in surface accessibility between 1N3R and 1N4R Tau fibrils, the sequences were numbered following the 1N4R Tau amino acid sequence for both 1N3R (a) and 1N4R (b). The fold changes correspond to the ratio of the intensity of peptides identified for fibrillary Tau to the intensity of the same peptides identified for monomeric Tau ( $F_c F/M$ ). The additional microtubule binding repeat R2, ranging from amino acid 249 to 282 is specific to 1N4R Tau is indicated hatched on the sequence. As a consequence of this difference in 1N3R and 1N4R sequence, Lys288 and 292 were only for 1N3R Tau, and Lys245 was detected only for 1N4R Tau. The dotted lines indicate the threshold of significative fold changes at 0.55 corresponding to regions or residues protected at least 1.8 times more in Tau fibrils compared to monomeric Tau. The scale has been cropped for fold change over 2,5 a.u.

**Figure 6. Comparison of different structural data obtained on distinct Tau filaments**

The sequence of 2N4R Tau, corresponding to the longest Tau isoform sequence is shown at the very top. For 1N3R Tau sequences the missing 4R-specific repeat is represented as a dotted line. Structural data obtained for recombinant Tau fibrils are shown at the top. First are the results obtained in the current study with a summary of the solvent protected regions that we have identified in our heparin-induced 1N3R and 1N4R Tau fibrils, using two

complementary protein proteomic surface accessibility approaches (red and blue boxes correspond to respectively highly and moderately protected regions identified by enzyme accessibility. Second, high-resolution structural data obtained by cryo-EM with recombinant heparin-induced tau filaments made of full-length Tau proteins are shown (top). Structural data reported for different Tau filaments purified from AD or PiD patient's brain are reported at the bottom. First are shown previous structural information obtained on Tau filaments using proteolysis experiments combined to identification of proteolysis products using immunodetection (Jakes 1991 (1); Hasegawa 2014 (2); Taniguchi-Watanabe 2019 (3). Second high-resolution structures obtained by cryo-electron microscopy of the pronase resistant core of Tau filaments are presented (Fitzpatrick 2017 (3); Falcon 2018a (4); Zhang 2019 (6))



Limited proteolysis of 1N3R and 1N4R Tau monomers and fibrils using endoproteinase GluC

**Figure 1**

Caroux E et al

**a** **1N3R**

```

1  mAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT
51  PTEDGSEEPG SETSDAKSTP TAEAEAGIG DTPSLEDEAA GHVTQARMVS
101 KSKDGTGSDD KKAKGADGKT KIATPRGAAP PGQKGQANAT RIPAKTPPAP
151 KTPPSSGEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV
201 RTPPKSPSSA KSRLQTAPVP MPDLKNVSK IGSTENLKHQ PGGGKVQIVY
251 KPVDLSKVT SKCGLGNIH KPGGGQVEV SEKLDFKDRV QSKIGSLDNI
301 THVPGGGNKK IETHKLTFR NAKAKTDHGA EIVYKSPVVS GDTSPRHLSN
351 VSSTGSIDMV DSPQLATLAD EVSASLAKQG L 381

```

**b** **1N4R**

```

1  mAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT
51  PTEDGSEEPG SETSDAKSTP TAEAEAGIG DTPSLEDEAA GHVTQARMVS
101 KSKDGTGSDD KKAKGADGKT KIATPRGAAP PGQKGQANAT RIPAKTPPAP
151 KTPPSSGEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV
201 RTPPKSPSSA KSRLQTAPVP MPDLKNVSK IGSTENLKHQ PGGGKVQIIN
251 KKLDSLNVQS KCGSKDNIKH VPGGGSVQIV YKPVDLSKVT SKCGLGNIH
301 HKPGGGQVEV KSEKLDFKDR VQSKIGSLDN ITHVPGGGNK KIETHKLTFR
351 ENAKAKTDHG AEIVYKSPVV SGGTSPRHLS NVSSTGSIDM VDSPQLATLA
401 DEVSASLAKQ GL 412

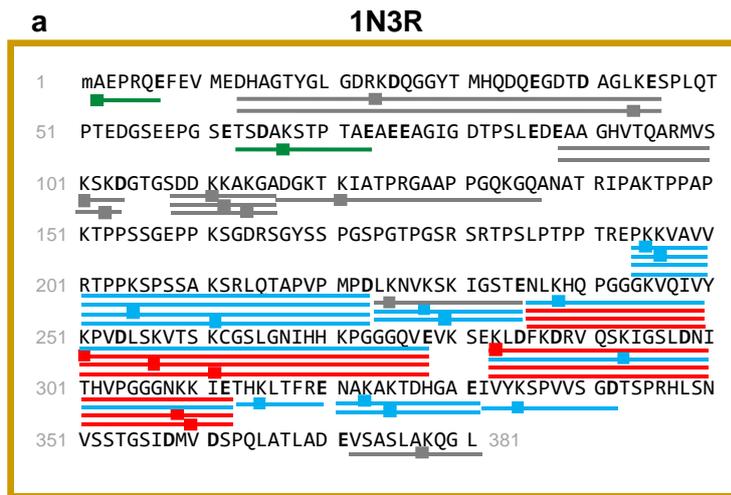
```

- Fib ≈ Mon
- Fib < Mon
- Fib << Mon

Changes in enzyme accessibility upon assembly of monomeric 1N3R and 1N4R into fibrils

## Figure 2

Caroux E et al

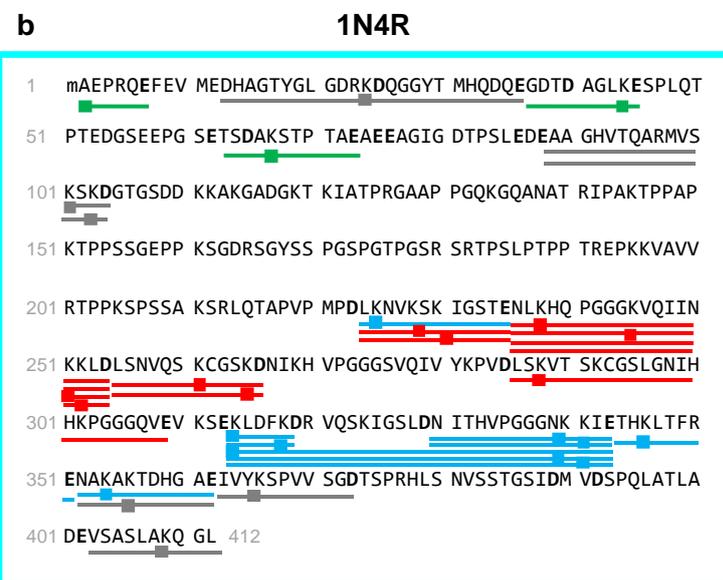
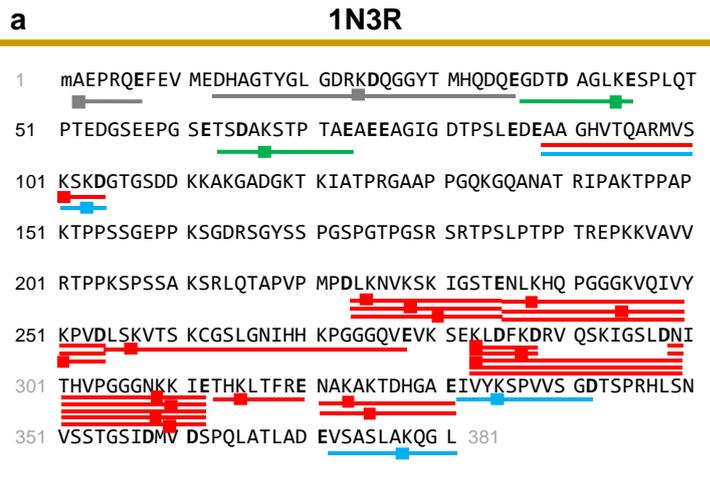


- Fib > Mon
- Fib ≈ Mon
- Fib < Mon
- Fib << Mon

Changes in lysine modification upon assembly of monomeric 1N3R and 1N4R into fibrils

**Figure 3**

Caroux E et al

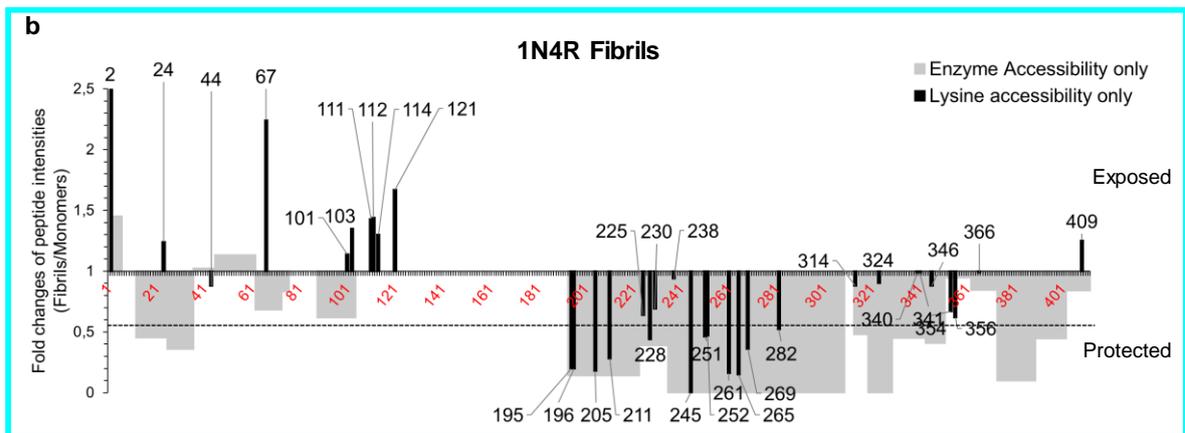
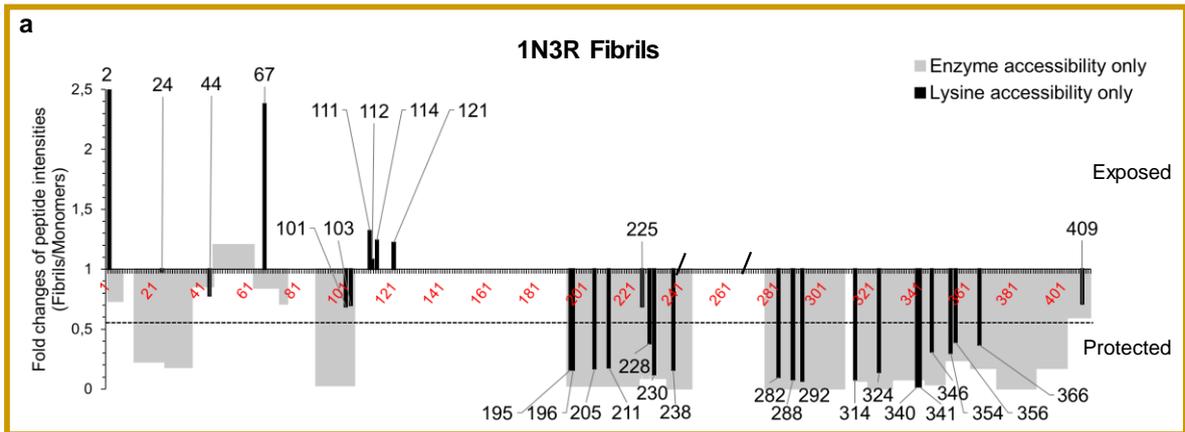


- Fib > Mon
- Fib ≈ Mon
- Fib < Mon
- Fib << Mon

Changes in enzyme accessibility combined to lysine modification upon assembly of monomeric 1N3R and 1N4R into fibrils

## Figure 4

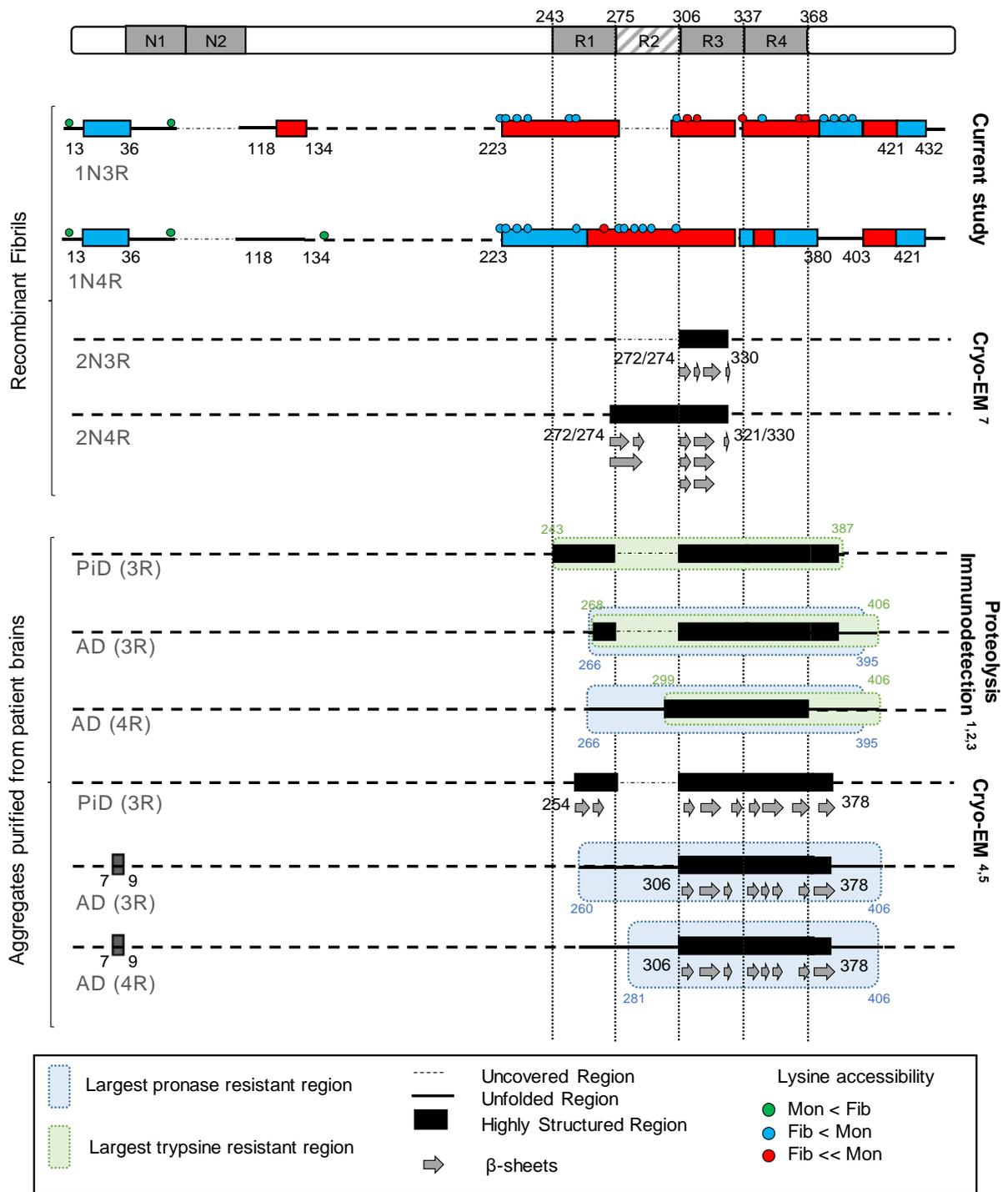
Caroux E et al



Map of 1N3R and 1N4R Tau surface accessibility and protection changes upon assembly of monomers into fibrils

**Figure 5**

Caroux E et al



Comparison of different structural data obtained on distinct Tau filaments

**Figure 6**

Caroux E et al

### Monomeric 1N3R (75%)

```

1 MAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT PTEGSEEPG SETSDAKSTP TAEAEAGIG DTPSLEDEAA GHVTQARMVS KSKDGTGSDD 110
111 KKAKGADGKT KIATPRGAAP PGQKGQANAT RIPAKTPPAP KTPPSSGEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV RPPKSPSSA KSRLQTAPVP 220
221 MPDLKNVRSK IGSTENLKHQ PGGGKVOIVY KPVDSLKVTG KCGSLGNIHH KPGGGQVEVK SEKLDLDFKDRV QSKIGSLDNI THVPGGGNKK IETHKLTFR NAKAKTDHGA 330
331 EIVYKSPVVS GDTSPRHLSN VSSTGSIDMV DSPQLATLAD EVSASLAKQG L 381

```

### Fibrillar 1N3R (48%)

```

1 MAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT PTEGSEEPG SETSDAKSTP TAEAEAGIG DTPSLEDEAA GHVTQARMVS KSKDGTGSDD 110
111 KKAKGADGKT KIATPRGAAP PGQKGQANAT RIPAKTPPAP KTPPSSGEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV RPPKSPSSA KSRLQTAPVP 220
221 MPDLKNVRSK IGSTENLKHQ PGGGKVOIVY KPVDSLKVTG KCGSLGNIHH KPGGGQVEVK SEKLDLDFKDRV QSKIGSLDNI THVPGGGNKK IETHKLTFR NAKAKTDHGA 330
331 EIVYKSPVVS GDTSPRHLSN VSSTGSIDMV DSPQLATLAD EVSASLAKQG L 381

```

### Monomeric 1N4R (74%)

```

1 MAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT PTEGSEEPG SETSDAKSTP TAEAEAGIG DTPSLEDEAA GHVTQARMVS KSKDGTGSDD 110
111 KKAKGADGKT KIATPRGAAP PGQKGQANAT RIPAKTPPAP KTPPSSGEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV RPPKSPSSA KSRLQTAPVP 220
221 MPDLKNVRSK IGSTENLKHQ PGGGKVOIIN KKLDSLNVQS KCGSKDNIHK VPGGGSVQIV YKPVDSLKVT SKCGSLGNIH HKPGGGQVEV KSEKLDLDFKDR VQSKIGSLDN 330
331 ITHVPGGGNK KIETHKLTFR ENAKAKTDHG AEIVYKSPVV SGTSPRHLS NVSSTGSIDM VDSPQLATLA DEVSASLAKQ GL 412

```

### Fibrillar 1N4R (62%)

```

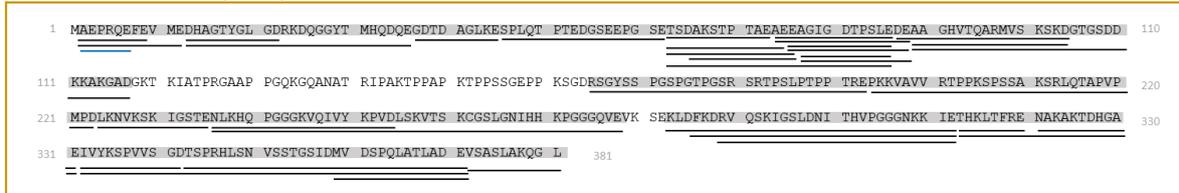
1 MAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT PTEGSEEPG SETSDAKSTP TAEAEAGIG DTPSLEDEAA GHVTQARMVS KSKDGTGSDD 110
111 KKAKGADGKT KIATPRGAAP PGQKGQANAT RIPAKTPPAP KTPPSSGEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV RPPKSPSSA KSRLQTAPVP 220
221 MPDLKNVRSK IGSTENLKHQ PGGGKVOIIN KKLDSLNVQS KCGSKDNIHK VPGGGSVQIV YKPVDSLKVT SKCGSLGNIH HKPGGGQVEV KSEKLDLDFKDR VQSKIGSLDN 330
331 ITHVPGGGNK KIETHKLTFR ENAKAKTDHG AEIVYKSPVV SGTSPRHLS NVSSTGSIDM VDSPQLATLA DEVSASLAKQ GL 412

```

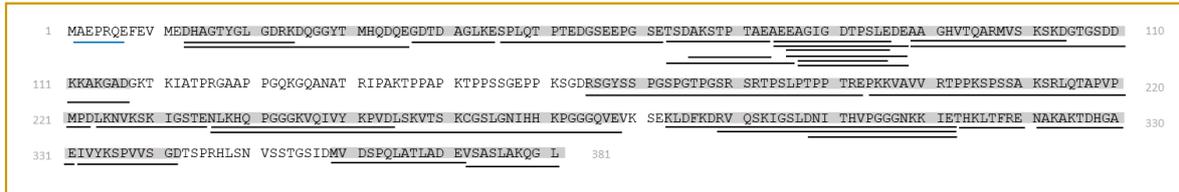
**Figure S1. Sequence coverage of Control Samples in-solution GluC digestion of monomeric and fibrillar Tau 1N3R and Tau 1N4R**

MASCOT MS/MS Ion Search Parameters : (Enzyme : Glu-C ; Fixed modification : Ø ; Variable modifications : Biotin (CSTY, K, Protein N-term), Oxidation (M) ; Mass values : Monoisotopic ; Peptide mass tolerance : ± 40 ppm; Fragment mass tolerance : ±0.05 Da ; Max missed cleavages : 5 ; Instrument type : ESI-QUAD-TOF ; Cut-off = 20) Peptides identified at least in two replicates in black ; Peptide manually found in blue. n=3

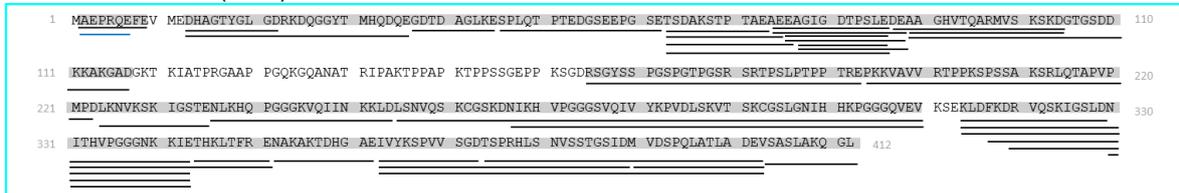
### Monomeric 1N3R (86%)



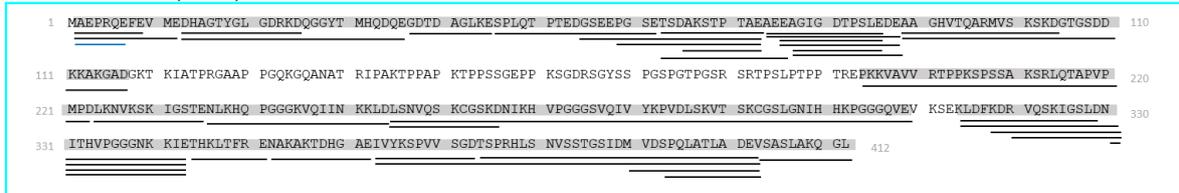
### Fibrillar 1N3R (79%)



### Monomeric 1N4R (87%)



### Fibrillar 1N4R (80%)



**Figure S2. Sequence coverage of Control Samples in-gel GluC digestion of monomeric and fibrillar Tau 1N3R and Tau 1N4R**

MASCOT MS/MS Ion Search Parameters : (Enzyme : Glu-C ; Fixed modification : Carbamidomethylation (C) ; Variable modifications : Biotin (CSTY, K, Protein N-term), Oxidation (M) ; Mass values : Monoisotopic ; Peptide mass tolerance :  $\pm 40$  ppm; Fragment mass tolerance :  $\pm 0.05$  Da ; Max missed cleavages : 5 ; Instrument type : ESI-QUAD-TOF ; Cut-off = 20)

Peptides identified at least in two replicates in black ; Peptide manually found in blue. n=3

### Monomeric 1N3R (92%)

```

1 MAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT PTEDGSEEPG SETSDAKSTP TAEAEAEAGIG DTPELEDEAA GHVTQARMVS KSKDGTGSDD 110
111 KKAKGADGKT KIATPRGAAP PGQKQANAT RIPAKTPPAP KTPSSGSEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV RTPPKSPSSA KSRLQTAPVP 220
221 MPDLKRVKSK IGSTENLKHQ PGGGKQIVY KPVDSLKVT KCGSLGNIH KPGGGQVEVK SEKLDKDRV QSKIGSLDNI THVPGGGNKK IETHKLTFR ENAKAKTDHGA 330
331 EIVYKSPVVS GDTSPRHLSN VSSTGSIDMV DSPQLATLAD EVSASLAKQ G 381

```

### Fibrillar 1N3R (92%)

```

1 MAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT PTEDGSEEPG SETSDAKSTP TAEAEAEAGIG DTPELEDEAA GHVTQARMVS KSKDGTGSDD 110
111 KKAKGADGKT KIATPRGAAP PGQKQANAT RIPAKTPPAP KTPSSGSEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV RTPPKSPSSA KSRLQTAPVP 220
221 MPDLKRVKSK IGSTENLKHQ PGGGKQIVY KPVDSLKVT KCGSLGNIH KPGGGQVEVK SEKLDKDRV QSKIGSLDNI THVPGGGNKK IETHKLTFR ENAKAKTDHGA 330
331 EIVYKSPVVS GDTSPRHLSN VSSTGSIDMV DSPQLATLAD EVSASLAKQ G 381

```

### Monomeric 1N4R (92%)

```

1 MAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT PTEDGSEEPG SETSDAKSTP TAEAEAEAGIG DTPELEDEAA GHVTQARMVS KSKDGTGSDD 110
111 KKAKGADGKT KIATPRGAAP PGQKQANAT RIPAKTPPAP KTPSSGSEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV RTPPKSPSSA KSRLQTAPVP 220
221 MPDLKRVKSK IGSTENLKHQ PGGGKQIIN KKLDSLNVQS KCGSKDNIKH VPGGGVQIV YKPVDSLKVT SKCGSLGNIH HKPGGGQVEV KSEKLDKDR VQSKIGSLDN 330
331 ITHVPGGGNK KIETHKLTFR ENAKAKTDHG AEIVYKSPV SGTSPRHLS NVSSTGSIDM VDSPQLATLA DEVSASLAKQ GL 412

```

### Fibrillar 1N4R (92%)

```

1 MAEPRQEFEV MEDHAGTYGL GDRKDQGGYT MHQDQEGDTD AGLKESPLQT PTEDGSEEPG SETSDAKSTP TAEAEAEAGIG DTPELEDEAA GHVTQARMVS KSKDGTGSDD 110
111 KKAKGADGKT KIATPRGAAP PGQKQANAT RIPAKTPPAP KTPSSGSEPP KSGDRSGYSS PGSPGTPGSR SRTPSLPTPP TREPKKVAVV RTPPKSPSSA KSRLQTAPVP 220
221 MPDLKRVKSK IGSTENLKHQ PGGGKQIIN KKLDSLNVQS KCGSKDNIKH VPGGGVQIV YKPVDSLKVT SKCGSLGNIH HKPGGGQVEV KSEKLDKDR VQSKIGSLDN 330
331 ITHVPGGGNK KIETHKLTFR ENAKAKTDHG AEIVYKSPV SGTSPRHLS NVSSTGSIDM VDSPQLATLA DEVSASLAKQ GL 412

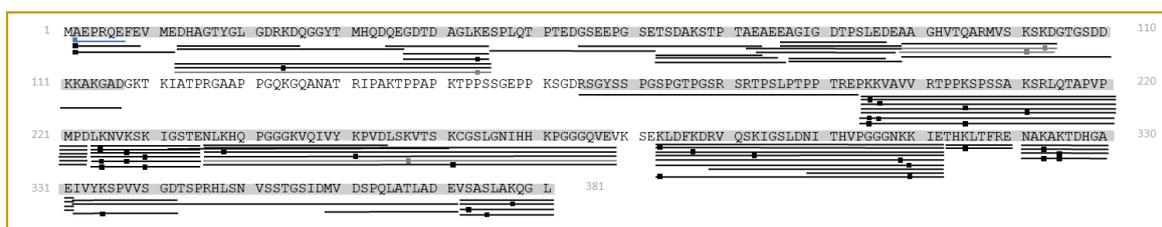
```

**Figure S3. Sequence coverage of Control Samples in-gel AspN digestion of monomeric and fibrillar Tau 1N3R and Tau 1N4R**

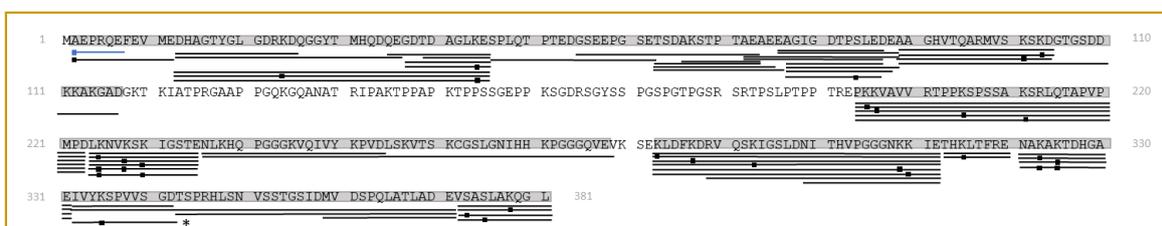
MASCOT MS/MS Ion Search Parameters : (Enzyme : AspN, Semi-AspN ; Fixed modification : Carbamidomethylation (C) ; Variable modifications : Biotin (CSTY, K, Protein N-term), Oxidation (M) ; Mass values : Monoisotopic ; Peptide mass tolerance :  $\pm 40$  ppm; Fragment mass tolerance :  $\pm 0.05$  Da ; Max missed cleavages : 5 ; Instrument type : ESI-QUAD-TOF ; Cut-off = 20)

Peptides identified at least in two replicates in black. n=3

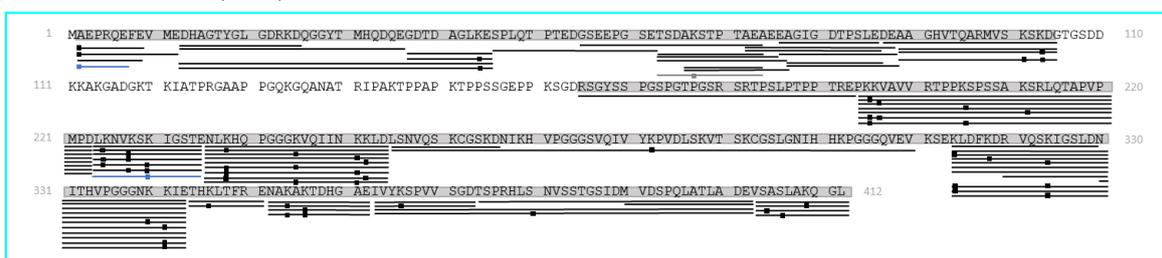
### Monomeric 1N3R (86%)



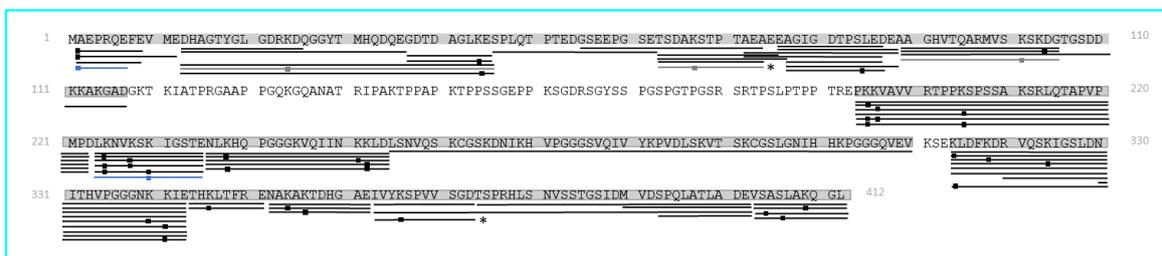
### Fibrillar 1N3R (78%)



### Monomeric 1N4R (84%)



### Fibrillar 1N4R (80%)

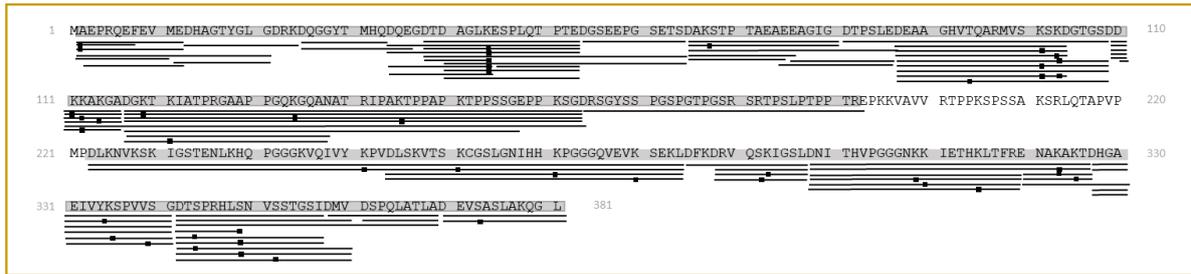


**Figure S4. Sequence coverage of Biotinylated Samples in-gel GluC digestion of monomeric and fibrillar Tau 1N3R and Tau 1N4R**

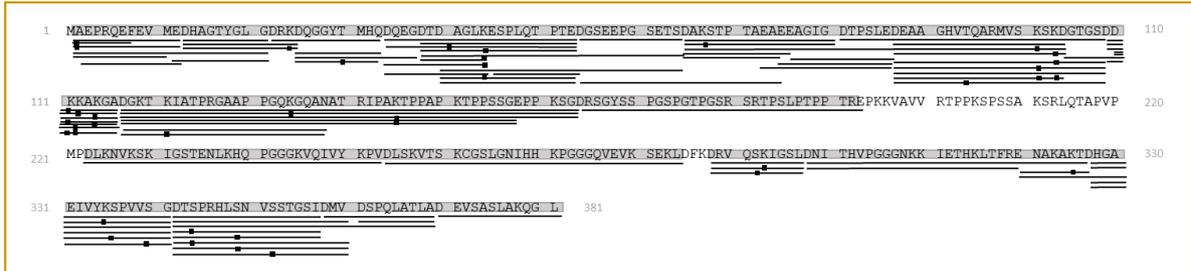
MASCOT MS/MS Ion Search Parameters : (Enzyme : Glu-C ; Fixed modification : Carbamidomethylation (C) ; Variable modifications : Biotin (CSTY, K, Protein N-term), Oxidation (M) ; Mass values : Monoisotopic ; Peptide mass tolerance :  $\pm 40$  ppm; Fragment mass tolerance :  $\pm 0.05$  Da ; Max missed cleavages : 5 ; Instrument type : ESI-QUAD-TOF ; Cut-off = 20 ; \* Cut-off <20 )

Peptides identified at least in two replicates in black ; Peptides identified at least in one replicate in grey ; Peptides manually found in blue. n=3

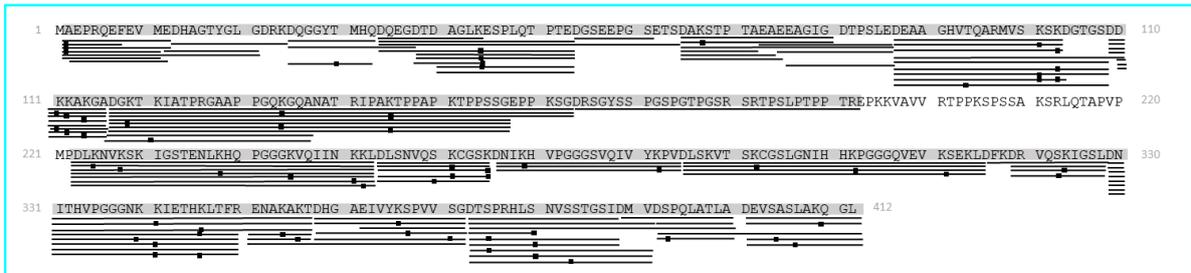
### Monomeric 1N3R (92%)



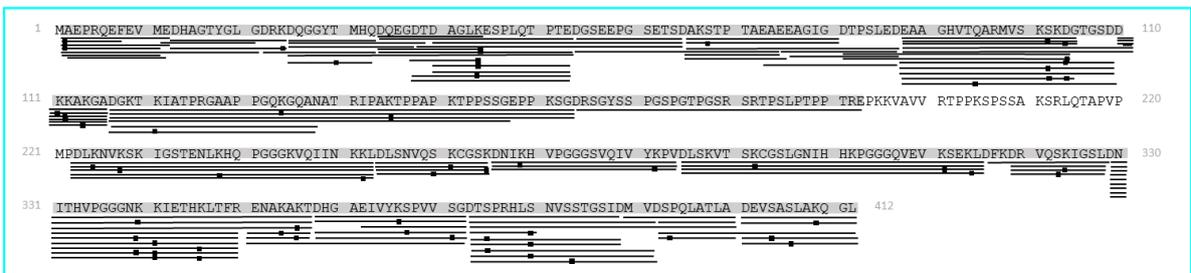
### Fibrillar 1N3R (91%)



### Monomeric 1N4R (92%)



### Fibrillar 1N4R (92%)

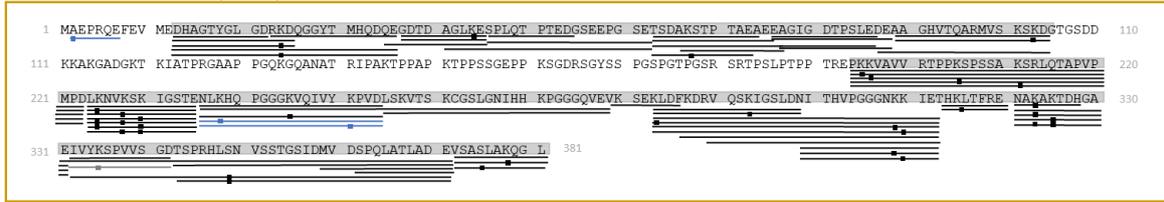


**Figure S5. Sequence coverage of Biotinylated Samples in-gel AspN digestion of monomeric and fibrillar Tau 1N3R and Tau 1N4R**

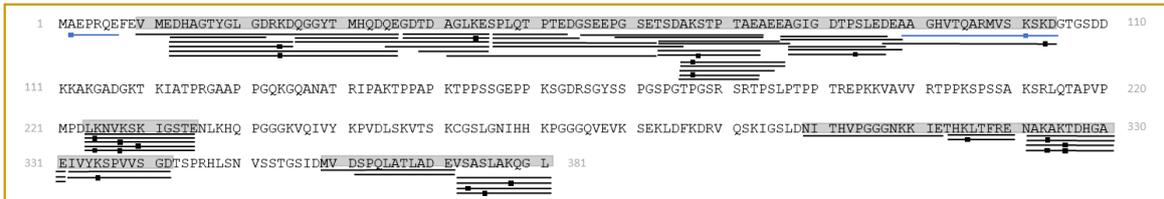
MASCOT MS/MS Ion Search Parameters : (Enzyme : AspN, Semi-AspN ; Fixed modification : Carbamidomethylation (C) ; Variable modifications : Biotin (CSTY, K, Protein N-term), Oxidation (M) ; Mass values : Monoisotopic ; Peptide mass tolerance :  $\pm 40$  ppm ; Fragment mass tolerance :  $\pm 0.05$  Da ; Max missed cleavages : 5 ; Instrument type : ESI-QUAD-TOF ; Cut-off = 20)

Peptides identified at least in two replicates in black. n=3

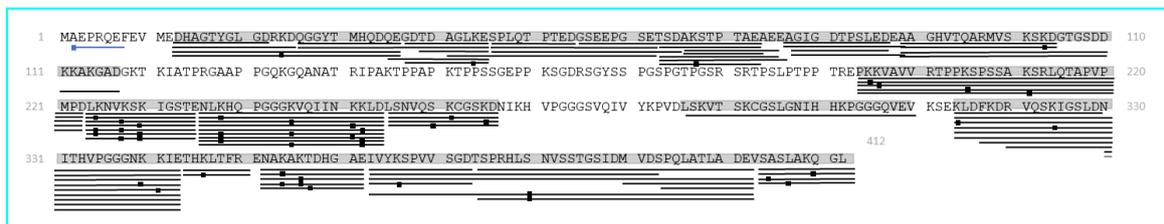
### Monomeric 1N3R (73%)



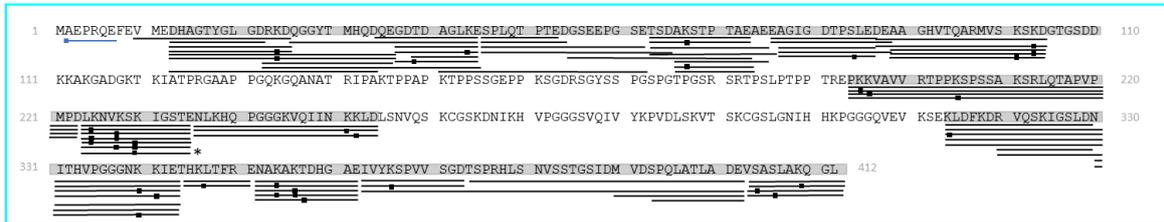
### Fibrillar 1N3R (45%)



### Monomeric 1N4R (73%)

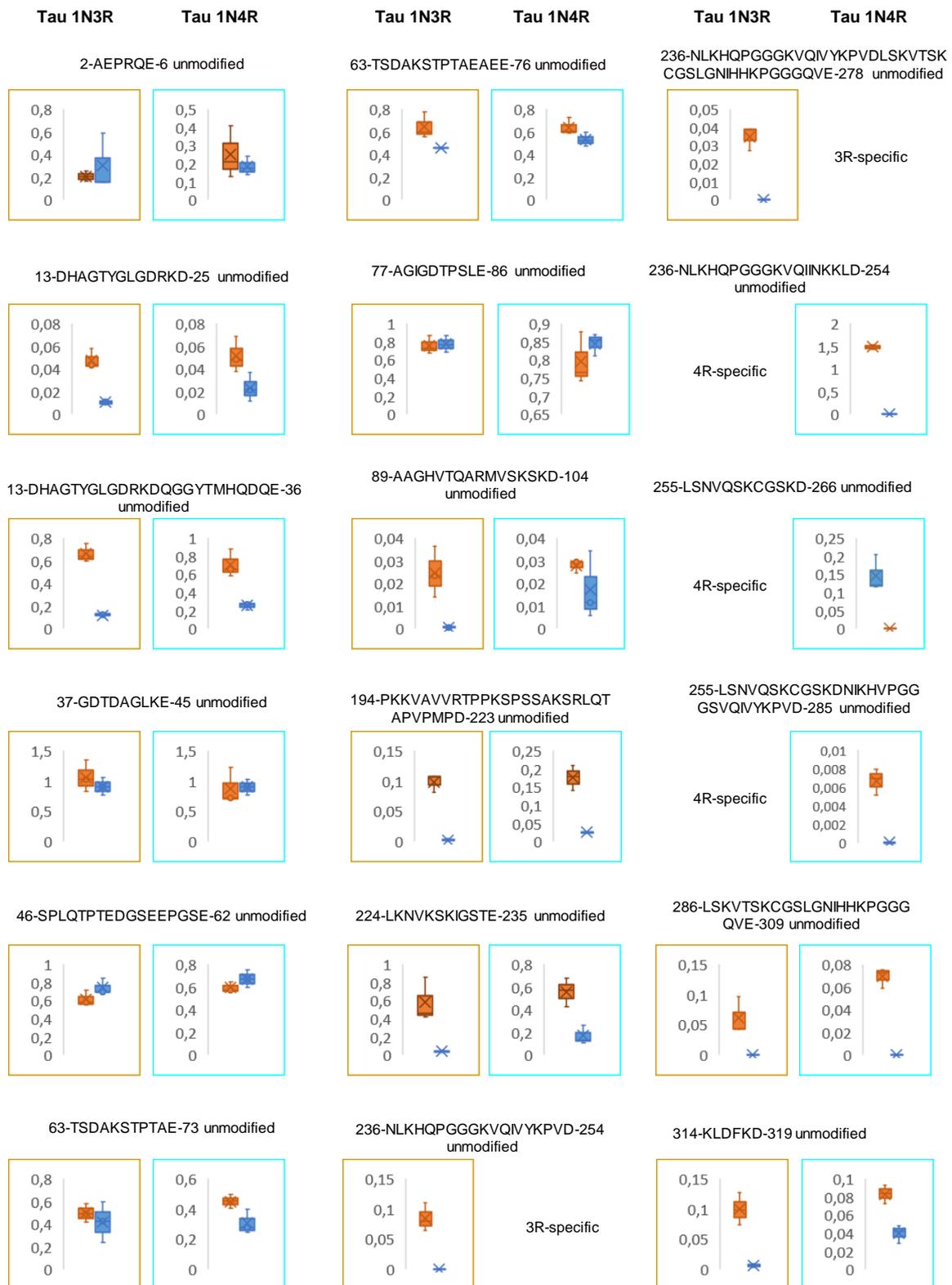


### Fibrillar 1N4R (62%)



**Figure S6. Sequence coverage of Biotinylated Samples in-solution GluC digestion of monomeric and fibrillar Tau 1N3R and Tau 1N4R**

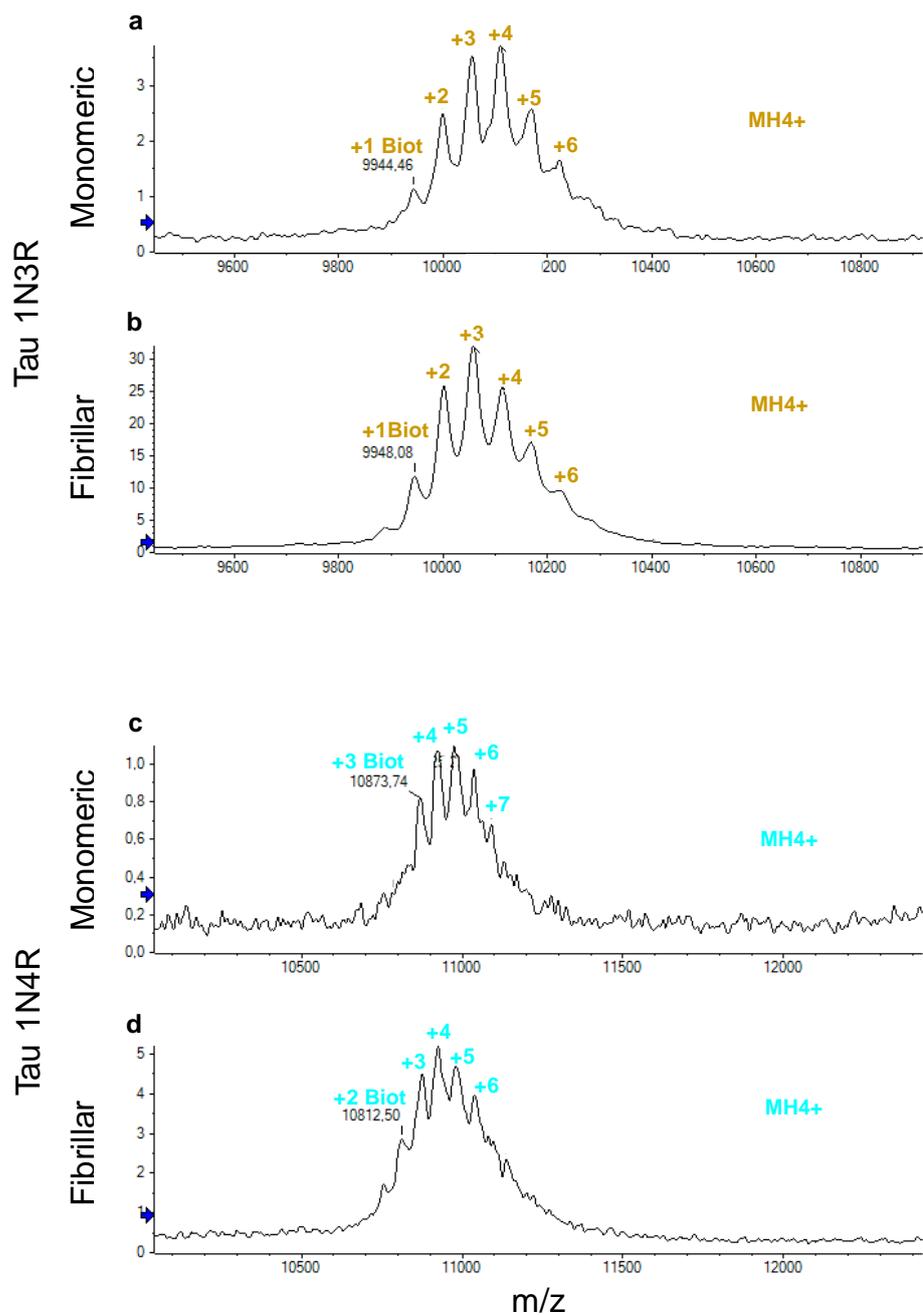
MASCOT MS/MS Ion Search Parameters : (Enzyme : Glu-C ; Fixed modification : Ø ; Variable modifications : Biotin (CSTY, K, Protein N-term), Oxidation (M) ; Mass values : Monoisotopic ; Peptide mass tolerance : ± 40 ppm; Fragment mass tolerance : ±0.05 Da ; Max missed cleavages : 5 ; Instrument type : ESI-QUAD-TOF ; Cut-off = 20 ; \* Cut-off <20) Peptides identified at least in two replicates in black ; Peptides identified at least in one replicate in grey ; Peptides manually found in blue. n=3



**Figure S7.a Changes in Enzyme Accessibility upon Tau 1N3R and 1N4R assemblies (Boxplots 1)**  
 GluC in-solution digestion. The average is represented by a cross and the median by a line. **Tau 1N4R** and **Tau 1N3R** Monomers (orange) and Fibrils (blue). n=3

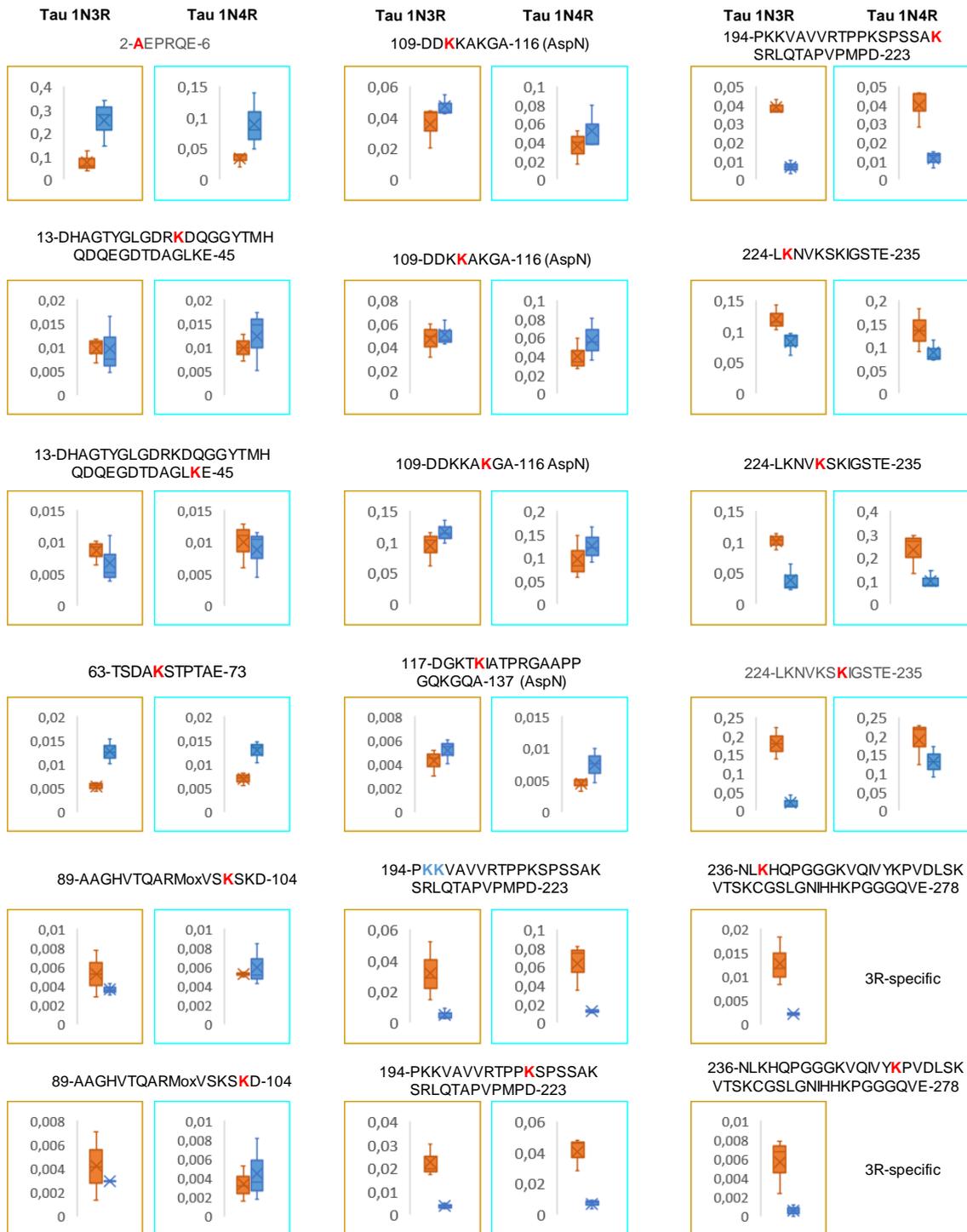


**Figure S7.b Changes in Enzyme Accessibility upon Tau 1N3R and 1N4R assemblies (Boxplots 2)**  
GluC in-solution digestion. The average is represented by a cross and the median by a line. **Tau 1N4R** and **Tau 1N3R** Monomers (orange) and Fibrils (blue). n=3



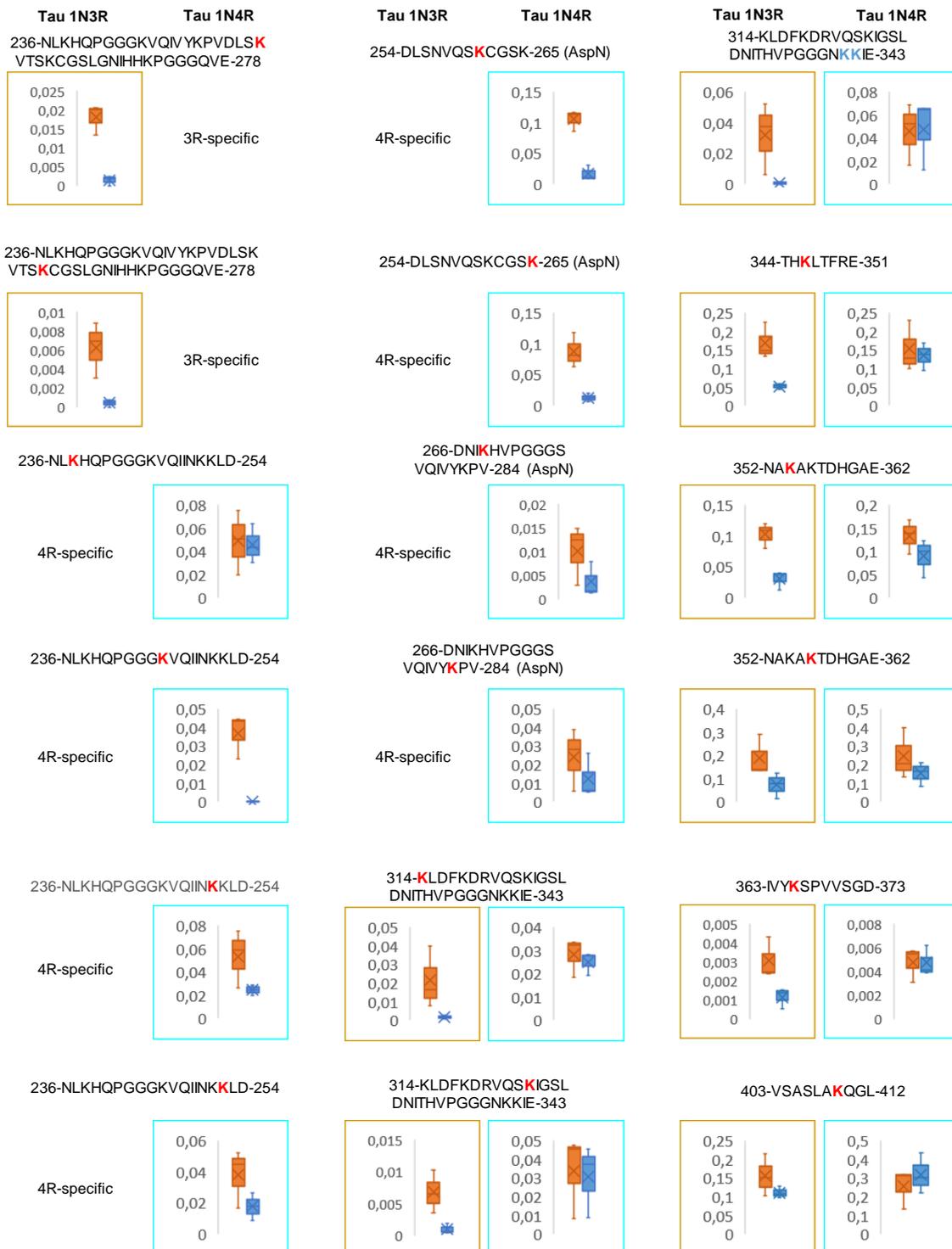
**Figure S8. Biotinylation Efficiency on full Protein**

MALDI-MS Spectra (Cf. Online Methods) of biotinylated 1N3R monomers (a) and fibrils (b); 1N4R monomers (c) and fibrils (d). Charge State ( $z$ ) used to determine the number of additional biotin molecules per tau unit (+ $n$ )



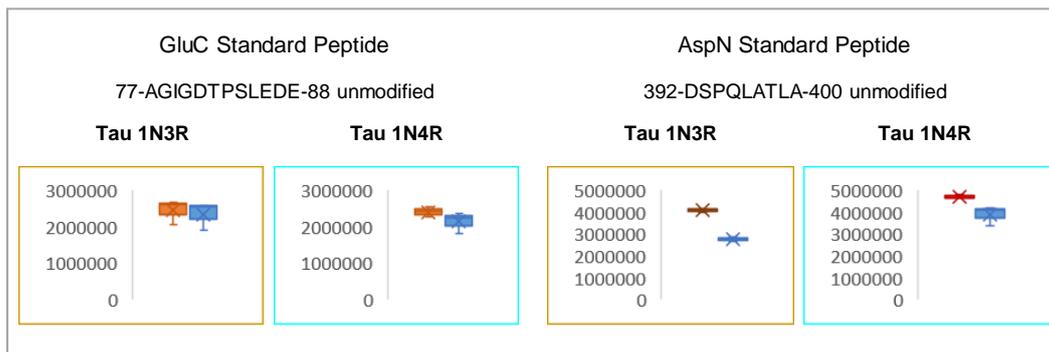
**Figure S9.a Changes in Surface labelling upon Tau 1N3R and 1N4R assemblies (Boxplots)**

The average is represented by a cross and the median by a line. **Tau 1N4R** and **Tau 1N3R** Monomers (orange) and Fibrils (blue). n=3



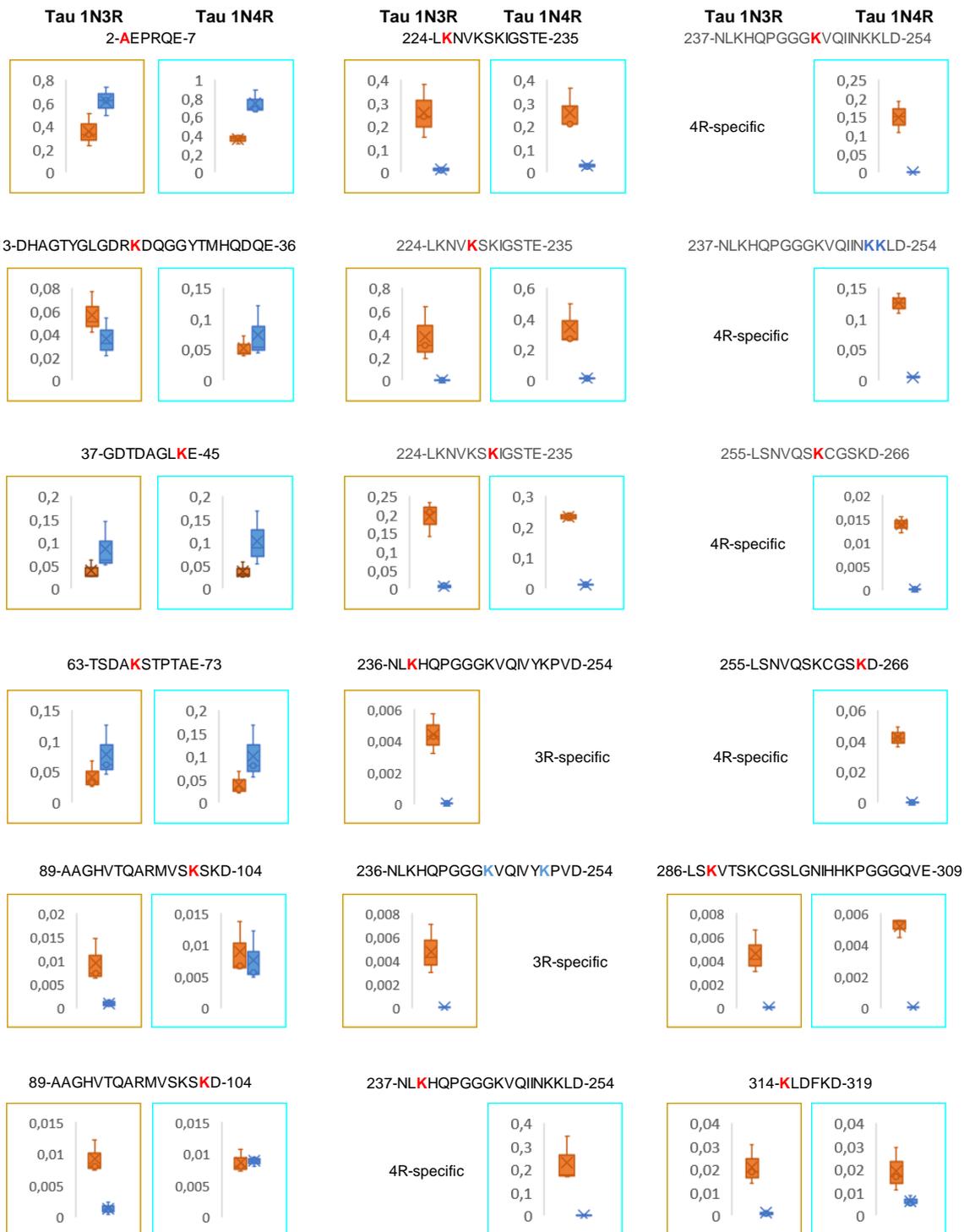
**Figure S9.b Changes in Surface labelling upon Tau 1N3R and 1N4R assemblies (Boxplots)**

The average is represented by a cross and the median by a line. **Tau 1N4R** and **Tau 1N3R** Monomers (orange) and Fibrils (blue). n=3



**Figure S9.c Changes in Surface labelling upon Tau 1N3R and 1N4R assemblies (Boxplots)**

The average is represented by a cross and the median by a line. **Tau 1N4R** and **Tau 1N3R** Monomers (orange) and Fibrils (blue). n=3



**Figure S10.a Changes in Enzyme Accessibility and Surface labelling upon Tau 1N3R and 1N4R assemblies (Boxplots)**

The average is represented by a cross and the median by a line. **Tau 1N4R** and **Tau 1N3R** Monomers (orange) and Fibrils (blue). n=3



**Figure S10.b Changes in Enzyme Accessibility and Surface labelling upon Tau 1N3R and 1N4R assemblies (Boxplots)**

The average is represented by a cross and the median by a line. **Tau 1N4R** and **Tau 1N3R** Monomers (orange) and Fibrils (blue). n=3

1N3R Peptide Sequence	m/z	z	Mr exp	Mth	Appm	Peptide Class	F c #/Mono	Significance (p-value)	1N3R Tau Fibrils Experiment 1		1N3R Tau Fibrils Experiment 2		1N3R Tau Fibrils Experiment 3		1N3R Tau Monomers Experiment 1		1N3R Tau Monomers Experiment 2		1N3R Tau Monomers Experiment 3		1N3R Tau Monomers Experiments			
									Intensity (XIC)	Normalized Intensity (XIC)	Intensity (XIC)	Normalized Intensity (XIC)	Intensity (XIC)	Normalized Intensity (XIC)	Intensity (XIC)	Normalized Intensity (XIC)	Intensity (XIC)	Normalized Intensity (XIC)	Intensity (XIC)	Normalized Intensity (XIC)	Intensity (XIC)	Normalized Intensity (XIC)	Intensity (XIC)	Normalized Intensity (XIC)
2-AEPRGE-7	365.1798	2	728.34508	728.3453	-0.30	1	0.73	0.8273	219333	0.1391	276971	0.1737	452336	0.2401	0.1843	0.0514	326339	0.1314	1093650	0.4081	608480	0.2133	0.2509	0.1422
13-DHAGTYLGDGRKD-25	468.6905	3	1403.6496	1403.643	4.70	2	0.23	0.0495	21476	0.0136	16848	0.0106	16325	0.0087	0.1110	0.0025	145029	0.0584	110002	0.0411	125219	0.0439	0.0478	0.0093
13-DHAGTYLGDGRKDDGYTMHQDQE-36	670.5451	4	2678.1416	2678.1416	3.62	2	0.18	0.0495	212457	0.1347	199443	0.1251	194892	0.1035	0.1211	0.0160	1874050	0.7543	1606764	0.5996	1821656	0.6395	0.6642	0.0805
37-GD TDAGLKE-45	453.2138	2	904.4131	904.4138	-0.77	1	0.85	0.5127	1684213	1.0678	1217262	0.7636	1707867	0.9067	0.1727	0.1522	3352689	1.3495	2226141	0.8308	2923405	1.0246	1.0683	0.2621
46-SPLPTIEDGSEEPGSE-62	860.3904	2	1756.7462	1756.7432	1.71	1	1.20	0.2752	1334094	0.8458	1109728	0.6962	1308965	0.6949	0.1456	0.0868	1759965	0.7205	1468777	0.5548	1660749	0.5821	0.6191	0.0888
63-TSDAKSTPTAE-73	554.2628	2	1106.511	1106.5091	1.72	1	0.84	0.8273	672975	0.4287	380766	0.2389	1127392	0.5985	0.4214	0.1799	1455976	0.5661	1120985	0.4183	1408798	0.4593	0.4994	0.0840
63-TSDAKSTPTAEAE-76	718.8234	2	1435.6322	1435.6314	0.56	1	0.71	0.0495	731452	0.4637	734071	0.4605	851167	0.4519	0.4587	0.0061	1927663	0.7789	1602345	0.5980	1599177	0.5605	0.6448	0.1157
77-AGIDITPSE-86	480.2386	2	958.4627	958.4607	2.09	1	1.02	0.5127	1206795	0.7651	1395993	0.8757	1297411	0.6886	0.7685	0.0940	2174270	0.8752	1962008	0.7322	1935361	0.6793	0.7619	0.1017
89-AAGHVTOARWVSKSKD-104	562.5336	3	1684.879	1684.8679	6.59	3	0.03	0.0495	1498	0.0009	889	0.0006	1166	0.0006	0.0007	0.0002	91038	0.0366	37164	0.0139	67390	0.0236	0.0247	0.0114
194-PKKVAVVTRPKSPSSAKSLQTPVPMPPD-223	634.7638	5	3168.7826	3168.7805	0.66	3	0.02	0.0495	1031	0.0007	3580	0.0022	8477	0.0045	0.0025	0.0019	267489	0.1077	289404	0.1080	231343	0.0811	0.0989	0.0154
224-LKNVKSQSTE-235	435.2572	3	1302.7496	1302.7507	-0.84	3	0.09	0.0495	117979	0.0747972	136069	0.0853589	106927	0.0567676	0.0723	0.0145	1598498	0.6434	2956731	1.1034	1558825	0.5464	0.7644	0.2976
236-NLKHQPGGKQVWYKVD-254	520.0455	4	2076.1529	2076.148	2.36	3	0.00	0.0495	72	0.0000	52	0.0000	0	0.0000	0.0000	0.0000	274403	0.1105	216041	0.0806	186082	0.0652	0.0854	0.0230
236-NLKHQPGGKQVWYKVD-254	749.4084	6	4480.4069	4480.3969	2.23	3	0.00	0.0495	0	0.0000	0	0.0000	0	0.0000	0.0000	0.0000	96437	0.0388	104651	0.0391	77439	0.0271	0.0350	0.0068
255-LSKVSKCGSLGNIHKPGGGQVE-278	609.0732	4	2432.2637	2432.2594	1.77	3	0.00	0.0495	7	0.0000	0	0.0000	0	0.0000	0.0000	0.0000	241487	0.0972	117129	0.0437	122245	0.0428	0.0613	0.0311
283-KLDFKID-288	383.2101	2	764.4057	764.4068	-1.44	3	0.07	0.0495	14169	0.0090	12897	0.0082	5726	0.0030	0.0067	0.0032	315266	0.1269	199849	0.0746	282498	0.0990	0.1002	0.0262
283-KLDFKIDVSKCGSLD-298	617.0121	3	1845.0143	1845.0105	2.06	3	0.00	0.0495	104	0.0001	54	0.0000	33	0.0000	0.0000	0.0000	94454	0.0380	70727	0.0264	86723	0.0304	0.0316	0.0059
298-NTHVPGGKNIKE-312	488.039	3	1462.7899	1462.7892	0.48	3	0.08	0.0495	16268	0.1013	11676	0.0073	4063	0.0022	0.0066	0.0041	259420	0.1008	151813	0.0567	289718	0.1012	0.0862	0.0256
313-THKLTFR-321	344.5254	3	1030.5543	1030.556	-1.66	3	0.04	0.0495	152444	0.0966	138982	0.0872	59966	0.0318	0.0719	0.0350	4976471	2.0031	5293544	1.9755	3564772	1.2564	1.7450	0.4234
321-NAKAKTDHGAE-331	521.2835	2	1140.5525	1140.5523	0.18	2	0.24	0.0495	104391	0.0662	125500	0.0787	40261	0.0214	0.0554	0.0302	363988	0.1465	1980267	0.4031	415445	0.1456	0.2318	0.1484
332-IVYKSPVYSGD-342	582.2312	2	1162.6278	1162.6234	3.87	2	0.18	0.0495	77423	0.0491	65881	0.0413	27661	0.0147	0.0350	0.0180	697947	0.2809	384831	0.1436	490218	0.1718	0.1988	0.0725
343-TSPRHLSVSTGSDMVDSPQLATLAD-371	1010.1598	3	3027.4575	3027.4455	2.74	2	0.00	0.0495	68	0.0000	80	0.0001	37	0.0000	0.0000	0.0000	269938	0.1087	170581	0.0637	194058	0.0680	0.0801	0.0248
359-MVDSPLATLAD-371	695.3338	2	1386.6531	1386.6493	2.74	2	0.18	0.0495	32366	0.0205	49121	0.0308	13250	0.0070	0.0195	0.0119	344165	0.1365	213526	0.0797	326572	0.1145	0.1109	0.0296
372-VSASIAKQGL-381	487.9688	2	972.5591	972.5604	-1.34	1	0.60	0.0495	394857	2.5035	3657142	2.2942	2869799	1.5236	2.1071	0.5161	9489458	3.8197	9691738	3.6168	8894028	3.1173	3.5179	0.3615
Standard Peptide 77-AGIDITPSE-88	602.2732	2	1202.5319	1202.5303	1.33	1	0.60	0.0495	1577318	-	1883591	-	1883591	-	-	-	2484338	-	2679615	-	2853123	-	-	-

Table S1: Changes in GluC protease accessibility for each proteolytic peptide of 1N3R Tau.

1N4R Peptide Sequence	m/z	z	Mr exp	Mth	Δppm	Peptide Class	F <sub>c</sub> R <sub>homo</sub>	Significance (p-value)	1N4R Tau Fibrils Experiment 1		1N4R Tau Fibrils Experiment 2		1N4R Tau Fibrils Experiment 3		1N4R Tau Monomers Experiment 1		1N4R Tau Monomers Experiment 2		1N4R Tau Monomers Experiment 3		1N4R Tau Monomers Experiments		
									Intensity (XIC)	Normalized Intensity	Intensity (XIC)	Normalized Intensity	Intensity (XIC)	Normalized Intensity	Intensity (XIC)	Normalized Intensity	Intensity (XIC)	Normalized Intensity	Intensity (XIC)	Normalized Intensity	Intensity (XIC)	Normalized Intensity	Intensity (XIC)
2-AEPRGE-7	365.1807	2	728.3468	728.3453	2.06	1	1.45	0.5127	334737	0.1574	1228619	0.8892	311929	0.1573	360564	0.1639	585550	0.2523	543148	0.2073	0.2073	0.2079	0.0442
13-DHAGTYGLGDRKD-25	468.8905	3	1403.6496	1403.6430	4.70	2	0.45	0.0495	78956	0.0371	24671	0.0119	42336	0.0234	151091	0.0687	87084	0.0375	127403	0.0486	0.0486	0.0516	0.0158
13-DHAGTYGLGDRKDDGGYTMHQDQE-36	670.5437	4	2678.1459	2678.1416	1.61	2	0.36	0.0495	625520	0.2941	425850	0.2045	531534	0.2881	1933447	0.8790	1364391	0.5879	1719307	0.6563	0.7077	0.7077	0.1522
37-GDTPDAQLKE-45	453.2142	2	904.4139	904.4138	0.11	1	1.02	0.5127	2207843	1.0380	1609137	0.7729	1789093	0.9024	2692001	1.2238	1678165	0.7231	1842547	0.7034	0.7034	0.6834	0.2950
46-SLPDTEIDGSEFPGE-62	860.3801	2	1758.7457	1758.7432	1.42	1	1.13	0.1266	1608971	0.7565	1242290	0.5967	1336720	0.6742	1432834	0.6514	1321414	0.5694	1497082	0.5715	0.5715	0.5974	0.0468
63-TSDAKSIPTAEAE-73	554.2627	2	1106.5109	1106.5091	1.63	1	0.68	0.0495	856760	0.4028	915547	0.2476	546555	0.2757	1001486	0.4553	936547	0.4035	1302253	0.4871	0.4871	0.4520	0.0469
63-TSDAKSIPTAEAE-76	718.8225	2	1435.6304	1435.6314	-0.70	1	0.83	0.1266	1278371	0.6010	995144	0.4780	1036521	0.5228	1601589	0.7281	1365404	0.5883	1593077	0.6081	0.6081	0.6415	0.0756
77-AGIGDTFSLLE-86	480.2377	2	958.4608	958.4607	-0.10	1	1.06	0.5127	1819923	0.8557	1697174	0.8126	1725071	0.8701	1932376	0.8785	1784881	0.7663	1944966	0.7424	0.7424	0.7957	0.0727
89-AAGHVYTOARVMVSKSKD-104	562.8310	3	1684.8711	1684.8679	1.90	1	0.62	0.5127	73370	0.0345	12333	0.0059	22839	0.0115	64663	0.0284	69009	0.0297	64663	0.0247	0.0247	0.0279	0.0028
194-PKKVAVVTRPPKSPSRSKSLRLOTAPVPIPD-223	634.7641	5	3168.7941	3168.7805	1.14	2	0.14	0.0495	50096	0.0236	57426	0.0276	51123	0.0258	462829	0.2104	328209	0.1414	468585	0.1793	0.1770	0.1770	0.0345
224-LNKKVKGSGTE-235	435.2573	3	1302.7500	1302.7507	-0.54	2	0.39	0.0495	961917	0.4523	424031	0.2037	486817	0.2456	1956339	0.8894	1460209	0.6292	2036806	0.7775	0.7654	0.7654	0.1305
236-NLKHQPGGKVIQNKILD-254 *	522.5595	4	2086.2090	2086.2011	3.79	3	0.00	0.0495	0	0.0000	236	0.0001	225	0.0001	3165242	1.4390	3587835	1.5459	3891030	1.4854	1.4901	1.4901	0.0536
255-LSNVQSKGSKVQNKHYVPGGSSVQVKKPVD-285 *	633.3126	2	1264.6106	1264.6081	1.98	3	0.00	0.0495	45	0.0000	32	0.0000	26	0.0000	449303	0.2043	267156	0.1151	315252	0.1203	0.1203	0.1466	0.0500
286-LSKVTSKCGSLGNHHKFGGQVE-309	814.1861	4	3252.7153	3252.6925	7.01	3	0.00	0.0495	69	0.0000	0	0.0000	0	0.0000	17326	0.0079	11847	0.0051	18243	0.0070	0.0070	0.0066	0.0014
314-KLDFKQ-319	609.0735	4	2432.2649	2432.2594	2.26	3	0.00	0.0495	0	0.0000	0	0.0000	0	0.0000	166197	0.0756	137863	0.0594	191459	0.0731	0.0731	0.0693	0.0087
314-KLDFKQVSKGSLD-329	383.3112	2	764.4078	764.4068	1.31	2	0.46	0.0495	103685	0.0487	59406	0.0285	84506	0.0426	205628	0.0935	168548	0.0726	220852	0.0843	0.0835	0.0835	0.1005
330-NIHFVGGGKIKIE-343	617.0122	3	1845.0746	1845.0705	2.33	3	0.00	0.0495	0	0.0000	167	0.0001	246	0.0001	132147	0.0601	83414	0.0359	132147	0.0504	0.0504	0.0468	0.0121
344-THKLTFRF-351	488.6046	3	1462.7920	1462.7892	1.91	2	0.45	0.0495	201336	0.0947	89736	0.0431	102161	0.0631	378667	0.1722	295377	0.1133	360271	0.1375	0.1404	0.1404	0.0306
352-NAKAKTDHGAE-362	344.5261	3	1030.5666	1030.5560	0.56	2	0.41	0.0495	1476967	0.6944	903113	0.4338	1006749	0.5078	3244082	1.4748	3340572	1.4394	2833656	1.0817	1.3320	1.3320	0.2174
352-NKYSFVSGD-373	571.2844	2	1140.5541	1140.5523	1.58	1	0.95	0.5127	265311	0.1247	519723	0.2496	264303	0.1333	340672	0.1549	323226	0.1393	631407	0.2410	0.2410	0.1784	0.0548
374-TSPRHLSNVSTGSDMVDSPQLATLADE-402	562.3217	3	1162.6288	1162.6234	4.64	1	0.84	0.8273	3283018	0.1546	243462	0.1169	319781	0.1613	40239	0.2251	309303	0.1333	404196	0.1543	0.1543	0.1709	0.0481
380-MVDSPQLATLADE-402	1010.1567	3	3027.4484	3027.4455	0.96	3	0.10	0.0495	36109	0.0170	21250	0.0102	2999	0.0015	285162	0.1296	152454	0.0657	237025	0.0905	0.0905	0.0322	
403-VSNLAKQGL-412	695.3335	2	1388.6525	1388.6493	2.30	2	0.45	0.0495	131719	0.0619	88878	0.0427	85375	0.0431	301212	0.1369	182200	0.0785	302453	0.1155	0.1103	0.1103	0.0296
Standard Peptide 777-AGIGDTFSLLEE-88	487.2885	2	972.5625	972.5604	2.16	1	0.84	0.2752	7011526	3.2966	5391134	2.5895	5786184	2.9186	8695054	3.9508	7570642	3.2620	8524168	3.2540	3.4890	3.4890	0.4000
	602.2736	2	1202.5327	1202.5303	2.00	2	0.00	0.0495	2126914	-	2081900	1.5895	1982531	-	2199688	-	2320843	-	2619582	-	-	-	-

Table S2: Changes in GluC protease accessibility for each proteolytic peptide of 1N4R Tau.

Labelled residue (IN3R)	1N3R Peptide Sequence	m/z obs	z	Mr exp	Mth	Δppm	Peptide Class	F. c. F <sub>obs</sub> /F <sub>theo</sub>	Significance (p-value)	Enzyme	1N3R Tau Fibrils Experiment 1			1N3R Tau Fibrils Experiment 2			1N3R Tau Fibrils Experiment 3			1N3R Tau Monomers Experiment 1			1N3R Tau Monomers Experiment 2			1N3R Tau Monomers Experiment 3		
											Intensity (XIC)	Normalized Intensity (XIC)	Standard deviation	Intensity (XIC)	Normalized Intensity (XIC)	Standard deviation	Intensity (XIC)	Normalized Intensity (XIC)	Standard deviation	Intensity (XIC)	Normalized Intensity (XIC)	Standard deviation	Intensity (XIC)	Normalized Intensity (XIC)	Standard deviation	Intensity (XIC)	Normalized Intensity (XIC)	Standard deviation
2-AEPDQE7	K22	476.21794	2	954.42132	354.4229	-1.66	0	3.44	0.0495	GHC	647510	0.3406	377239	0.1468	713346	0.2792	0.2552	0.0966	158033	0.0692	0.0692	78810	0.0383	333581	0.1242	0.0742	0.0446	
13-DHAGTVLGDGRDGGVYMHQDQEGDIDAGLKE-45	K24	948.669	4	3790.6469	3790.6224	6.46	1	0.99	0.8273	GHC	31439	0.0165	12220	0.0047	19130	0.0075	0.0096	0.0062	17482	0.0067	0.0117	28933	0.0117	28933	0.0108	0.0097	0.0027	
83-TSDAGSTPTAE-73	K44	946.669	4	3790.6473	3790.6224	6.57	1	0.78	0.5127	GHC	20839	0.0110	9973	0.0039	13010	0.0051	0.0066	0.0038	16656	0.0063	0.0092	18889	0.0092	27144	0.0101	0.0085	0.0020	
89-AAAGHTVQARIMVSKD-104	K67	667.2988	2	1326.5871	1332.5867	-1.20	0	2.38	0.0495	GHC	28845	0.0152	26312	0.0102	31877	0.0125	0.0126	0.0025	11329	0.0043	0.0056	11504	0.0056	16173	0.0060	0.0053	0.0009	
108-DLKKAKGA-116	K101	482.7442	2	1926.9475	1926.9404	3.68	1	0.69	0.5127	GHC	8068	0.0043	7933	0.0031	9004	0.0030	0.0036	0.0006	13610	0.0052	0.0052	5776	0.0028	20830	0.0078	0.0052	0.0025	
108-DLKKAKGA-116	K103	643.3242	3	1926.9506	1926.9404	5.29	1	0.70	0.5127	GHC	5907	0.0029	7394	0.0029	7654	0.0030	0.0029	0.0001	10801	0.0041	0.0041	2715	0.0013	18864	0.0070	0.0042	0.0029	
108-DLKKAKGA-116	K111	529.7615	2	1057.5118	1057.5226	-13.43	1	1.32	0.5127	AspN	158354	0.0550	113539	0.0428	121109	0.0439	0.0472	0.0067	183300	0.0440	0.0440	40441	0.0013	84060	0.0205	0.0358	0.0133	
108-DLKKAKGA-116	K114	529.7615	2	1057.5131	1057.5226	-8.98	1	1.24	0.8273	AspN	182691	0.0634	121141	0.0456	117139	0.0425	0.0505	0.0113	205762	0.0484	0.0484	46163	0.0059	128449	0.0313	0.0469	0.0144	
117-DSKIKATPRGAPGGKGA-137	K121	759.07301	3	2274.19719	2274.1903	3.03	1	1.22	0.2752	AspN	389120	0.1351	263146	0.0991	314843	0.1142	0.1162	0.0181	431026	0.1036	0.1036	461714	0.1151	252282	0.0615	0.0934	0.0282	
194-PKVAIVRTIPKSPSSAKSRLOIAPVMPD-223	K195 or K196	849.723	4	3394.8628	3394.8691	1.36	2	0.16	0.0495	GHC	19265	0.0096	8745	0.0034	7377	0.0029	0.0053	0.0037	38945	0.0148	0.0148	108305	0.0526	78603	0.0293	0.0322	0.0180	
194-PKVAIVRTIPKSPSSAKSRLOIAPVMPD-223	K205	849.7241	4	3394.86728	3394.8691	2.70	2	0.17	0.0495	GHC	9883	0.0052	6475	0.0025	9012	0.0035	0.0037	0.0014	45765	0.0174	0.0174	62831	0.0305	53900	0.0198	0.0226	0.0070	
194-PKVAIVRTIPKSPSSAKSRLOIAPVMPD-223	K211	849.72623	4	3394.8718	3394.8691	4.04	2	0.17	0.0495	GHC	13255	0.0070	8223	0.0032	26096	0.0102	0.0068	0.0035	34904	0.0362	0.0362	69155	0.0433	95369	0.0570	0.0386	0.0039	
224-LKNVSKGSTE-235	K225	765.4213	2	1528.8281	1528.8283	-0.13	1	0.69	0.0495	GHC	116373	0.0644	239829	0.0923	24974	0.0070	0.0035	0.0095	268242	0.1026	0.1026	238716	0.1149	384596	0.1432	0.1202	0.0208	
224-LKNVSKGSTE-235	K228	765.4215	2	1528.8279	1528.8283	-0.26	2	0.38	0.0495	GHC	122519	0.0644	73608	0.0284	60118	0.0055	0.0088	0.0223	272055	0.1037	0.1037	181231	0.0980	365227	0.1196	0.1018	0.0129	
236-NLKHQPGGKVVYKVPVDSKVTSK	K230	796.5941	6	4773.5212	4773.4959	5.30	2	0.16	0.0495	GHC	60443	0.0423	25360	0.0098	27783	0.0109	0.0210	0.0165	468793	0.1786	0.1786	283889	0.1378	596897	0.2229	0.1796	0.0425	
236-NLKHQPGGKVVYKVPVDSKVTSK	K238	796.5941	6	4773.5212	4773.4959	5.30	2	0.16	0.0495	GHC	4310	0.0023	4727	0.0018	5539	0.0022	0.0021	0.0021	21683	0.0083	0.0083	23983	0.0116	48862	0.0182	0.0127	0.0060	
236-NLKHQPGGKVVYKVPVDSKVTSK	K282	796.5941	6	4773.5208	4773.4959	5.22	3	0.10	0.0495	GHC	2197	0.0012	1512	0.0006	0	0.0000	0.0006	0.0006	20963	0.0078	0.0078	4768	0.0023	18093	0.0087	0.0096	0.0029	
236-NLKHQPGGKVVYKVPVDSKVTSK	K288	796.5960	6	4773.5325	4773.4959	7.67	3	0.08	0.0495	GHC	4310	0.0023	4727	0.0018	0	0.0000	0.0014	0.0014	52415	0.0200	0.0200	27416	0.0133	55404	0.0206	0.0180	0.0040	
236-NLKHQPGGKVVYKVPVDSKVTSK	K292	796.596	6	4773.5265	4773.4959	6.41	3	0.07	0.0495	GHC	1464	0.0008	1189	0.0005	0	0.0000	0.0004	0.0004	18165	0.0069	0.0069	6179	0.0030	23627	0.0088	0.0062	0.0030	
283-KLDFKRVYKSGSLDNIHPPGGNKKE-312	K314	704.7849	5	3518.8883	3518.8667	6.14	3	0.08	0.0495	GHC	5240	0.0038	1849	0.0007	4577	0.0018	0.0018	0.0018	43334	0.0165	0.0165	15778	0.0077	108025	0.0402	0.0215	0.0168	
283-KLDFKRVYKSGSLDNIHPPGGNKKE-312	K320	704.785	5	3518.8886	3518.8667	6.20	2	0.14	0.0495	GHC	1075	0.0006	951	0.0004	4779	0.0019	0.0009	0.0008	17087	0.0085	0.0085	7296	0.0035	27456	0.0102	0.0088	0.0033	
283-KLDFKRVYKSGSLDNIHPPGGNKKE-312	K340 or K341	704.7847	5	3518.8847	3518.8667	5.12	3	0.02	0.0495	GHC	756	0.0004	492	0.0002	4422	0.0017	0.0008	0.0008	98193	0.0374	0.0374	13607	0.0066	141192	0.0526	0.0322	0.0234	
313-TKLTRE-320	K346	628.3261	2	1256.6377	1256.6336	3.26	0.31	0.0495	GHC	117752	0.0588	113404	0.0438	137217	0.0537	0.0521	0.0076	350751	0.1337	0.1337	307593	0.1483	603422	0.2246	0.1692	0.0486		
321-NAKATDHAE-331	K354	684.3219	2	1366.6293	1366.6298	-0.44	2	0.30	0.0495	GHC	23573	0.0124	102485	0.0386	100276	0.0392	0.0304	0.0156	208663	0.0799	0.0799	246149	0.1195	293559	0.1083	0.1029	0.0206	
321-NAKATDHAE-331	K356	684.3224	2	1366.6303	1366.6298	0.29	2	0.39	0.0495	GHC	26153	0.0138	209198	0.0808	319847	0.0861	0.0733	0.0661	368887	0.1409	0.1409	278388	0.1351	785084	0.2923	0.1884	0.0881	
321-NKSPVSGD-342	K366	695.5567	2	1388.7029	1388.701	1.37	2	0.37	0.0495	GHC	1034	0.0095	3905	0.0015	3553	0.0014	0.0011	0.0005	6287	0.0024	0.0024	5104	0.0025	11645	0.0043	0.0031	0.0011	
372-VSASLAKGL-381	K409	600.3276	2	1196.6366	1196.638	0.50	1	0.71	0.2752	GHC	186163	0.0979	280191	0.1063	327850	0.1283	0.1115	0.0165	270287	0.1030	0.1030	310019	0.1505	576500	0.2146	0.1560	0.0560	
372-VSASLAKGL-381	K406	600.3270	2	1196.6394	1196.638	1.17	1	1.11	0.5127	GHC	10198	0.0054	23963	0.0092	26359	0.0103	0.0083	0.0026	20995	0.0078	0.0078	12771	0.0062	22739	0.0085	0.0075	0.0012	
Standard Peptide 77-AGDTPSLEDE-88		455.2372	2	1202.5334	1202.5303	-1.41				GHC	1901260	-	2587387	-	2654745	-	-	-	2624229	-	-	2060091	-	2686193	-	-	-	
Standard Peptide 361-DSPQLATLA-869		455.2372	2	914.4559	914.4709	-12.03				AspN	2880216	-	2654745	-	2756178	-	-	-	4161989	-	-	4022273	-	4026077	-	-	-	

Table S3: Changes in lysine biotinylation for each proteolytic GluC or AspN peptide of 1N3R Tau.

1N4R Peptide Sequence	Labelled residue	m/z obs	z	Mr exp	Mth	Δppm	Peptide Class	F.C. FibMon	Significant (p-value)	Enzyme	1N4R Tau Fibrils Experiment 1			1N4R Tau Fibrils Experiment 2			1N4R Tau Fibrils Experiment 3			1N4R Tau Monomers Experiment 1			1N4R Tau Monomers Experiment 2			1N4R Tau Monomers Experiment 3			Averaged Normalized Intensity	Standard deviation
											Intensity (XIC)	Normalized Intensity	Standard deviation	Intensity (XIC)	Normalized Intensity	Standard deviation	Intensity (XIC)	Normalized Intensity	Standard deviation	Intensity (XIC)	Normalized Intensity	Standard deviation	Intensity (XIC)	Normalized Intensity	Standard deviation	Intensity (XIC)	Normalized Intensity	Standard deviation		
2- <b>A</b> ERREZ	A2	478.2181	2	954.4216	954.4239	-1.40	0	2.63	0.0495	GluC	316787	0.1403	89879	0.0493	187536	0.0795	0.0397	102800	0.0405	99296	0.0410	47018	0.0286	0.0340	0.0116					
13-DHAGTVELGDRK	K24	946.6635	4	3790.6250	3790.6224	0.69	1	1.24	0.1527	GluC	38752	0.0712	9161	0.0050	34431	0.0146	0.0123	17309	0.0071	30905	0.0128	22420	0.0098	0.0099	0.0028					
13-DHAGTVELGDRRGGSTTHMQDEGDTDAGLK	K44	946.6626	4	3790.6248	3790.6224	-1.43	1	0.88	0.1527	GluC	25855	0.0114	8058	0.0044	24056	0.0050	0.0050	15997	0.0059	30905	0.0128	25018	0.0110	0.0099	0.0035					
63-TSDAKTPTFAE-73	K67	667.2997	2	1332.5913	1332.5867	-1.43	0	2.24	0.0495	GluC	202376	0.0896	46587	0.0447	100766	0.0427	0.0530	43746	0.0172	71298	0.0284	54590	0.0243	0.0237	0.0061					
89-AAHGHTDARRMoxVSKSD-104	K101	482.7432	4	1926.9438	1926.9404	1.66	1	1.14	0.8273	GluC	19116	0.0095	7815	0.0043	12128	0.0051	0.0060	13332	0.0053	13180	0.0054	11314	0.0060	0.0052	0.0002					
89-AAHGHTDARRMoxVSKSD-104	K103	482.7432	4	1926.9421	1926.9404	0.88	1	1.35	0.5127	GluC	19435	0.0092	3204	0.0018	8267	0.0035	0.0045	7854	0.0031	12786	0.0053	3641	0.0016	0.0033	0.0019					
108-DDKAKGA-116	K111	529.1681	2	1057.5216	1057.5228	-0.95	1	1.43	0.8273	AspN	271009	0.0801	155499	0.0378	161609	0.0365	0.0521	0.0242	187488	0.0400	255288	0.0527	77502	0.0167	0.0385	0.0183				
108-DDKAKGA-116	K112	529.1682	2	1057.5218	1057.5228	-0.76	1	1.44	0.2752	AspN	274845	0.0812	147413	0.0358	234412	0.0588	0.0576	0.0228	157143	0.0336	289455	0.0598	125252	0.0270	0.0401	0.0174				
108-DDKAKGA-116	K114	529.1676	2	1057.5206	1057.5228	-1.89	1	1.30	0.2752	AspN	588836	0.1651	371107	0.0901	506682	0.1205	0.1253	381690	0.0815	718412	0.1484	270814	0.0584	0.0961	0.0468					
117-DGKATPRGAPQKGOA-137	K121	759.0753	3	2274.2039	2274.1903	5.98	1	1.67	0.2752	AspN	33963	0.0100	18922	0.0046	31559	0.0074	0.0074	23317	0.0050	24070	0.0050	15239	0.0033	0.0044	0.0010					
184-PKVAVTRPPSPSSAKSRLQTPVMPD-223	K196 or K198	849.7176	4	3394.8411	3394.8561	-5.01	2	0.20	0.0495	GluC	30171	0.0134	18756	0.0103	33136	0.0140	0.0126	88510	0.0349	160239	0.0744	186592	0.0637	0.0637	0.0262					
184-PKVAVTRPPSPSSAKSRLQTPVMPD-223	K205	849.7182	4	3394.8478	3394.8561	-3.03	2	0.18	0.0495	GluC	17987	0.0080	7483	0.0041	22040	0.0093	0.0071	71471	0.0282	111170	0.0459	108349	0.0479	0.0407	0.0109					
194-PKVAVTRPPSPSSAKSRLQTPVMPD-223	K211	849.7177	4	3394.8418	3394.8561	-4.80	2	0.28	0.0495	GluC	33940	0.0150	11324	0.0062	30720	0.0130	0.0114	10046	0.0046	71471	0.0282	112609	0.0465	0.0465	0.0104					
224-LANVKSIGSTE-235	K225	765.4166	2	1528.8186	1528.8283	-6.34	1	0.64	0.1266	GluC	258920	0.1147	140838	0.0773	186894	0.0715	0.0978	0.0234	231753	0.0913	329029	0.1358	413904	0.1814	0.1362	0.0451				
224-LANVKSIGSTE-235	K228	765.4166	2	1528.8175	1528.8283	-7.06	2	0.44	0.1266	GluC	322736	0.1429	151638	0.0832	184274	0.1014	0.1014	0.0360	339183	0.1333	714513	0.2949	615639	0.2689	0.2327	0.0870				
224-LANVKSIGSTE-235	K230	765.4166	2	1528.8186	1528.8283	-6.34	1	0.69	0.2752	GluC	385565	0.1707	163330	0.0886	310432	0.1315	0.1306	0.0406	312412	0.1231	528829	0.2183	517958	0.2271	0.1895	0.0576				
236-NKHQPGGGVQIKLID-264*	K238	771.7640	3	2312.2702	2312.2767	-3.72	1	0.94	0.8273	GluC	142651	0.0632	56075	0.0308	101943	0.0432	0.0432	50530	0.0199	123171	0.0508	171550	0.0752	0.0487	0.0277					
236-NKHQPGGGVQIKLID-264*	K245	771.7639	3	2312.2692	2312.2767	-3.89	3	0.00	0.0495	GluC	0	0.0000	0	0.0000	0	0.0000	0.0000	59026	0.0233	107375	0.0443	98724	0.0453	0.0370	0.0119					
236-NKHQPGGGVQIKLID-264*	K251	771.7660	3	2312.2762	2312.2767	-1.08	2	0.46	0.1266	GluC	65307	0.0289	45452	0.0249	45859	0.0195	0.0244	0.0047	65865	0.0260	142392	0.0588	170614	0.0748	0.0532	0.0249				
236-NKHQPGGGVQIKLID-264*	K252	771.7645	3	2312.2715	2312.2767	-3.11	2	0.47	0.2752	GluC	19447	0.0086	48613	0.0268	42068	0.0178	0.0177	0.0091	42711	0.0168	109213	0.0461	117959	0.0517	0.0379	0.0185				
254-DLSNVQSGSK-365*	K261	774.8646	2	1547.7147	1547.7072	4.86	2	0.15	0.0495	AspN	107487	0.0318	38209	0.0093	45216	0.0108	0.0173	0.0226	534857	0.1142	953897	0.1165	403717	0.0870	0.1059	0.0164				
266-DNKHVPGGGVQIKLID-284*	K265	745.0710	3	2332.1912	2332.1725	8.38	2	0.36	0.1266	AspN	26443	0.0203	29054	0.0070	54119	0.0129	0.0134	0.0066	380070	0.0812	570314	0.1178	292898	0.0652	0.0874	0.0279				
266-DNKHVPGGGVQIKLID-284*	K269	745.0708	3	2332.1906	2332.1725	8.11	2	0.52	0.2752	AspN	88194	0.0261	22213	0.0054	26374	0.0063	0.0126	0.0117	131341	0.0281	187933	0.0388	26979	0.0068	0.0242	0.0168				
314-KLDFKRVQSKGSLNTHVPGGKKE-343	K314	704.7790	5	3518.8584	3518.8667	-2.36	1	0.88	0.5127	GluC	63692	0.0282	34985	0.0079	65808	0.0251	0.0051	0.0051	47554	0.0187	78764	0.0325	77200	0.0338	0.0284	0.0084				
314-KLDFKRVQSKGSLNTHVPGGKKE-343	K324	704.7791	5	3518.8590	3518.8667	-2.19	1	0.90	0.5127	GluC	102676	0.0455	16249	0.0089	88384	0.0375	0.0306	0.0192	21217	0.0084	115549	0.0477	104205	0.0457	0.0339	0.0222				
314-KLDFKRVQSKGSLNTHVPGGKKE-343	K340 or K341	704.7791	5	3518.8591	3518.8667	-2.16	1	0.99	0.8273	GluC	146834	0.0650	15365	0.0074	155904	0.0657	0.0461	0.0334	42994	0.0169	168022	0.0694	120946	0.0530	0.0464	0.0286				
344-THKLTRE-351	K346	629.3206	2	1256.6266	1256.6336	-5.57	1	0.88	0.8273	GluC	313895	0.1390	175064	0.0961	400648	0.1698	0.1548	0.0370	252886	0.0997	558898	0.2307	292347	0.1282	0.1528	0.0689				
352-NAKAKTDHGAE-382	K354	684.3193	2	1366.6241	1366.6299	-4.24	1	0.67	0.2752	GluC	280475	0.1242	78672	0.0432	239510	0.1015	0.0986	0.0418	238776	0.0941	405536	0.1674	318132	0.1395	0.1337	0.0370				
363-VYKSPVSGD-373	K366	695.3532	2	1386.6239	1386.6299	-4.39	1	0.82	0.5127	GluC	479580	0.2124	150719	0.0827	386716	0.1639	0.1530	0.0655	338438	0.1338	866655	0.3980	470133	0.2061	0.2463	0.1371				
403-VSALAKQL-412	K409	600.3250	2	1198.6354	1198.6390	-2.17	1	1.25	0.8273	GluC	98154	0.0062	7526	0.0041	9066	0.0038	0.0047	0.0013	7782	0.0031	13891	0.0057	12440	0.0055	0.0048	0.0015				
Standard Peptide 77-AGIGDTPSLEDE-88		602.2699	2	1202.5252	1202.5303	-4.24	1	1.25	0.8273	GluC	2280803	0.4380	401037	0.2200	716118	0.3035	0.3205	0.1100	343168	0.1360	2422886	0.4677	2281125	0.3133	0.2562	0.1041				
Standard Peptide 892-DSPQALTA-400		458.245	2	914.475	914.471	4.16	1	0.99	0.8273	AspN	3384603	0.4183	4202084	0.4183	4682051	0.4183	0.4183	0.4183	4682051	0.4183	4682051	0.4183	4682051	0.4183	4682051	0.4183	4682051			

Table S4: Changes in lysine biotinylation for each proteolytic GluC or AspN peptide of 1N4R Tau.





## IV. Conclusion et perspectives

### I- Conclusion générale

Au cours de mes travaux de thèse, j'ai pu optimiser et mettre en place un ensemble de stratégies d'analyse permettant de documenter les surfaces accessibles et protégées d'assemblages polymériques insolubles. Je me suis focalisée sur les fibres de Tau 1N3R et 1N4R pour effectuer ces mises au point qui m'ont permis dans un second temps de d'obtenir des informations structurales sur 75% de la séquence de Tau pour des fibres constituées d'une part de Tau 1N3R, et d'autre part de Tau 1N4R.

Pour ce faire, étant donné que la nature de nos objets d'étude contraint grandement les techniques d'étude disponibles, j'ai mis en place deux stratégies complémentaires et les méthodes associées (traitement chimique/biochimique, préparation d'échantillon, spectrométrie de masse, analyses de données) basées sur la reconnaissance des surfaces des fibres, soit par des protéases, soit par un réactif de petite taille et l'identification et la quantification des produits générés, dans le but de cartographier les surfaces des fibres.

#### I.1- Apport des méthodes mises en place

D'une part, j'ai pu obtenir un certain nombre de profils de protéolyses ménagées suivies par SDS-PAGE dont deux d'entre elles (i.e GluC et Chymotrypsine) nous ont permis de disposer d'un code barre moléculaire des fibres de Tau 1N3R et 1N4R. Alors que nous avons pu définir les bornes régionales de ces fragments protéolytiques caractéristiques par western-blot, l'identification des sites de coupure spécifiques par différentes méthodes associées à la spectrométrie de masse a mis en évidence certaines limitations de ces techniques. En effet, une analyse du recouvrement de séquence par une approche Bottom-up, bien que classique en protéomique, s'est révélée insuffisante dans le cadre de notre étude. En effet, malgré la séparation sur gel SDS-PAGE, la diffusion des bandes et l'excellente capacité d'ionisation de certains peptides contaminants conduisaient à de potentielles erreurs d'attribution de séquences. L'optimisation d'une approche de N-terminomique par marquage par le TMPP du résidu N-terminal a permis d'identifier l'extrémité N-terminale de plusieurs fragments protéolytiques. Toutefois, cette dérivatisation à pH 8,2, décrite comme spécifique de l'amine terminale (Bertaccini *et al.* 2013, Ayoub *et al.* 2012) s'est révélée, malgré toutes les précautions entreprises, également marquer de manière substantielle des chaînes latérales de certains acides aminés (e.g sérines, lysines, thréonines, tyrosines) ce qui a considérablement compliqué l'analyse. De plus, en accord avec des observations antérieures (ABRF. 2014), l'efficacité de la dérivatisation du N-terminal par le TMPP semble être

séquence-dépendante, à l'origine d'identifications parfois infructueuses. De façon analogue, le résidu C-terminal des fragments protéolytiques pourrait être identifié par une dérivation spécifique du groupement carboxylique C-terminal par la chimie oxazolone couplée à une approche Bottom-up (Kim *et al.* 2011). Toutefois, la communauté s'accorde à dire que ce marquage est difficile à mettre en place et aurait certainement nécessité un long développement (Nakazawa *et al.* 2008).

L'analyse de la masse entière de ces fragments insolubles de haut poids moléculaire est difficile. En effet, la nature fibrillaire, agrégée, des fragments protéolytiques constituait une contrainte majeure dans le cadre de notre étude, pour l'application de la méthodologie Top-down. Cette dernière repose sur la séparation chromatographique de fragments protéolytiques/protéines entier(ère)s, leur analyse MS et MS/MS, sans digestion préalable (Donnelly *et al.* 2019) et est très appréciée dans le cadre de la biologie structurale et de la dégradomique (Tholey and Becker. 2017). Elle nécessite toutefois la solubilisation des assemblages de Tau, à l'aide de détergents compatibles avec la spectrométrie de masse couplée à une chromatographie liquide. A la fin de ma thèse, j'ai essayé plusieurs protocoles de solubilisation basés notamment sur l'optimisation dans notre équipe de protocoles efficaces de décontamination de surfaces contaminées par les agrégats protéiques de Tau, d'A $\beta$  et de synucléine (Fenyi *et al.* 2018) Ainsi, j'ai pu optimiser un protocole (données non montrées) qui pourra ultérieurement être utilisé au sein du laboratoire.

En conditions natives, une protéolyse avancée des formes monomérique et fibrillaire de Tau 1N3R et 1N4R, générant des peptides de petite taille qui ont pu être identifiés et quantifiés, nous a permis d'observer des recouvrements de séquence différentiels, dépendants de la conformation. Une observation similaire avait déjà été rapportée (Schopper *et al.* 2017) En effet, les populations peptidiques issues des fibres diffèrent nettement de celles issues des monomères. Nous avons donc quantifié relativement les différents peptides générés (i.e intensité du pic MS) que nous avons normalisé par rapport à la quantité totale de protéines (i.e intensité du pic MS d'un peptide de référence), à l'image de la normalisation utilisée précédemment (Barthelemy *et al.* 2016). Une quantification relative similaire (intégration du pic MS sans marquage préalable dans un mélange peptidique) avait déjà été utilisée dans la littérature (Resing and Ahn. 2005) et montrait une corrélation directe entre l'intensité du pic MS et la concentration protéique et l'absence de nécessité de marquage isotopique pour la normalisation (Bondarenko *et al.* 2002, Wang *et al.* 2003). D'autre part, nous avons développé une approche de marquage de surface permettant de décrire une accessibilité à l'échelle des résidus lysines. Pour ce faire, nous avons choisi d'opter pour un marquage covalent par la NHS-Biotine en postulant qu'un marquage différentiel résulterait d'une différence d'accessibilité des lysines entre les formes monomériques et fibrillaires, à l'image des travaux

de Dreger *et al.* 2009. Nous avons donc optimisé ce marquage de surface à travers des gammes de ratios molaires (Lysines/NHS-Biotine) et des cinétiques de marquage. Cette mise au point nous a permis de définir un temps de marquage et un ratio [Lysines/NHS-Biotine] pour lesquels le nombre de biotines ajoutées par molécule de Tau n'était pas susceptible d'altérer la structure des fibres. Suite à une digestion en conditions dénaturantes, nous avons pu quantifier, de façon relative, les peptides biotinylés et les comparer entre les formes fibrillaires et les formes monomériques. Cette approche a complété et validé nos précédentes observations. Enfin, une combinaison de ces deux approches (i.e protéolyse avancée et marquage de surface) a permis de tirer des conclusions similaires quant à l'accessibilité et la protection de régions au sein des fibres et a également confirmé que le marquage à la biotine ne perturbait pas les propriétés structurales des assemblages. Ainsi, le marquage à la biotine combiné la protéolyse avancée constituernt une stratégie de choix pour l'étude d'assemblages insolubles de haut poids moléculaire impliqués dans d'autres protéinopathies.

## I.2. Apport des résultats de cartographie de surface obtenues sur les fibres de Tau 1N3R et 1N4R : Principales surfaces accessibles ou protégées identifiées et différences structurales entre les fibres 1N3R et 1N4R

Par protéolyse ménagée des fibres par la GluC et par la Chymotrypsine, nous avons dressé des code-barres moléculaires distincts entre les fibres de Tau 1N3R et de Tau 1N4R, conformationnel-spécifiques. Ces fragments générés restent fibrillaires, même après 300 minutes de protéolyse ménagée, traduisant une véritable propriété de résistance à la protéolyse, caractéristique essentielle des espèces amyloïdes. L'identification de ces espèces montre (1) une accessibilité de la région N-terminale de Tau similaire pour les fibres 1N3R et 1N4R, (2) une différence d'accessibilité de la région C-terminale de Tau pour les fibres de Tau 1N3R et 1N4R, en aval de la seconde répétition (R2).

Plus particulièrement, nous avons en effet pu mettre en évidence que les fibres 1N3R et les fibres 1N4R partageaient une accessibilité de la région N-terminale, notamment au niveau des acides aminés F8, A74, A77, D87 et A89 dans les premiers temps de protéolyse par la GluC. De plus, la tyrosine 168 semble être un site préférentiel de protéolyse et serait donc hautement accessible au sein des fibres 1N3R et 1N4R.

En C-terminal, nous avons pu montrer que la région 340-365 (numérotation 1N4R) constitue une région accessible aux enzymes uniquement au sein des fibres 1N4R, plus spécifiquement

au niveau des acides aminés E343, F349 et Y365. Cette même région C-terminale au sein des fibres 1N3R est quant à elle hautement protégée.

Par protéolyse avancée, nous avons mis en évidence par protéolyse avancée, (1) une structuration partielle de certaines séquences (13-36 et 89-104) dans une région N-terminale globalement accessible pour les deux isoformes assemblées (davantage marquée pour Tau 1N3R) et (2) une accessibilité différentielle de la région C-terminale, principalement dans la région bornée par les acides aminés N330- D373, puis du peptide C-terminal V403-L412. Ces informations concordent avec les données de protéolyses ménagées et n'avaient jamais été décrites auparavant.

Par marquage de surface, nous avons pu observer que l'extrémité N-terminale et les lysines situées dans la région N-terminale étaient accessibles au solvant, voire pour certaines nettement exposées. A l'inverse, le marquage de surface indique une diminution de l'accessibilité au solvant des acides aminés composant les régions les moins accessibles à l'enzyme soit une séquence de 20 acides aminés en amont du MTBR, les répétitions R1, R3, R4 et une séquence de 30 acides aminés en aval du C-terminal pour les fibres 1N3R. Cette diminution de l'accessibilité des lysines au solvant est moins marquée pour les fibres 1N4R et ne semble concerner que la séquence de 20 acides aminés en amont du MTBR, les répétitions R1 et, pour 1N4R, surtout la répétition 4R-spécifique R2. Ainsi, là encore, les lysines en C-terminal de la protéine semblent davantage accessibles après assemblage pour Tau 1N4R que pour Tau 1N3R. Toutes ces données appuient que la différence conformationnelle isoforme-spécifique majeure entre les fibres de Tau 1N3R et 1N4R est l'accessibilité de la partie C-terminale pour 1N4R, et réciproquement, sa protection au sein des fibres 1N3R.

L'exposition de l'alanine N-terminale au sein des assemblages concorde avec son acétylation au sein des agrégats issus de patients atteints de la maladie d'Alzheimer (Hasegawa *et al.* 1992). Une structuration partielle de la région N-terminale, justifiée ici par les expériences de protéolyse, en accord avec les prédictions de structures secondaires (Gamblin *et al.* 2005), explique que ces lysines puissent être singulièrement exposées au sein des fibres, ce qui n'avait jusqu'à présent jamais été montré. Notre étude, diffère de celles réalisées auparavant en ce sens qu'elle permet d'une part de mettre en évidence des régions accessibles mais également des régions protégées, et résulte en la construction d'une cartographie complète des fibres. Ainsi, le cœur hautement protégé des fibres 1N3R et 1N4R assemblées *in vitro* à partir de protéines recombinantes que nous avons pu décrire dans cette étude semble plus étendu que ce qui a été décrit par Zhang et ses collaborateurs (Zhang *et al.* 2019). Cette divergence d'observation peut s'expliquer d'une part par les propriétés intrinsèques des

techniques utilisées pour l'étude structurale. En effet, la cryo-microscopie électronique ne peut décrire que des régions extrêmement rigides, ultra-structurées, au sein de la macromolécule. Ces différences peuvent également résulter des différentes conditions d'assemblage utilisées (PBS 1X et DTT 1mM dans le cadre de notre étude ; 30 mM MOPS, pH 7.2, 1mM AEBSF et 4mM TCEP dans le cadre de l'étude menée par Zhang et collaborateurs publiée en 2019). Les régions identifiées dans notre étude comme hautement protégées, i.e constituant le cœur des fibres, sont toutefois d'avantage comparables aux cœurs protégés mis en évidence par les expériences de protéolyse ménagée réalisées sur les agrégats issus de cerveaux de patients (Jakes *et al.* 1991 ; Novak *et al.* 1993 ; Taniguchi-Watanabe *et al.* 2016).

## II- Perspectives

Cette thèse a permis de cartographier les surfaces accessibles et les régions protégées de fibres de Tau entière par deux méthodes complémentaires qui associent la spectrométrie de masse. Deux perspectives intéressantes à ce projet concernent l'exploitation des surfaces accessibles de Tau identifiées, qui peuvent servir de cibles pour concevoir de nouveaux outils moléculaires capables d'interagir avec cette surface pour soit détecter et marquer cette surface à des fins diagnostiques, soit inhiber cette surface pour empêcher l'interaction avec certains partenaires.

Ainsi, une étude similaire pourrait être réalisée sur des agrégats issus d'échantillons post mortem de patients atteints de tauopathies différentes (e.g composition en isoformes des agrégats) afin d'y rechercher les signatures C-terminales (région 343-373) révélées par les protéolyse ménagée et avancée et le marquage de surface ; et plus exactement les peptides spécifiques d'une conformation, encore appelés « peptides conformotypiques », terme proposé récemment (Feng *et al.* 2014 ; Boersema *et al.* 2018). Une autre approche pourrait consister en la production d'anticorps monoclonaux ciblant des peptides composant la région C-terminale divergente et d'en tester la réactivité sur des échantillons post mortem provenant de patients. Les surfaces accessibles mises en évidence pourraient constituer des régions cibles d'intérêt pour le développement de traceurs TEP par exemple (Leuzy *et al.* 2019). Il pourrait être intéressant de déterminer si ces signatures d'accessibilité peuvent être reliées à des événements de troncation ou de phosphorylation propres à chacune des types de fibres chez les patients, comme les connaissances actuelles semblent le suggérer (Mondragón-Rodríguez *et al.* 2008 ; Probst *et al.* 1996).

Enfin, ces informations d'accessibilité différentielle suggèrent que les partenaires d'interaction des fibres 1N3R et 1N4R, participant au gain de fonction toxique de ces agrégats, pourraient être profondément différents. Une identification différentielle de ces interactants permettrait donc de mieux comprendre le processus pathologique sous-jacent aux différentes tauopathies (Lafferrière *et al.* 2019). Les régions protégées pourraient être la cible de petites molécules ou de peptides visant à freiner le processus d'agrégation et celles accessibles la cible de molécules empêchant l'interaction avec les partenaires dans le but d'interférer dans le processus pathologique.

## Références

1. Site EM Consulte. Available at: <https://www.em-consulte.com/>.
2. Abraha, A. *et al.* C-terminal inhibition of tau assembly in vitro and in Alzheimer's disease. *J. Cell Sci.* **113**, 3737–3745 (2000).
3. ABRF. Protein Sequencing Research Group ABRF 2014 annual meeting. (2014).
4. Adams, S. J., de Ture, M. A., McBride, M., Dickson, D. W. & Petrucelli, L. Three repeat isoforms of tau inhibit assembly of four repeat tau filaments. *PLoS One* **5**, (2010).
5. ADI. World Alzheimer Report. (2015).
6. Agarwal, S. & Gilbert, R. *Progressive Supranuclear Palsy. StatPearls* (2019).
7. Aguzzi, A. Beyond the prion principle. *Nature* **459**, 924–925 (2009).
8. Ahmed, M. *et al.* Structural conversion of neurotoxic amyloid-B 1-42 oligomers to fibrils. *Nat. Struct. Mol. Biol.* **17**, 561–567 (2010).
9. Al-Hilaly, Y. K. *et al.* Alzheimer's Disease-like Paired Helical Filament Assembly from Truncated Tau Protein Is Independent of Disulfide Crosslinking. *J. Mol. Biol.* **429**, 3650–3665 (2017).
10. Allen, B. *et al.* Abundant tau filaments and nonapoptotic neurodegeneration in transgenic mice expressing human P301S tau protein. *J. Neurosci.* **22**, 9340–51 (2002).
11. Alzforum. Tau Antibodies. Available at: [https://www.alzforum.org/antibodies/search?category\[616\]=Tau](https://www.alzforum.org/antibodies/search?category[616]=Tau). (Accessed: 16th September 2019)
12. Andreadis, A. Tau gene alternative splicing: Expression patterns, regulation and modulation of function in normal brain and neurodegenerative diseases. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1739**, 91–103 (2005).
13. Andronesi, O. C. *et al.* Characterization of Alzheimer's-like paired helical filaments from the core domain of tau protein using solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **130**, 5922–5928 (2008).
14. Ankri, J. Prévalence, incidence et facteurs de risque de la maladie d'Alzheimer. *Gérontologie et société* **32 / n° 12**, 129 (2009).
15. Aoyagi, H., Hasegawa, M. & Tamaoka, A. Fibrillogenic nuclei composed of P301L mutant tau induce elongation of P301L tau but not wild-type tau. *J. Biol. Chem.* **282**, 20309–20318 (2007).
16. Arai T, A. *et al.* Distinct isoforms of tau aggregated in neurons and glial cells in brains of patients with Pick's disease, corticobasal degeneration and progressive supranuclear palsy. *Acta Neuropathol.* **101**, 167–173 (2001).
17. Armengaud, J., Dedieu, A., Solques, O., Pellequer, J. L. & Quemeneur, E. Deciphering structure and topology of conserved COG2042 orphan proteins. *BMC Struct. Biol.* **5**, (2005).
18. Armstrong, M. J. *et al.* Criteria for the diagnosis of corticobasal degeneration. *Neurology* **80**, 496–503 (2013).
19. Armstrong, R. A. Visual signs and symptoms of corticobasal degeneration. *Clin. Exp. Optom.* **99**, 498–506 (2016).
20. Asai, H. *et al.* Depletion of microglia and inhibition of exosome synthesis halt tau propagation. *Nat. Neurosci.* **18**, 1584–1593 (2015).
21. Audouard, E. *et al.* High-Molecular-Weight Paired Helical Filaments from Alzheimer Brain Induces Seeding of Wild-Type Mouse Tau into an Argyrophilic 4R Tau Pathology in Vivo. *Am. J. Pathol.* **186**, 2709–2722 (2016).
22. Avila, J., Lucas, J. J., Pérez, M. & Hernández, F. Role of Tau Protein in Both Physiological and Pathological Conditions. *Physiol. Rev.* **84**, 361–384 (2004).
23. Ayoub, D. Vers une étude approfondie des protéomes: caractérisation des extrémités N-terminales des protéines. (2013).
24. Ayoub, D. *et al.* Characterization of the N-Terminal Heterogeneities of Monoclonal Antibodies Using In-Gel Charge Derivatization of  $\alpha$ -Amines and LC-MS/MS. *Anal. Chem.* **87**, 3784–3790 (2015).
25. Azim-Zadeh, O. *et al.* Use of biotin derivatives to probe conformational changes in proteins. *J. Biol. Chem.* **282**, 21609–21617 (2007).
26. Baker, M. Cryo-electron microscopy shapes up. *Nature* **561**, 565–567 (2018).
27. Bancher, C. *et al.* Accumulation of abnormally phosphorylated  $\tau$  precedes the formation of neurofibrillary tangles in Alzheimer's disease. *Brain Res.* **477**, 90–99 (1989).
28. Bantscheff, M., Lemeer, S., Savitski, M. M. & Kuster, B. Quantitative mass spectrometry in proteomics: Critical review update from 2007 to the present. *Anal. Bioanal. Chem.* **404**, 939–965 (2012).
29. Barbier, P. *et al.* Role of Tau as a Microtubule-Associated Protein: Structural and Functional Aspects. *Front. Aging Neurosci.* **11**, 1–14 (2019).
30. Barghorn, S. & Mandelkow, E. Toward a unified scheme for the aggregation of tau into Alzheimer paired helical filaments. *Biochemistry* **41**, 14885–14896 (2002).

31. Barghorn, S., Biernat, J. & Mandelkow, E. Purification of Recombinant Tau Protein and Preparation of Alzheimer-Paired Helical Filaments In Vitro. *Methods Mol. Biol.* 035–052 (2005). doi:10.1385/1-59259-874-9:035
32. Barrachina, M. & Ferrer, I. DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain. *J. Neuropathol. Exp. Neurol.* **68**, 880–891 (2009).
33. Barthélemy, N. R. *et al.* Differential Mass Spectrometry Profiles of Tau Protein in the Cerebrospinal Fluid of Patients with Alzheimer's Disease, Progressive Supranuclear Palsy, and Dementia with Lewy Bodies. *J. Alzheimers. Dis.* **51**, 1033–1043 (2016).
34. Baud, A. *et al.* SOLEIL shining on the solution-state structure of biomacromolecules by synchrotron X-ray footprinting at the Metrology beamline. *J. Synchrotron Radiat.* **24**, 576–585 (2017).
35. Bendor, J. T., Logan, T. P. & Edwards, R. H. The function of  $\alpha$ -synuclein. *Neuron* **79**, 1044–1066 (2013).
36. Bertaccini, D. *et al.* An improved stable isotope N-terminal labeling approach with light/heavy TMPP to automate proteogenomics data validation: DN-TOP. *J. Proteome Res.* **12**, 3063–3070 (2013).
37. Bertini, I., Gonnelli, L., Luchinat, C., Mao, J. & Nesi, A. A new structural model of A $\beta$  40 fibrils. *J. Am. Chem. Soc.* **133**, 16013–16022 (2011).
38. Biernat, J., Gustke, N., Drewes, G., Mandelkow, E. & Mandelkow, E. Phosphorylation of Ser262 strongly reduces binding of tau to microtubules: Distinction between PHF-like immunoreactivity and microtubule binding. *Neuron* **11**, 153–163 (1993).
39. Boersema, P. J. *et al.* Biology/Disease-Driven Initiative on Protein-Aggregation Diseases of the Human Proteome Project: Goals and Progress to Date. *J. Proteome Res.* **17**, 4072–4084 (2018).
40. Bondarenko, P. V., Chelius, D. & Shaler, T. A. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography - Tandem mass spectrometry. *Anal. Chem.* **74**, 4741–4749 (2002).
41. Borroni, B. *et al.* Tau forms in CSF as a reliable biomarker for progressive supranuclear palsy. *Neurology* (2008). doi:10.1212/01.wnl.0000335941.68602.39
42. Boulicault, J. Mécanismes d'ionisation MALDI en mode négatif à travers l'étude de polyoxométallates HAL Id : tel-01307139. (2015).
43. Bousset, L. *et al.* Structural and functional characterization of two alpha-synuclein strains. *Nat. Commun.* **4**, 2575 (2013).
44. Bousset, L. *et al.* Structural Characterization of the Fibrillar Form of the Yeast *Saccharomyces cerevisiae* Prion Ure2p. *Biochemistry* **43**, 5022–5032 (2004).
45. Bowen, D. M., Smith, C. B., White, P. & Davison, A. N. Neurotransmitter-related enzymes and indices of hypoxia in senile dementia and other abiotrophies. *Brain* **99**, 459–496 (1976).
46. Braak, H. B. and E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* 1–4 (1991). doi:10.1007/bf00308809
47. Braak, H. & Braak, E. Staging of alzheimer's disease-related neurofibrillary changes. *Neurobiol. Aging* **16**, 271–278 (1995).
48. Brandt, R., Léger, J. & Lee, G. Interaction of tau with the neural plasma membrane mediated by tau's amino-terminal projection domain. *J. Cell Biol.* **131**, 1327–1340 (1995).
49. Brown, P. *et al.* Iatrogenic creutzfeldt-Jakob disease, final assessment. *Emerg. Infect. Dis.* **18**, 901–907 (2012).
50. Brundin, P., Melki, R. & Kopito, R. Prion-like transmission of protein aggregates in neurodegenerative diseases. *Nat. Rev. Mol. Cell Biol.* **11**, 301–307 (2010).
51. Brüne, D., Andrade-Navarro, M. A. & Mier, P. Proteome-wide comparison between the amino acid composition of domains and linkers. *BMC Res. Notes* **11**, 1–6 (2018).
52. Buée, Luc ; Delacourte, A. Tauopathy and Alzheimer disease: a full degenerating process. *Psychol NeuroPsychiatr Vieil* **4**, 261–273 (2006).
53. Burré, J. The synaptic function of  $\alpha$ -synuclein. *J. Parkinsons. Dis.* **5**, 699–713 (2015).
54. Cadene, M. & Chait, B. T. A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins. *Anal. Chem.* **72**, 5655–5658 (2000).
55. Caffrey, T. M. & Wade-Martins, R. Functional MAPT haplotypes: Bridging the gap between genotype and neuropathology. *Neurobiol. Dis.* **27**, 1–10 (2007).
56. Caillet-Boudin, M. L., Buée, L., Sergeant, N. & Lefebvre, B. Regulation of human MAPT gene expression. *Mol. Neurodegener.* **10**, 1–14 (2015).
57. Canul-Tec, J. C. *et al.* Structure and allosteric inhibition of excitatory amino acid transporter 1. *Nature* **544**, 446–451 (2017).
58. Carrier, M. F., Simon, C., Cassoly, R. & Pradel, L. A. Interaction between microtubule-associated protein tau and spectrin. *Biochimie* **66**, 305–311 (1984).

59. Cattaneo, E., Zuccato, C. & Tartari, M. Normal huntingtin function: An alternative approach to Huntington's disease. *Nat. Rev. Neurosci.* **6**, 919–930 (2005).
60. Cavallini, A. *et al.* An unbiased approach to identifying tau kinases that phosphorylate tau at sites associated with alzheimer disease. *J. Biol. Chem.* **288**, 23331–23347 (2013).
61. Chambers, C. B., Lee, J. M., Troncoso, J. C., Reich, S. & Muma, N. A. Overexpression of four-repeat tau mRNA isoforms in progressive supranuclear palsy but not in Alzheimer's disease. *Ann. Neurol.* **46**, 325–332 (1999).
62. Chatani, E. & Yamamoto, N. Recent progress on understanding the mechanisms of amyloid nucleation. *Biophys. Rev.* **10**, 527–534 (2018).
63. Chea, E. E. & Jones, L. M. Modifications generated by fast photochemical oxidation of proteins reflect the native conformations of proteins. *Protein Sci.* **27**, 1047–1056 (2018).
64. Chen, D. *et al.* Tau local structure shields an amyloid-forming motif and controls aggregation propensity. *Nat. Commun.* **10**, (2019).
65. Chen, J., Kanai, Y., Cowan, N. J. & Hirokawa, N. Projection domains of MAP2 and tau determine spacings between microtubules in dendrites and axons. *Nature* **360**, 674–677 (1992).
66. Chirita, C. N., Congdon, E. E., Yin, H. & Kuret, J. Triggers of full-length tau aggregation: A role for partially folded intermediates. *Biochemistry* **44**, 5862–5872 (2005).
67. Cicognola, C. *et al.* Novel tau fragments in cerebrospinal fluid: relation to tangle pathology and cognitive decline in Alzheimer's disease. *Acta Neuropathol.* **137**, 279–296 (2019).
68. Clavaguera, F. *et al.* Brain homogenates from human tauopathies induce tau inclusions in mouse brain. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 9535–9540 (2013).
69. Cliffe, R. *et al.* Filamentous Aggregates Are Fragmented by the Proteasome Holoenzyme. *Cell Rep.* **26**, 2140–2149.e3 (2019).
70. Congdon, E. E. *et al.* Nucleation-dependent tau filament formation: The importance of dimerization and an estimation of elementary rate constants. *J. Biol. Chem.* **283**, 13806–13816 (2008).
71. Connell, J. W. *et al.* Quantitative analysis of tau isoform transcripts in sporadic tauopathies. *Mol. Brain Res.* **137**, 104–109 (2005).
72. Constantijnidis, J., Richard, J. & Tissot, R. Pick's disease: Histological and clinical correlations. *Eur. Neurol.* (1974). doi:10.1159/000114319
73. Cornwell, O., Radford, S. E., Ashcroft, A. E. & Ault, J. R. Comparing Hydrogen Deuterium Exchange and Fast Photochemical Oxidation of Proteins: a Structural Characterisation of Wild-Type and  $\Delta N6$   $\beta$ -Microglobulin. *J. Am. Soc. Mass Spectrom.* **29**, 2413–2426 (2018).
74. Cortés, R. *et al.* Relative expression of 3R and 4R tau isoforms in Alzheimer's disease and argyrophilic grain disease brains. (2012).
75. Coughlin, D. & Irwin, D. J. Emerging Diagnostic and Therapeutic Strategies for Tauopathies. *Current Neurology and Neuroscience Reports* **17**, (2017).
76. Crary, J. F. *et al.* Primary age-related tauopathy (PART): a common pathology associated with human aging. *Acta Neuropathol.* **128**, 755–766 (2014).
77. Cripps, D. *et al.* Alzheimer disease-specific conformation of hyperphosphorylated paired helical filament-Tau is polyubiquitinated through Lys-48, Lys-11, and Lys-6 ubiquitin conjugation. *J. Biol. Chem.* **281**, 10825–10838 (2006).
78. Crowther, R. A. Straight and paired helical filaments in Alzheimer disease have a common structural unit. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2288–2292 (1991).
79. Cuillé, J. & Chelle, P. L. Scrapie in Sheep is Certainly Transmissible. *Compte Rendu l'Academie des Sci.* (1938).
80. Cummings, J. L., Morstorf, T. & Zhong, K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimers. Res. Ther.* **6**, 37 (2014).
81. Cummings, J., Lee, G., Ritter, A. & Zhong, K. Alzheimer's disease drug development pipeline: 2018. *Alzheimer's Dement. Transl. Res. Clin. Interv.* **4**, 195–214 (2018).
82. Daebel, V. *et al.*  $\beta$ -sheet core of tau paired helical filaments revealed by solid-state NMR. *J. Am. Chem. Soc.* **134**, 13982–13989 (2012).
83. Dan, A. *et al.* Extensive deamidation at asparagine residue 279 accounts for weak immunoreactivity of tau with RD4 antibody in Alzheimer's disease brain. *Acta Neuropathol. Commun.* **1**, 54 (2013).
84. De Silva, R. *et al.* Pathological inclusion bodies in tauopathies contain distinct complements of tau with three or four microtubule-binding repeat domains as demonstrated by new specific monoclonal antibodies. *Neuropathol. Appl. Neurobiol.* **29**, 288–302 (2003).
85. Delacourte, A., Sergeant, N., Wattez, A., Gauvreau, D. & Robitaille, Y. Vulnerable neuronal subsets in Alzheimer's and Pick's disease are distinguished by their  $\tau$  isoform distribution and phosphorylation. *Ann. Neurol.* **43**, 193–204 (1998).
86. Delacourte, A. Progressive-Supranuclear-Palsy-PSP @ [www.alzheimer-europe.org](http://www.alzheimer-europe.org). (2009).

87. Delacourte, A. ; J. K. Argyrophilic-Grain-Disease-AGD @ www.alzheimer-europe.org. (2009).
88. Delrieu, J. *Traitement médicamenteux de la maladie d'Alzheimer Actualités et perspectives*. (2017).
89. Derisbourg, M. *et al.* Role of the Tau N-terminal region in microtubule stabilization revealed by new endogenous truncated forms. *Sci. Rep.* **5**, 1–10 (2015).
90. Dickson, D. W. Pick's disease: a modern approach. *Brain Pathol.* **8**, 339–54 (1998).
91. Dickson, D. W., Rademakers, R. & Hutton, M. L. Progressive supranuclear palsy: Pathology and genetics. *Brain Pathol.* **17**, 74–82 (2007).
92. Dinkel, P. D., Siddiqua, A., Huynh, H., Shah, M. & Margittai, M. Variations in filament conformation dictate seeding barrier between three- and four-repeat tau. *Biochemistry* **50**, 4330–4336 (2011).
93. Dixit, R., Ross, J. L., Goldman, Y. E. & Holzbour, E. L. F. Differential Regulation of Dynein and Kinesin Motor Proteins by Tau. *Science (80- )*. **319**, 1086–1089 (2008).
94. Donnelly, D. P. *et al.* Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nat. Methods* **16**, 587–594 (2019).
95. Dorval, V. & Fraser, P. E. Small ubiquitin-like modifier (SUMO) modification of natively unfolded proteins tau and  $\alpha$ -synuclein. *J. Biol. Chem.* **281**, 9919–9924 (2006).
96. Dr Voisin, F. de médecine de T. Syndrome Parkinsonien Le Syndrome Parkinsonien. (2015).
97. Drechsel, D. N., Hyman, A. A., Cobb, M. H. & Kirschner, M. W. Modulation of the dynamic instability of tubulin assembly by the microtubule-associated protein tau. *Mol. Biol. Cell* **3**, 1141–1154 (1992).
98. Dreger, M., Bo, W. L., Brownlee, G. G. & Deng, T. A quantitative strategy to detect changes in accessibility of protein regions to chemical modification on heterodimerization. *Protein Sci.* **18**, 1448–1458 (2009).
99. Drubin, D. G. & Kirschner, M. W. Tau protein function in living cells. *J. Cell Biol.* **103**, 2739–2746 (1986).
100. Dujardin, S. *et al.* Ectosomes: A new mechanism for non-exosomal secretion of Tau protein. *PLoS One* **9**, 28–31 (2014).
101. Duyckaerts, C. *et al.* PART is part of Alzheimer disease. *Acta Neuropathol.* **129**, 749–756 (2015).
102. Eanes, E. D. & Glenner, G. G. X-ray diffraction studies on amyloid filaments. *J. Histochem. Cytochem.* **16**, 673–677 (1968).
103. Editorial. Challenges for cryo-EM. *Nat. Methods* **15**, 985 (2018).
104. El Moustaine, D., Perrier, V., Smeller, L., Lange, R. & Torrent, J. Full-length prion protein aggregates to amyloid fibrils and spherical particles by distinct pathways. *FEBS J.* **275**, 2021–2031 (2008).
105. Elbaum-Garfinkle, S. & Rhoades, E. Identification of an Aggregation-Prone Structure of Tau. *J. Am. Chem. Soc.* **134**, 16607–16613 (2012).
106. Elie, A. *et al.* Tau co-organizes dynamic microtubule and actin networks. *Sci. Rep.* **5**, 1–10 (2015).
107. Englander, S. W., Mayne, L., Bai, Y. & Sosnick, T. R. Hydrogen exchange: The modern legacy of Linderstrom-Lang. *Protein Sci.* **6**, 1101–1109 (1997).
108. Evans, L. D. *et al.* Extracellular Monomeric and Aggregated Tau Efficiently Enter Human Neurons through Overlapping but Distinct Pathways. *Cell Rep.* **22**, 3612–3624 (2018).
109. Evidente, V. G. H. & Gwinn, K. A. Post-encephalitic parkinsonism. *J. Neurol. Neurosurg. Psychiatry* **64**, 5–5 (1998).
110. Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* **13**, 397–406 (2014).
111. Falcon, B. *et al.* Conformation determines the seeding potencies of native and recombinant Tau aggregates. *J. Biol. Chem.* **290**, 1049–1065 (2015).
112. Falcon, B. *et al.* Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **561**, 137–140 (2018).
113. Falcon, B. *et al.* Novel tau filament fold in chronic traumatic encephalopathy encloses hydrophobic molecules. *Nature* **568**, 420–423 (2019).
114. Feng, Y. *et al.* Global analysis of protein structural changes in complex proteomes. *Nat. Biotechnol.* **32**, 1036–1044 (2014).
115. Fenyi, A., Coens, A., Bellande, T., Melki, R. & Bousset, L. Assessment of the efficacy of different procedures that remove and disassemble alpha-synuclein, tau and A-beta fibrils from laboratory material and surfaces. *Sci. Rep.* **8**, 1–9 (2018).
116. Fenyi, A. *et al.* Detection of alpha-synuclein aggregates in gastrointestinal biopsies by protein misfolding cyclic amplification. *Neurobiol. Dis.* **129**, 38–43 (2019).
117. Ferrer, I., Santpere, G. & van Leeuwen, F. W. Argyrophilic grain disease. *Brain* **131**, 1416–1432 (2008).
118. Field, E. J. Scrapie: a review of its relation to human disease and ageing. *J. Med. Genet.* **13**, 479–495 (1976).

119. Fink, A. L. Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **3**, 9–23 (1998).
120. Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185–190 (2017).
121. Flach, K. *et al.* Tau oligomers impair artificial membrane integrity and cellular viability. *J. Biol. Chem.* **287**, 43223–43233 (2012).
122. Fontana, A. *et al.* Probing protein structure by limited proteolysis. *Acta Biochim. Pol.* **51**, 299–321 (2004).
123. Fontana, A. *et al.* Probing the conformational state of apomyoglobin by limited proteolysis. *J. Mol. Biol.* **266**, 223–230 (1997).
124. Foote, M. & Zhou, Y. 14-3-3 Proteins in Neurological Disorders. *Int. J. Biochem. Mol. Biol.* **3**, 152–164 (2012).
125. Formichi, P., Battisti, C., Radi, E. & Federico, A. Cerebrospinal fluid tau, A $\beta$ , and phosphorylated tau protein for the diagnosis of Alzheimer's disease. *J. Cell. Physiol.* **208**, 39–46 (2006).
126. Forrest, S. L. *et al.* Heritability in frontotemporal tauopathies. *Alzheimer's Dement. Diagnosis, Assess. Dis. Monit.* **11**, 115–124 (2019).
127. Frecklington, D. General Method for MALDI-MS Analysis of Proteins and Peptides. *Cold Spring Harb. Protoc.* **2007**, pdb.prot4679 (2007).
128. Frederick, J. Pick disease: a brief overview. *Arch. Pathol. Lab. Med.* **130**, 1063–6 (2006).
129. Friedhoff, P., Von Bergen, M., Mandelkow, E. M., Davies, P. & Mandelkow, E. A nucleated assembly mechanism of Alzheimer paired helical filaments. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 15712–15717 (1998).
130. Friedhoff, P., Schneider, A., Mandelkow, E. M. & Mandelkow, E. Rapid assembly of Alzheimer-like paired helical filaments from microtubule-associated protein tau monitored by fluorescence in solution. *Biochemistry* **37**, 10223–10230 (1998).
131. Frost, B. & Diamond, M. I. Prion-like mechanisms in neurodegenerative diseases. *Nat. Rev. Neurosci.* **11**, 155–159 (2010).
132. Frost, B. & Diamond, M. I. Prion-like mechanisms in neurodegenerative diseases. *Nat. Rev. Neurosci.* **11**, 155–159 (2010).
133. Frost, B., Ollesch, J., Wille, H. & Diamond, M. I. Conformational diversity of wild-type tau fibrils specified by templated conformation change. *J. Biol. Chem.* **284**, 3546–3551 (2009).
134. Funk, K. E. *et al.* Lysine methylation is an endogenous post-translational modification of tau protein in human brain and a modulator of aggregation propensity. *Biochem. J.* **462**, 77–88 (2014).
135. Gabant, G., Augier, J. & Armengaud, J. Assessment of solvent residues accessibility using three Sulfo-NHS-biotin reagents in parallel: Application to footprint changes of a methyltransferase upon binding its substrate. *J. Mass Spectrom.* (2008). doi:10.1002/jms.1328
136. Gaber, A., Gunčar, G. & Pavšič, M. Proper evaluation of chemical cross-linking-based spatial restraints improves the precision of modeling homo-oligomeric protein complexes. *BMC Bioinformatics* **20**, 1–12 (2019).
137. Gallagher, W. *FTIR\_of\_proteins.* (2009).
138. Gamblin, T. C. Potential structure/function relationships of predicted secondary structural elements of tau. *Biochimica et Biophysica Acta - Molecular Basis of Disease* **1739**, 140–149 (2005).
139. Gamblin, T. C., Berry, R. W. & Binder, L. I. Tau polymerization: Role of the amino terminus. *Biochemistry* **42**, 2252–2257 (2003).
140. Gamblin, T. C. *et al.* Caspase cleavage of tau: Linking amyloid and neurofibrillary tangles in Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10032–10037 (2003).
141. Gao, J., Friedrichs, M. S., Dongre, A. R. & Opiteck, G. J. Guidelines for the routine application of the peptide hits technique. *J. Am. Soc. Mass Spectrom.* **16**, 1231–1238 (2005).
142. Gath, J. *et al.* Unlike twins: An NMR comparison of two  $\alpha$ -synuclein polymorphs featuring different toxicity. *PLoS One* **9**, 1–11 (2014).
143. Gau, B. C., Chen, J. & Gross, M. L. Fast photochemical oxidation of proteins for comparing solvent-accessibility changes accompanying protein folding: Data processing and application to barstar. *Biochim. Biophys. Acta - Proteins Proteomics* **1834**, 1230–1238 (2013).
144. Giannetti, A. M. *et al.* Fibers of tau fragments, but not full length tau, exhibit a cross  $\beta$ -structure: Implications for the formation of paired helical filaments. *Protein Sci.* **9**, 2427–2435 (2000).
145. Glabe, C. G. Structural classification of toxic amyloid oligomers. *J. Biol. Chem.* **283**, 29639–29643 (2008).
146. Goedert, M., Spillantini, M. G., Potier, M. C., Ulrich, J. & Crowther, R. A. Cloning and sequencing of the cDNA encoding an isoform of microtubule-associated protein tau containing four tandem repeats: differential expression of tau protein mRNAs in human brain. *EMBO J.* **8**, 393–399 (1989).

147. Goedert, M. *et al.* Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans. *Nature* **383**, 550–553 (1996).
148. Goedert, M. Alzheimer's and Parkinson's diseases: The prion concept in relation to assembled A $\beta$ , tau, and  $\alpha$ -synuclein. *Science (80- )*. **349**, 61–69 (2015).
149. Goedert, M., Falcon, B., Zhang, W., Ghetti, B. & Scheres, S. H. W. Distinct Conformers of Assembled Tau in Alzheimer's and Pick's Diseases. *Cold Spring Harb. Symp. Quant. Biol.* **LXXXIII**, 037580 (2019).
150. Goedert, M. & Jakes, R. Expression of separate isoforms of human tau protein: correlation with the tau pattern in brain and effects on tubulin polymerization. *EMBO J.* **9**, 4225–30 (1990).
151. Golde, T. E. *et al.* Thinking laterally about neurodegenerative proteinopathies Find the latest version : Science in medicine Thinking laterally about neurodegenerative proteinopathies. **123**, 1847–1855 (2013).
152. Golghalyani, V., Neupärtl, M., Wittig, I., Bahr, U. & Karas, M. ArgC-Like Digestion: Complementary or Alternative to Tryptic Digestion? *J. Proteome Res.* (2017). doi:10.1021/acs.jproteome.6b00921
153. Goode, B. L., Chau, M., Denis, P. E. & Feinstein, S. C. Structural and functional differences between 3-repeat and 4-repeat tau isoforms: Implications for normal tau function and the onset of neurodegenerative disease. *J. Biol. Chem.* **275**, 38182–38189 (2000).
154. Gorsky, M. K., Burnouf, S., Dols, J., Mandelkow, E. & Partridge, L. Acetylation mimic of lysine 280 exacerbates human Tau neurotoxicity in vivo. *Sci. Rep.* **6**, 1–12 (2016).
155. Graham, W. V., Bonito-Oliva, A. & Sakmar, T. P. Update on Alzheimer's Disease Therapy and Prevention Strategies. *Annu. Rev. Med.* **68**, 413–430 (2017).
156. Grdadolnik, J. ATR-FTIR spectroscopy: Its advantages and limitations. *Acta Chim. Slov.* **49**, 631–642 (2002).
157. Gremer, L. *et al.* Fibril structure of amyloid- $\beta$ (1–42) by cryo–electron microscopy. *Science (80- )*. **358**, 116–119 (2017).
158. Griffin, T. J. *et al.* iTRAQ Reagent-Based Quantitative Proteomic Analysis on a Linear Ion Trap Mass Spectrometer. *J. Proteome Res.* **6**, 4200–4209 (2007).
159. Grimes, J. M. *et al.* Where is crystallography going? *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 152–166 (2018).
160. Grimsley, G. R., Scholtz, J. M. & Pace, C. N. A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci.* **18**, 247–251 (2009).
161. Grinberg, L. T. & Heinsen, H. Argyrophilic grain disease: An update on a frequent cause of dementia. *Dement. Neuropsychol.* **3**, 2–7 (2009).
162. Grundke-Iqbal, I. *et al.* Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc. Natl. Acad. Sci.* **83**, 4913–4917 (1986).
163. Guerrero-Ferreira, R. *et al.* Two new polymorphic structures of alpha-synuclein solved by cryo-electron microscopy. *bioRxiv* 654582 (2019). doi:10.1101/654582
164. Guerrero-Ferreira, R. *et al.* Cryo-EM structure of alpha-synuclein fibrils. *Elife* **7**, 1–18 (2018).
165. Guo, J. L. *et al.* Unique pathological tau conformers from alzheimer's brains transmit tau pathology in nontransgenic mice. *J. Exp. Med.* **213**, 2635–2654 (2016).
166. Hamuro, Y. *et al.* Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry. *J. Biomol. Tech.* **14**, 171–82 (2003).
167. Han, X., Aslanian, A. & Yates, J. R. Mass spectrometry for proteomics. *Curr. Opin. Chem. Biol.* **12**, 483–490 (2008).
168. Hardy, J. & Higgins, G. Alzheimer's disease: the amyloid cascade hypothesis. *Science (80- )*. **256**, 184–185 (1992).
169. Hasegawa, M. *et al.* Protein sequence and mass spectrometric analyses of tau in the Alzheimer's disease brain. *J. Biol. Chem.* **267**, 17047–17054 (1992).
170. He, H. J. *et al.* The proline-rich domain of tau plays a role in interactions with actin. *BMC Cell Biol.* **10**, (2009).
171. Hefti, M. M. *et al.* High-resolution temporal and regional mapping of MAPT expression and splicing in human brain development. *PLoS One* **13**, 1–14 (2018).
172. Heinisch, J. J. & Brandt, R. Signaling pathways and posttranslational modifications of tau in Alzheimer's disease: The humanization of yeast cells. *Microb. Cell* **3**, 135–146 (2016).
173. Hering, J. A., Innocent, P. R. & Haris, P. I. An alternative method for rapid quantification of protein secondary structure from FTIR spectra using neural networks. *Spectroscopy* **16**, 53–69 (2002).
174. Hikosou, R. *et al.* Effect of DNA on filament formation of tau microtubule-binding domain: Structural dependence of DNA. *Chem. Pharm. Bull.* **55**, 1030–1033 (2007).
175. Hill AF, Desbruslais M, Joiner S, Sidle KC, Gowland I, Collinge J, Doey LJ, L. P. The same prion strain causes vCJD and BSE. *Nature* **26**, 1996–1998 (1997).

176. Holmes, B. B. *et al.* Heparan sulfate proteoglycans mediate internalization and propagation of specific proteopathic seeds. *Proc. Natl. Acad. Sci. U. S. A.* **110**, (2013).
177. Holmes, W. M., Klaips, C. L. & Serio, T. R. Defining the limits: Protein aggregation and toxicity in vivo. *Crit. Rev. Biochem. Mol. Biol.* **49**, 294–303 (2014).
178. Huang, R. Y. C. *et al.* Probing Conformational Dynamics of Tau Protein by Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **29**, 174–182 (2018).
179. Hung, S. Y. & Fu, W. M. Drug candidates in clinical trials for Alzheimer's disease. *J. Biomed. Sci.* **24**, 1–12 (2017).
180. Huynh, M. B. *et al.* Glycosaminoglycans from Alzheimer's disease hippocampus have altered capacities to bind and regulate growth factors activities and to bind tau. *PLoS One* **14**, 1–18 (2019).
181. Hwang, S. C., Jhon, D.-Y., Bae, Y. S., Kim, J. H. & Rhee, S. G. Activation of Phospholipase C- $\gamma$  by the Concerted Action of Tau Proteins and Arachidonic Acid. *J. Biol. Chem.* **271**, 18342–18349 (1996).
182. Im, S. Y., Kim, Y. E. & Kim, Y. J. Genetics of Progressive Supranuclear Palsy. *J. Mov. Disord.* **8**, 122–129 (2015).
183. Inekci, D., Jonesco, D. S., Kennard, S., Karsdal, M. A. & Henriksen, K. The potential of pathological protein fragmentation in blood-based biomarker development for dementia - with emphasis on Alzheimer's disease. *Front. Neurol.* **6**, 1–14 (2015).
184. Ingelsson, M. *et al.* Increase in the relative expression of tau with four microtubule binding repeat regions in frontotemporal lobar degeneration and progressive supranuclear palsy brains. *Acta Neuropathol.* **114**, 471–479 (2007).
185. Irwin, D. J. Tauopathies as clinicopathological entities. *Parkinsonism Relat. Disord.* **22**, S29–S33 (2016).
186. Irwin, D. J. *et al.* Deep clinical and neuropathological phenotyping of Pick disease. *Ann. Neurol.* **79**, 272–287 (2016).
187. J Z, W. Abnormal Hyperphosphorylation of Tau. in *Alzheimer's Disease: Advances for a New Century* (ed. Perry, G., Zhu, X., Smith†, M.A., Sorensen, A., Avila, J.) page 128 (IOS Press, 2003).
188. Jakes, R., Novak, M., Davison, M. & Wischik, C. M. Identification of 3- and 4-repeat tau isoforms within the PHF in Alzheimer's disease. *EMBO J.* **10**, 2725–2729 (1991).
189. Jarero-Basulto, J. J. *et al.* Proteolytic cleavage of polymeric tau protein by caspase-3: Implications for Alzheimer disease. *J. Neuropathol. Exp. Neurol.* **72**, 1145–1161 (2013).
190. Jellinger, K. A. Postencephalitic Parkinsonism. in *Neurodegeneration: The Molecular Pathology of Dementia and Movement Disorders: Second Edition* (2011). doi:10.1002/9781444341256.ch19
191. Johnson, G. V. W. & Stoothoff, W. H. Tau phosphorylation in neuronal cell function and dysfunction. *J. Cell Sci.* **117**, 5721–5729 (2004).
192. Johnston, J. A., Ward, C. L. & Kopito, R. R. Aggresomes: a cellular response to misfolded proteins. *J. Cell Biol.* **143**, 1883–98 (1998).
193. Jones, A. X. *et al.* Improving mass spectrometry analysis of protein structures with arginine-selective chemical cross-linkers. *Nat. Commun.* **10**, (2019).
194. Josephs, K. A. Current Understanding of Neurodegenerative Diseases Associated With the Protein Tau. *Mayo Clinic Proceedings* **92**, 1291–1303 (2017).
195. Kalinderi, K., Fidani, L. & Bostantjopoulou, S. From 1997 to 2007: a decade journey through the H1 haplotype on 17q21 chromosome. *Parkinsonism Relat. Disord.* **15**, 2–5 (2009).
196. Kampers, T., Friedhoff, P., Biernat, J., Mandelkow, E. M. & Mandelkow, E. RNA stimulates aggregation of microtubule-associated protein tau into Alzheimer-like paired helical filaments. *FEBS Lett.* **399**, 344–349 (1996).
197. Karamanos, T. K., Kalverda, A. P., Thompson, G. S. & Radford, S. E. Mechanisms of amyloid formation revealed by solution NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **88–89**, 86–104 (2015).
198. Katayama, T. *et al.* Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. *Nat. Cell Biol.* **1**, 479–485 (1999).
199. Kaufman, S. K. *et al.* Tau Prion Strains Dictate Patterns of Cell Pathology, Progression Rate, and Regional Vulnerability In Vivo. *Neuron* **92**, 796–812 (2016).
200. Kfoury, N., Holmes, B. B., Jiang, H., Holtzman, D. M. & Diamond, M. I. Trans-cellular propagation of Tau aggregation by fibrillar species. *J. Biol. Chem.* **287**, 19440–19451 (2012).
201. Khanna, M. R., Kovalevich, J., Lee, V. M., Trojanowski, J. Q. & Brunden, K. R. Therapeutic strategies for the treatment of tauopathies: Hopes and challenges. *Alzheimer's Dement.* **12**, 1051–1065 (2016).
202. Kheterpal, I., Williams, A., Murphy, C., Bledsoe, B. & Wetzel, R. Structural features of the A $\beta$  amyloid fibril elucidated by limited proteolysis. *Biochemistry* **40**, 11757–11767 (2001).
203. Kheterpal, I., Cook, K. D. & Wetzel, R. Hydrogen/Deuterium Exchange Mass Spectrometry Analysis of Protein Aggregates. *Methods Enzymol.* **413**, 140–166 (2006).

204. Kheterpal, I., Zhou, S., Cook, K. D. & Wetzel, R. A $\beta$  amyloid fibrils possess a core structure highly resistant to hydrogen exchange. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13597–13601 (2000).
205. Kidd, M. Paired Helical Filaments in Electron Microscopy of Alzheimer's Disease. *Nature* **197**, 192–193 (1963).
206. Kim, D. *et al.* Primary age-related tauopathy: An elderly brain pathology frequently encountered during autopsy. *J. Pathol. Transl. Med.* **53**, 159–163 (2019).
207. Kim, D. *et al.* Identification of disulfide cross-linked tau dimer responsible for tau propagation. *Sci. Rep.* **5**, 1–10 (2015).
208. Kim, G. H., Kim, J. E., Rhie, S. J. & Yoon, S. The Role of Oxidative Stress in Neurodegenerative Diseases. *Exp. Neurobiol.* **24**, 325 (2015).
209. Kim, J. S., Shin, M., Song, J. S., An, S. & Kim, H. J. C-terminal de novo sequencing of peptides using oxazolone-based derivatization with bromine signature. *Anal. Biochem.* **419**, 211–216 (2011).
210. King, M. E., Gamblin, T. C., Kuret, J. & Binder, L. I. Differential assembly of human tau isoforms in the presence of arachidonic acid. *J. Neurochem.* **74**, 1749–1757 (2000).
211. King, M. E., Ghoshal, N., Wall, J. S., Binder, L. I. & Ksiezak-Reding, H. Structural analysis of pick's disease-derived and in vitro-assembled tau filaments. *Am. J. Pathol.* **158**, 1481–1490 (2001).
212. Kizuka, Y., Kitazume, S. & Taniguchi, N. N-glycan and Alzheimer's disease. *Biochim. Biophys. Acta - Gen. Subj.* **1861**, 2447–2454 (2017).
213. Kluger, R. & Alagic, A. Chemical cross-linking and protein-protein interactions—a review with illustrative protocols. *Bioorg. Chem.* **32**, 451–472 (2004).
214. Kolarova, M., García-Sierra, F., Bartos, A., Ricny, J. & Ripova, D. Structure and pathology of tau protein in Alzheimer disease. *Int. J. Alzheimers. Dis.* (2012). doi:10.1155/2012/731526
215. Kordower, J. H., Chu, Y., Hauser, R. A., Freeman, T. B. & Olanow, C. W. Lewy body-like pathology in long-term embryonic nigral transplants in Parkinson's disease. *Nat. Med.* **14**, 504–506 (2008).
216. Kouri, N. *et al.* Genome-wide association study of corticobasal degeneration identifies risk variants shared with progressive supranuclear palsy. *Nat. Commun.* **6**, 1–7 (2015).
217. Kovacs, G. G. Invited review: Neuropathology of tauopathies: Principles and practice. *Neuropathol. Appl. Neurobiol.* **41**, 3–23 (2015).
218. Kremen, S. A., Solis, O. E., Shapira, J. S., Vinters, H. V. & Mendez, M. F. 'fantastic Thinking' in pathologically proven pick disease. *Cogn. Behav. Neurol.* **23**, 130–134 (2010).
219. Ksiezak-Reding, H. *et al.* Ultrastructure and biochemical composition of paired helical filaments in corticobasal degeneration. *Am. J. Pathol.* **145**, 1496–1508 (1994).
220. Kuret, J. *et al.* Pathways of tau fibrillization. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1739**, 167–178 (2005).
221. Kurowska, Z. *et al.* Signs of degeneration in 12-22-year old grafts of mesencephalic dopamine neurons in patients with Parkinson's disease. *J. Parkinsons. Dis.* **1**, 83–92 (2011).
222. Laferrière, F. *et al.* TDP-43 extracted from frontotemporal lobar degeneration subject brains displays distinct aggregate assemblies and neurotoxic effects reflecting disease progression rates. *Nat. Neurosci.* **22**, 65–77 (2019).
223. Landau, M. *et al.* Towards a Pharmacophore for Amyloid. *PLoS Biol.* **9**, 25–27 (2011).
224. Landino, L. M., Skreslet, T. E. & Alston, J. A. Cysteine oxidation of tau and microtubule-associated protein-2 by peroxynitrite: Modulation of microtubule assembly kinetics by the thioredoxin reductase system. *J. Biol. Chem.* **279**, 35101–35105 (2004).
225. Lasagna-Reeves, C. A. *et al.* Alzheimer brain-derived tau oligomers propagate pathology from endogenous tau. *Sci. Rep.* **2**, (2012).
226. Lasagna-Reeves, C. A. *et al.* Identification of oligomers at early stages of tau aggregation in Alzheimer's disease. *FASEB J.* **26**, 1946–1959 (2012).
227. Lee, J. C., Kim, S. J., Hong, S. & Kim, Y. S. Diagnosis of Alzheimer's disease utilizing amyloid and tau as fluid biomarkers. *Exp. Mol. Med.* **51**, (2019).
228. Lee, V. M. Y. Pathological Tau Aggregation. *Nat. Commun.* **2**, 1–18 (2011).
229. Lee, V. M.-Y., Goedert, M. & Trojanowski, J. Q. Neurodegenerative Tauopathies. *Annu. Rev. Neurosci.* **24**, 1121–159 (2001).
230. Lefaki, M., Papaevgeniou, N. & Chondrogianni, N. Redox regulation of proteasome function. *Redox Biol.* **13**, 452–458 (2017).
231. Leuzy, A. *et al.* Tau PET imaging in neurodegenerative tauopathies—still a challenge. *Molecular Psychiatry* (2019). doi:10.1038/s41380-018-0342-8
232. Li, B. *et al.* Cryo-EM of full-length  $\alpha$ -synuclein reveals fibril polymorphs with a common structural kernel. *Nat. Commun.* **9**, 1–10 (2018).
233. Li, D. *et al.* Structure-based design of functional amyloid materials. *J. Am. Chem. Soc.* **136**, 18044–18051 (2014).

234. Li, H. L. *et al.* Phosphorylation of tau antagonizes apoptosis by stabilizing  $\beta$ -catenin, a mechanism involved in Alzheimer's neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3591–3596 (2007).
235. Li, K. S., Rempel, D. L. & Gross, M. L. Conformational-Sensitive Fast Photochemical Oxidation of Proteins and Mass Spectrometry Characterize Amyloid Beta 1–42 Aggregation. *J. Am. Chem. Soc.* **138**, 12090–12098 (2016).
236. Li, Q. *et al.* Structural attributes of mammalian prion infectivity: Insights from studies with synthetic prions. *J. Biol. Chem.* **293**, 18494–18503 (2018).
237. Liberski, P. P. Kuru: A journey back in time from Papua New Guinea to the Neanderthals' extinction. *Pathogens* **2**, 472–505 (2013).
238. Lim, J. & Yue, Z. Neuronal Aggregates: Formation, Clearance, and Spreading. *Dev. Cell* **32**, 491–501 (2015).
239. Lima, D. B. *et al.* Characterization of homodimer interfaces with cross-linking mass spectrometry and isotopically labeled proteins. *Nat. Protoc.* (2018). doi:10.1038/nprot.2017.113
240. Liu, C. *et al.* Characteristics of Amyloid-Related Oligomers Revealed by Crystal Structures of Macrocyclic  $\beta$ -Sheet Mimics. *J. Am. Chem. Soc.* **133**, 6736–6744 (2011).
241. Liu, F. & Gong, C. X. Tau exon 10 alternative splicing and tauopathies. *Mol. Neurodegener.* **3**, 1–10 (2008).
242. Liu, F., Iqbal, K., Grundke-Iqbal, I., Hart, G. W. & Gong, C. X. O-GlcNAcylation regulates phosphorylation of tau: A mechanism involved in Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 10804–10809 (2004).
243. Liu, T., Marcinko, T. M., Kiefer, P. A. & Vachet, R. W. Using Covalent Labeling and Mass Spectrometry To Study Protein Binding Sites of Amyloid Inhibiting Molecules. *Anal. Chem.* **89**, 11583–11591 (2017).
244. Lu, P. J., Wulf, G., Zhou, X. Z., Davies, P. & Lu, K. P. The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* **399**, 784–788 (1999).
245. Luk, K. C. *et al.* Pathological  $\alpha$ -Synuclein Transmission Initiates Parkinson-like Neurodegeneration in Nontransgenic Mice. *Science (80- )*. **338**, 949–953 (2012).
246. Luk, K. C. *et al.* Exogenous  $\alpha$ -synuclein fibrils seed the formation of Lewy body-like intracellular inclusions in cultured cells. *Proc. Natl. Acad. Sci.* **106**, 20051–20056 (2009).
247. Lyumkis, D. Challenges and opportunities in cryo-EM single-particle analysis. *J. Biol. Chem.* **294**, 5181–5197 (2019).
248. Maeda, S. *et al.* Granular tau oligomers as intermediates of tau filaments. *Biochemistry* **46**, 3856–3861 (2007).
249. Mahapatra, R. K., Edwards, M. J., Schott, J. M. & Bhatia, K. P. Corticobasal degeneration. *Lancet Neurol.* **3**, 736–743 (2004).
250. Mair, W. *et al.* FLEXITau: Quantifying Post-translational Modifications of Tau Protein in Vitro and in Human Disease. *Anal. Chem.* **88**, 3704–3714 (2016).
251. Makrides, V. *et al.* Microtubule-dependent oligomerization of tau: Implications for physiological tau function and tauopathies. *J. Biol. Chem.* **278**, 33298–33304 (2003).
252. Malinchik, S. B., Inouye, H., Szumowski, K. E. & Kirschner, D. A. Structural analysis of Alzheimer's  $\beta$ (1-40) amyloid: Protofilament assembly of tubular fibrils. *Biophys. J.* **74**, 537–545 (1998).
253. Marsh, A. P. Molecular mechanisms of proteinopathies across neurodegenerative disease: a review. **8**, 1–7 (2019).
254. Masuda-Suzukake, M. *et al.* Prion-like spreading of pathological  $\alpha$ -synuclein in brain. *Brain* **136**, 1128–1138 (2013).
255. Mazzoni, M. R., Porchia, F. & Hamm, H. E. Proteolytic Fragmentation for Epitope Mapping. in *Epitope Mapping Protocols* **524**, 77–86 (2009).
256. Medina, M. An overview on the clinical development of tau-based therapeutics. *Int. J. Mol. Sci.* **19**, (2018).
257. Melki, R. How the shapes of seeds can influence pathology. *Neurobiol. Dis.* **109**, 201–208 (2018).
258. Melková, K. *et al.* Structure and functions of microtubule associated proteins tau and map2c: Similarities and differences. *Biomolecules* **9**, 1–32 (2019).
259. Mendoza, V. L. & Vachet, R. W. Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom. Rev.* **28**, 785–815 (2009).
260. Merezko, M. *et al.* Secretion of Tau via an Unconventional Non-vesicular Mechanism. *Cell Rep.* **25**, 2027–2035.e4 (2018).
261. Meyer-Luehmann, M. *et al.* Exogenous induction of cerebral  $\beta$ -amyloidogenesis is governed by agent and host. *Science (80- )*. **313**, 1781–1784 (2006).
262. Mirbaha, H. *et al.* Inert and seed-competent tau monomers suggest structural origins of aggregation. *Elife* **7**, 1–29 (2018).

263. Misra, U. K. & Kalita, J. Spectrum of movement disorders in encephalitis. *J. Neurol.* **257**, 2052–2058 (2010).
264. Molinuevo, J. L. *et al.* *Current state of Alzheimer's fluid biomarkers.* *Acta Neuropathologica* **136**, (Springer Berlin Heidelberg, 2018).
265. Mondragón-Rodríguez, S. *et al.* Conformational changes and cleavage of tau in Pick bodies parallel the early processing of tau found in Alzheimer pathology. *Neuropathol. Appl. Neurobiol.* **34**, 62–75 (2008).
266. Moore, B. D., Rangachari, V., Tay, W. M., Milkovic, N. M. & Rosenberry, T. L. Biophysical analyses of synthetic amyloid- $\beta$ (1-42) aggregates before and after covalent cross-linking. Implications for deducing the structure of endogenous amyloid- $\beta$  oligomers. *Biochemistry* **48**, 11796–11806 (2009).
267. Morales, R., Abid, K. & Soto, C. The prion strain phenomenon: Molecular basis and unprecedented features. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1772**, 681–691 (2007).
268. Moreaud. Les formes “sémantiques” des aphasies primaires progressives : variant sémantique et démence sémantique. *La Lett. du Neurol.* **XVII**, 46–50 (2013).
269. Moreno-Herrero, F., Pérez, M., Baró, A. M. & Avila, J. Characterization by Atomic Force Microscopy of Alzheimer Paired Helical Filaments under Physiological Conditions. *Biophys. J.* **86**, 517–525 (2004).
270. Morris, H. R., Paxton, T., Panico, M., McDowell, R. & Dell, A. A novel geometry mass spectrometer, the Q-TOF, for low- femtomole/attomole-range biopolymer sequencing. in *Journal of Protein Chemistry* (1997). doi:10.1023/A:1026309410737
271. Mudher, A. *et al.* What is the evidence that tau pathology spreads through prion-like propagation? *Acta Neuropathol. Commun.* **5**, 99 (2017).
272. Mukrasch, M. D. *et al.* Structural polymorphism of 441-residue Tau at single residue resolution. *PLoS Biol.* **7**, 0399–0414 (2009).
273. Müller, U. C. & Zheng, H. Physiological functions of APP family proteins. *Cold Spring Harb. Perspect. Med.* **2**, 1–17 (2012).
274. Münch, G., Thome, J., Foley, P., Schinzel, R. & Riederer, P. Advanced glycation endproducts in ageing and Alzheimer's disease. *Brain Res. Rev.* **23**, 134–143 (1997).
275. Nakazawa, T. *et al.* Terminal proteomics: N- and C-terminal analyses for high-fidelity identification of proteins using MS. *Proteomics* **8**, 673–685 (2008).
276. Narasimhan, S. *et al.* Pathological tau strains from human brains recapitulate the diversity of tauopathies in nontransgenic mouse brain. *J. Neurosci.* **37**, 11406–11423 (2017).
277. Nazabal, A., Dos Reis, S. D., Bonneu, M., Saube, S. J. & Schmitter, J. M. Conformational transition occurring upon amyloid aggregation of the HET-s prion protein of *Podospora anserina* analyzed by hydrogen/deuterium exchange and mass spectrometry. *Biochemistry* **42**, 8852–8861 (2003).
278. Novak, M., Jakes, R., Edwards, P. C., Milstein, C. & Wischik, C. M. Difference between the tau protein of Alzheimer paired helical core and normal tau revealed by epitope analysis of monoclonal antibodies 423 and 7.51. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 5837–5841 (1991).
279. Novak, M., Kabat, J. & Wischik, C. M. Molecular characterization of the minimal protease resistant tau unit of the Alzheimer's disease paired helical filament. *EMBO J.* **12**, 365–370 (1993).
280. Nury, C. *et al.* A novel bio-orthogonal cross-linker for improved protein/protein interaction analysis. *Anal. Chem.* **87**, 1853–1860 (2015).
281. O'Connor, C. M., DiMaggio, P. A., Shenk, T. & Garcia, B. A. Quantitative proteomic discovery of dynamic epigenome changes that control human cytomegalovirus (HCMV) infection. *Mol. Cell. Proteomics* **13**, 2399–2410 (2014).
282. Orr, M. E., Sullivan, A. C. & Frost, B. A Brief Overview of Tauopathy: Causes, Consequences, and Therapeutic Strategies. *Trends in Pharmacological Sciences* **38**, 637–648 (2017).
283. Paramelle, D., Miralles, G., Subra, G. & Martinez, J. Chemical cross-linkers for protein structure studies by mass spectrometry. *Proteomics* **13**, 438–456 (2013).
284. Park, S., Lee, J. H., Jeon, J. H. & Lee, M. J. Degradation or aggregation: The ramifications of post-translational modifications on tau. *BMB Rep.* **51**, 265–273 (2018).
285. Park, S. Y. & Ferreira, A. The generation of a 17 kDa neurotoxic fragment: An alternative mechanism by which tau mediates  $\beta$ -amyloid-induced neurodegeneration. *J. Neurosci.* **25**, 5365–5375 (2005).
286. Parthimos, T. P. & Schulpis, K. H. The significant characteristics of corticobasal syndrome. *Ann. Res. Hosp.* **3**, 4–4 (2019).
287. Passamonti, L. *et al.* 18F-AV-1451 positron emission tomography in Alzheimer's disease and progressive supranuclear palsy. *Brain* **140**, 781–791 (2017).
288. Patterson, K. R. *et al.* Characterization of prefibrillar tau oligomers in vitro and in Alzheimer disease. *J. Biol. Chem.* (2011). doi:10.1074/jbc.M111.237974
289. Patterson, K. R. *et al.* Characterization of prefibrillar tau oligomers in vitro and in Alzheimer disease. *J. Biol. Chem.* **286**, 23063–23076 (2011).

290. Peelaerts, W. *et al.*  $\alpha$ -Synuclein strains cause distinct synucleinopathies after local and systemic administration. *Nature* **522**, 340–344 (2015).
291. Peeraer, E. *et al.* Intracerebral injection of preformed synthetic tau fibrils initiates widespread tauopathy and neuronal loss in the brains of tau transgenic mice. *Neurobiol. Dis.* **73**, 83–95 (2015).
292. Pérez, M., Valpuesta, J. M., Medina, M., Montejo de Garcini, E. & Avila, J. Polymerization of tau into filaments in the presence of heparin: the minimal sequence required for tau-tau interaction. *J. Neurochem.* **67**, 1183–90 (1996).
293. Petrotchenko, E. V. & Borchers, C. H. Crosslinking combined with mass spectrometry for structural proteomics. *Mass Spectrom. Rev.* **29**, 862–876 (2010).
294. Pickering-Brown, S. *et al.* Pick's disease is associated with mutations in the tau gene. *Ann. Neurol.* **48**, 859–867 (2000).
295. Polanco, J. C. *et al.* Amyloid- $\beta$  and tau complexity - Towards improved biomarkers and targeted therapies. *Nat. Rev. Neurol.* **14**, 22–40 (2018).
296. Polverino De Laureto, P. *et al.* Protein aggregation and amyloid fibril formation by an SH3 domain probed by limited proteolysis. *J. Mol. Biol.* **334**, 129–141 (2003).
297. Prescrire. Pour mieux soigner, des médicaments à écarter : bilan 2015. *Prescrire* **35**, 144–151 (2015).
298. Probst, A. *et al.* Axonopathy and amyotrophy in mice transgenic for human four-repeat tau protein. *Acta Neuropathol.* **99**, 469–481 (2000).
299. Prusiner, S. B. Novel proteinaceous infectious particles cause scrapie. *Science (80-. )*. **216**, 136–144 (1982).
300. Qian, W. & Liu, F. Regulation of alternative splicing of tau exon 10. *Neurosci. Bull.* **30**, 367–377 (2014).
301. Quinn, J. P., Corbett, N. J., Kellett, K. A. B. & Hooper, N. M. Tau Proteolysis in the Pathogenesis of Tauopathies: Neurotoxic Fragments and Novel Biomarkers. *J. Alzheimers. Dis.* **63**, 13–33 (2018).
302. Rail, D., Scholtz, C. & Swash, M. Post-encephalitic Parkinsonism: Current experience. *J. Neurol. Neurosurg. Psychiatry* **44**, 670–676 (1981).
303. Ramachandran, G. & Udgaonkar, J. B. Difference in fibril core stability between two tau four-repeat domain proteins: A hydrogen-deuterium exchange coupled to mass spectrometry study. *Biochemistry* **52**, 8787–8789 (2013).
304. Ramachandran, G. & Udgaonkar, J. B. Mechanistic studies unravel the complexity inherent in tau aggregation leading to Alzheimer's disease and the tauopathies. *Biochemistry* **52**, 4107–4126 (2013).
305. Ramakrishnan, P., Dickson, D. W. & Davies, P. Pin1 colocalization with phosphorylated tau in Alzheimer's disease and other tauopathies. *Neurobiol. Dis.* **14**, 251–264 (2003).
306. Rane, J. S., Kumari, A. & Panda, D. An acetylation mimicking mutation, K274Q, in tau imparts neurotoxicity by enhancing tau aggregation and inhibiting tubulin polymerization. *Biochem. J.* **476**, 1401–1417 (2019).
307. Rappsilber, J. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173**, 530–540 (2011).
308. Redeker, V. *et al.* Hydrogen/Deuterium Exchange Mass Spectrometric Analysis of Conformational Changes Accompanying the Assembly of the Yeast Prion Ure2p into Protein Fibrils. *J. Mol. Biol.* **369**, 1113–1125 (2007).
309. Redeker, V., Pemberton, S., Bienvenut, W., Bousset, L. & Melki, R. Identification of protein interfaces between  $\alpha$ -synuclein, the principal component of Lewy bodies in Parkinson disease, and the molecular chaperones human Hsc70 and the yeast Ssa1p. *J. Biol. Chem.* **287**, 32630–32639 (2012).
310. Renner, M. & Melki, R. Protein aggregation and prionopathies. *Pathol. Biol.* **62**, 162–168 (2014).
311. Resing, K. A. & Ahn, N. G. Proteomics strategies for protein identification. *FEBS Lett.* **579**, 885–889 (2005).
312. Rey, M., Dupré, M., Lopez-Neira, I., Duchateau, M. & Chamot-Rooke, J. EXL-MS: An Enhanced Cross-Linking Mass Spectrometry Workflow to Study Protein Complexes. *Anal. Chem.* **90**, 10707–10714 (2018).
313. Reynolds, C. H. *et al.* Phosphorylation regulates tau interactions with Src homology 3 domains of phosphatidylinositol 3-kinase, phospholipase Cy1, Grb2, and Src family kinases. *J. Biol. Chem.* **283**, 18177–18186 (2008).
314. Rhein, V. *et al.* Amyloid- $\beta$  and tau synergistically impair the oxidative phosphorylation system in triple transgenic Alzheimer's disease mice. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 20057–20062 (2009).
315. Ricciarelli, R. & Fedele, E. The Amyloid Cascade Hypothesis in Alzheimer's Disease: It's Time to Change Our Mind. *Curr. Neuropharmacol.* **15**, 926–935 (2017).
316. Robertson, L. A., Moya, K. L. & Breen, K. C. The potential role of tau protein O-glycosylation in Alzheimer's disease. *J. Alzheimer's Dis.* **6**, 489–495 (2004).

317. Rodriguez, J. A. *et al.* Structure of the toxic core of  $\alpha$ -synuclein from invisible crystals. *Nature* **525**, 486–490 (2015).
318. Rodriguez, R. D. & Grinberg, L. T. Argrophilic grain disease: An underestimated tauopathy. *Dement. Neuropsychol.* **9**, 2–8 (2015).
319. Roepstorff, P. & Fohlman, J. Letter to the editors - Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biol. Mass Spectrom.* (1984). doi:10.1002/bms.1200111109
320. Ruesink, H. *et al.* Stabilization of  $\alpha$ -synuclein oligomers using formaldehyde. *PLoS One* **14**, 1–11 (2019).
321. S, G. *et al.* Ortho-proteogenomics : Multiple proteomes investigation through orthology and a new MS-based protocol. *Genome Res.* 128–135 (2009). doi:10.1101/gr.081901.108.128
322. Sachse, C., Fändrich, M. & Nikolaus Grigorieff. Paired beta-sheet structure of an A $\beta$  (1-40) amyloid fibril revealed by electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 7462–7466 (2008).
323. Sadqi, M. *et al.*  $\alpha$ -helix structure in Alzheimer's disease aggregates of tau-protein. *Biochemistry* **41**, 7150–7155 (2002).
324. Saijo, E. *et al.* Ultrasensitive and selective detection of 3-repeat tau seeding activity in Pick disease brain and cerebrospinal fluid. *Acta Neuropathol.* **133**, 751–765 (2017).
325. Saman, S. *et al.* Exosome-associated tau is secreted in tauopathy models and is selectively phosphorylated in cerebrospinal fluid in early Alzheimer disease. *J. Biol. Chem.* **287**, 3842–3849 (2012).
326. Sanders, D. W. *et al.* Distinct tau prion strains propagate in cells and mice and define different tauopathies. *Neuron* **82**, 1271–1288 (2014).
327. Santa-Maria, I. *et al.* The MAPT H1 haplotype is associated with tangle-predominant dementia. *Acta Neuropathol.* **124**, 693–704 (2012).
328. Sarazin, M., Cruz De Souza, L. & Dubois, B. Maladie d'Alzheimer : nouveaux critères de diagnostic clinique. *La Lett. du Neurol.* **XVI**, 301–304 (2012).
329. Sato, Y., Naito, Y., Grundke-Iqbal, I., Iqbal, K. & Endo, T. Analysis of N-glycans of pathological tau: Possible occurrence of aberrant processing of tau in Alzheimer's disease. *FEBS Lett.* **496**, 152–160 (2001).
330. Sawaya, M. R. *et al.* Atomic structures of amyloid cross- $\beta$  spines reveal varied steric zippers. *Nature* **447**, 453–457 (2007).
331. Scaravilli, T., Tolosa, E. & Ferrer, I. Progressive supranuclear palsy and corticobasal degeneration: Lumping versus splitting. *Mov. Disord.* **20**, (2005).
332. Scarff, C. A. *et al.* Examination of ataxin-3 (atx-3) aggregation by structural mass spectrometry techniques: A rationale for expedited aggregation upon polyglutamine (polyQ) expansion. *Mol. Cell. Proteomics* **14**, 1241–1253 (2015).
333. Scheckel, C. & Aguzzi, A. Prions, prionoids and protein misfolding disorders. *Nat. Rev. Genet.* **19**, 405–418 (2018).
334. Schedin-Weiss, S., Winblad, B. & Tjernberg, L. O. The role of protein glycosylation in Alzheimer disease. *FEBS J.* **281**, 46–62 (2014).
335. Schmidt, M. *et al.* Peptide dimer structure in an A $\beta$ (1-42) fibril visualized with cryo-EM. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11858–11863 (2015).
336. Schneider, A., Biernat, J., Von Bergen, M., Mandelkow, E. & Mandelkow, E. M. Phosphorylation that detaches tau protein from microtubules (Ser262, Ser214) also protects it against aggregation into Alzheimer paired helical filaments. *Biochemistry* **38**, 3549–3558 (1999).
337. Schöll, M. *et al.* Biomarkers for tau pathology. *Mol. Cell. Neurosci.* **97**, 18–33 (2019).
338. Schopper, S. *et al.* Measuring protein structural changes on a proteome-wide scale using limited proteolysis-coupled mass spectrometry. *Nat. Protoc.* **12**, 2391–2410 (2017).
339. Schweers, O., Mandelkow, E. M., Biernat, J. & Mandelkow, E. Oxidation of cysteine-322 in the repeat domain of microtubule-associated protein tau controls the in vitro assembly of paired helical filaments. *Proc. Natl. Acad. Sci.* **92**, 8463–8467 (1995).
340. Scialò, C., De Cecco, E., Manganotti, P. & Legname, G. Prion and Prion-Like Protein Strains: Deciphering the Molecular Basis of Heterogeneity in Neurodegeneration. *Viruses* **11**, 1–37 (2019).
341. Sergeant, N., Delacourte, A. & Buée, L. Tau protein as a differential biomarker of tauopathies. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1739**, 179–197 (2005).
342. Serpa, J. J. *et al.* Using multiple structural proteomics approaches for the characterization of prion proteins. *J. Proteomics* **81**, 31–42 (2013).
343. Serpell, L. C., Berriman, J., Jakes, R., Goedert, M. & Crowther, R. A. Fiber diffraction of synthetic  $\alpha$ -synuclein filaments shows amyloid-like cross- $\beta$  conformation. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4897–4902 (2000).

344. Serrano-Pozo, A., Frosch, M. P., Masliah, E. & Hyman, B. T. Neuropathological alterations in Alzheimer disease. *Cold Spring Harb. Perspect. Med.* **1**, 1–23 (2011).
345. Sheng, Y., Capri, J., Waring, A., Valentine, J. S. & Whitelegge, J. Exposure of Solvent-Inaccessible Regions in the Amyloidogenic Protein Human SOD1 Determined by Hydroxyl Radical Footprinting. *J. Am. Soc. Mass Spectrom.* **30**, 218–226 (2019).
346. Sherman, M. Y. & Goldberg, A. L. Cellular Defenses against Unfolded Proteins. *Neuron* **29**, 15–32 (2001).
347. Shivu, B. *et al.* Distinct  $\beta$ -sheet structure in protein aggregates determined by ATR-FTIR spectroscopy. *Biochemistry* **52**, 5176–5183 (2013).
348. Shrivastava, A. N. *et al.* Clustering of Tau fibrils impairs the synaptic composition of  $\alpha$ 3-Na<sup>+</sup>/K<sup>+</sup>-ATPase and AMPA receptors. *EMBO J.* **38**, 1–22 (2019).
349. Sibille, N. *et al.* Structural impact of heparin binding to full-length Tau as studied by NMR spectroscopy. *Biochemistry* **45**, 12560–12572 (2006).
350. Siddiqua, A. *et al.* Conformational basis for asymmetric seeding barrier in filaments of three- and four-repeat tau. *J. Am. Chem. Soc.* **134**, 10271–10278 (2012).
351. Siddiqua, A. & Margittai, M. Three- and four-repeat tau coassemble into heterogeneous filaments: An implication for Alzheimer disease. *J. Biol. Chem.* **285**, 37920–37926 (2010).
352. Signor, L. & Erba, E. B. Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometric analysis of intact proteins larger than 100 kDa. *J. Vis. Exp.* 1–7 (2013). doi:10.3791/50635
353. Sillen, A. *et al.* Regions of tau implicated in the paired helical fragment core as defined by NMR. *ChemBioChem* **6**, 1849–1856 (2005).
354. Silva, J. C. *et al.* Quantitative proteomic analysis by accurate mass retention time pairs. *Anal. Chem.* **77**, 2187–2200 (2005).
355. Šimić, G. *et al.* Tau Protein Hyperphosphorylation and Aggregation in Alzheimer's Disease and Other Tauopathies, and Possible Neuroprotective Strategies. *Biomolecules* **6**, 6 (2016).
356. Snow, A. D., Mar, H., Nochlin, D., Kresse, H. & Wight, T. N. Peripheral distribution of dermatan sulfate proteoglycans (decorin) in amyloid-containing plaques and their presence in neurofibrillary tangles of Alzheimer's disease. *J. Histochem. Cytochem.* **40**, 105–113 (1992).
357. Souirti, Z. *et al.* D - 15 Syndrome parkinsonien aigu post encéphalitique. *Rev. Neurol. (Paris)*. **163**, 166 (2007).
358. Speicher, K. D., Gorman, N. & Speicher, D. W. N-Terminal Sequence Analysis of Proteins and Peptides. in *Current Protocols in Protein Science* **23**, 11.10.1-11.10.38 (John Wiley & Sons, Inc., 1997).
359. Spillantini, M. G. & Goedert, M. Tau pathology and neurodegeneration. *The Lancet Neurology* (2013). doi:10.1016/S1474-4422(13)70090-5
360. Squier, T. C. Oxidative stress and protein aggregation during biological aging. *Exp. Gerontol.* **36**, 1539–1550 (2001).
361. Stefani, M. & Dobson, C. M. Protein aggregation and aggregate toxicity: New insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* **81**, 678–699 (2003).
362. Stelzmann, R. A., Norman Schnitzlein, H. & Murrilagh, F. R. *An English Translation of Alzheimer's 1907 Paper, 'über eine eigenartige Erkrankung der Hirnrinde'.* *Clinical Anatomy* **8**, (1995).
363. Stöhr, J. *et al.* Purified and synthetic Alzheimer's amyloid beta (A $\beta$ ) prions. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11025–11030 (2012).
364. Stöhr, J. *et al.* A 31-residue peptide induces aggregation of tau's microtubule-binding region in cells. *Nat. Chem.* **9**, 874–881 (2017).
365. Strang, K. H., Golde, T. E. & Giasson, B. I. MAPT mutations, tauopathy, and mechanisms of neurodegeneration. *Lab. Invest.* **99**, 912–928 (2019).
366. Sunde, M. *et al.* Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273**, 729–739 (1997).
367. Sweeney, P. *et al.* Protein misfolding in neurodegenerative diseases: Implications and strategies. *Transl. Neurodegener.* **6**, 1–13 (2017).
368. Takanashi, M., Mori, H., Arima, K., Mizuno, Y. & Hattori, N. Expression patterns of tau mRNA isoforms correlate with susceptible lesions in progressive supranuclear palsy and corticobasal degeneration. *Brain Res. Mol. Brain Res.* **104**, 210–219 (2002).
369. Takei, Y., Teng, J., Harada, A. & Hirokawa, N. Defects in Axonal Elongation and Neuronal Migration in Mice with Disrupted tau and map1b Genes. *J. Cell Biol.* **150**, 989–1000 (2000).
370. Taniguchi-Watanabe, S. *et al.* Biochemical classification of tauopathies by immunoblot, protein sequence and mass spectrometric analyses of sarkosyl-insoluble and trypsin-resistant tau. *Acta Neuropathol.* **131**, 267–280 (2016).

371. Tardivel, M. *et al.* Tunneling nanotube (TNT)-mediated neuron-to neuron transfer of pathological Tau protein assemblies. *Acta Neuropathol. Commun.* **4**, 117 (2016).
372. The Nobel Foundation. *Press release: The Nobel Prize in Chemistry 2017.* (2017).
373. Thibodeau, M., Chertkow, H. & Léger, G. Une évolution des critères diagnostiques des tauopathies. *Can. Rev. Alzheimer's Dis. Other Dementias* 4–9 (2009).
374. Tholey, A. & Becker, A. Top-down proteomics for the analysis of proteolytic events - Methods, applications and perspectives. *Biochim. Biophys. Acta - Mol. Cell Res.* **1864**, 2191–2199 (2017).
375. Thomas, S. N. *et al.* Dual modification of Alzheimer's disease PHF-tau protein by lysine methylation and ubiquitylation: A mass spectrometry approach. *Acta Neuropathol.* **123**, 105–117 (2012).
376. Togo, T. *et al.* Argyrophilic grain disease is a sporadic 4-repeat tauopathy. *J. Neuropathol. Exp. Neurol.* **61**, 547–556 (2002).
377. Trabzuni, D. *et al.* MAPT expression and splicing is differentially regulated by brain region: Relation to genotype and implication for tauopathies. *Hum. Mol. Genet.* **21**, 4094–4103 (2012).
378. Trinczek, B., Biernat, J., Baumann, K., Mandelkow, E. & Mandelkow, E. M. Domains of tau protein, differential phosphorylation, and dynamic instability of microtubules. *Mol. Biol. Cell* **6**, 1887–1902 (1995).
379. Tsiatsiani, L., Akeroyd, M., Olsthoorn, M. & Heck, A. J. R. Aspergillus Niger Prolyl Endoprotease for Hydrogen-Deuterium Exchange Mass Spectrometry and Protein Structural Studies. *Anal. Chem.* **89**, 7966–7973 (2017).
380. Tuttle, M. D. *et al.* Solid-state NMR structure of a pathogenic fibril of full-length human  $\alpha$ -synuclein. *Nat. Struct. Mol. Biol.* **23**, 409–415 (2016).
381. Uchihara, T. Silver diagnosis in neuropathology: Principles, practice and revised interpretation. *Acta Neuropathol.* **113**, 483–499 (2007).
382. Udeshi, N. D. *et al.* Antibodies to biotin enable large-scale detection of biotinylation sites on proteins. *Nat. Methods* **14**, 1167–1170 (2017).
383. Van Dam, D., Vloeberghs, E., Abramowski, D., Staufenbiel, M. & De Deyn, P. P. APP23 mice as a model of Alzheimer's disease: An example of a transgenic approach to modeling a CNS disorder. *CNS Spectr.* **10**, 207–222 (2005).
384. Van Swieten, J. C. *et al.* The  $\Delta$ K280 mutation in MAP tau favors exon 10 skipping in vivo. *J. Neuropathol. Exp. Neurol.* **66**, 17–25 (2007).
385. Vanderstichele, H. *et al.* Analytical performance and clinical utility of the INNOTEST® PHOSPHO-TAU (181P) assay for discrimination between Alzheimer's disease and dementia with Lewy bodies. *Clin. Chem. Lab. Med.* **44**, 1472–1480 (2006).
386. Vasili, E., Dominguez-Meijide, A. & Outeiro, T. F. Spreading of  $\alpha$ -synuclein and tau: A systematic comparison of the mechanisms involved. *Front. Mol. Neurosci.* **12**, 1–23 (2019).
387. Vlassenko, A. G., Benzinger, T. L. S. & Morris, J. C. PET amyloid-beta imaging in preclinical Alzheimer's disease. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1822**, 370–379 (2012).
388. von Bergen, M. *et al.* Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif (306VQIVYK311) forming beta structure. *Proc. Natl. Acad. Sci.* **97**, 5129–5134 (2000).
389. Von Bergen, M., Barghorn, S., Biernat, J., Mandelkow, E. M. & Mandelkow, E. Tau aggregation is driven by a transition from random coil to beta sheet structure. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1739**, 158–166 (2005).
390. Von Bergen, M. *et al.* Mutations of Tau Protein in Frontotemporal Dementia Promote Aggregation of Paired Helical Filaments by Enhancing Local  $\beta$ -Structure. *J. Biol. Chem.* **276**, 48165–48174 (2001).
391. Von Bergen, M. *et al.* The core of tau-paired helical filaments studied by scanning transmission electron microscopy and limited proteolysis. *Biochemistry* **45**, 6446–6457 (2006).
392. Wang, J. Z., Grundke-Iqbal, I. & Iqbal, K. Glycosylation of microtubule-associated protein tau: an abnormal posttranslational modification in Alzheimer's disease. *Nat. Med.* **2**, 871–5 (1996).
393. Wang, K. K. W., Villalobo, A. & Roufogalis, B. D. Calmodulin-binding proteins as calpain substrates. *Biochem. J.* **262**, 693–706 (1989).
394. Wang, S. S. S., Tobler, S. A., Good, T. A. & Fernandez, E. J. Hydrogen exchange-mass spectrometry analysis of  $\beta$ -amyloid peptide structure. *Biochemistry* **42**, 9507–9514 (2003).
395. Wang, W. *et al.* Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal. Chem.* **75**, 4818–4826 (2003).
396. Wang, Y. *et al.* The release and trans-synaptic transmission of Tau via exosomes. *Mol. Neurodegener.* **12**, 1–25 (2017).
397. Wang, Y., Garg, S., Mandelkow, E. M. & Mandelkow, E. Proteolytic processing of tau. *Biochem. Soc. Trans.* **38**, 955–961 (2010).

398. Wedemeyer, W. J., Welker, E. & Scheraga, H. A. Proline cis-trans isomerization and protein folding. *Biochemistry* **41**, 14637–14644 (2002).
399. Wegmann, S. *et al.* Tau protein liquid–liquid phase separation can initiate tau aggregation. *EMBO J.* **37**, 1–21 (2018).
400. Wei, H. *et al.* Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications. *Drug Discov. Today* **19**, 95–102 (2014).
401. Weingarten, M. D., Lockwood, A. H., Hwo, S. Y. & Kirschner, M. W. A protein factor essential for microtubule assembly. *Proc. Natl. Acad. Sci.* **72**, 1858–1862 (1975).
402. Weismiller, H. A. *et al.* Structural disorder in four-repeat Tau fibrils reveals a new mechanism for barriers to cross-seeding of Tau isoforms. *J. Biol. Chem.* **293**, 17336–17348 (2018).
403. Wemheuer, W. M., Wrede, A. & Schulz-Schaeffer, W. J. Types and strains: Their essential role in understanding protein aggregation in neurodegenerative diseases. *Front. Aging Neurosci.* **9**, 1–13 (2017).
404. Whitehouse, P. *et al.* Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain. *Science (80- )*. **215**, 1237–1239 (1982).
405. Whitwell, J. L. *et al.* Midbrain atrophy is not a biomarker of progressive supranuclear palsy pathology. *Eur. J. Neurol.* **20**, 1417–1422 (2013).
406. Williams, D. R. Tauopathies: Classification and clinical update on neurodegenerative diseases associated with microtubule-associated protein tau. *Internal Medicine Journal* **36**, 652–660 (2006).
407. Williams, D. R. *et al.* Characteristics of two distinct clinical phenotypes in pathologically proven progressive supranuclear palsy: Richardson's syndrome and PSP-parkinsonism. *Brain* **128**, 1247–1258 (2005).
408. Wilson, D. M. & Binder, L. I. Free fatty acids stimulate the polymerization of tau and amyloid  $\beta$  peptides: In vitro evidence for a common effector of pathogenesis in Alzheimer's disease. *Am. J. Pathol.* **150**, 2181–2195 (1997).
409. Wilson, D. M. & Binder, L. I. Polymerization of microtubule-associated protein tau under near-physiological conditions. *J. Biol. Chem.* **270**, 24306–24314 (1995).
410. Wiltzius, J. J. W. *et al.* Molecular mechanisms for protein-encoded inheritance. *Nat. Struct. Mol. Biol.* **16**, 973–978 (2009).
411. Wischik, C. M. *et al.* Isolation of a fragment of tau derived from the core of the paired helical filament of Alzheimer disease. *Proc. Natl. Acad. Sci.* **85**, 4506–4510 (1988).
412. Wood, S. J., Maleeff, B., Hart, T. & Wetzel, R. Physical, morphological and functional differences between pH 5.8 and 7.4 aggregates of the Alzheimer's amyloid peptide A $\beta$ . *J. Mol. Biol.* **256**, 870–877 (1996).
413. Wszolek, Z. K. *et al.* Frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17). *Orphanet J. Rare Dis.* **1**, 1–9 (2006).
414. Wu, G., Huang, H., Abreu, J. G. & He, X. Inhibition of GSK3 phosphorylation of  $\beta$ -catenin via phosphorylated PPPSPXS motifs of Wnt coreceptor LRP6. *PLoS One* **4**, (2009).
415. Xiao, Y. *et al.* A $\beta$ (1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nat. Struct. Mol. Biol.* **22**, 499–505 (2015).
416. Xu, S., Brunden, K. R., Trojanowski, J. Q. & Lee, V. M.-Y. Characterization of tau fibrillization in vitro. *Alzheimer's Dement.* **6**, 110–117 (2010).
417. Yamada, M. *et al.* Senile Dementia of the Neurofibrillary Tangle Type: A Comparison with Alzheimer's Disease. *Dement. Geriatr. Cogn. Disord.* **12**, 117–126 (2001).
418. Yang, L. -S & Ksiazek-Reding, H. Calpain-Induced Proteolysis of Normal Human Tau and Tau Associated with Paired Helical Filaments. *Eur. J. Biochem.* **233**, 9–17 (1995).
419. Zee, B. M. & Garcia, B. A. Discovery of lysine post-translational modifications through mass spectrometric detection. *Essays Biochem.* **52**, 147–163 (2012).
420. Zettlmeissl, G., Rudolph, R. & Jaenicke, R. Limited proteolysis as a tool to study the kinetics of protein folding: Conformational rearrangements in acid-dissociated lactic dehydrogenase as determined by pepsin digestion. *Arch. Biochem. Biophys.* **224**, 161–168 (1983).
421. Zhang, H., Wen, J., Huang, R. Y.-C., Blankenship, R. E. & Gross, M. L. Mass spectrometry-based carboxyl footprinting of proteins: Method evaluation. *Int. J. Mass Spectrom.* **312**, 78–86 (2012).
422. Zhang, W. *et al.* Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife* **8**, 1–24 (2019).
423. Zhang, W. *et al.* Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife* **8**, (2019).
424. Zhang, Y. *et al.* An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–11947 (2014).

425. Zhong, Q., Congdon, E. E., Nagaraja, H. N. & Kuret, J. Tau isoform composition influences rate and extent of filament formation. *J. Biol. Chem.* **287**, 20711–20719 (2012).
426. Zhou, X. Z. *et al.* Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and Tau proteins. *Mol. Cell* **6**, 873–883 (2000).
427. Zhou, Y. *et al.* Relevance of phosphorylation and truncation of tau to the etiopathogenesis of Alzheimer's disease. *Front. Aging Neurosci.* **10**, 1–10 (2018).
428. Zhu, S. *et al.* Hyperphosphorylation of intrinsically disordered tau protein induces an amyloidogenic shift in its conformational ensemble. *PLoS One* **10**, 1–15 (2015).
429. Zhukareva, V. *et al.* Sporadic Pick's disease: A tauopathy characterized by a spectrum of pathological  $\tau$  isoforms in gray and white matter. *Ann. Neurol.* **51**, 730–739 (2002).

## Matériel et méthodes

### Expression et purification de Tau 1N3R et 1N4R recombinante humaine

#### Séquences primaires de Tau 1N3R et 1N4R humaines

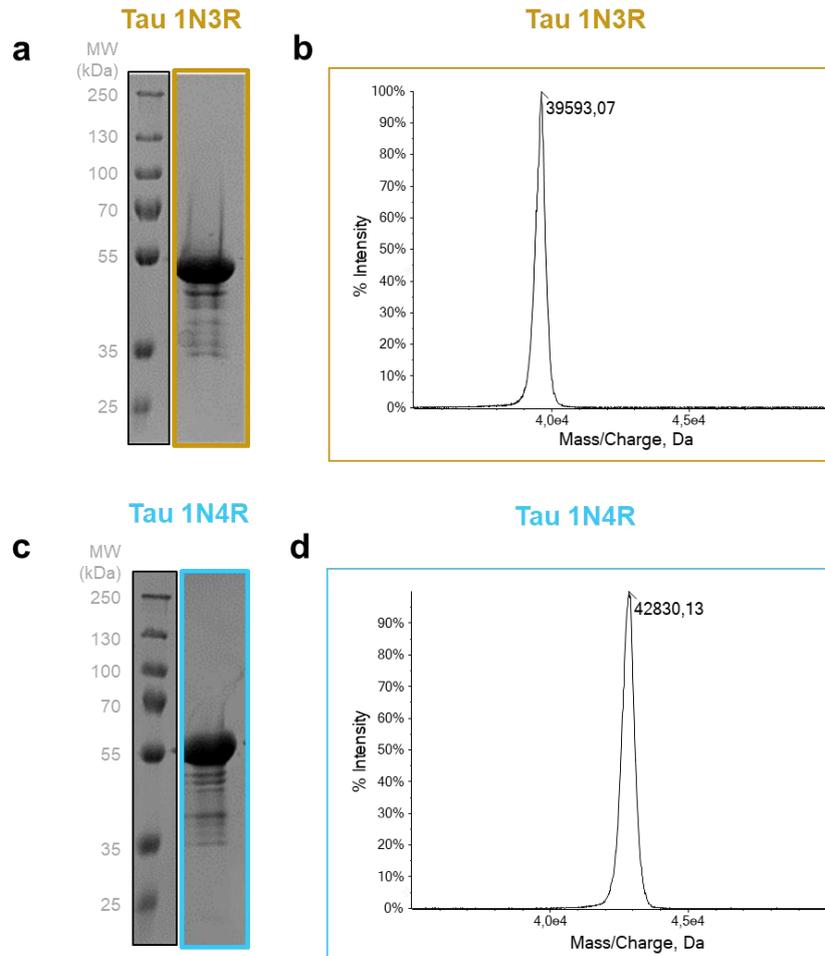
##### Tau 1N3R (381 acides aminés)

<u>10</u>	<u>20</u>	<u>30</u>	<u>40</u>	<u>50</u>	<u>60</u>
MAEPRQEFEV	MEDHAGTYGL	GDRKDQGGYT	MHQDQEGDTD	AGLKESPLQT	PTEDGSEEPG
<u>70</u>	<u>80</u>	<u>90</u>	<u>100</u>	<u>110</u>	<u>120</u>
SETSDAKSTP	TAEAEAEAGIG	DTPSLEDEAA	GHVTQARMVS	KSKDGTGSDD	KKAKGADGKT
<u>130</u>	<u>140</u>	<u>150</u>	<u>160</u>	<u>170</u>	<u>180</u>
KIATPRGAAP	PGQKGQANAT	RIPAKTPPAP	KTPPSSGEPP	KSGDRSGYSS	PGSPGTPGSR
<u>190</u>	<u>200</u>	<u>210</u>	<u>220</u>	<u>230</u>	<u>240</u>
SRTPSLPTPP	TREPKKVAVV	RTPPKSPSSA	KSRLQTAPVP	MPDLKNVSKS	IGSTENLKHQ
<u>250</u>	<u>260</u>	<u>270</u>	<u>280</u>	<u>290</u>	<u>300</u>
PGGGKVQIVY	KPVDLSKVT	KCGSLGNIHH	KPGGGQVEVK	SEKLDKFDKRV	QSKIGSLDNI
<u>310</u>	<u>320</u>	<u>330</u>	<u>340</u>	<u>350</u>	<u>360</u>
THVPGGGNKK	IETHKLTFR	NAKAKTDHGA	EIVYKSPVVS	GDTSPRHLSN	VSSTGSIDMV
<u>370</u>	<u>380</u>				
DSPQLATLAD	EVSASLAKQG	L			

##### Tau 1N4R (412 acides aminés)

<u>10</u>	<u>20</u>	<u>30</u>	<u>40</u>	<u>50</u>	<u>60</u>
MAEPRQEFEV	MEDHAGTYGL	GDRKDQGGYT	MHQDQEGDTD	AGLKESPLQT	PTEDGSEEPG
<u>70</u>	<u>80</u>	<u>90</u>	<u>100</u>	<u>110</u>	<u>120</u>
SETSDAKSTP	TAEAEAEAGIG	DTPSLEDEAA	GHVTQARMVS	KSKDGTGSDD	KKAKGADGKT
<u>130</u>	<u>140</u>	<u>150</u>	<u>160</u>	<u>170</u>	<u>180</u>
KIATPRGAAP	PGQKGQANAT	RIPAKTPPAP	KTPPSSGEPP	KSGDRSGYSS	PGSPGTPGSR
<u>190</u>	<u>200</u>	<u>210</u>	<u>220</u>	<u>230</u>	<u>240</u>
SRTPSLPTPP	TREPKKVAVV	RTPPKSPSSA	KSRLQTAPVP	MPDLKNVSKS	IGSTENLKHQ
<u>250</u>	<u>260</u>	<u>270</u>	<u>280</u>	<u>290</u>	<u>300</u>
PGGGKVQIIN	KKLDLSNVQS	KCGSKDNIKH	VPGGGSVQIV	YKPVDLSKVT	SKCGSLGNIH
<u>310</u>	<u>320</u>	<u>330</u>	<u>340</u>	<u>350</u>	<u>360</u>
HKPGGGQVEV	KSEKLDKFDKDR	VQSKIGSLDN	ITHVPGGGNK	KIETHKLTFR	ENAKAKTDHG
<u>370</u>	<u>380</u>	<u>390</u>	<u>400</u>	<u>410</u>	
AEIVYKSPVV	SGDTSPRHLS	NVSSTGSIDM	VDSPQLATLA	DEVSASLAKQ	GL

Les protéines Tau 1N3R et de Tau 1N4R recombinantes humaines ont été exprimées et purifiées d'après le protocole décrit dans Tardivel *et al.* 2016. On constate que la forme entière de Tau 1N3R et 1N4R est très majoritaire dans la préparation protéique ; quelques produits de dégradation sont observés à l'état de traces (Figure 3 a et c). L'analyse MALDI-MS indique une excision de la méthionine N-terminale (Figure 3 b et d). L'absence de la méthionine initiatrice avait déjà été observée dans la fraction sakosyl-insoluble d'homogénats de cerveaux de patients atteints de la maladie d'Alzheimer (Hasegawa *et al.* 1992).



SDS PAGE de Tau 1N3R (a) et Tau 1N4R (c) purifiées. Analyse MALDI-MS de Tau 1N3R (b) et 1N4R (d) purifiées (MH+(2-381),  $\Delta m=5,18$  Da ; MH+(2-412),  $\Delta m=-3,49$  Da).

### Assemblage de Tau 1N3R et 1N4R

Les protéines Tau 1N3R et de Tau 1N4R recombinantes humaines ont été assemblées en fibres en présence d'héparine au ratio molaire [Tau : Héparine] de 4 :1 en PBS 1X DTT 1mM, sous agitation à 37°C, comme défini précédemment dans Tardivel *et al.* 2016.

La polymérisation a pu être suivie à la fois par lecture de fluorescence de la thioflavine T et par sédimentation suivie par SDS-PAGE (Figure 4).

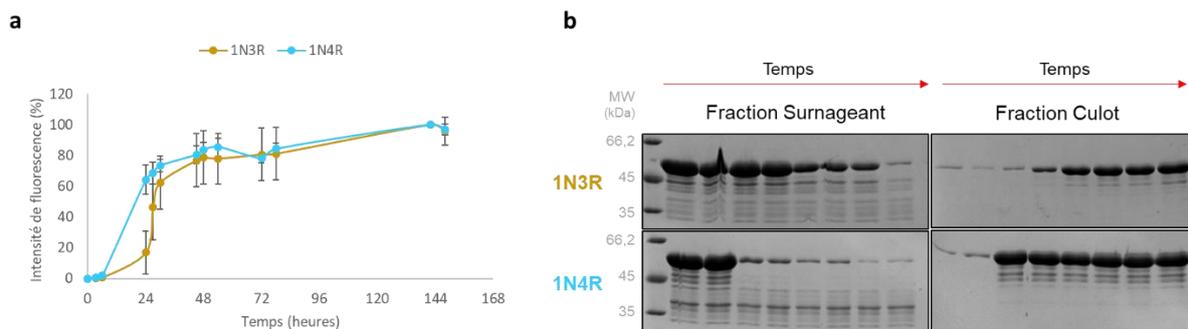
### Suivi de la polymérisation par ultracentrifugation suivie d'un SDS-PAGE

Aux différents points de cinétique, la solution protéique est ultracentrifugée (50000 rpm, 30min, 20°C), le surnageant (formes monomériques) est prélevé et le culot resuspendu dans un volume de PBS 1X DTT 1mM équivalent au volume initial de l'échantillon. Le ratio de protéines

solubles (fraction surnageant) et de fibres (fraction culot) est évalué par SDS-PAGE suivi d'une coloration au Bleu de Coomassie à l'aide d'un ChemiDoc™ MP (BioRad). La quantification est réalisée sur le logiciel sur ImageLab version 5.2.1.

### Suivi de la polymérisation par liaison à la Thioflavine T

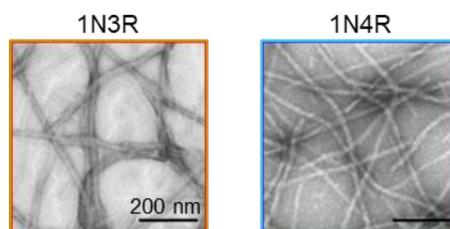
Aux différents points de cinétique, 10µL de la solution protéique est prélevée et mélangée à 400µL d'une solution de Thioflavine T à 10 µM final. L'ensemble est vortexé et une lecture de fluorescence est réalisée ( $\lambda$  Excitation = 420nm ;  $\lambda$  Emission = 480nm) à l'aide d'un spectrofluorimètre Cary Eclipse (Varian Medical Systems Inc).



Cinétique d'assemblage de Tau 1N3R et Tau 1N4R suivie par (a) la fluorescence de la liaison de la Thioflavine T aux fibres et (b) sédimentation par ultracentrifugation suivie d'un gel SDS-PAGE des fractions culot et surnageant.

Comme attendu, on constate que la phase de latence est plus longue pour Tau 1N3R que pour Tau 1N4R et que les deux isoformes ont un rendement d'assemblage de 100% après cinq jours d'incubation.

Les fibres de Tau concentrées par ultracentrifugation sont observées par MET : elles partagent une morphologie de fibres allongées. Aucune différence structurale n'est mise en évidence par MET.



Clichés de MET des fibres de Tau 1N3R et Tau 1N4R après 5 jours de polymérisation (Agent de contraste : Acétate d'uranyle) Grossissement x10000

## Protéolyses ménagées et identification des fragments protéolytiques

### *Protéolyses ménagées*

#### Cinétique de protéolyses ménagées suivie par SDS-PAGE

40µM de fibres de Tau 1N3R et Tau 1N4R sont incubées en présence d'enzyme à différents ratios E/S (w/w) en fonction de l'endoprotéase considérée dans une solution de PBS DTT 1mM à 37°C sous agitation (300 rpm).

Enzymes	Ratio (E/S) w/w
Trypsine	1/7200
Chymotrypsine	1/550
GluC	1/150
AspN	1/120
LysC	1/4800
LysN	1/4800
ArgC	1/480

Table des ratios (E/S)p/p utilisés pour les différentes protéolyses ménagées présentées en Figure 25.

A différents points de la cinétique (2, 10, 45, 120 et 300 minutes), l'échantillon est prélevé : un premier aliquot est immédiatement congelé rapidement à l'azote et stocké à -80°C ; un second aliquot est mélangé volume à volume avec un tampon Laemmli Blue 3X, chauffé 5 minutes à 80°C puis congelé rapidement à l'azote et stocké à -80°C .

5µL d'échantillon en Laemmli Blue 1.5X sont déposés sur un gel d'acrylamide 12%, puis mis à migrer à 200 V pendant 45 minutes. Le gel SDS-PAGE est finalement coloré au Bleu de Coomassie, imagé à l'aide d'un ChemiDoc™ MP (BioRad) et analysé sur ImageLab version 5.2.1.

Le protocole de protéolyse ménagée de Tau 1N3R et de Tau 1N4R monomériques par la GluC est strictement identique à celui appliqué aux fibres.

### Test de sédimentation des fragments protéolytiques

La protéolyse ménagée par la GluC est arrêtée par addition d'AEBSF à la concentration finale de 10mM puis les fragments protéolytiques sont sédimentés par ultracentrifugation pendant 30 minutes à 50000 rpm à 20°C. Les fractions surnageant et culots sont analysées sur un gel SDS-PAGE 12% coloré au Bleu de Coomassie, imagé à l'aide d'un ChemiDoc™ MP (BioRad) et analysé sur ImageLab version 5.2.1.

### Suivi de la protéolyse ménagée des fibres de Tau par la GluC par microscopie électronique à transmission

Après 0, 45 et 300 minutes de protéolyse ménagée des fibres de Tau 1N3R et 1N4R par la GluC, le mélange réactionnel est dilué en Tris 5mM pH 7.5 jusqu'à une concentration finale de 10µM. Cette solution est absorbée sur des grilles de cuivre recouvertes d'un film fin de carbone (200-mesh) qui sont ensuite colorées à l'acétate d'uranyle 1%. Les images sont ensuite acquises grâce à un microscope électronique à transmission JEOL 1400 équipé d'une caméra Gatan Orius CCD (Gatan).

### *Analyse des fragments protéolytiques associant la spectrométrie de masse*

#### Analyse Bottom-up sans marquage par LC-MS/MS des fragments protéolytiques GluC et Chymotrypsine

7µL (20µM, ≈ 5µg) d'échantillon de Tau 1N3R et 1N4R fibrillaire à T0 et aux différents temps de protéolyse ménagée GluC et à T2 minutes de la protéolyse ménagée chymotryptique ont été déposés sur un gel Tris-Glycine de 12% mis à migrer pendant 45 minutes (200V) puis coloré au Bleu de Coomassie. Les bandes ont été excisées (une demi-bande par condition de digestion). Certaines bandes ont ensuite été lavées par trois cycles de lavage (acétonitrile 100% et 25 mM NH<sub>4</sub>HCO<sub>3</sub>). Après réduction (10 mM DTT) et alkylation (55 mM Iodoacetamide) des cystéines, une digestion par la trypsine (Promega Gold; Promega, Madison, WI) en tampon 50 mM NH<sub>4</sub>HCO<sub>3</sub> a été réalisée sur la nuit à 37°C sous agitation (300rpm). Le volume de la solution enzymatique (10 ng/µl) a été adaptée à la taille de la bande protéique (e.g 25 µL d'une solution à 4 ng/µL, (E/S)<sub>p/p</sub> = 1/50). Les peptides ont ensuite été extraits par incubation d'une heure à température ambiante dans 1,2 volumes de 60% CH<sub>3</sub>CN, 0.1% HCOOH suivie d'une incubation de 2 volumes de CH<sub>3</sub>CN 100%. Les peptides extraits ont ensuite été lyophilisés et stockés à -20°C jusqu'à l'analyse MS. Les peptides sont repris dans 20µL de

TFA 0.1% et 2  $\mu$ L ont été analysés par nanoLC-MS (ESI-LTQ-Orbitrap) suivant le protocole publié par Shrivastava *et al.* 2015 (à l'exception d'une modification du gradient : 70 min de 5 à 60% de B). Les données de nanoLC-MS/MS ont été traitées automatiquement par le logiciel Thermo Proteome Discoverer (version 1.3) et le moteur de recherche SEQUEST avec la modification fixe Carbamidométhylation des cystéines et la modification variable Oxydation des méthionines. Seuls les peptides identifiés avec une confiance haute ont été retenus pour établir le pourcentage de recouvrement de séquence. Les fichiers raw ont également été convertis en mgf pour l'analyse MASCOT (spécifications identiques à celles utilisées pour l'analyse Proteome Discoverer) des recouvrements de séquence des différents fragments protéolytiques.

#### Analyse Bottom-up avec marquage spécifique du résidu N-terminal par LC-MS/MS

*Marquage du N-terminal de la protéine en gel suivie d'une digestion totale par la trypsine et d'une analyse par nanoLC-MS/MS*

Le protocole de traitement des échantillons en gel utilisé est le même que celui énoncé précédemment à l'exception de l'étape de marquage et de cycles de lavage supplémentaires réalisés en aval du mrquage comme explicité ci-après. Les résidus N-terminaux ont été marqués par addition d'1,5  $\mu$ L d'une solution de TMPP (N-succinimidylloxycarbonylmethyl)-tris(2,4,6-trimethoxyphenyl)- phosphonium bromide) à différentes concentrations croissantes (0 mM ; 1 mM ; 10 mM et 100 mM) sur la protéine entière (mise au point) puis uniquement à 100 mM pour les fragments protéolytiques en Tris HCl 100 mM pH 8.2 pendant une heure sous agitation (600rpm) à température ambiante. La réaction a été arrêtée par addition de 5  $\mu$ L de 16M hydroxylamine (30 minutes d'incubation) et les bandes ont à nouveau été lavées trois fois (acétonitrile 100% et 25 mM  $\text{NH}_4\text{HCO}_3$ ).

Les peptides tryptiques sont repris dans 20 $\mu$ L de TFA 0.1% et 2  $\mu$ L ont été analysés par nanoLC-MS suivant la même méthode que celle décrite pour l'analyse Bottom-Up sans marquage. Les données de nanoLC-MS/MS ont été traitées automatiquement par le logiciel Thermo Proteome Discoverer (version 1.3) et le Sequest search engine avec la modification fixe Carbamidométhylation des cystéines et les modifications variables suivantes : oxydation des méthionines, marquage des peptides N-ter, des lysines, des sérines, des thréonines et des tyrosines. Les spectres  $\text{MS}^2$  des peptides marqués en N-ter ont été vérifiés manuellement.

### *Analyse en masse entière par MALDI-MS*

Les échantillons fibrillaires à T0, T10 et T300 minutes de protéolyse ménagées par la GluC sont purifiés sur Zip-Tip C18 d'après le protocole classique. Les échantillons sont acidifiés à 1% de TFA final pour inactiver la GluC avant d'être chargé sur la micro-colonne à hauteur de 120 pmol. Le zip-tip est d'abord avec 20 µL de TFA 0,1% avant l'élution par paliers des échantillons protéiques à 10% d'ACN puis 70% d'ACN en TFA 0,1% (3 fois 2µL). Les échantillons purifiés sont mélangés volume à volume avec la matrice (Acide Sinapinique, 10 mg/mL en ACN 50%, TFA 0,1%) et déposés sur une plaque MALDI. L'analyse par spectrométrie de masse a été réalisée à l'aide d'un MALDI-TOF-TOF 5800 (AB Sciex) en mode linéaire classique. Les spectres T2D ont été analysés sur Data Explorer version 4.11. AB SCIEX.

### *Marquage par la NHS-Biotine*

#### *Analyse du marquage sur protéine entière par MALDI-MS*

Après optimisation, le protocole standardisé comprend une sonication rapide (10 coups) pour les échantillons fibrillaires. Les échantillons Contrôle et marqués aux ratios molaires [K/NHS-Biotine] 40/1 et 10/1 sont dilués en en TFA 0.1% à la concentration finale de 4µM tandis que les échantillons marqués aux ratios 2,5/1 et 1/1 sont dilués à la concentration finale de 8 µM avant co-cristallisation en HCCA (70% ACN, 0,1% AF).

Tous les dépôts (0.8µL d'échantillon et 0.8µL matrice) sont réalisés sur une plaque d'argent x396. L'analyse par spectrométrie de masse a été réalisée à l'aide d'un MALDI-TOF 5800 (AB Sciex) en mode linéaire avec une intensité laser de 3000 après calibration externe. Les spectres T2D ont été sujets à lissage gaussien (5 points) sur Data Explorer version 4.11. AB SCIEX.

#### *Gamme de ratios molaires [K/NHS-Biotine] et cinétique de marquage*

Les protéines Tau 1N3R et 1N4R monomérique et fibrillaire ont été incubées dans un solution de Sulfosuccinimidyl biotin (Ref.B5161, Sigma) repris en PBS 1X DTT 1mM à un ratio molaire (K/NHS-Biotine) final de 40/1, 10/1, 2,5/1 et 1/1 pendant une durée définie de 20 secondes, 1 minute, 10 minutes et 30 minutes à température ambiante sans agitation. La réaction a été arrêtée par addition de Tris pH 7.5 à la concentration finale de 40-50 mM. Les échantillons

monomériques ont alors été rapidement congelés par incubation dans l'azote liquide et stockés à  $-80^{\circ}\text{C}$  jusqu'à analyse. Les échantillons fibrillaires sont centrifugés pendant 30 minutes à 50000 rpm à  $20^{\circ}\text{C}$  puis le surnageant est remplacé par une solution de Tris 50 mM et les échantillons sont congelés et stockés suivant le même protocole que celui appliqué aux protéines monomériques.

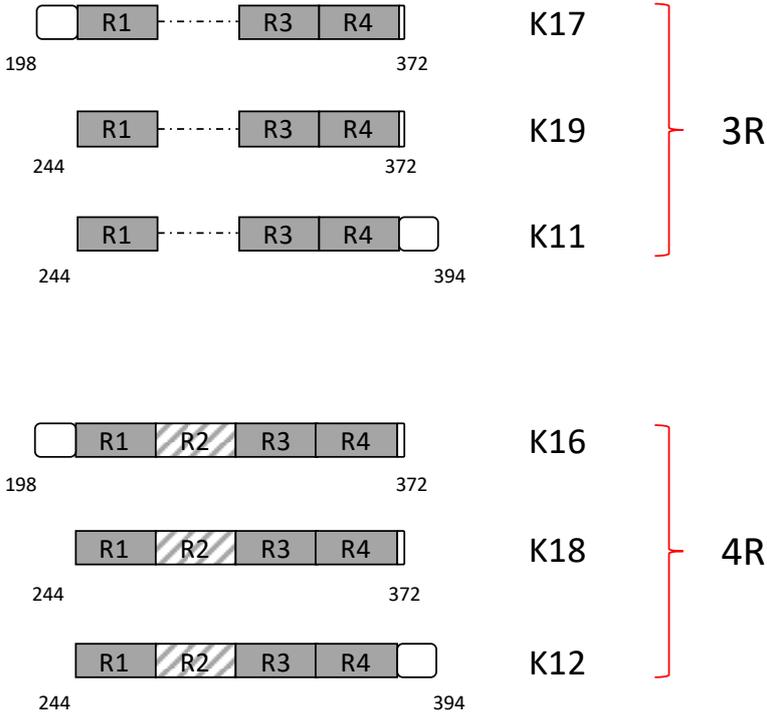
## Annexes

### Annexe 1. Liste des tauopathies actuellement répertoriées

D'après Spillantini and Goedert. 2013.

Alzheimer's disease
Amyotrophic lateral sclerosis and parkinsonism-dementia complex
Argyrophilic grain disease
Chronic traumatic encephalopathy
Corticobasal degeneration
Diffuse neurofibrillary tangles with calcification
Down's syndrome
Familial British dementia
Familial Danish dementia
Frontotemporal dementia and parkinsonism linked to chromosome 17 (caused by MAPT mutations)
Frontotemporal lobar degeneration (some cases caused by C9ORF72 mutations)
Gerstmann-Sträussler-Scheinker disease
Guadeloupean parkinsonism
Myotonic dystrophy
Neurodegeneration with brain iron accumulation
Niemann-Pick disease type C
Non-Guamanian motor neuron disease with neurofibrillary tangles
Primary age-related tauopathy (Tangle-only dementia)
Pick's disease
Postencephalitic parkinsonism
Prion protein cerebral amyloid angiopathy
Progressive subcortical gliosis
Progressive supranuclear palsy
SLC9A6-related mental retardation
Subacute sclerosing panencephalitis
White matter tauopathy with globular glial inclusions

**Annexe 2. Formes tronquées de Tau couramment utilisées en laboratoire**



### Annexe 3. La spectrométrie de masse

La spectrométrie de masse est une technique analytique qui permet de séparer des molécules en phase gazeuse et d'en mesurer de façon extrêmement précise le rapport masse sur charge ( $m/z$ ).

Cette méthode repose sur l'association séquentielle de (1) l'ionisation, (2) l'analyse et (3) la détection des molécules. Vient ensuite le traitement des données spectrales à l'aide d'outils informatiques.



Au cours de ma thèse, j'ai utilisé trois types d'instruments :

- MALDI-TOF
- nanoLC-nano ESI LTQ Orbitrap Velos
- nano LC-nanoESI Triple TOF

Je vais donc brièvement décrire les sources et analyseurs de ces instruments.

#### I. Ionisation des molécules et passage en phase gazeuse

##### *MALDI (Ionisation/Désorption Laser assistée par matrice)*

Cette méthode d'ionisation repose sur la cocrystallisation de l'échantillon, ici protéique ou peptidique, avec une solution de matrice sur un support solide. Le choix de la matrice dépend de la nature de l'échantillon mais aussi du contexte expérimental. Après évaporation, ce réseau cristallin va être irradié par un laser, entraînant ainsi un transfert de charges des molécules de matrice vers les molécules d'intérêt, d'une désorption et d'une accélération des ions dans la source. Les espèces ionisées par transfert de charge sont faiblement chargées. Cependant, l'efficacité d'ionisation dépend de la nature de la matrice et des propriétés physico-chimiques des espèces composant l'échantillon (poids moléculaires, composition en acides aminés).

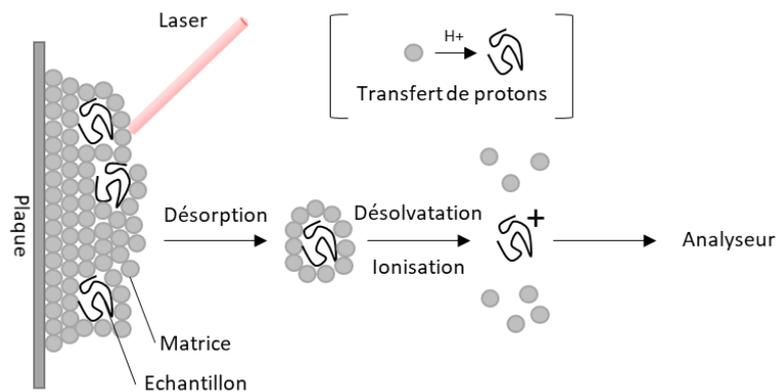


Schéma illustrant l'ionisation / désorption laser assisté par matrice

### *Ionisation par Electronébulisation (ESI)*

Cette méthode d'ionisation repose sur la création d'un champ électrique (grâce à une différence de potentiel élevé entre l'émetteur et la contre-électrode à l'entrée de l'analyseur) appliqué sur l'échantillon en solution dénaturante et acide (H<sub>2</sub>O, ACN, AF). Ainsi, les peptides et protéines sont protonés et généralement multiprotonés en solution. Lorsque le champ électrique est suffisamment élevé, la solution chargée va former des gouttelettes chargées en sortie de l'émetteur pour former le « cône de Taylor ». Lorsque la charge de la solution atteint la limite de Rayleigh, le cône va former des gouttelettes chargées en surface, de plus en plus désolvatées jusqu'à atteindre la limite de stabilité de Rayleigh (forces de répulsion coulombienne supérieures à la tension de surface) et leur fission par explosion coulombienne. Les nanogouttelettes subséquentes vont subir une désolvatation complète par application d'un vide poussé et ainsi entrer dans l'analyseur sous phase gazeuse. Cette méthode d'ionisation occasionne davantage de réarrangements ioniques et la formation d'espèces multichargées contrairement à la méthode d'ionisation dite « MALDI ».

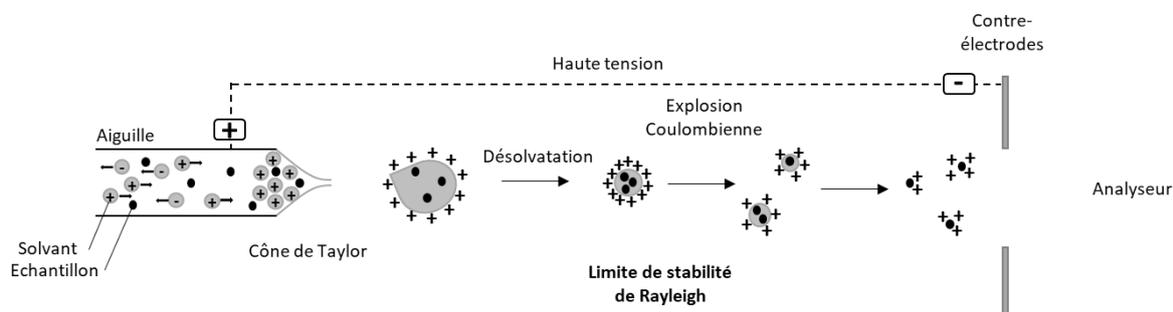


Schéma illustrant l'ionisation par électronébulisation

## II. Analyseurs

### II.1 Analyseurs simples

#### *Analyseur Temps de vol (TOF)*

L'analyseur en temps de vol permet d'établir le rapport masse sur charge, en fonction du temps mis par la molécule ionisée, accélérée par un champ électrique, à atteindre le détecteur. Ainsi, les molécules de plus haut poids moléculaires auront un temps de vol allongé, plus long que les molécules de petit poids moléculaire. En raison de l'hétérogénéité du temps de départ et de la dispersion en énergie cinétique initiale au moment de l'ionisation MALDI, à l'origine d'une dispersion de temps de vol et par conséquent d'un manque de résolution dans la mesure de masse, deux améliorations techniques sont venues compléter l'analyseur initial : l'extraction retardée et le réflectron. L'extraction retardée dans la source permet de refocaliser les ions avant qu'ils ne soient réaccélérés. Le réflectron, dans la seconde partie du temps de vol, permet quant à lui de créer un champ électrique opposé à la trajectoire des ions qui va permettre aux ions de plus haute énergie d'effectuer un trajet plus long, contrebalançant ainsi leur vitesse plus élevée. Ceci permettra aux ions de même  $m/z$  mais dont l'énergie de départ diffère d'arriver en même temps au détecteur et donc de gagner en résolution. Toutefois, le système réflectron n'est pas adapté à l'analyse de molécules de haut poids moléculaires pour lesquelles on lui préférera le mode linéaire.

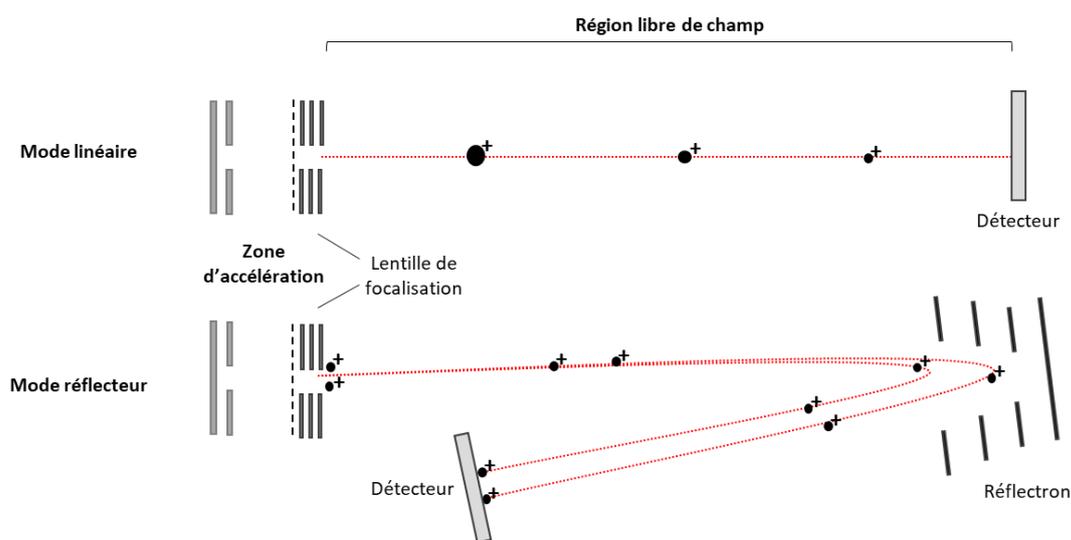


Schéma illustrant le principe de l'analyseur Temps de vol en mode linéaire (haut) et en mode réflectron (bas)

## Analyseur Quadripôle (Q)

Un quadripôle est un analyseur composé de quatre électrodes cylindriques couplées deux à deux et générant ainsi deux champs électriques. Les ions vont ainsi traverser le quadripôle en suivant une trajectoire régie par les équations de Mathieu. La variation de la tension appliquée aura pour résultante une instabilité des ions propre à leur rapport  $m/z$ .

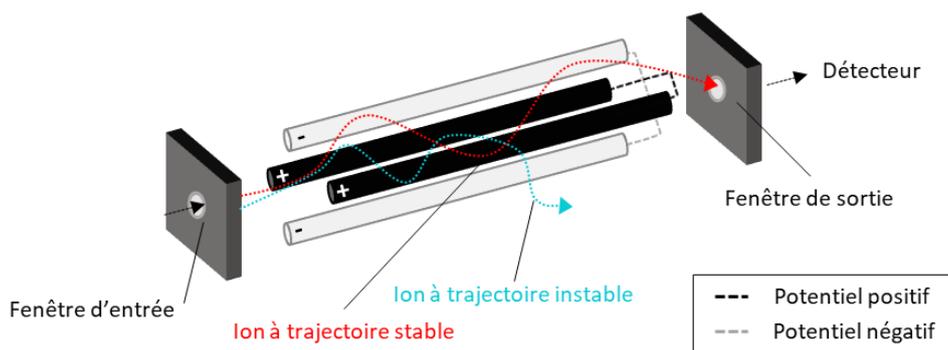


Schéma illustrant le principe de l'analyseur quadripolaire

## II.2 Analyseurs hybrides

### Analyseur hybride Quadripôle-Temps de vol (Q-TOF)

Cet analyseur composite associe deux ou trois quadripôles en série suivis d'un analyseur temps de vol ; il permet de réunir les excellentes performances de séparation et sélection des ions du quadripôle et celles de détection et de résolution du temps de vol. Les ions vont donc être sélectionnés, focalisés et fragmentés dans la partie quadripolaire puis les ions fragments analysés dans l'analyseur temps de vol orthogonal. La géométrie particulière de l'appareil permet une concentration des ions fragments à l'origine de spectres  $MS^2$  fournis (Morris *et al.* 1997).

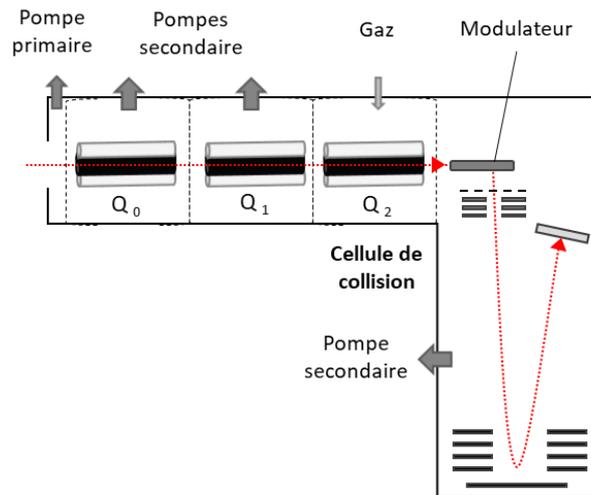


Schéma illustrant le principe de l'analyseur hybride Q-TOF

### Analyseur hybride LTQ-Orbitrap

Le principe de l'Orbitrap repose sur le piégeage et la rotation orbitale des ions autour d'une électrode centrale en forme de fuseau. Les ions oscillent de façon harmonique le long de l'axe avec une fréquence caractéristique de leur rapport masse sur charge. Ce signal est traité par les électrodes extérieures (transformée de Fourier) produisant ainsi un spectre de masse associé. Le principe du LTQ-Orbitrap, technique hybride, est basé sur le couplage entre un LTQ (Linear Trap Quadripole), une C-Trap et l'Orbitrap. Ainsi, le LTQ-Orbitrap combine d'excellentes performances en terme de reproductibilité, de sensibilité et de résolution ; ce qui en fait un instrument de choix en protéomique. Cependant, le principe d'analyse et la nécessité d'accumuler des ions pour obtenir une excellente résolution de masse allonge considérablement le temps d'analyse de l'orbitrap par rapport à d'autres analyseurs (Han *et al.* 2008).

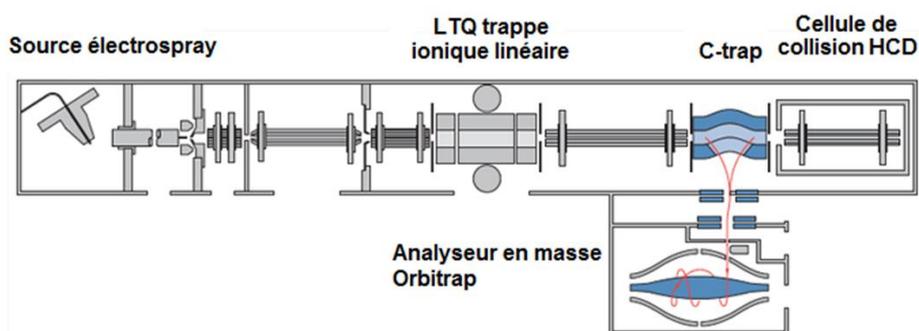
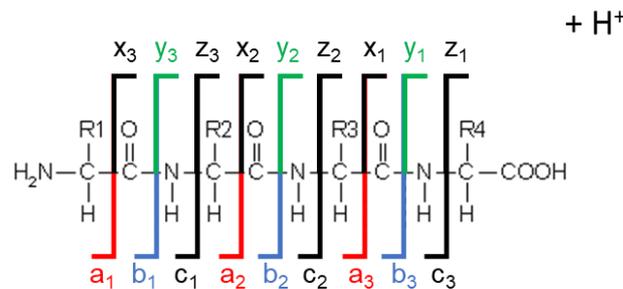


Schéma illustrant le principe du LTQ-Orbitrap. Issu de Boulicault. 2015

#### IV. Fragmentation ou spectrométrie de masse en tandem MS<sup>2</sup>

Le type de fragments obtenus dans un spectre MS/MS dépend de différents facteurs dont font partie la séquence primaire, la charge interne, l'énergie de collision, le mode de fragmentation ou encore la nature de l'éventuel adduit. Les principaux fragments générés sont annotés dans la figure suivante.



Principaux ions fragments générés en mode positif. Issu de Roepstorff and Fohlman. 1984.

Seuls les fragments portant au moins une charge seront détectés. Si la charge est portée par le fragment N-terminal, on obtiendra des ions a, b et c. A l'inverse, si la charge est portée par le fragment C-terminal, on obtiendra des ions x, y et z. D'autres ions fragments internes peuvent être identifiés : les ions amino-acylium, les ions amino-immonium et les ions immonium. Ces derniers ne sont constitués que de la chaîne latérale d'un acide aminé. Comme évoqué précédemment, le type d'ions obtenus dans le spectre MS<sup>2</sup> dépend également du mode de fragmentation. Ainsi, en mode CID (i.e Collision Induced Dissociation), on retrouve essentiellement les séries d'ions b et y, avec possible perte d'ammoniac (-17 Da) ou d'eau (-18 Da). A l'inverse en mode ETD (i.e Electron Transfer Dissociation), les spectres sont peuplés d'ions c, y et z. Certains adduits peuvent également modifier le type d'ions fragments obtenus. Ainsi, lors de la recherche automatique, il est important de considérer l'ensemble de ces paramètres et d'en mesurer l'effet sur les scores des spectres identifiés. Au cours de ma thèse, je n'ai utilisé qu'un seul mode de fragmentation : le mode CID.

L'utilisation de ces différents appareils au cours de mon travail de thèse est brièvement résumée dans le tableau ci-suitant.

Instrument	Applications dans l'étude
MALDI-TOF/TOF 5800 (AB Sciex)	Contrôle qualité des protéines produites Mise au point des conditions expérimentales Evaluation de l'efficacité de biotinylation sur protéine entière
LTQ-Orbitrap Vélos (ThermoFisher Scientific)	Analyse du recouvrement de séquence après digestion des fragments protéolytiques Peptides marqués par le TMPP
NanoESI-Triple TOF 4600 (AB Sciex)	Peptides issus de la protéolyse avancée Peptides marqués par la NHS-Biotine

Annexe 4. Recouvrement de séquence obtenu sur la protéine entière Tau 1N3R et Tau 1N4R après digestion en gel par la trypsine et analyse LC-MS/MS

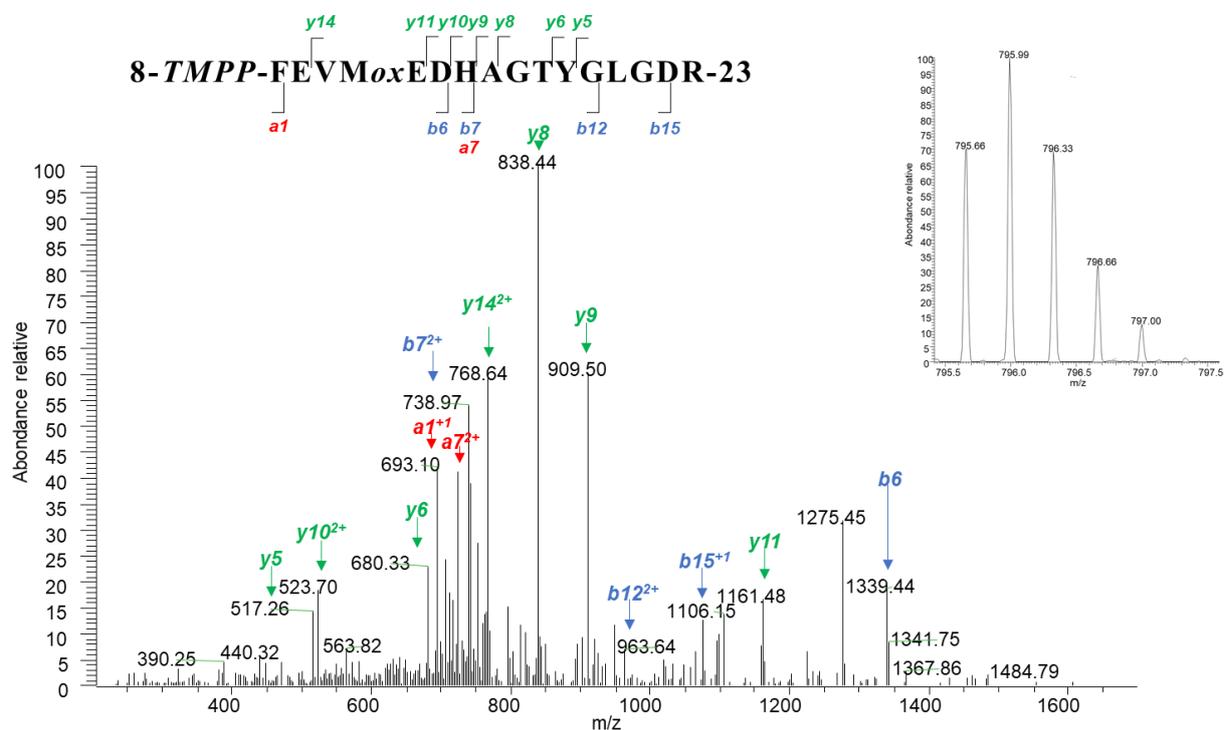
**Tau 1N3R**

10	20	30	40	50	60
MAEPRQEFEV	MEDHAGTYGL	GDRKDQGGYT	MHQDQEGDTD	AGLKESPLQT	PTEDGSEEPG
70	80	90	100	110	120
SETSDAKSTP	TAEAEAEAGIG	DTPSLEDEAA	GHVTQARMVS	KSKDGTGSDD	KKAKGADGKT
130	140	150	160	170	180
KIATPRGAAP	PGQKGQANAT	RIPAKTPPAP	KTPPSSGEPP	KSGDRSGYSS	PGSPGTPGSR
190	200	210	220	230	240
SRTPSLPTPP	TREP <sup>K</sup> VAVV	RTPPKSPSSA	KSRLQTAPVP	MPDL <sup>K</sup> NV <sup>K</sup> SK	IGSTENLKHQ
250	260	270	280	290	300
PGGGKVQIVY	KPVDLSK <sup>V</sup> TS	KCGSLGNIHH	KPGGGQVEVK	SEKLDFKDRV	QSKIGSLDNI
310	320	330	340	350	360
THVPGGGNKK	IE <sup>T</sup> HKL <sup>T</sup> FR	NAKAKTDHGA	EIVYKSPVVS	GDTSPRHLSN	VSSTGSIDMV
370	380				
DSPQLATLAD	EVSASLAK <sup>Q</sup> G	L			

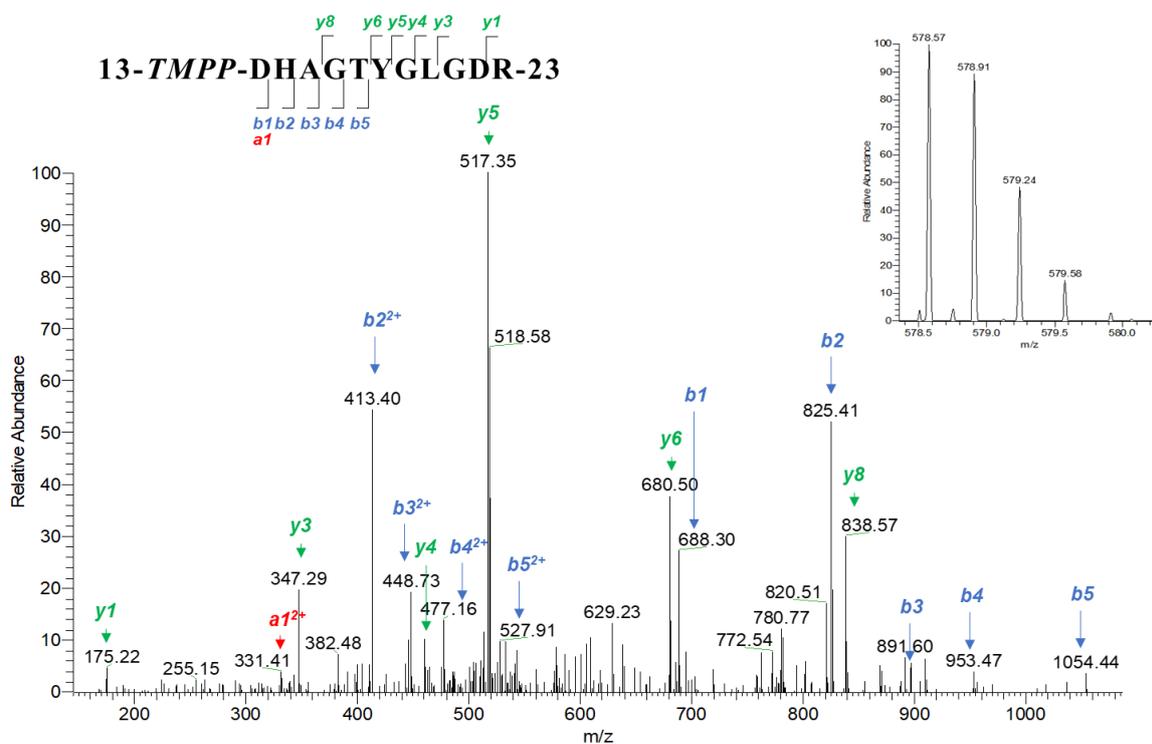
**Tau 1N4R**

10	20	30	40	50	60
MAEPRQEFEV	MEDHAGTYGL	GDRKDQGGYT	MHQDQEGDTD	AGLKESPLQT	PTEDGSEEPG
70	80	90	100	110	120
SETSDAKSTP	TAEAEAEAGIG	DTPSLEDEAA	GHVTQARMVS	KSKDGTGSDD	KKAKGADGKT
130	140	150	160	170	180
KIATPRGAAP	PGQKGQANAT	RIPAKTPPAP	KTPPSSGEPP	KSGDRSGYSS	PGSPGTPGSR
190	200	210	220	230	240
SRTPSLPTPP	TREP <sup>K</sup> VAVV	RTPPKSPSSA	KSRLQTAPVP	MPDL <sup>K</sup> NV <sup>K</sup> SK	IGSTENLKHQ
250	260	270	280	290	300
PGGGKVQIIN	KKLDLSNVQS	KCGSK <sup>D</sup> NIKH	VPGGGSVQIV	YKPVDLSK <sup>V</sup> T	SKCGSLGNIH
310	320	330	340	350	360
HKPGGGQVEV	KSEKLDFKDR	VQSKIGSLDN	ITHVPGGGNK	KIETHKL <sup>T</sup> FR	ENAKAKTDHG
370	380	390	400	410	
AEIVYKSPV	SGDTSPRHLS	NVSSTGSIDM	VDSPQLATLA	DEVSASLAK <sup>Q</sup>	GL

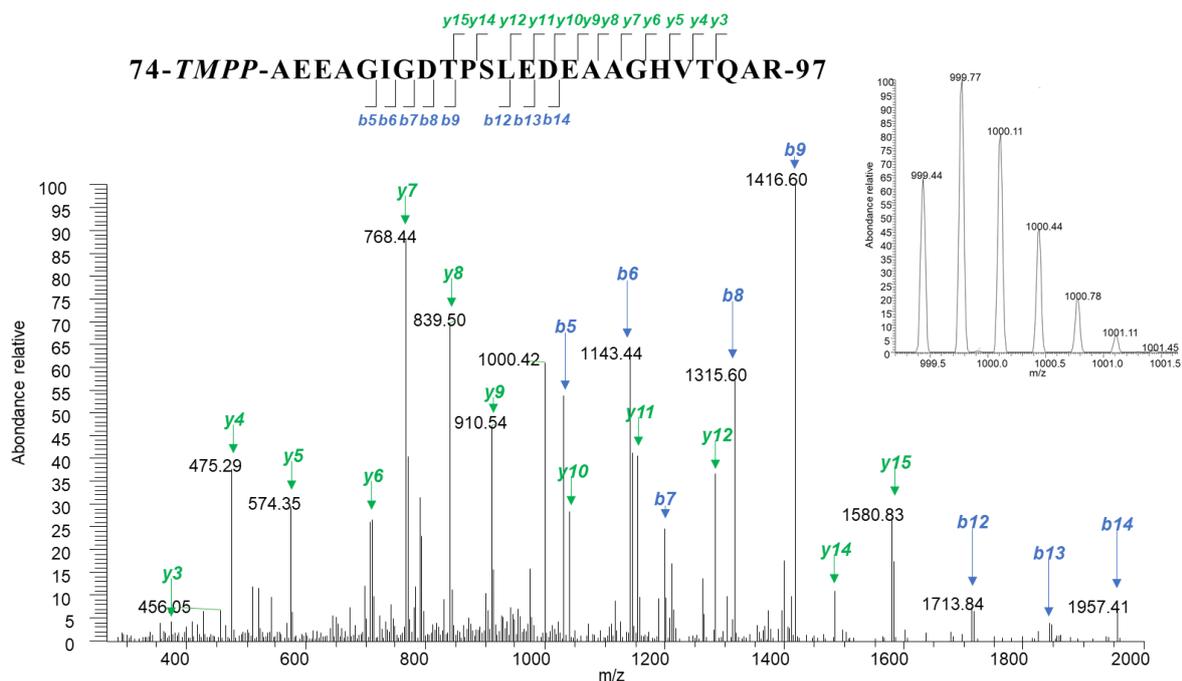
## Annexe 5. Identification du peptide N-terminal dérivatisé au TMPP des fragments protéolytiques



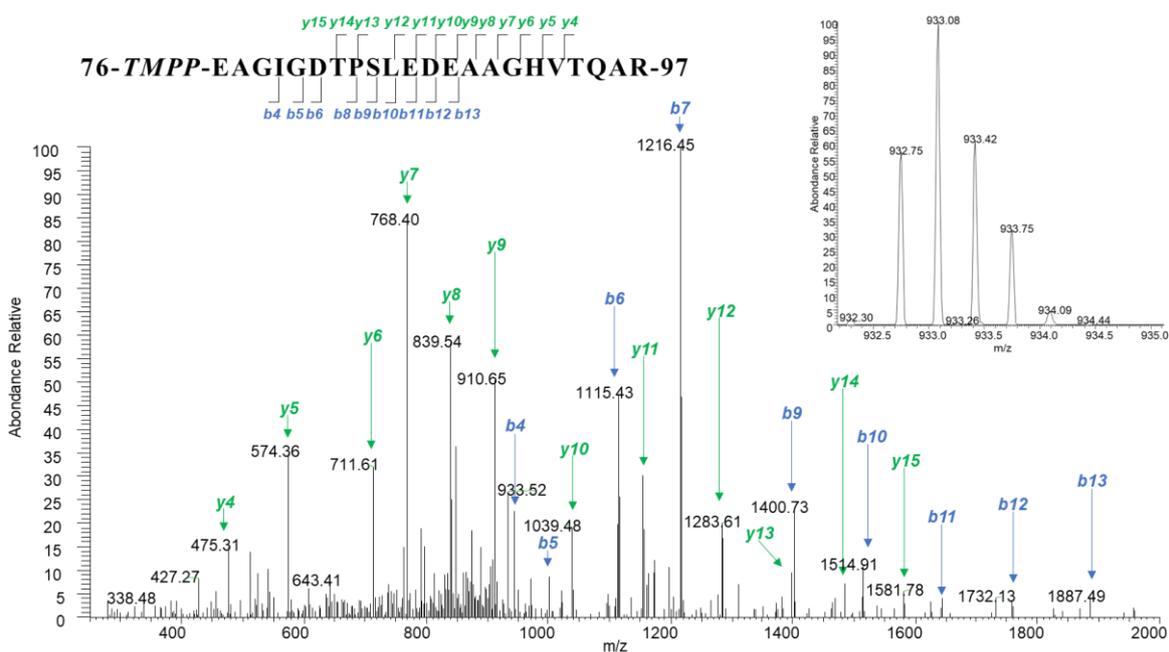
Spectres MS et MS<sup>2</sup> du précurseur MH<sup>3+</sup> du peptide 8-23 oxydé sur la méthionine 11 et marqué au TMPP issu de l'analyse LC-MS/MS (ESI-LTQ-Orbitrap) de la protéolyse ménagée par la GluC des fibres 1N3R et 1N4R Bandes 1



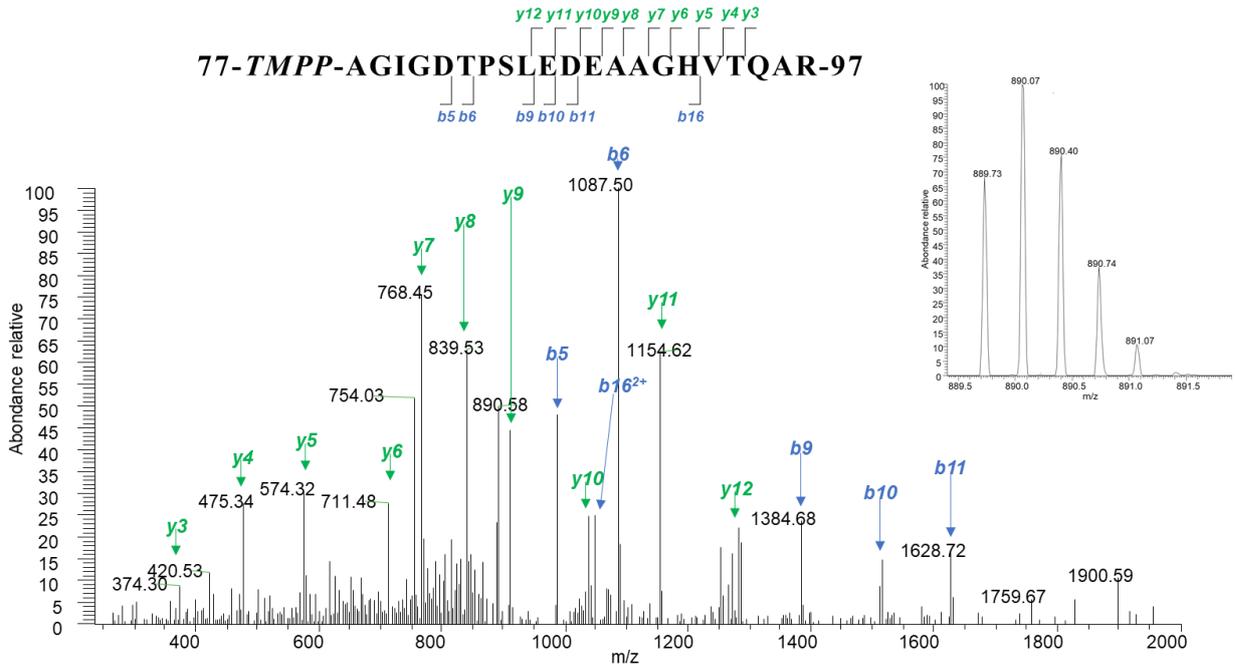
Spectres MS et MS<sup>2</sup> du précurseur MH<sup>3+</sup> du peptide 13-23 marqué au TMPP issu de l'analyse LC-MS/MS (ESI-LTQ-Orbitrap) de la protéolyse ménagée par la GluC des fibres 1N3R et 1N4R Bandes 1



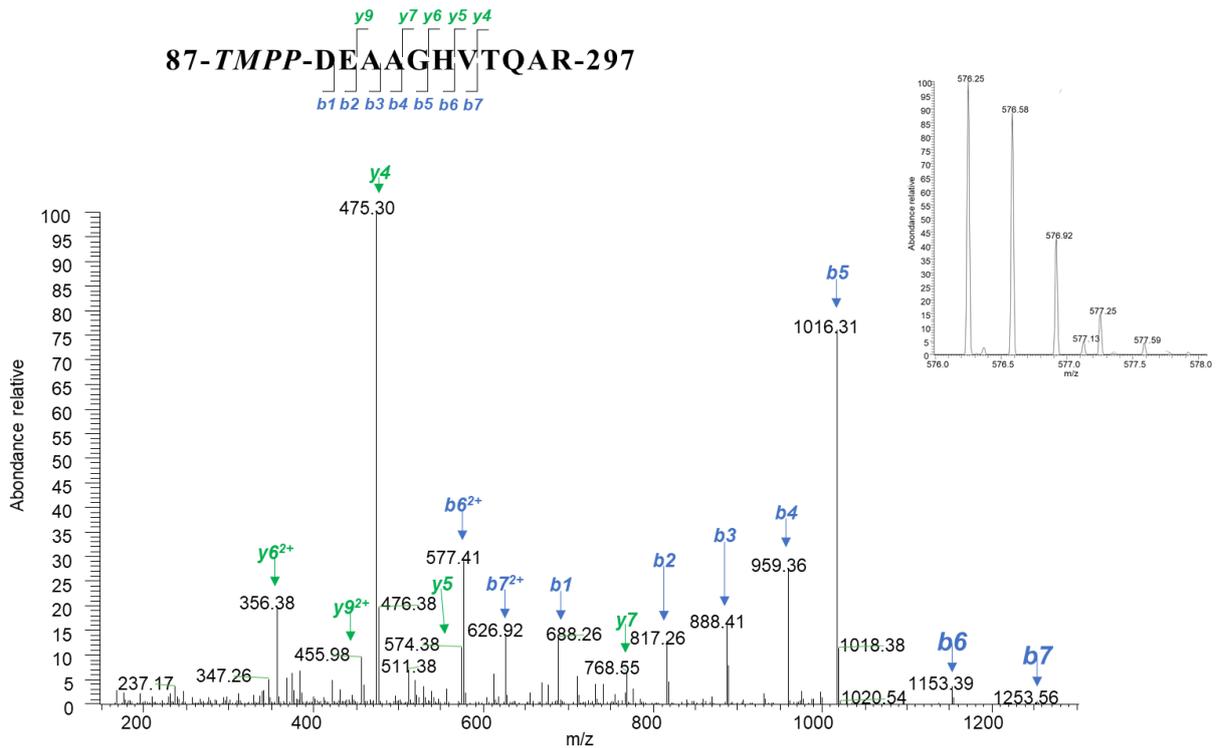
Spectres MS (gauche) et MS<sup>2</sup> (droite) du précurseur MH<sup>3+</sup> du peptide 74-97 marqué au TMPP issu de l'analyse LC-MS/MS (ESI-LTQ-Orbitrap) de la protéolyse ménagée par la GluC des fibres 1N4R Bande 3



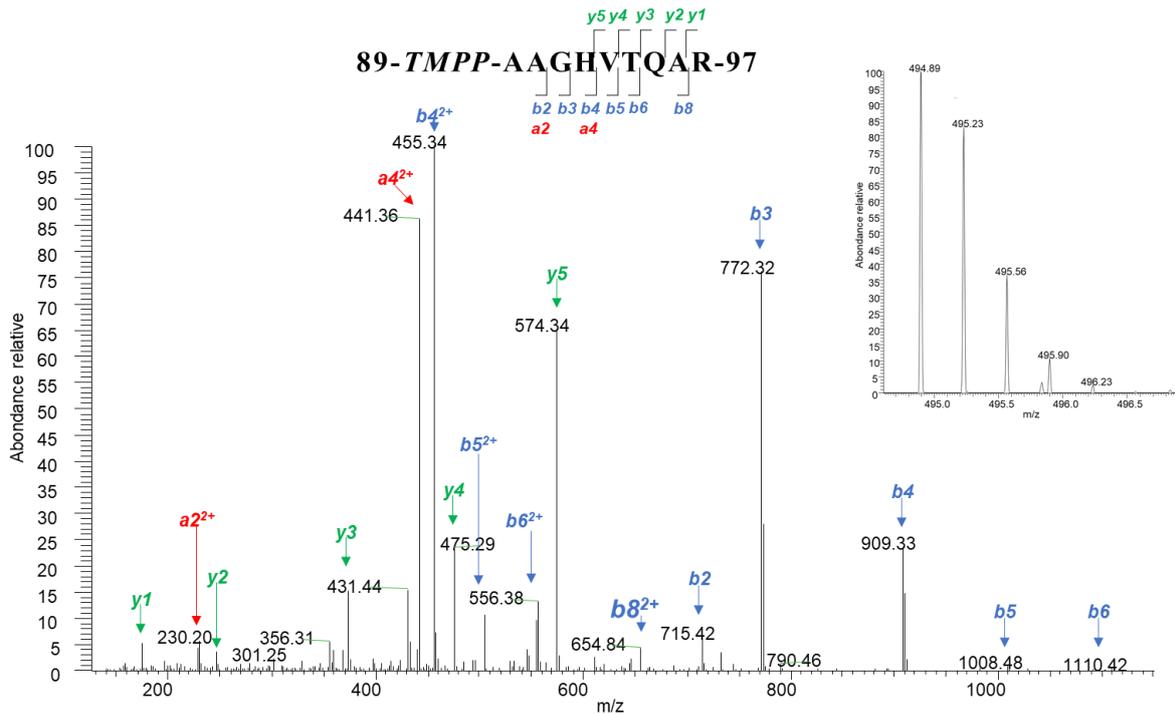
Spectres MS (gauche) et MS<sup>2</sup> (droite) du précurseur MH<sup>3+</sup> du peptide 76-97 marqué au TMPP issu de l'analyse LC-MS/MS (ESI-LTQ-Orbitrap) de la protéolyse ménagée par la GluC des fibres 1N4R Bande 3



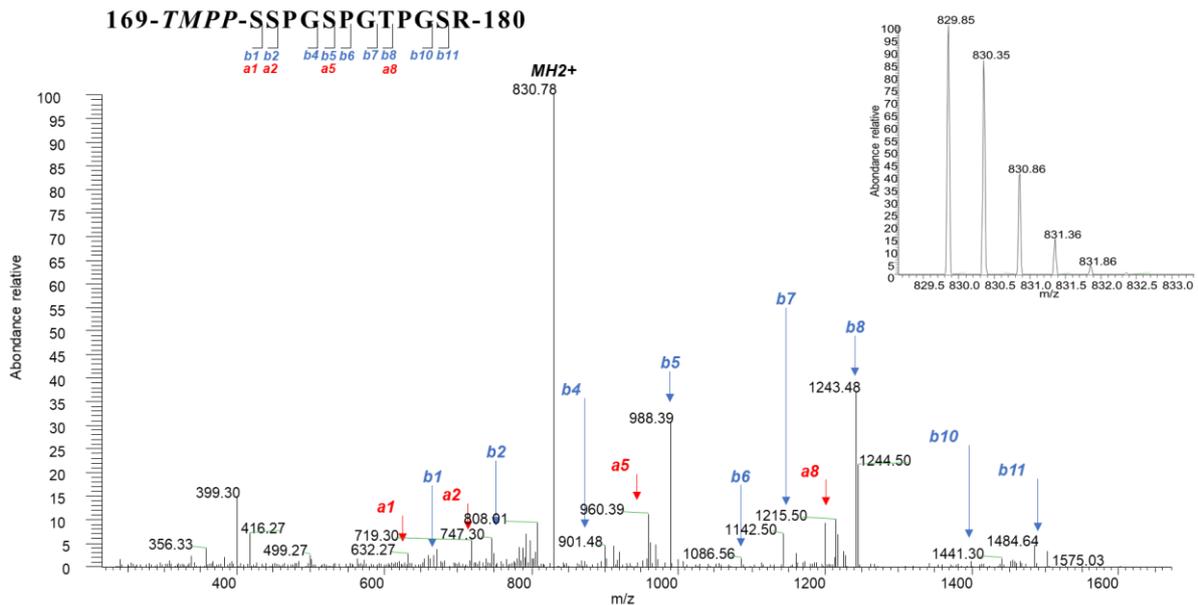
Spectres MS (gauche) et MS<sup>2</sup> (droite) du précurseur MH<sup>3+</sup> du peptide 77-97 marqué au TMPP issu de l'analyse LC-MS/MS (ESI-LTQ-Orbitrap) de la protéolyse ménagée par la GluC des fibres 1N4R Bande 3



Spectres MS (gauche) et MS<sup>2</sup> (droite) du précurseur MH<sup>3+</sup> du peptide 87-97 marqué au TMPP issu de l'analyse LC-MS/MS (ESI-LTQ-Orbitrap) de la protéolyse ménagée par la GluC des fibres 1N4R Bandes 3, 4 et 5



Spectres MS (gauche) et MS<sup>2</sup> (droite) du précurseur MH<sup>3+</sup> du peptide 89-97 marqué au TMPP issu de l'analyse LC-MS/MS (ESI-LTQ-Orbitrap) de la protéolyse ménagée par la GluC des fibres 1N3R Bande 4 et 1N4R Bandes 3, 4 et 5



Spectres MS (gauche) et MS<sup>2</sup> (droite) du précurseur MH<sup>2+</sup> du peptide 169-180 marqué au TMPP issu de l'analyse LC-MS/MS (ESI-LTQ-Orbitrap) de la protéolyse ménagée par la Chymotrypsine des fibres 1N3R Bande 1 et 1N4R Bandes 1, 2 et 3



## Résumé

Les dépôts intracellulaires de la protéine Tau agrégée sont la caractéristique commune des tauopathies, une famille de maladies neurodégénératives dont fait partie la maladie d'Alzheimer. Alors que les isoformes de Tau contenant trois (3R) ou quatre (4R) domaines de liaison aux microtubules sont retrouvées à des niveaux similaires dans le cerveau des individus sains, elles diffèrent au sein des inclusions intracellulaires en fonction des tauopathies. Notre étude repose sur l'identification de déterminants structuraux communs et distincts de fibres de Tau 3R et 4R. Pour cela deux approches de protéomique structurale complémentaires ont été mises au point à partir de fibres de Tau 1N3R et 1N4R produites *in vitro*. La première stratégie, reposant sur l'utilisation de protéolyses ménagées, nous a permis d'identifier les fragments protéolytiques qui composent un « code-barre » moléculaire propre à chaque assemblage. La seconde stratégie a utilisé un marquage chimique covalent des lysines accessibles par la NHS-Biotine suivi de l'analyse qualitative et quantitative des acides aminés marqués par spectrométrie de masse. Nous avons ainsi pu montrer que la partie N-terminale de la protéine était accessible au sein des fibres 1N3R et 1N4R tandis que la région C-terminale de la protéine est protégée pour Tau 1N3R et accessible au solvant pour Tau 1N4R. Nos résultats ouvrent la voie à de nouveaux outils moléculaires pour discriminer les tauopathies.

## Abstract

Intracellular deposits of Tau protein aggregates are the common hallmark of tauopathies, a range of neurodegenerative diseases including Alzheimer's disease. Levels of tau with three (3R) or four (4R) microtubule binding repeats are found similar in the normal adult brain, whereas they differ in neuropathological intracellular Tau inclusions, according to the type of tauopathy. Our study consists of the identification of common and different structural molecular determinants of 3R and 4R Tau fibrils. To this end, two proteomic approaches were optimized using 1N3R and 1N4R recombinant fibrils. The first strategy, using limited proteolysis, allowed us to identify the proteolytic fragments composing the molecular "bar-code" for each type of fibril. The second strategy we optimized used chemical covalent surface labelling of accessible lysines with NHS-Biotin, and qualitative and quantitative analysis of the biotinylated residues using mass spectrometry. We show that, while the N-terminal part of the protein remains accessible within 1N3R and 1N4R fibrils, the C-terminal region is protected within 1N3R yet solvent accessible for 1N4R assemblies. Our results pave the way to new molecular tools to discriminate tauopathies.