

THESIS / THÈSE

MASTER EN SCIENCES CHIMIQUE À FINALITÉ SPÉCIALISÉE EN CHIMIE EN ENTREPRISE

Caractérisation structurale de nouvelles protéines amyloïdogènes intrinsèquement désordonnées DPF3b et DPF3a

LEYDER, Tanguy

Award date:
2023

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Université de Namur
Faculté des Sciences
Laboratoire de Chimie Physique des Biomolécules



Caractérisation structurale de nouvelles protéines amyloïdogènes intrinsèquement désordonnées DPF3b et DPF3a

Mémoire présenté pour l'obtention

du grade académique de Master Chimie «Chimie du Vivant et des Matériaux» : Finalité Spécialisée

Tanguy LEYDER

Janvier 2023

Table des matières

Propos liminaire	1
I. Résumé	2
II. Remerciements	3
III. Liste des abréviations	5
Partie n°1 – Introduction et contextualisation	9
1.1. La protéine à double doigt PHD 3 : DPF3	10
1.1.1. Domaine et structure de la protéine à doigts de zinc DPF3	10
1.1.2. Fonctions biologiques de DPF3	12
1.1.3. Pathologies associées à DPF3	15
1.1.4. Propriétés structurales de DPF3a, une protéine intrinsèquement désordonnée	17
1.2. Protéines intrinsèquement désordonnées	18
1.2.1. Description générale et propriétés structurales	18
1.2.2. Modes d'action et fonctions biologiques	21
1.2.3. Pathologies et voies de ciblage thérapeutique	23
1.3. Protéines amyloïdogènes	25
1.3.1. Fibrillation amyloïde	25
1.3.2. L' α -synucléine dans la maladie de Parkinson	28
1.3.3. La protéine tau et le peptide amyloïde β dans la maladie d'Alzheimer	29
1.3.4. Stratégies de ciblage thérapeutique	30
1.3.5. Influence des ions métalliques sur le phénomène d'agrégation	31
Partie n°2 – Objectifs et stratégies	33
2.1. Objectifs de recherche	34
2.2. Stratégies de production et de purification	34
Partie n°3 – Résultats et discussions	37
Partie n°3.1 – Purification et analyses préliminaires	38
3.1.1. Purification par FPLC	39
3.1.2. Comportement singulier par SDS-PAGE	40
3.1.3. Identification protéique par spectrométrie de masse (LC-ESI-MS/MS)	41
3.1.4. Comportement singulier par spectroscopie d'absorption UV-visible	42
Partie n°3.2 – Caractérisation des isoformes de DPF3 : DPF3b FL et DPF3a FL	44
3.2.1. Prédiction de propriétés associées au désordre	45

3.2.2. Analyses expérimentales de la nature désordonnée des isoformes de DPF3	49
3.2.3. Investigation de la propension à l'agrégation des isoformes de DPF3	54
3.2.4. Morphologie des agrégats des isoformes de DPF3	62
Partie n° 3.3 – Caractérisation de la région C-terminale des isoformes de DPF3	66
3.3.1. Analyses expérimentales de la nature désordonnée des régions C-terminales	67
3.3.2. Investigation de la propension à l'agrégation des régions C-terminales.....	69
3.3.3. Morphologie des agrégats des régions C-terminales	74
Partie n°3.4 – Influence de cations divalents sur l'agrégation de la région C-terminale des isoformes de DPF3.....	77
3.4.1. Influence des cations divalents sur la structure secondaire des régions C-terminales	78
3.4.2. Influence des cations divalents sur la structure tertiaire des régions C-terminales.....	79
3.4.3. Influence des cations divalents sur la cinétique d'agrégation des régions C-terminales	83
3.4.4. Analyse de séquence des régions C-terminales	84
3.4.5. Influence de cations divalents sur la morphologie des agrégats des régions C-terminales.....	85
Partie n°4 – Conclusions et perspectives	88
4.1. Conclusions.....	89
4.2. Perspectives.....	95
Partie n°5 – Matériels et méthodes	98
5.1. Référence et composition des tampons de purification.....	99
5.2. Transformation de souches compétentes de <i>E. coli</i> pour l'expression d'une protéine de fusion marquée à la GST.....	99
5.3. Surexpression hétérologue dans <i>E. coli</i> d'une protéine de fusion marquée à la GST	100
5.4. Lyse par sonication de souches <i>E. coli</i> induites à l'IPTG	101
5.5. Purification et clivage par chromatographie d'affinité (FPLC) d'une protéine de fusion marquée à la GST sur colonnes pré-packées	101
5.6. Analyse par SDS-PAGE d'échantillons protéiques	102
5.7. Reconcentration d'un échantillon protéique à l'aide du PEG-20000.....	103
5.8. Préparation des solutions protéiques contenant des cations divalents	104
5.9. Prédiction des propriétés structurales et analyses bioinformatiques	104
5.10. Détermination de la concentration protéique	105
5.11. Spectrométrie de masse (LC-ESI-MS/MS).....	105
5.12. Dichroïsme circulaire dans l'UV lointain (FCD).....	106
5.13. Spectroscopies de fluorescence intrinsèque (ITF, ITyrF et dbAF)	106

5.14. Fluorescence extrinsèque de la thioflavine T (ThT)	107
5.15. Test de liaison au rouge Congo (RC).....	107
5.16. Diffusion dynamique de la lumière (DLS)	107
5.17. Microscopie électronique à transmission (TEM).....	107
Partie n°6 – Bibliographie	108

Propos liminaire

Caractérisation structurale de nouvelles protéines amyloïdogènes intrinsèquement désordonnées DPF3b et DPF3a

LEYDER Tanguy

I. Résumé

La protéine à doigt de zinc DPF3 est un régulateur épigénétique appartenant au complexe de remodelage de la chromatine BAF. Au-delà de son implication dans des hypertrophies cardiaques et différentes formes de cancer, DPF3 est surexprimée dans les maladies neurodégénératives telles que les maladies de Parkinson et d'Alzheimer. DPF3 existe sous deux isoformes, à savoir DPF3b et DPF3a. Il a déjà été montré que DPF3a est une protéine intrinsèquement désordonnée (IDP) et qu'elle s'agrège spontanément en fibrilles amyloïdes. Les IDPs défient le paradigme structure-fonction classique étant donné qu'elles existent sous la forme d'ensembles conformationnels dynamiques. Cette flexibilité leur permet d'interagir avec de nombreux partenaires différents et d'être multifonctionnelle.

Afin de comprendre leur implication dans les processus de neurodégénérescence, ce projet de recherche s'est focalisé sur la caractérisation des propriétés désordre-agrégation des isoformes de DPF3 ainsi que de leur région C-terminale (C-TERb et C-TERa), qui les différencient, en combinant une série de prédicteurs et de techniques biophysiques. A l'instar de DPF3a, DPF3b est une IDP de par une grande proportion de désordre, des résidus partiellement exposés au solvant et des conformations étendues. DPF3b est aussi capable de s'agréger en fibrilles amyloïdes en suivant un mécanisme d'agrégation similaire mais plus lent que DPF3a. Concernant les régions C-terminales, C-TERb montre une propension à l'agrégation spontanée plus rapide que C-TERa, qui s'avère être plus désordonnée, malgré des morphologies de fibrilles similaires. Aussi, étant donné la prépondérance des métaux dans les maladies neurodégénératives, l'influence de cations métalliques divalents (Cu^{2+} , Mg^{2+} , Ni^{2+} et Zn^{2+}) a été évaluée sur l'agrégation des régions C-terminales. Tous les métaux ralentissent la fibrillation et provoquent la rigidification et la fragmentation des fibrilles de C-TERb. Alors qu'il en va de même pour C-TERa en présence de Ni^{2+} et Cu^{2+} , l'addition de Mg^{2+} et Zn^{2+} respectivement accélère et génère des fibrilles structurellement différentes.

Mémoire de master en Sciences Chimiques à Finalité Spécialisée

Janvier 2023

Promoteur : C. Michaux

Co-promoteur : E. A. Perpète

II. Remerciements

Parce que ces années d'études, qui touchent presque à leur fin, n'auraient pas été les mêmes sans vous, je tenais, par ces quelques lignes, à vous remercier.

J'aimerais, dans un premier temps, remercier ma promotrice Catherine Michaux pour sa confiance, son investissement, sa disponibilité et sa bienveillance. La véritable maman du CPB aura réussi avec le plus grand des succès l'objectif le plus important. On a fondé une véritable famille qui a su se soutenir dans les bons et les mauvais moments. Cette année de mémoire aurait perdu beaucoup de magie sans toi. Je souhaite évidemment remercier très chaleureusement mon co-promoteur Éric Perpète. En quelques mois seulement, tu as su transmettre toutes les valeurs importantes à tes yeux. Que ce soit au Chapitre autour d'une Pater ou dans les tribunes de Bossière, tu as toujours été présent pour nous, les mémorants.

Je ne peux pas écrire ces remerciements sans adresser quelques mots à tous les membres du CPB avec qui je n'ai partagé que de bons souvenirs cette année. A commencer par Virgile, ou le seul homme sur Terre à pouvoir être bourré de dos, qui est toujours là pour me faire perdre aux cartes ou pour m'enlever des points en cas de tricherie. La team IDP ne serait pas au complet sans Blinera, qui, au-delà de ses choix sportifs discutables et de son incapacité à faire une pompe, est toujours présente pour rigoler à mes vidéos. Mais aussi Laurelenn, aka Tata Laulau, avec qui j'ai pu partager un de mes plus grands fou rire après un (trop) long chapitre et une danse TikTok plus qu'approximative. A partir de maintenant, mes repas de midi ne seront plus aussi tristes qu'avant grâce à ma belle boîte à tartines. Je pense aussi à Laurane, Eloïse et Maude avec qui j'aurai pu rigoler un bon nombre de fois. Sans oublier mes collègues mémorants (Jean, Hugo et Ugo) qui ont partagé cette année extraordinaire avec moi. C'était « l'Italie dans mon assiette ». On en aura eu des délires dans ce bureau qui s'est transformé en terrain de basket, foot, tennis ou baseball selon les occasions. Je terminerai cette partie par Cyril, avec qui j'ai pu échanger sur un bon nombre de débats, footballistiques ou autres, et grâce à qui j'ai pu passer une semaine de folie en Normandie. Alors pour tout ça, merci à vous d'être vous !

Je tiens tout particulièrement à remercier mon encadrant qui m'a donné goût au monde de la recherche, aux IDPs et à DPF3. Comme tu m'as dit à mon arrivée, « Garde bien cette idée dans un coin de ta tête ». Je l'ai gardée et j'espère que mon aventure au CPB ne fait que commencer. Alors merci Julien, aka docteur Mignon, aka Jumi, aka Dzumi pour ton temps investi, ta confiance et tout ce que tu m'as transmis cette année. J'en ai débloqué des compétences qu'elles soient professionnelles ou un peu moins. Je n'aurais pas pu espérer un

meilleur encadrant et grâce à toi, chaque journée était accompagnée de ses fous rires que ce soit à propos d'Hasbulla (Hasbi pour les intimes), du CPB que je n'aurai enfreint qu'une seule fois, de Sylvain ou encore de Michel. Le tout autour d'un (ou plusieurs) verre(s) d'Urine ou pendant de longues heures passées en fluo.

Je n'oublie évidemment pas tous les autres étudiants, qui au fil des années, m'ont accompagné dans mes aventures. Je pense notamment à William avec qui j'ai passé d'innombrables bons souvenirs, soirées, fous rires et j'en passe. Mais également à ceux avec qui j'ai traversé ces dernières années remplies de galères. On aura eu raison des études de chimie. Alors merci à Romain, Charles, Baptiste, Maria et tous ceux avec qui j'ai partagé un bout de chemin.

Dans le cadre de ce mémoire, je tiens à également à remercier Corry Charlier pour l'utilisation de la microscopie électronique à transmission. C'était un vrai plaisir de pouvoir observer des bactéries et des nanotubes dans un froid polaire alors qu'il faisait 40 °C à l'extérieur.

Pour terminer, j'aimerais remercier mes parents et mon frère pour leur présence et disponibilité quand j'en avais besoin. Cette aventure n'aurait pas été aussi facile et ne se serait pas passée dans d'aussi bonnes conditions sans vous.

Bossière, le 18 décembre 2022

Tanguy Leyder

III. Liste des abréviations

A des fins de confort de lecture, la liste des abréviations, décrites ci-après, peut être retrouvée condensée au sein d'un dépliant en fin de manuscrit.

α -syn : α -synucléine

A β : peptide amyloïde β

AD : maladie d'Alzheimer (*Alzheimer's disease*)

ADN : acide désoxyribonucléique

AFM-IR : microscopie à force atomique couplée à la spectroscopie infrarouge (*atomic force microscopy-infrared spectroscopy*)

ANS : acide 8-anilino-1-naphtalènesulfonique

APR : région promotrice d'agrégation (*aggregation prone region*)

APS : persulfate d'ammonium (*ammonium persulfate*)

ATP : adénosine triphosphate

BAF : facteurs associés à BRG1 (*BRG1-associated factors*)

BB : tampon de fixation (*binding buffer*)

CD : dichroïsme circulaire (*circular dichroism*)

CPB : laboratoire de Chimie Physique des Biomolécules

C-TER : région C-terminale

C-TERa : région C-terminale de DPF3a

C-TERb : région C-terminale de DPF3b

dbAF : autofluorescence dans le bleu profond (*deep-blue autofluorescence*)

DLS : diffusion dynamique de la lumière (*dynamic light scattering*)

DPF3 : double doigt PHD 3 (*double PHD fingers 3*)

DPR : résidu promoteur de désordre (*disorder-promoting residue*)

DRX : diffraction de rayons X

DSF : fibrille droite double (*double straight fibril*)

EB : tampon d'éluion (*elution buffer*)

FCD : dichroïsme circulaire dans l'UV lointain (*far-UV circular dichroism*)

FRC : fraction de résidus chargés

FI : indice de repliement (*foldedness index*)

FPLC : chromatographie en phase liquide de protéine rapide (*fast protein liquid chromatography*)

FRET : transfert d'énergie par résonance de fluorescence (*fluorescence resonance energy transfer*)

FT : facteur de transcription

GSH : glutathion réduit

GST : glutathion S-transférase

IDP : protéine intrinsèquement désordonnée (*intrinsically disordered protein*)

IDPR : protéine avec des régions intrinsèquement désordonnées (*protein with intrinsically disordered regions*)

IDR : région intrinsèquement désordonnée (*intrinsically disordered region*)

IPTG : isopropyl- β -D-1-thiogalactopyranoside

ITF : fluorescence intrinsèque du tryptophane (*intrinsic tryptophan fluorescence*)

ITyrF : fluorescence intrinsèque de la tyrosine (*intrinsic tyrosine fluorescence*)

LB : tampon de lyse (*lysis buffer*)

LCD : domaine de faible complexité (*low complexity domain*)

LCR : région de faible complexité (*low complexity region*)

LLPS : séparation de phase liquide-liquide (*liquid-liquid phase separation*)

MN : maladie neurodégénérative

MoRF : caractéristique de reconnaissance moléculaire (*molecular recognition feature*)

MSF : fibrille droite multiple (*multiple straight fibril*)

NCD : dichroïsme circulaire dans l'UV proche (*near-UV circular dichroism*)

NLS : signal de localisation nucléaire (*nuclear localisation signal*)

N-TER : région N-terminale de DPF3

OPR : résidu promoteur d'ordre (*order-promoting residue*)

ORDP : protéine ordonnée (*ordered protein*)

PcTS : porphyrine phtalocyanine tétrasulfonate

PD : maladie de Parkinson (*Parkinson's disease*)

PDB : *Protein Data Bank*

PEG : polyéthylène glycol

PHD : homéodomaine végétal (*plant homeodomain*)

PMSF : fluorure de phénylméthylsulfonyl (*phenylmethylsulfonyl fluoride*)

PONDR : prédicteurs de régions désordonnées naturelles (*predictors of natural disordered regions*)

PPI : interaction protéine-protéine (*protein-protein interaction*)

PPID : pourcentage moyen de désordre intrinsèque prédit par résidu (*average per-residue predicted percentage of intrinsic disorder*)

PTM : modification post-traductionnelle (*posttranslational modification*)

RC : rouge Congo

RMN : résonance magnétique nucléaire

SLiM : motif linéaire court (*short linear motif*)

SDS : dodécylsulfate de sodium (*sodium dodecyl sulphate*)

SDS-PAGE : électrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium (*sodium dodecyl sulfate-polyacrylamide gel electrophoresis*)

SF : filament droit (*straight filament*)

STRF : fibrille striée et torsadée en forme de ruban (*striated twisted ribbon fibril*)

SNU : unité de nucléation sphérique (*spherical nucleation unit*)

SSF : fibrille droite simple (*simple straight fibril*)

sw : largeur de fente (*slit width*)

TEM : microscopie électronique à transmission (*transmission electron microscopy*)

TEMED : tétraméthyléthylènediamine

ThT : thioflavine T

TPB : tampon d'élution associé à la protéase TEV (*TEV protease buffer*)

ZnF : doigt de zinc (*zinc finger*)

Partie n°1

—

Introduction et contextualisation

1.1. La protéine à double doigt PHD 3 : DPF3

1.1.1. Domaine et structure de la protéine à doigts de zinc DPF3

DPF3 est une protéine à double doigt de zinc dont l'acronyme signifie plus exactement double doigt PHD (*plant homeodomain*) 3 (*double PHD fingers 3*). Les protéines à doigts de zinc sont répandues chez les humains, animaux, plantes, virus et procaryotes. Originellement, cette protéine était appelée CER-D4. Il s'agit du troisième membre de la famille protéique D4 dont les principaux sites d'expression connus sont la rétine et le cervelet. La famille D4 comprend trois protéines, à savoir NEURO-D4 ou DPF1, UBI-D4 ou DPF2 et CER-D4 ou DPF3. NEURO-D4 a été la première à être mise en évidence dans le cortex cérébral du rat en 1992. Alors que DPF1, majoritairement retrouvée dans le cerveau, exerce un rôle dans le développement des neurones, DPF2 est exprimée de manière ubiquitaire dans les tissus humains et inhibe la différenciation des cellules sanguines.¹

Un élément commun à toutes ces protéines est le domaine D4 qui est un motif en tandem composé de deux doigts de zinc (ZnFs pour *zinc fingers*) de type PHD.²⁻⁴ Il consiste en une hélice α et une paire de deux feuillets β antiparallèles coordonnant quatre ions Zn^{2+} (deux pour chaque domaine PHD) par l'intermédiaire de cystéines et d'histidines. Le terme PHD vient du fait que ces homéodomaines ont été identifiés en premier lieu dans des protéines végétales.⁵⁻⁷ Ces homéodomaines reconnaissent des régions régulatrices de gènes sur lesquelles ils se fixent. Plus précisément, le doigt PHD a été découvert en 1993 sur la protéine végétale HAT3.1 de par ses similarités de séquence avec des domaines de liaison métallique, tel que le doigt RING, possédant une signature typique C3HC4 (trois cystéines-une histidine-quatre cystéines) qui permet la coordination de deux Zn^{2+} . Au niveau séquentiel, ces doigts PHD contiennent environ 60 résidus.

Chez les eucaryotes, les doigts de type PHD jouent un rôle dans le remodelage de la chromatine ainsi que dans les interactions protéine-protéine (PPIs pour *protein-protein interactions*). La première description de leur implication dans la transcription des gènes concerne la reconnaissance de l'histone H3 méthylée au niveau d'une lysine. Depuis, des études ont montré que les ZnFs PHD ont une capacité de lecture des séquences d'histone. Elles soulignent la polyvalence fonctionnelle des domaines PHD en tant que lecteurs épigénétiques qui contrôlent l'expression des gènes par le recrutement de complexes multiprotéiques régulateurs de la chromatine et de facteurs de transcription.⁶

La famille de protéine D4 est caractérisée par une similarité de structure et de séquence élevée des différents domaines structuraux (Fig. 1). A partir de l'extrémité N-terminale, se

trouvent un domaine 2/3, contenant un signal de localisation nucléaire (NLS pour *nuclear localisation signal*), qui est nécessaire à la pénétration de la protéine dans le noyau cellulaire, ainsi qu'un ZnF de type Krüppel (C_2H_2). Quant à l'extrémité C-terminale, y est retrouvée la paire de ZnFs de type PHD (PHD-1 et PHD-2), situés l'un à la suite de l'autre.^{8,9}

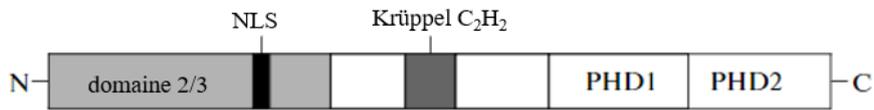


Figure 1 – Représentation schématique de l'organisation des domaines des protéines de la famille D4.⁸

La protéine d'intérêt, DPF3 (entrée UniProt : Q92784), existe sous deux isoformes, à savoir DPF3b (entrée UniProt : Q92784-1), l'isoforme canonique découverte historiquement en premier lieu, et DPF3a (entrée UniProt : Q92784-2). Elles se différencient au niveau de leur composition en acides aminés et au niveau de leur longueur de séquence (Fig. 2). DPF3b est composée de 378 résidus avec une masse de 43.08 kDa, alors que DPF3a possède 357 acides aminés pour une masse moléculaire de 40.24 kDa. Leurs différences proviennent du côté C-terminal étant donné que la composition en acides aminés est identique pour les deux isoformes du premier résidu jusqu'au 292^{ème}. Parmi les domaines communs, se trouvent le domaine 2/3 long de 85 résidus et le domaine en ZnF de type Krüppel (C_2H_2) de 22 acides aminés. Ce domaine, qui contient un Zn^{2+} coordonné par deux cystéines et deux histidines, est caractérisé, à l'instar des doigts PHD, d'une hélice α et de deux feuillets β antiparallèles.

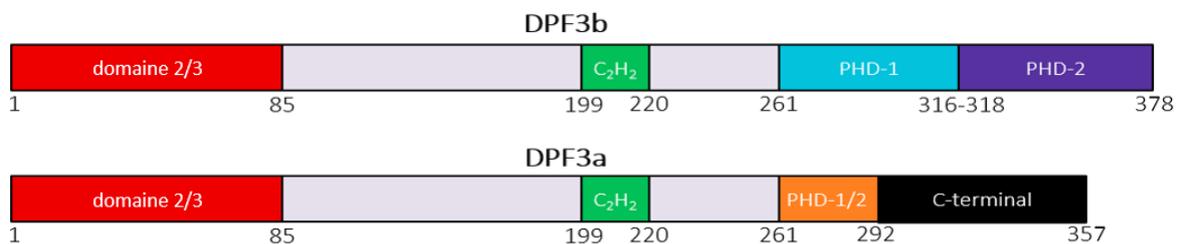


Figure 2 – Séquence et organisation des domaines des isoformes de DPF3.

Alors que DPF3b est composée des deux ZnFs (PHD-1 et PHD-2) typiques de la famille D4, DPF3a ne contient qu'un domaine PHD tronqué (PHD-1/2) suivi d'un domaine C-terminal de fonction inconnue. Les seuls domaines dont la structure a été résolue jusqu'à présent sont les deux ZnFs PHD-1 et PHD-2 de DPF3b (Fig. 3). Il est important de noter que ces données

ont été obtenues dans le cadre d'un complexe avec l'histone H3 acétylée au niveau de la lysine 14 (H3K14ac).¹⁰⁻¹²



Figure 3 – Structures cristallographiques (DRX) du tandem PHD-1 (bleu à gauche) et PHD-2 (mauve à droite) de DPF3b (entrée PDB : 5I3L). Les ions Zn^{2+} coordonnés par les résidus cystéines et histidines (représentés en bâtonnets) sont indiqués par des sphères grises.

1.1.2. Fonctions biologiques de DPF3

Les deux isoformes de DPF3 appartiennent au complexe BAF (pour *BRG1-associated factors*), qui est l'analogue humain et mammalien du complexe protéique SWI/SNF (pour *switch/sucrose non-fermentable*) (Fig. 4). De ce fait, DPF3 est également désignée par BAF45C (pour *BRG1-associated factor 45C*). Les protéines DPF1 et DPF2 sont également présentes dans le complexe BAF et sont respectivement nommées BAF45B et BAF45D. Ledit complexe est un assemblage polymorphique, encodé par 29 gènes et pouvant être constitué jusqu'à 15 sous-unités, qui joue un rôle dans le remodelage de la chromatine, la régulation des gènes, la différenciation cellulaire et le développement de l'organisme.¹³⁻¹⁷

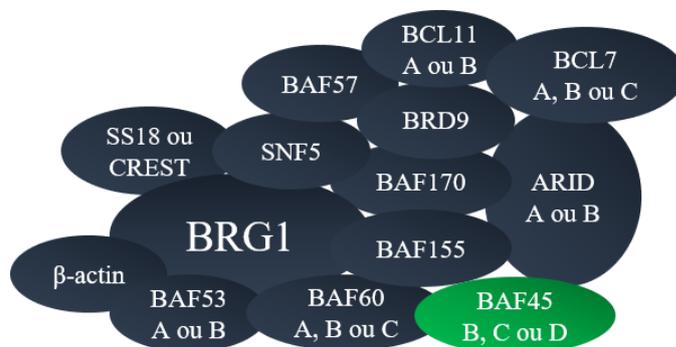


Figure 4 – Représentation schématique du complexe BAF et de ses sous-unités. La sous-unité associée à DPF3 est représentée en vert.¹⁴

La chromatine est une forme compactée de l'acide désoxyribonucléique (ADN). Plus exactement, la double hélice d'ADN s'enroule autour d'octamères d'histone pour former des nucléosomes. Ces derniers se condensent et se rassemblent pour constituer la fibre de chromatine présente dans le noyau des cellules. Il existe cinq membres chez les histones, classés en deux catégories : les histones de cœur (H2A, H2B, H3 et H4) et celles de liaison (H1 et H5). Les premières ont une structure avec un domaine central conservé (*histone fold*) ainsi qu'une queue N-terminale déstructurée sujette à de nombreuses modifications post-traductionnelles (PTMs pour *post-translational modifications*). Par ailleurs, l'histone H1 fait office de lien entre les octamères constitués des histones de cœur, ce qui facilite l'assemblage des nucléosomes. H5 consiste en une variante de H1 présente dans les globules rouges.¹⁸

Lorsque les cellules se divisent, la chromatine se condense davantage en chromosomes. La chromatine sous sa forme la plus condensée est désignée par hétérochromatine et repasse sous sa forme décondensée, l'euchromatine, lors de la transcription de l'ADN. Il existe deux mécanismes majeurs responsables de la réorganisation de la chromatine. Le premier consiste en des modifications épigénétiques comme l'acétylation des lysines présentes sur les queues des histones, provoquant la neutralisation de l'amine chargée et diminuant l'affinité avec l'ADN chargé négativement. L'acétylation et la désacétylation favorisent respectivement la formation de l'euchromatine et de l'hétérochromatine.^{19,20} Le second mécanisme exploite l'utilisation de complexes de remodelage de la chromatine, tel que le complexe BAF, dépendant de l'adénosine triphosphate (ATP) dont l'énergie libérée lors de l'hydrolyse de cette dernière sert à perturber les interactions ADN-histones afin d'augmenter l'accessibilité de l'ADN.^{17,21}

Dans ce contexte, DPF3b agit en tant que régulateur épigénétique de la transcription en se liant aux lysines acétylées et méthylées des histones H3. En effet, contrairement à DPF3a, l'isoforme DPF3b possède les domaines PHD-1 et PHD-2 capables de reconnaître les PTMs sur les lysines (Fig. 5). Plus exactement, le domaine PHD-1 reconnaît la lysine 14 acétylée sur l'histone H3 (H3K14ac), ce qui permet le recrutement du complexe BAF et bloque la transcription. Le ZnF PHD-2 reconnaît la méthylation de la lysine 4 (H3K4me3) menant à la dissociation du complexe et à l'association du complexe de la machinerie transcriptionnelle (CMT), ce qui conduit à l'initiation de la transcription génique. Par conséquent, l'interaction avec l'histone H3, modulée par différentes modifications spécifiques, est cruciale pour la fonction de DPF3b dans la régulation de la transcription des gènes, notamment au cours du développement des cellules musculaires du cœur et du squelette.²²

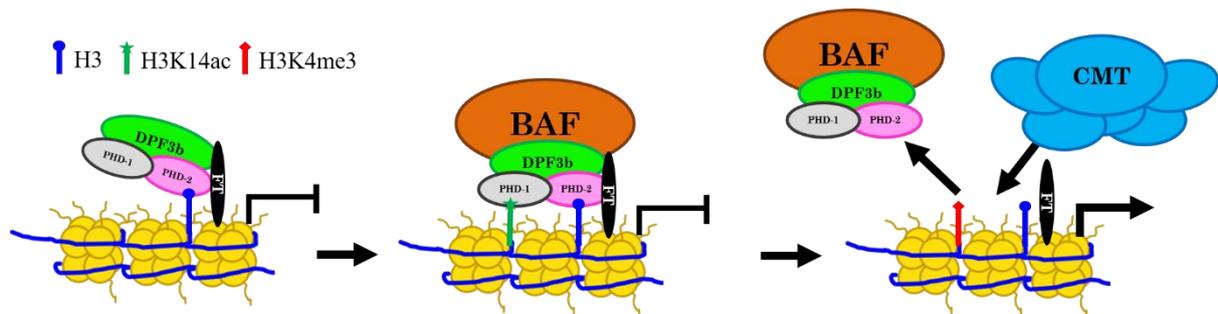


Figure 5 – Représentation schématique de la modulation de la liaison des domaines PHD-1 (gris) et PHD-2 (rose) de DPF3b (vert) au sein du complexe BAF (orange) à l'histone H3 par l'acétylation de la lysine 14 (H3K14ac) et la méthylation de la lysine 4 (H3K4me3). FT (noir) représente un facteur de transcription spécifique à un gène et CMT (bleu) correspond au complexe de la machinerie transcriptionnelle.²²

Récemment, une nouvelle fonction de DPF3, indépendante du complexe BAF, a été identifiée. L'équipe du Dr. Denis Mottet de la Grappe Interdisciplinaire de Génoprotéomique Appliquée (GIGA, Université de Liège) a montré l'implication de DPF3 en tant que nouveau régulateur de la mitose de par l'émergence d'erreurs mitotiques à la suite de la déplétion des deux isoformes de DPF3 (Fig. 6). Cela a été mis en évidence par des protéines fluorescentes spécifiques aux histones H2 (rouge) et à l'actine (vert).

Dans des conditions normales, autrement dit quand DPF3 est exprimée, la mitose commence par la prophase au cours de laquelle les chromosomes se condensent et le fuseau mitotique se construit. Lors de la prométaphase, l'enveloppe nucléaire se rompt et les chromosomes s'attachent au fuseau mitotique. La métaphase correspond à l'alignement des chromosomes sur la plaque équatoriale. Lors de l'anaphase, chaque chromosome se sépare en deux chromatides-sœurs qui vont migrer vers les pôles. Ceux-ci deviennent alors deux nouveaux centrosomes indépendants. Enfin, lors de la télophase, les enveloppes nucléaires se reforment autour des deux nouveaux noyaux.²³ A la suite de la déplétion de DPF3, dès les 20 premières minutes, au moment de la métaphase, une aberration d'alignement des chromosomes au niveau de la plaque équatoriale est observée. Aussi, la non-ségrégation et les erreurs de séparation des chromosomes au cours de l'anaphase et de la télophase, au-delà de 300 minutes, provoquent une réponse apoptotique, c'est-à-dire la mort cellulaire. Ces résultats préliminaires révèlent le caractère essentiel de DPF3 comme régulateur de la division cellulaire.

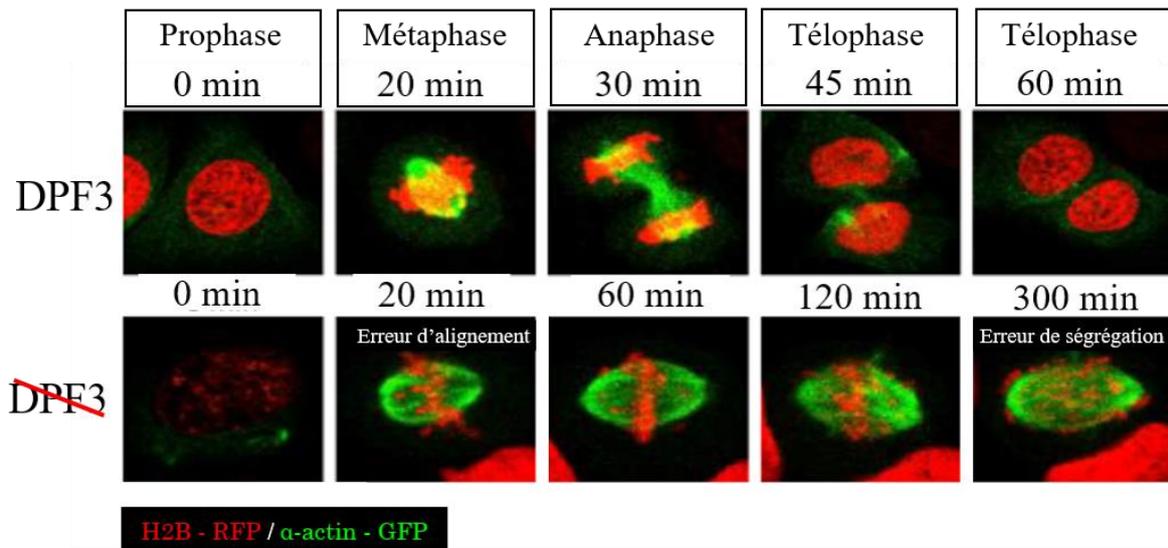


Figure 6 – Imagerie par immunofluorescence du cycle mitotique à la suite de la déplétion d'une cellule des deux isoformes de DPF3 (en bas) par rapport à une cellule de contrôle non déplétée (en haut). Des erreurs d'alignement et de ségrégation apparaissent dans les cellules déplétées. Les histones H2 et l'actine sont respectivement marquées à la RFP (red fluorescent protein) et à la GFP (green fluorescent protein).

1.1.3. Pathologies associées à DPF3

Du point de vue pathologique, DPF3 est retrouvée impliquée dans diverses maladies humaines. En effet, DPF3 a été identifiée dans différentes formes de cancer et, notamment, sa régulation négative est associée à la prolifération et à la motilité des cellules cancéreuses du sein.^{24,25} A l'inverse, dans la leucémie lymphocytaire chronique, un cancer du sang et de la moelle osseuse, la régulation positive de DPF3 entraîne la prolifération des cellules malignes.²⁶ DPF3 maintient également le caractère souche des cellules initiatrices de gliomes, ce qui contribue à la résistance aux médicaments, à la récurrence et à la croissance tumorale dans les glioblastomes. Il s'agit de la tumeur cérébrale maligne la plus fréquente chez l'adulte.²⁷ En outre, DPF3 est un marqueur pronostique dans le cancer du poumon et la maladie pulmonaire obstructive chronique, une obstruction progressive du flux d'air dans les poumons responsable d'une mortalité élevée dans le monde.²⁸ Plus récemment, DPF3 s'est révélée particulièrement essentielle aux voies de signalisation impliquées dans l'oncogenèse du carcinome des cellules rénales (RCC pour *renal cell carcinoma*) en inhibant l'apoptose ainsi qu'en favorisant la croissance cellulaire et la formation de métastases. Chaque année, plus de 300000 nouveaux cas de RCC sont diagnostiqués. Il a été découvert que la région C-terminale de DPF3a interagit avec la région intrinsèquement désordonnée (IDR pour *intrinsically disordered region*) N-terminale de la protéine SNIP1 dans le RCC à cellules claires.²⁹⁻³¹ DPF3 jouant un rôle dans la régulation de la division cellulaire, cela fait de cette protéine une alternative prometteuse

pour le développement d'antimitotiques et, plus exactement, en tant que nouvelle cible thérapeutique dans la lutte contre le cancer.

Etant donné le rôle de régulateur dans la différenciation et le développement des cellules musculaires du cœur, DPF3 est surexprimée chez les personnes atteintes d'hypertrophie cardiaque à l'image de la Tétralogie de Fallot. Celle-ci se manifeste par une hypertrophie du myocarde ventriculaire droit, résultant en la circulation d'un sang peu oxygéné. Elle se produit dans trois cas sur 10000 naissances et représente 7 à 10 % de toutes les malformations cardiaques congénitales. Le développement du cœur au stade embryonnaire nécessite l'expression de gènes cardiaques qui ne doivent plus l'être une fois l'organe formé. La dérégulation desdits gènes, causée par des mutations et des modifications épigénétiques, peut conduire à l'apparition de maladies cardiaques congénitales.

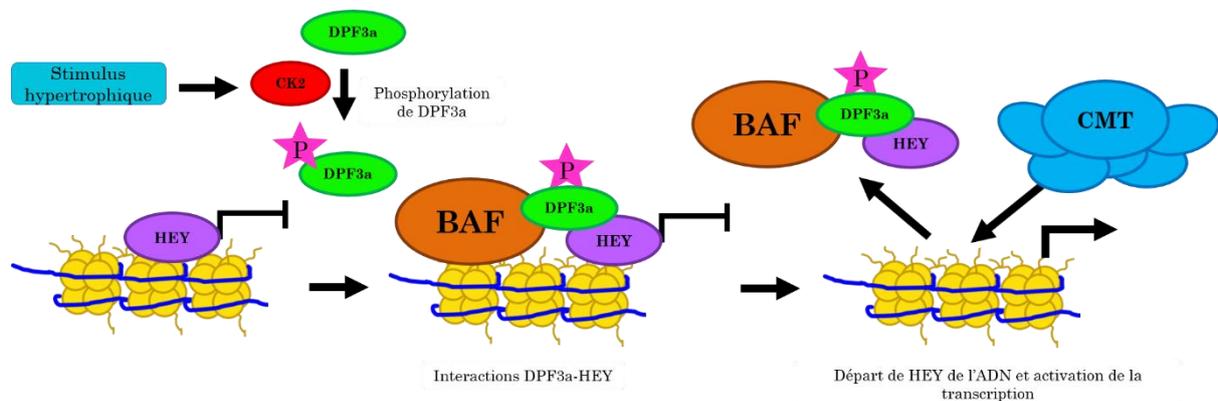


Figure 7 – Représentation schématique de la régulation des gènes cibles de HEY. Sous un stimulus hypertrophique, la caséine kinase 2 (CK2 en rouge) phosphoryle (rose) DPF3a (vert) conduisant au recrutement du complexe BAF (orange) et à l'interaction avec les répresseurs de transcription HEY (mauve). DPF3a libère alors HEY de ses cibles géniques et la transcription est activée par le complexe de la machinerie transcriptionnelle CMT (bleu).³³

DPF3a, bien qu'incapable de se lier aux histones modifiées en raison de son tandem PHD incomplet, est également impliquée dans la différenciation myogénique *via* le complexe BAF. En effet, elle interagit avec un autre régulateur épigénétique, la protéine HRP2 (pour *hepatoma-derived growth factor-related protein 2*), qui va lier les histones modifiées (H3K36me2), conduisant à la transcription du gène myogénique.³² Dans ce contexte, sous un stimulus hypertrophique, la caséine kinase 2 phosphoryle DPF3a au niveau de sa sérine 348. Cette PTM amorce l'interaction de DPF3a avec les répresseurs transcriptionnels HEY (*HES-related repressor protein*), ce qui est suivi par la libération de HEY de l'ADN (Fig. 7). Le complexe BAF étant lié à DPF3a, il est alors recruté par les cibles génomiques de HEY lors de l'interaction entre les deux partenaires, entraînant la transcription de gènes de différenciation

cardiaque et, par conséquent, une hypertrophie cardiaque. En résumé, la phosphorylation de DPF3a lors de stimuli hypertrophiques réactive l'expression des gènes cardiaques fœtaux silencés par l'intermédiaire du complexe BAF.³³⁻³⁵

Par ailleurs, DPF3 est également fortement associée à l'infertilité masculine, conduisant à des altérations de la qualité du sperme et des risques d'azoospermie et d'oligozoospermie.^{36,37} DPF3 est aussi impliquée dans les anomalies dentaires, notamment les fentes labiales et palatines.³⁸ La maladie de Hirschsprung est une malformation gastro-intestinale congénitale représentative provoquant l'absence de cellules ganglionnaires des plexus myentériques faisant partie du système nerveux entérique s'occupant de la digestion. Il a été montré que DPF3 est surexprimée dans différents segments du côlon et participe à la maturation et à la différenciation de ces cellules ganglionnaires.³⁹

Enfin, il a été montré par un *screening* de gènes que DPF3 est aussi impliquée dans les maladies neurodégénératives, en particulier les maladies de Parkinson (PD pour *Parkinson's disease*) et d'Alzheimer (AD pour *Alzheimer's disease*). En effet, cette protéine contribue au développement de la neurodégénérescence et est surexprimé dans l'AD.^{40,41} Des niveaux d'expression élevés de DPF3 ont récemment été détectés dans des groupes de neurones qui sont impliqués dans les dommages cellulaires associés à la PD.⁴²

1.1.4. Propriétés structurales de DPF3a, une protéine intrinsèquement désordonnée

Des prédictions basées sur la séquence ont dévoilé que la grande majorité des sous-unités du complexe BAF sont des protéines intrinsèquement désordonnées (IDPs pour *intrinsically disordered proteins*).¹⁷ La nature désordonnée de DPF3a a dès lors été investiguée au sein du laboratoire de Chimie Physique des Biomolécules (CPB). Selon les prédicteurs de désordre, DPF3a possède un pourcentage de désordre intrinsèque prédit par résidu (PPID pour *predicted per-residue percentage of intrinsic disorder*) de ~ 62 %. Ce pourcentage élevé s'explique par les trois IDRs de la protéine. La première entre les domaines 2/3 et C₂H₂ faisant 114 résidus de long, la deuxième entre les domaines C₂H₂ et PHD-1/2 longue de 41 acides aminés et la troisième, de 65 résidus, du côté C-terminal. Expérimentalement, DPF3a a révélé des signatures typiques des IDPs obtenues par différentes techniques spectroscopiques. Les IDPs ayant tendance à s'agréger, il a également été mis en évidence que l'isoforme DPF3a est susceptible de s'agréger spontanément en oligomères de haut ordre (*high-order oligomers*) ainsi qu'en structures fibrillaires de type amyloïde riches en feuillets β.⁴³

1.2. Protéines intrinsèquement désordonnées

1.2.1. Description générale et propriétés structurales

Depuis la fin des années 1990, un changement radical s'est opéré dans la définition des relations structure-fonction des protéines. En effet, alors que les protéines répondaient au paradigme classique « à une séquence correspond une structure déterminant une fonction », une nouvelle classe de protéines a été identifiée et est apparue essentielle dans la compréhension des processus biologiques. Il s'agit des IDPs qui obéissent au paradigme récent « désordre-fonctions ». Contrairement aux protéines globulaires devant se replier dans la conformation la plus stable pour exercer leur fonction, les IDPs ne possèdent pas de structure tertiaire fixe et rigide.⁴⁴ Au regard du paysage énergétique, les IDPs diffèrent des protéines ordonnées, qui ne possèdent qu'un seul et unique minimum énergétique associé à l'état replié le plus stable et fonctionnel. En effet, les protéines désordonnées existent sous la forme d'un ensemble conformationnel dynamique, permettant la transition d'un conformère à un autre sans que cela ne coûte trop d'énergie étant donné les petits écarts énergétiques entre les différentes conformations (Fig. 8).⁴⁵⁻⁴⁷ Les IDPs sont également caractérisées par une faible diversité de séquence, à savoir qu'elles sont principalement composées de résidus chargés et hydrophiles. Cela résulte en l'absence d'un cœur hydrophobe stable, typiquement retrouvé chez les protéines globulaires. La haute proportion en acides aminés chargés empêche en effet la formation d'états conformationnels compacts.⁴⁵ La plupart des IDPs ne sont néanmoins pas constituées à 100 % de désordre et apparaissent plus fréquemment sous la forme de protéines désordonnées hybrides, contenant à la fois des domaines ordonnés et des régions intrinsèquement désordonnées (IDRs).^{44,46}

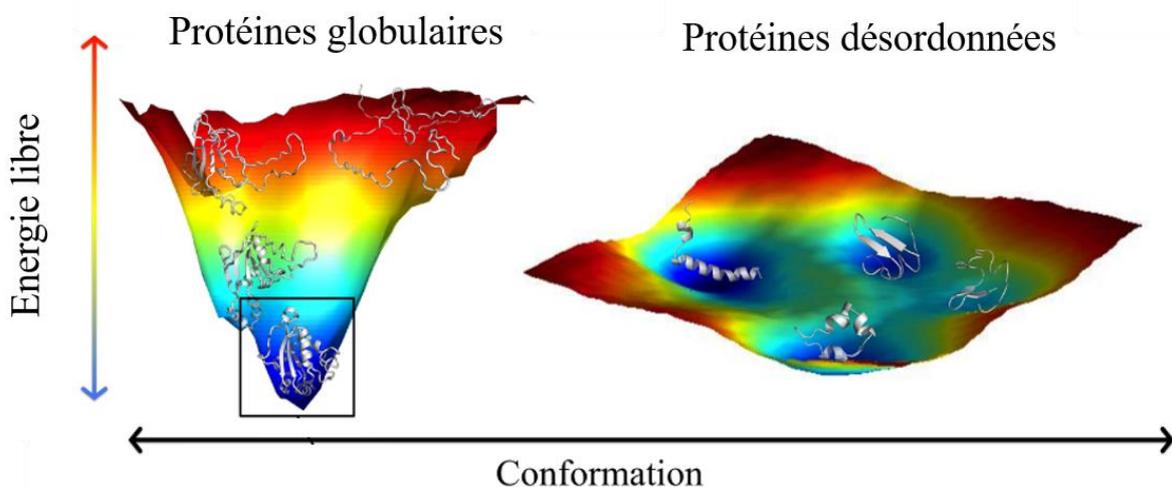


Figure 8 – Paysage énergétique conformationnel des protéines globulaires (gauche) et désordonnées (droite).⁴⁷

Les IDPs ont gagné un intérêt croissant dans la littérature de par les rôles centraux qu'elles occupent dans la cellule. Il est notamment estimé qu'environ un tiers du protéome eucaryote est composé de protéines contenant des IDRs d'au moins 30 acides aminés de long.⁴⁴ Nonobstant le faible nombre d'études sur les IDPs au sein du protéome viral, il est reconnu qu'elles sont impliquées dans les interactions entre l'hôte et le virus ainsi que dans la dérégulation de la réponse immunitaire antivirale de l'hôte. De plus, afin de détourner le réseau de signaux cellulaires, les virus sont connus pour mimer les IDRs, ce qui leur permet d'augmenter l'affinité avec leurs cibles.^{48,49} Concernant les procaryotes, seulement 14 % du protéome des archées et des bactéries sont considérés désordonnés.⁵⁰ Cette différence d'abondance en termes de désordre intrinsèque se base sur les séquences en acides aminés.

En effet, un enrichissement en résidus promoteurs de désordre (DPRs pour *disorder-promoting residues*) est particulièrement observé chez les protéines eucaryotes.⁵¹ Parmi les DPRs sont retrouvés trois résidus chargés (Arg, Glu et Lys), empêchant la formation de noyau hydrophobe, deux apolaires à courte chaîne latérale (Ala et Gly), deux polaires non-chargés (Gln et Ser) ainsi qu'un casseur de structure secondaire (Pro). A l'inverse, certains résidus sont considérés comme promoteurs d'ordre (OPRs pour *order-promoting residues*). Il s'agit des trois acides aminés aromatiques (Phe, Trp et Tyr), de trois apolaires (Ile, Leu et Val), de la cystéine (Cys) de par ses possibilités de formation de ponts disulfure et de l'asparagine (Asn). Enfin, le dernier groupe, regroupant les acides aminés non-promoteurs (NPRs pour *non-promoting residues*), contient un résidu chargé (Asp), un apolaire (Met) et deux polaires (His et Thr).⁵²

Une étude statistique a révélé que le protéome humain peut être divisé en trois catégories, se distinguant sur la base de critères de séquence et de fonction protéiques. Tout d'abord, 49 % des protéines correspondent aux ORDPs (pour *ordered proteins*), qui possèdent moins de 30 % de DPRs. Au contraire, les IDPs, représentant 32 % du protéome, sont associées aux protéines contenant plus de 30 % de DPRs dans la chaîne polypeptidique. Enfin, les 19 % restants sont occupés par les IDPRs (pour *proteins with intrinsically disordered regions*), qui ont moins de 30 % de DPRs dans leur séquence mais au moins un segment désordonné de 40 résidus consécutifs trouvé en dehors des positions C- ou N-terminales ou un segment désordonné de 30 résidus consécutifs faisant partie des extrémités de la chaîne.⁵³

Les IDPs peuvent être classées en trois catégories en fonction de leur composition en acides aminés et, plus précisément, en résidus chargés et polaires. Celles avec des traces

polaires (appelées *polar tracts*) sont déficientes en résidus chargés, hydrophobes et prolines mais enrichies en résidus polaires (Asn, Gly, Gln, His, Ser et Thr). En revanche, les polyampholytes et polyélectrolytes possèdent des acides aminés chargés en grand nombre. Ils peuvent être caractérisés par les paramètres f_+ et f_- , représentant respectivement la quantité de fractions de résidus chargés (FRCs) positivement et négativement dans la séquence. Les polyampholytes possèdent des fractions plus ou moins équivalentes ($f_+ \approx f_-$), alors que les polyélectrolytes ont une fraction largement supérieure à l'autre ($f_+ > f_-$ ou $f_+ < f_-$).

A l'aide de ces paramètres, le diagramme de phase de Das-Pappu, permettant de rationaliser les relations qui existent entre la composition et la conformation des IDPs, est divisé en quatre compartiments relatifs à quatre classes conformationnelles, désignées par R1, R2, R3 et R4 (Fig. 9). La région R1, caractérisée par de faibles f_+ et f_- (FRC < 0.25), correspond aux *polar tracts* et aux polyampholytes et polyélectrolytes faibles. La faible proportion en résidus chargés explique la présence de protéines globulaires dans cette région et, notamment, les protéines de type « têtard » (*tadpole-like*) possédant une « tête » globulaire ordonnée et une « queue » désordonnée. La zone R1 comporte 25 % des IDPs. La région R3 contient les polyampholytes forts à conformation enroulée (*coil*) ou en forme d'épingle à cheveux (*hairpin*) en fonction des valeurs de FRC. Environ 30 % des IDPs se retrouvent dans cette zone. La région R2 est une zone de transition entre les compartiments R1 et R3 ($0.25 < \text{FRC} < 0.35$). Les conformations désordonnées et celles dépendantes du contexte (*context-dependent*) sont retrouvées dans cette partie du diagramme. Ces structures, appelées « séquence de Janus », sont soit effondrées (semblable à l'état de globule fondu) par effet hydrophobe soit étendues (semblable à l'état désordonné) par une solvatation favorable de la chaîne. Environ 40 % des IDPs se trouvent dans cette région, faisant de cette dernière la zone où les IDPs sont le plus représentées. Enfin, la zone R4 est divisée en deux parties : l'une pour les électrolytes chargés positivement, l'autre pour ceux chargés négativement (FRC > 0.35). Seulement 5 % des IDPs composent cette région pour laquelle les répulsions électrostatiques donnent lieu à des conformations étendues enroulées (*coils*) et semi-flexibles (*semi-flexible rods*).⁵⁴

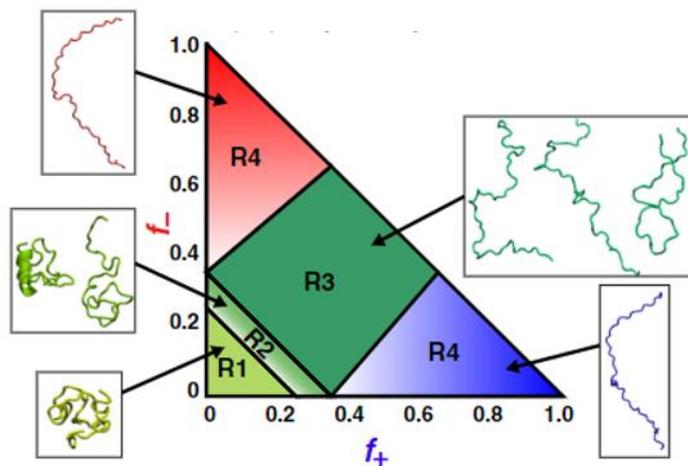


Figure 9 – Diagramme de phase de Das-Pappu et état conformationnel type des IDPs associé chaque région.⁵⁴

Un autre paramètre décrivant le désordre intrinsèque, issu de l'analyse de séquence et complémentaire au FRC, est le facteur κ . Ce dernier rend compte de la distribution des acides aminés chargés positivement et négativement le long de la séquence. En effet, une protéine n'aura pas tendance à se replier de la même manière et n'aura pas la même compacité si les résidus chargés sont contigus ou espacés dans la structure primaire. La valeur du facteur κ est définie sur une échelle allant de 0 à 1. Une faible valeur correspond à une séquence où les acides aminés de charges opposées sont uniformément distribués le long de la séquence. Cette disposition résulte en des conformations protéiques très étendues. Au contraire, une valeur de κ élevée, proche de 1, signifie que les charges positives et négatives sont ségréguées de part et d'autre de la séquence. De ce fait, les charges opposées vont s'attirer et engendrer un certain repliement de la protéine.⁵⁵ En outre, le paramètre Ω rend compte, sur une échelle de 0 à 1, non seulement de la disposition des résidus chargés mais également des prolines par rapport à tous les autres acides aminés. Comme les résidus prolines et chargés sont des DPRs et favorisent les conformations dépliées, une séquence uniformément mélangée donne une valeur Ω faible, ce qui entraîne des conformères de protéine étendus.⁵⁶

1.2.2. Modes d'action et fonctions biologiques

De par la grande flexibilité de leur chaîne désordonnée et leur dynamique conformationnelle, les IDPs peuvent interagir avec une grande variété de partenaires (protéines, lipides, acides nucléiques, métaux, ...) au sein de la cellule. Des études ont montré que les

protéines contenant des IDRs jouent des rôles clés dans les PPIs. En effet, les IDPs sont considérées comme des centres (*hubs*) au sein des réseaux de PPIs et elles se lient souvent de façon promiscue ainsi qu'avec une faible affinité à leurs partenaires étant donné que les contributions enthalpiques sont rarement suffisantes pour compenser la perte d'entropie. La promiscuité des IDPs rend compte du caractère de proximité et de multiplicité de leur liaison. Les protéines *hubs* sont séparées en deux catégories. D'une part, celles formant simultanément des complexes stables avec leurs différents partenaires d'interaction, d'autre part celles qui interagissent de manière transitoire avec leurs partenaires à des moments distincts. Les IDPs appartiennent à cette deuxième catégorie.^{46,57,58}

Trois types de motifs sont responsables du caractère promiscue de la liaison entre l'IDP et ses partenaires : les MoRFs (pour *molecular recognition features*), les SLiMs (pour *short linear motifs*) et les LCRs (pour *low complexity regions*). Les MoRFs sont des segments désordonnés de 10 à 70 résidus capables de subir des transitions désordre-ordre à la suite d'une liaison à un partenaire. Il en existe quatre catégories se différenciant par le type de structure secondaire formé : les α -MoRFs formant des hélices α , les β -MoRFs formant des feuilletts β , les *i*-MoRFs et les *complex*-MoRFs formant respectivement des structures irrégulières et un mélange de structures secondaires. Il existe deux modes de liaison des MoRFs, et des IDPs en général, à leur(s) partenaire(s) d'interaction. Le premier, appelé sélection conformationnelle, consiste en la liaison de la protéine à son partenaire, cette dernière possédant une structure complémentaire au site de liaison. Le second mode, appelé *folding upon binding* ou *induced fit*, est un processus commençant par des contacts non-spécifiques entre la protéine et son partenaire. La partie désordonnée se replie ensuite dans une structure définie, résultant en des contacts plus spécifiques.⁵⁹ Les MoRFs suivent plus souvent ce second mécanisme, ce qui leur permet d'interagir avec des partenaires ayant des structures différentes.

Les SLiMs, à l'instar des MoRFs, agissent comme des interfaces d'interactions dans les régions désordonnées et subissent aussi des transitions désordre-ordre à la suite d'une liaison à un partenaire. Contrairement aux MoRFs, les SLiMs consistent en de plus courts segments, de 3 à 10 résidus, et leur liaison se base davantage sur la séquence que sur la structure, ce qui assure une plus grande spécificité. Également retrouvées chez des protéines plus ordonnées, les SLiMs lient les différents domaines ordonnés entre eux et donnent à la protéine la possibilité d'interagir avec différents partenaires.^{46,60}

Enfin, les LCRs aussi désignées par LCDs (pour *low complexity domains*) correspondent à des séquences de 10 à 40 acides aminés avec une faible diversité de séquence. Autrement dit, ces domaines sont particulièrement enrichis en un ou deux types d'acide aminé et leurs propriétés dépendent du ou des résidu(s) répété(s). Il a également été montré que les LCRs correspondent généralement à un domaine désordonné et qu'elles sont retrouvées au sein des IDRs. Comparativement aux MoRFs et aux SliMs, ces LCDs sont plutôt impliqués dans les mécanismes d'agrégation et de séparation de phase liquide-liquide (LLPS pour *liquid-liquid phase separation*).⁶¹

Les IDPs ayant un grand nombre de partenaires d'interaction, elles sont impliquées dans de nombreuses fonctions biologiques. Parmi elles, sont retrouvés la régulation de la transcription, traduction et de l'autoassemblage de complexes multiprotéiques (comme le ribosome, par exemple), la transduction du signal cellulaire, le remodelage de la chromatine ou encore la phosphorylation de protéines.⁶² De par l'absence de structure tertiaire fixe et rigide, augmentant l'accessibilité des chaînes désordonnées, les IDPs sont particulièrement sensibles aux PTMs telles que la phosphorylation, l'acétylation, la méthylation, la bromination, la carboxylation ou encore l'ubiquitination. Les PTMs affectent la stabilité, les potentielles interactions et la localisation des protéines au sein de la cellule. De plus, par modification de certains acides aminés, les PTMs permettent davantage de fonctions protéiques pour une même structure primaire en modulant la structure tertiaire de la protéine, ce qui est particulièrement vrai pour les IDPs.^{50,62,63}

1.2.3. Pathologies et voies de ciblage thérapeutique

La flexibilité des IDPs, la grande diversité de leurs partenaires d'interaction ainsi que leur promiscuité de liaison ne sont pas sans conséquences. En effet, les protéines désordonnées sont particulièrement sensibles aux dérégulations et sont impliquées dans certaines pathologies humaines graves. Il est estimé qu'environ 79 % des protéines associées aux cancers sont des IDPs ou contiennent au moins une IDR d'au moins 30 résidus consécutifs.⁶⁴ Les estimations s'élèvent à ~ 57 % pour les maladies cardiovasculaires.⁶⁵ Ces chiffres font des IDPs de nouvelles cibles thérapeutiques essentielles. Cependant, les méthodes traditionnelles de conception médicamenteuse, se basant sur la structure de la protéine impliquée, ne peuvent être appliquées étant donné l'aspect très dynamique des chaînes désordonnées. Il est donc très difficile de concevoir un principe actif qui pourra se lier à une chaîne très flexible. C'est pourquoi, de nouvelles stratégies spécifiques aux IDPs ont été développées ou sont en cours de

développement (Fig. 10). La première consiste à bloquer les PPIs entre l'IDP et son partenaire ordonné en ciblant et mimant ce dernier. La transition structurale engendrée par le principe actif empêchera l'assemblage des protéines en un complexe fonctionnel. La difficulté de cette stratégie réside dans le fait qu'une IDP interagit avec une multitude de partenaires. La seconde vise directement l'IDP afin de déplacer son ensemble conformationnel vers une distribution plus ordonnée et stable. Bien que cette méthode semble la plus prometteuse, elle demeure la plus complexe au regard des propriétés dynamiques des IDPs. Il est également possible de viser la partie ordonnée de l'IDP afin de provoquer une transition désordre-ordre de toute la protéine même si cela est plus rare. La troisième stratégie cible le complexe formé entre l'IDP et son partenaire dans son ensemble de sorte à empêcher sa fonction biologique.^{66,67}

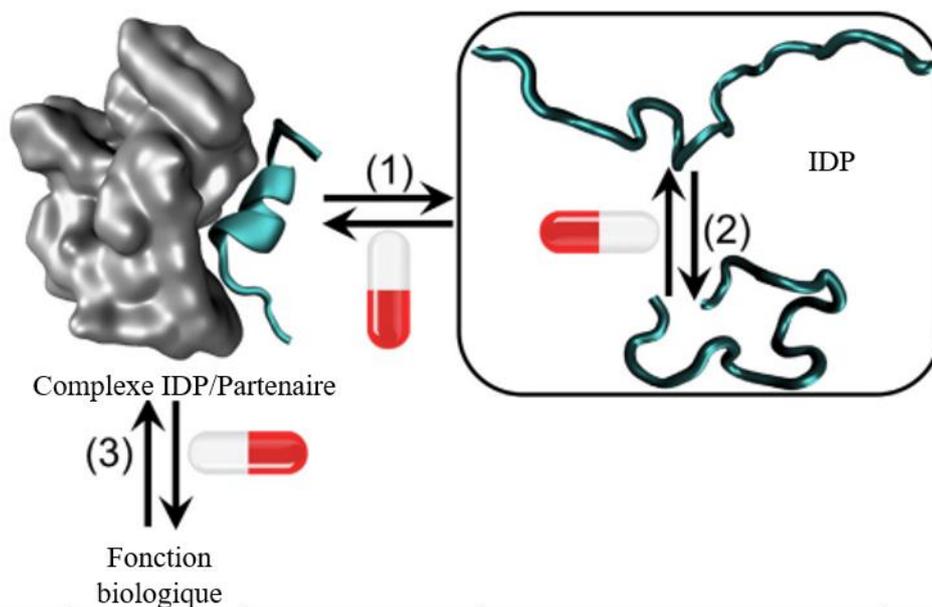


Figure 10 – Représentation schématique des stratégies de ciblage des IDPs. (1) Inhiber les PPIs entre l'IDP et son partenaire. (2) Viser directement l'IDP en déplaçant son ensemble conformationnel. (3) Bloquer le complexe IDP/partenaire pour empêcher sa fonction biologique.⁶⁶

Un exemple relatif à la première stratégie concerne le facteur de transcription p53 impliqué dans la régulation du cycle cellulaire, l'apoptose, la réparation de l'ADN et la réponse aux stress cellulaires. Contenant deux IDRs reliées par une région ordonnée, il fait office de gardien du génome étant donné que la perte de ses fonctions engendre des transformations cellulaires pro-tumorales. La liaison de la ligase MDM2 à p53 bloque ce dernier et empêche ses fonctions transcriptionnelles antitumorales. Dès lors, il est possible de cibler MDM2, à savoir le partenaire ordonné, afin de prévenir leur liaison. Pour ce faire, de petites molécules

ont été développées pour s'insérer dans la poche de liaison responsable des interactions entre p53 et MDM2. Certains médicaments, comme RGT388, sont d'ores et déjà en phase clinique III.^{66,68,69}

La deuxième approche est illustrée par l'IDP c-MYC, retrouvée dérégulée dans 50 % des cancers humains. c-MYC, qui est un régulateur central des processus physiologiques comme l'apoptose ou la prolifération cellulaire, interagit avec MAX (*Myc-associated factor X*), une autre IDP, pour former un hétérodimère capable de réguler l'expression des gènes. Thérapeutiquement, il est possible de cibler uniquement c-MYC (ou MAX) en utilisant une petite molécule qui va s'y lier et la bloquer dans une conformation donnée. Malgré la difficulté de développer une molécule spécifique étant donné la multitude de partenaires d'interaction de c-MYC, trois composés ont tout de même été retenus. Parmi ceux-ci, plusieurs études précliniques ont démontré les effets anticancéreux d'une petite molécule sur le cancer des ovaires soit en tant qu'agent unique soit combinée avec d'autres agents chimiothérapeutiques.^{66,70}

La dernière méthode consiste à bloquer, par exemple, l'interaction entre l'ADN et le complexe c-MYC-MAX, une fois formé, sans affecter l'hétérodimérisation. Récemment, un criblage virtuel a permis de trouver un nouveau composé qui bloque l'activité de c-MYC en entrant en compétition avec l'ADN pour la liaison au complexe c-MYC-MAX. Des études d'optimisation doivent encore être réalisées pour de futures applications cliniques.^{66,71}

1.3. Protéines amyloïdogènes

1.3.1. Fibrillation amyloïde

Les protéines, ordonnées ou désordonnées, sont susceptibles de mal se replier (*misfolding*) à cause, notamment, de stress oxydatifs ou de PTMs, entraînant la formation d'agrégats associés à des maladies neurodégénératives (MNs) comme les maladies d'Alzheimer (AD), de Parkinson (PD) ou encore d'Huntington (HD pour *Huntington's disease*). Pour éviter le mauvais repliement des protéines, les cellules sont dotées d'un système de réparation, localisé dans le réticulum endoplasmique, qui détecte et replie les protéines problématiques, ce qui permet de contrôler leur homéostasie. Il est connu que les IDPs sont prompts à s'agréger, ce qui affecte, en conséquence, un certain nombre de voies métaboliques donnant lieu à des MNs ainsi qu'à différentes formes de cancer. En effet, les PTMs et en particulier les phosphorylations affectent les propriétés physico-chimiques des protéines,

entraînant des modifications structurales qui, à leur tour, sont susceptibles d'altérer les propriétés biologiques des IDPs.⁷²

Au-delà de l'état natif d'une protéine, il existe de multiples états conformationnels, incluant les états partiellement ou totalement dépliés et les différentes formes d'agrégats, reliés entre eux par des équilibres thermodynamiques et des barrières cinétiques (Fig. 11). Généralement, les protéines sont dans leur état natif et soluble pour être fonctionnelle (voie non amyloïdogène ou non amyloïde) mais, dans certaines conditions, elles peuvent passer dans un état non fonctionnel sous forme d'agrégat (voie amyloïdogène ou amyloïde).

Le paysage énergétique de la voie amyloïde est plus complexe car il contient une plus grande diversité d'états conformationnels, à savoir les états repliés et (partiellement) dépliés ainsi que les différents agrégats, c'est-à-dire la phase amorphe, les oligomères, les fibrilles amyloïdes et les agrégats natifs. Ces derniers sont formés à partir d'états monomériques partiellement repliés. Les états agrégés sont thermodynamiquement favorables. La complexité de cette voie provient également de la grande hétérogénéité morphologique de ces derniers.⁷³

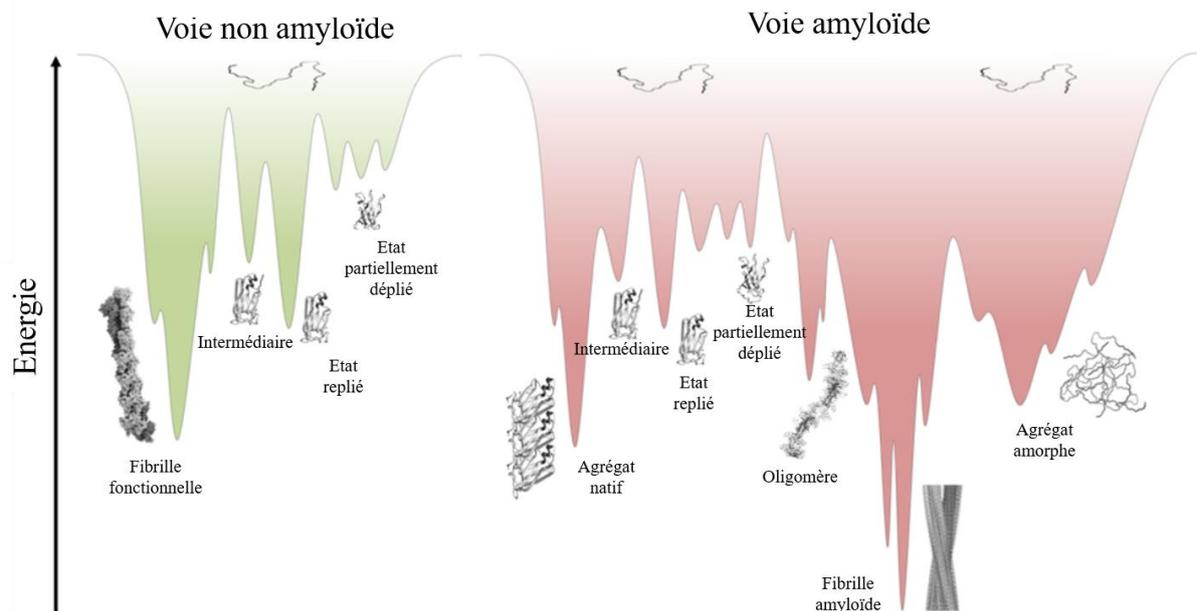


Figure 11 – Paysage énergétique du repliement d'une protéine (voie non amyloïde en vert et amyloïde en rouge) et de ses états agrégés associés.⁷³

Bien que les IDPs s'agrègent selon des mécanismes différents, conduisant à diverses formes agrégées, elles forment préférentiellement des fibrilles amyloïdes dans les MNs. Il

s'agit des agrégats protéiques les plus stables énergétiquement et qui sont caractérisés par une morphologie en feuillets β croisés (*cross β -sheets*). Ces structures sont perpendiculaires à l'axe de la fibrille avec une distance inter-brins d'habituellement $\sim 4.8 \text{ \AA}$ et inter-feuillets de $\sim 10 \text{ \AA}$ (Fig. 12). Ce motif structural a été identifié pour la première fois en 1950 par diffraction de rayons X (DRX) et microscopie électronique.^{74,75}

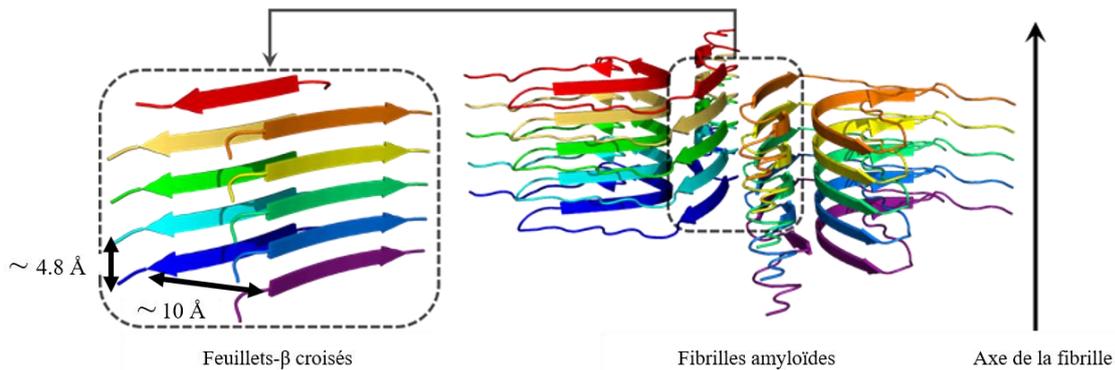


Figure 12 – Structure d'une fibrille amyloïde de l' α -synucléine (entrée PDB : 6OSJ) avec un zoom (à gauche) sur le motif en feuillets β croisés avec une distance inter-brins de $\sim 4.8 \text{ \AA}$ et inter-feuillets de $\sim 10 \text{ \AA}$.

Même si les interactions des chaînes latérales jouent un rôle dans la stabilisation, les fibrilles amyloïdes sont majoritairement constituées de ponts hydrogène le long de l'axe permettant de maintenir leur grande stabilité. Il en existe constituées de courts segments peptidiques voire de protéines entières. Les plus courtes possèdent un motif fondamental, consistant en une fermeture éclair stérique (*steric zipper*) comprenant une paire de feuillets β . Pour les fibrilles contenant de plus longs segments, chaque polypeptide se plie en une structure plane composée de brins β reliés entre eux par des tournants (*turns*) et des boucles (*loops*). Des fermetures éclair stériques sont également présentes pour stabiliser la structure. Celles-ci se forment souvent entre deux feuillets n'ayant pas la même séquence (*hetero-steric zipper*).⁷⁶

Malgré la variété des agrégats, le mécanisme de fibrillation implique souvent des changements conformationnels de la protéine dans l'état natif menant à la formation d'oligomères qui s'assemblent en protofibrilles et, enfin, en des fibrilles amyloïdes matures. Bien qu'il ait longtemps été considéré que ces dernières sont responsables de différentes pathologies, il a été montré que les intermédiaires oligomériques formés durant la fibrillation peuvent se révéler plus toxiques pour plusieurs raisons. La plus petite taille des oligomères

facilite leur diffusion dans les tissus et leur plus grand nombre d'extrémités actives permet plus d'interactions avec les cibles cellulaires.⁷⁷

A côté de ces agrégats structurés, des phases amorphes peuvent aussi être formées. Ces agrégats ont une morphologie différente des fibrilles et incluent différents types d'assemblages sans interactions intermoléculaires, qui permettraient à ces agrégats d'être ordonnés. Ces phases sont souvent retrouvées chez des protéines qui forment aussi des fibrilles amyloïdes. Aussi, les agrégats amorphes peuvent être impliqués dans les MNs.⁷⁸

Les fibrilles formées au sein de la voie non amyloïdogène sont dites fonctionnelles. Alors que certaines agissent comme des biofilms en jouant un rôle de structure dans les microorganismes, d'autres fonctionnent comme des adhésines bactériennes, régulant ainsi l'activité des toxines ou de facteurs de virulence et composant des couches de protéines extracellulaires.⁷⁹ Chez les champignons filamenteux, les fibrilles forment une couche amphipatique en s'assemblant parallèlement à la surface d'une interphase hydrophobe/hydrophile, permettant l'inversion de la mouillabilité. Elles sont souvent retrouvées sous la forme de structure solénoïdale- β . Contrairement aux fibrilles retrouvées dans les MNs, les résidus hydrophiles et chargés des chaînes latérales permettent le contrôle de la fibrillation ainsi qu'une grande réversibilité.^{74,80}

1.3.2. L' α -synucléine dans la maladie de Parkinson

L' α -synucléine (α -syn) est une IDP hybride abondamment exprimée dans le système nerveux et se trouvant dans les neurones présynaptiques.⁸¹ L' α -syn agit en tant que modulateur de la transmission synaptique. La particularité de cette protéine est son enrichissement en hélices α une fois liée à une bicouche lipidique. De plus, il semblerait qu'elle agisse comme un frein à la libération des neurotransmetteurs. Autrement dit, l' α -syn est un contrôleur du relargage des neurotransmetteurs. Cette IDP est liée génétiquement et neuropathologiquement aux synucléinopathies, comme la PD, de par sa région hydrophobe promotrice d'agrégation dans sa séquence.⁸²

Les premiers liens établis entre l' α -syn et la PD datent de 1997 où une mutation dans le gène codant pour l' α -syn a été observée chez des patients touchés par la maladie.⁸³ L' α -syn a tendance à s'agréger en formant des fibrilles amyloïdes riches en feuillet β . Ces fibrilles constituent notamment les corps de Lewy formés dans les synucléinopathies. Le processus d'agrégation de l' α -syn passe par différents intermédiaires. Dans un premier temps, des

oligomères sphériques solubles sont formés. Ils permettent la formation de protofibrilles qui deviennent insolubles et sont à la base de la formation des fibrilles amyloïdes. L' α -syn contribue à la PD de nombreuses manières mais le mécanisme le plus accepté est celui impliquant les protofibrilles, qui sont des espèces cytotoxiques responsables de la rupture de l'homéostasie et de la mort neuronal, ce qui affecte les fonctions synaptiques.⁸² L'agrégation de l' α -syn et donc, la formation de corps de Lewy, entraîne une perte du contrôle de la motricité fine, ce qui conduit finalement à l'apparition des symptômes cliniques de la PD.⁸⁴ De plus, l' α -syn contribue à la fibrillation du peptide amyloïde β ($A\beta$) et de la protéine tau, deux protéines clés de la maladie d'Alzheimer, ce qui suggère un rôle central de la toxicité de l' α -syn dans la neurodégénérescence.⁸⁵

1.3.3. La protéine tau et le peptide amyloïde β dans la maladie d'Alzheimer

La protéine tau est une IDP associée aux microtubules jouant un rôle crucial dans leur stabilisation au niveau des axones, ce qui permet de diminuer leur instabilité et de favoriser leur assemblage ainsi que le développement neuronal. Il existe en tout six isoformes de tau dans le cerveau, se différenciant par la présence, ou non, de certains exons. L'isoforme la plus longue comprend 441 résidus dont 80 sont phosphorylables. La protéine tau ne contient que peu de résidus hydrophobes, comme cela est attendu de son caractère désordonné. L'agrégation de tau peut être accélérée par des polyanions comme des polysaccharides, polyglutamates ou acides nucléiques. Les interactions entre tau et ces composés provoquent son réarrangement conformationnel avec un enrichissement marqué en feuillets β , typique des substances amyloïdes. L'une des principales caractéristiques de l'AD est la formation d'enchevêtrements neurofibrillaires. Il s'agit d'agrégats amyloïdes riches en feuillets β de tau hyperphosphorylée, également appelés filaments hélicoïdaux appariés.⁸⁶⁻⁸⁸

L'AD est également caractérisée par la formation des plaques amyloïdes extracellulaires, constituées de filaments de peptide $A\beta$. Les peptides $A\beta$ sont formés à partir du clivage d'une glycoprotéine appelée APP (pour *amyloid precursor protein*), exprimée dans de nombreux tissus et, en particulier, dans les synapses neuronales. Cette glycoprotéine joue différents rôles importants dans toute une série d'activités biologiques, notamment le développement des neurones, la signalisation, le transport intracellulaire et d'autres aspects de l'homéostasie neuronale. L'APP peut être clivée de deux façons différentes. La première concerne le processus non amyloïdogène, alors que la deuxième implique un processus

amyloïdogène qui résulte en la formation des peptides A β 40 et A β 42, composés respectivement de 40 et 42 acides aminés.⁸⁹⁻⁹²

Les monomères de peptide A β peuvent s'agréger sous différentes formes, ce qui inclut des oligomères amyloïdogènes solubles, composés de feuillets β parallèles et antiparallèles pouvant se propager dans le cerveau, ainsi que des protofibrilles et fibrilles amyloïdes, caractérisées par des feuillets β croisés s'assemblant en plaques amyloïdes. Cependant, il a été démontré qu'il n'y a pas de corrélation directe entre la formation de ces dernières et la perte de synapses et de neurones dans le cerveau des personnes atteintes de l'AD.^{92,93}

1.3.4. Stratégies de ciblage thérapeutique

Les IDPs ayant une forte propension à s'agréger en suivant la voie amyloïdogène, elles font l'objet de plus en plus d'études afin de trouver des anti-agrégants pour lutter contre les MNs. Les thérapies anti-amyloïdes ciblant directement les IDPs se basent sur trois mécanismes principaux : la stabilisation des monomères, la désagrégation des oligomères préformés et la génération d'agrégats non toxiques. A titre d'exemple, il a été montré que l'épigallocatechine gallate, qui est actuellement en phase clinique III⁷⁷, permet la transition des monomères désordonnés du peptide A β vers des oligomères hors voie (*off-pathway*) avant qu'ils ne deviennent amyloïdogènes. Cela signifie que les oligomères hors-voie ne mènent pas à l'assemblage en fibrilles amyloïdes.⁹⁴ De plus, il a été démontré que la porphyrine phtalocyanine tétrasulfonate (PcTS) se lie directement aux résidus aromatiques de l' α -syn liée à la membrane, stabilisant ainsi son état hélicoïdal, ce qui permet de retarder son agrégation et, par conséquent, d'empêcher ses effets neurotoxiques.⁹⁵ La PcTS possède aussi un effet sur tau, empêchant ainsi son agrégation en favorisant la formation d'oligomères solubles non toxiques.⁹⁶

Au-delà des MNs où sont retrouvées des protéines agrégées en structures riches en feuillets β , certaines protéines, impliquées dans le cancer, le sont également. Il a été montré que les cancers pouvaient avoir des caractéristiques communes avec les MNs et les maladies à prion. Cependant, l'agrégation des protéines dans le cancer entraîne une croissance cellulaire incontrôlée ainsi que le maintien du caractère tumoral, alors que les agrégats associés aux MNs conduisent à la mort cellulaire. Un des exemples le plus connu est celui du répresseur de tumeur p53. L'agrégation de cette protéine en oligomères et fibrilles amyloïdes a été détectée dans certaines tumeurs. Ces recherches montrent que l'agrégation est un phénomène crucial dans le

développement des cancers étant donné que p53 perd ses fonctions sous sa forme agrégée. Cette perte de fonction est présente dans plus de 50 % des cancers. Ces études ont poussé au développement de nouvelles stratégies anti-cancéreuses basées sur la recherche d'anti-agrégants. Le criblage d'une bibliothèque d'oligopyridylamides, des inhibiteurs de fibrillation dans l'AD et dans le diabète de type II, a permis d'identifier l'atripyridylamide (ADH-6), qui empêche l'auto-assemblage et l'agrégation de p53. De plus, ADH-6 cible et dissocie les agrégats de p53 dans les cellules cancéreuses humaines, ce qui rétablit l'activité transcriptionnelle de p53 entraînant l'arrêt du cycle cellulaire et l'apoptose.⁹⁷⁻¹⁰¹

1.3.5. Influence des ions métalliques sur le phénomène d'agrégation

La progression de la neurodégénérescence dans les MNs est fortement influencée par différents paramètres (température, pH, force ionique, ...) et, notamment, par une altération de l'homéostasie des métaux. En effet, la dérégulation des niveaux de Cu^{2+} , Mg^{2+} ou Zn^{2+} dans le cerveau provoque une réponse inflammatoire, un stress oxydatif et l'accumulation de dépôts protéiques insolubles, ce qui favorise la neurodégénérescence. D'un point de vue neurologique, le magnésium joue un rôle essentiel dans la transmission nerveuse et la conduction neuromusculaire, ainsi que dans la protection contre les excitations excessives qui peuvent conduire à la mort des cellules neuronales.^{102,103} L'exposition à certains métaux, comme le Ni^{2+} , retrouvé dans l'air, la nourriture ou l'eau, peut également augmenter l'effet des troubles neurodégénératifs. Il a été montré que l'exposition au nickel provoque des dommages sur la mémoire, l'apprentissage, la locomotion ainsi que les processus affectifs et cognitifs. En outre, l'exposition précoce au nickel peut altérer le développement normal du cerveau et accroître la susceptibilité aux MNs à un âge plus avancé.¹⁰⁴

La liaison des ions métalliques aux IDPs stabilise des états repliés, notamment en diminuant la répulsion électrostatique entre les charges négatives au sein de la protéine. L'agrégation provenant de conformations partiellement repliées, la liaison métallique peut donc promouvoir des conformères propices à l'agrégation qui soit l'accéléreront soit stabiliseront des états agrégés non-amyloïdogènes.¹⁰⁵⁻¹⁰⁷ L'influence de la liaison métallique sur le processus d'agrégation dépend aussi de plusieurs facteurs extérieurs, tels que le pH, la concentration de la protéine ou des ions métalliques ainsi que la nature de ces derniers.¹⁰⁸ De plus, la coordination des métaux aux protéines dépend de nombreux paramètres, tels que l'environnement diélectrique, l'accessibilité au solvant et les propriétés des cations. Seuls

quelques acides aminés, dont les résidus histidine, cystéine, aspartate et glutamate, sont majoritairement impliqués dans les liaisons métal-protéine.¹⁰⁶

Plus particulièrement, l'agrégation de tau et du peptide A β , impliqués dans l'AD, est fortement influencée par la présence de métaux. En effet, les interactions avec les cations métalliques, tels que Cu²⁺ et Zn²⁺, présents en forte concentration au niveau des synapses (de 10 à 50 μ M), ont un impact sur le mécanisme d'agrégation du peptide A β . La fixation des ions Cu²⁺ et Zn²⁺ accélère ce processus et tend à favoriser la formation d'oligomères plutôt que des fibrilles.^{105,109-111} De même, la liaison du Zn²⁺, ainsi que du Mg²⁺, à la protéine tau accélère significativement sa fibrillation, contrairement au Cu²⁺, qui n'affecte pas directement son agrégation. En outre, il a été démontré que la conformation de tau est modulée par le Ni²⁺, empêchant sa fibrillation et promouvant la formation de courts agrégats fragmentés.¹¹² Dans des conditions réductrices et en présence de peroxyde d'hydrogène, les ions Cu²⁺ liés aux fibrilles de tau et du peptide A β participent à la génération d'espèces réactives oxygénées (ROS pour *reactive oxygen species*), ce qui entraîne un stress oxydatif, des dommages cellulaires et la production de petits agrégats toxiques contribuant à la mort neuronale. De plus, la formation de ROS peut conduire à l'oxydation de peptides, dont A β , influençant également les mécanismes d'agrégation.^{105,113-115}

A l'instar de l'AD, les métaux ont également un impact sur la PD par modification de la conformation de l' α -syn et de sa voie d'agrégation. Les ions Cu²⁺ ayant une forte affinité pour l' α -syn, la protéine est liée à ces derniers *in vivo*. Ils accélèrent son agrégation, sans pour autant modifier la morphologie des fibrilles, et favorisent la formation d'oligomères neurotoxiques.^{116,117} En revanche, à faible concentration, comme c'est le cas dans le cerveau, la liaison au Mg²⁺ empêche l'agrégation de l' α -syn, contrebalançant les effets des autres métaux et jouant ainsi un rôle neuroprotecteur.¹¹⁸ Concernant le Zn²⁺, il favorise, d'une part, la formation de fibrilles de l' α -syn plutôt que d'oligomères neurotoxiques,^{105,119} et, d'autre part, renforce l'activité anti-agrégante de chaperonne qui protège les régions promotrices d'agrégation (APRs pour *aggregation prone regions*) de l' α -syn.¹²⁰

Partie n°2

—

Objectifs et stratégies

2.1. Objectifs de recherche

Ce mémoire a pour objectif général la caractérisation structurale, par combinaison de techniques prédictives et biophysiques, de nouvelles protéines amyloïdogènes intrinsèquement désordonnées, à savoir les isoformes DPF3b et DPF3a. Il s'agit, dans un premier temps, d'étudier et de prouver la nature désordonnée de DPF3b, d'approfondir ce qui est déjà connu sur DPF3a et de comparer leurs propriétés. L'élucidation de ces propriétés structurales est en effet nécessaire à la compréhension des spécificités de fonction et de réseaux d'interactions de chaque isoforme. Une étude approfondie des différents domaines N- et C-terminal de DPF3 permettra d'analyser leurs caractéristiques structurales. Dans un second temps, la propension à l'agrégation des deux isoformes sera examinée. DPF3a ayant déjà montré une capacité à se fibriller, la tendance à s'agréger des deux isoformes et de leurs domaines N- et C-terminal ainsi que leur comportement amyloïdogène seront comparés. L'influence de cations métalliques divalents associés à la neurodégénérescence (Cu^{2+} , Mg^{2+} , Ni^{2+} et Zn^{2+}) sur les propriétés d'agrégation des régions C-terminales sera par ailleurs réalisée étant donné que les deux isoformes se différencient par ces régions.

DPF3 étant impliquée dans différentes formes de cancer, les résultats de cette étude permettront, à long terme, de mieux appréhender le réseau d'interactions des isoformes de DPF3 au cours de la mitose afin de développer de nouveaux antimitotiques. Aussi, la compréhension du mécanisme et du comportement d'agrégation de DPF3 et de ses domaines est essentielle à la conception de médicaments qui empêcheront sa fibrillation amyloïde dans les MNs étant donné son implication dans la PD et l'AD.

2.2. Stratégies de production et de purification

Avant les différentes étapes de caractérisation, il sera nécessaire de produire le matériel protéique, à savoir les protéines complètes DPF3b et DPF3a (abrégées FL pour *full-length*) ainsi que les régions C-terminales (C-TER correspondant aux régions entre le 200^{ème} et le dernier résidu) comprenant les domaines C₂H₂, PHD-1 et PHD-2 pour DPF3b (C-TERb) et les domaines C₂H₂ et PHD-1/2 pour DPF3a (C-TERa). Pour ce faire, ces protéines seront surexprimées dans des systèmes bactériens (*Escherichia coli* abrégée *E. coli*) compétents sous la forme de protéines recombinantes (Fig. 13). Les protéines d'intérêt sont fusionnées à un *tag* protéique de glutathion S-transférase (GST) à leur extrémité N-terminale afin de faciliter le processus de purification. Les protéines de fusion seront récupérées par lyse mécanique avant d'être purifiées par chromatographie d'affinité en phase liquide de protéine rapide (FPLC pour

fast protein liquid chromatography). Elles seront alors clivées de leur *tag* GST à l'aide d'une protéase spécifique. La pureté et la présence des protéines purifiées seront vérifiées par électrophorèse sur gel d'acrylamide en présence de dodécylsulfate de sodium (SDS-PAGE pour *sodium dodecyl sulphate*) et confirmées par spectrométrie de masse en tandem par ionisation en électrospray couplée à une chromatographie en phase liquide (LC-ESI-MS/MS pour *liquid chromatography-electrospray ionisation-tandem mass spectrometry*). Leur concentration sera déterminée par spectroscopie d'absorption UV-visible.

Le caractère désordonné et la tendance à l'agrégation des protéines seront analysés par une série de techniques biophysiques et de prédiction (Fig. 13). En se basant sur la séquence de chaque isoforme, le pourcentage moyen de désordre intrinsèque par résidu, la classe conformationnelle ainsi que l'état d'enfouissement des résidus aromatiques pourront être estimés. Les MoRFs et les LCDs seront également localisés. La spectroscopie de fluorescence sera utilisée pour l'identification de structures locales (non) repliées et la détermination de l'état d'exposition au solvant de résidus aromatiques. Les techniques de fluorescence comprennent la fluorescence intrinsèque du tryptophane (ITF pour *intrinsic tryptophan fluorescence*) et de la tyrosine (ITyrF pour *intrinsic tyrosine fluorescence*), l'autofluorescence dans le bleu profond (dbAF pour *deep-blue autofluorescence*) ainsi que le test de liaison à la thioflavine T (ThT). Cette dernière est un fluorophore extrinsèque permettant de suivre spécifiquement la fibrillation.¹²¹ Le test de liaison au rouge Congo par spectroscopie UV-visible y sera complémentaire. La tailles des monomères et des espèces agrégées sera évaluée par diffusion dynamique de la lumière (DLS). Le contenu en structure secondaire et les changements de conformation seront mis en évidence par spectroscopie de dichroïsme circulaire dans l'UV lointain (FCD pour *far-UV circular dichroism*). La morphologie des fibrilles et des agrégats sera visualisée par microscopie électronique à transmission (TEM pour *transmission electron microscopy*).

Toutes ces étapes seront également appliquées aux régions C-TER afin de discriminer lesquelles sont impliquées dans le processus d'agrégation et lesquelles sont spécifiques à la fonctionnalité de chaque isoforme. Afin d'investiguer plus en profondeur l'agrégation de la région C-TER de chaque isoforme, l'influence de métaux divalents (Cu^{2+} , Mg^{2+} , Ni^{2+} et Zn^{2+}) sera analysée par spectroscopies de fluorescence, FCD et par TEM (Fig. 13).

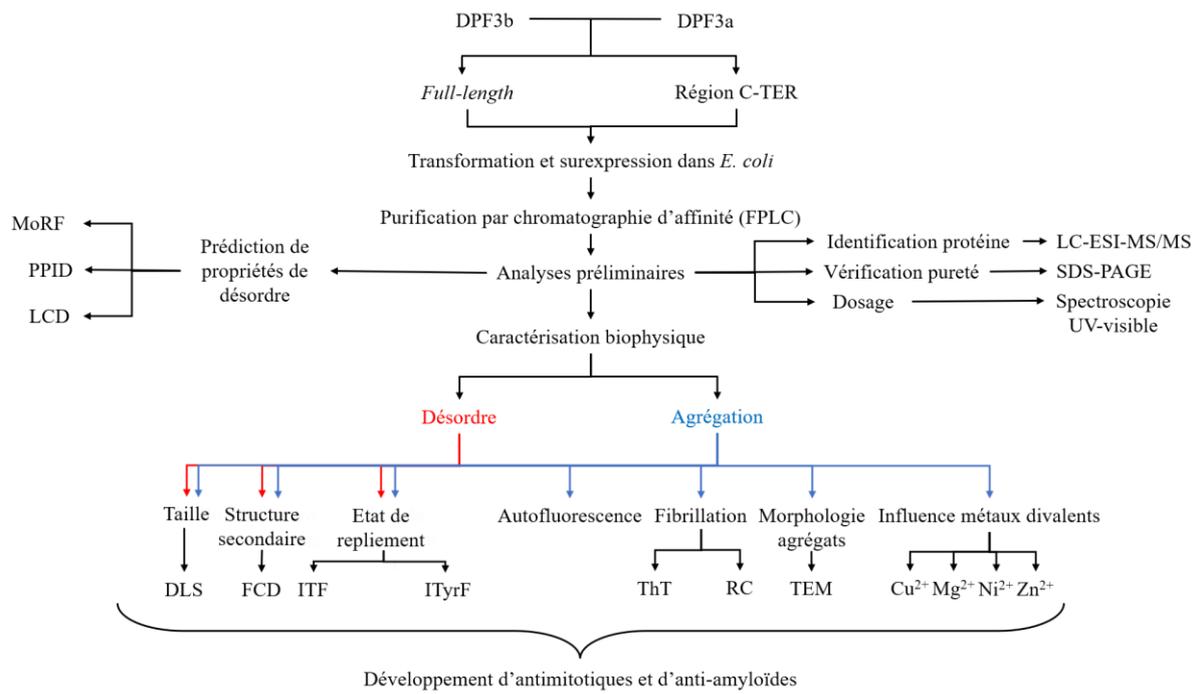


Figure 13 – Représentation schématique des stratégies de production et de caractérisation des isoformes de DPF3 et de leur région C-terminale (C-TER) respective.

Partie n°3

—

Résultats et discussions

Partie n°3.1

—

Purification et analyses préliminaires

3.1.1. Purification par FPLC

A l'instar des étapes de transformation et de surexpression, la méthode de purification de DPF3 par chromatographie d'affinité en phase liquide de protéine rapide (FPLC) sur des colonnes pré-packées a déjà été optimisée dans le cadre du mémoire de Julien Mignon.⁴³ Il s'agit, dans un premier temps, de fixer la protéine de fusion (DPF3 couplée à son *tag* GST) sur la matrice de la colonne contenant du glutathion réduct (GSH) à sa surface étant donné que le GSH est le substrat de la GST. Cette étape permet de purifier la protéine en se débarrassant des protéines non désirées, à savoir non-recombinantes, qui passent dans le *flow-through* (Fig. 14). Dans un second temps, la protéine d'intérêt (DPF3) est clivée de son *tag* à l'aide de la protéase TEV, éluée et récupérée pour les caractérisations. Afin de régénérer la matrice de la colonne, la GST couplée au GSH est décrochée de la matrice par passage d'un tampon concentré en GSH qui entre en compétition avec le couple GST-GSH.¹²² Cette procédure a été utilisée pour les purifications de chaque construction, à savoir DPF3b FL, DPF3a FL, C-TERb et C-TERa.

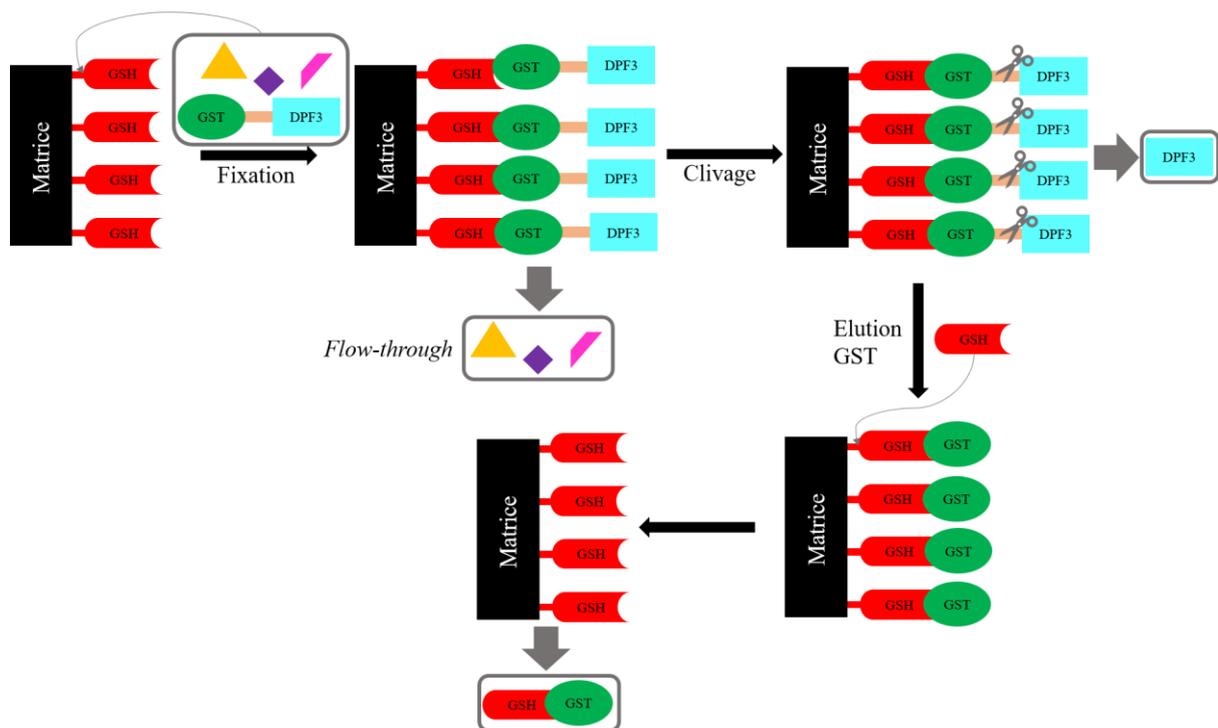


Figure 14 – Représentation schématique du procédé de purification par chromatographie d'affinité de DPF3 couplée à la GST.

Les chromatogrammes de chaque purification étant identiques au regard des conditions de purification, celui associé à la purification de DPF3b FL est représenté à titre exemplatif (Fig. 15). Les trois premières bandes, indiquées en noir, correspondent aux différents *flow-through* à la suite des trois injections du lysat (Fig. 15A). Le signal UV à 280 nm augmente

fortement et est beaucoup plus intense que celui associé au passage de DPF3b FL (en rouge). La protéine d'intérêt, clivée de son *tag* GST, est récupérée en 8 mL (Fig. 15B). L'absorbance augmente une dernière fois lors du passage de la GST (en vert). Cette dernière bande confirme que la protéine de fusion a bien été clivée ainsi que le bon fonctionnement du procédé de purification.

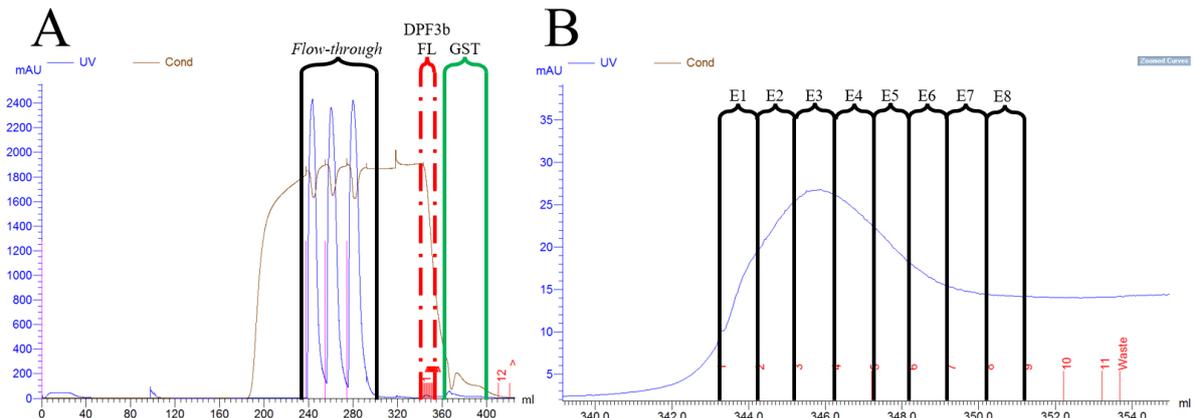


Figure 15 – Chromatogramme entier (A) de la protéine de fusion DPF3b FL-GST clivée à l'aide de la protéase TEV et purifiée par FPLC ainsi qu'un zoom (B) sur la bande d'éluion de DPF3b FL.

3.1.2. Comportement singulier par SDS-PAGE

Afin de vérifier le clivage et la pureté, des gels SDS-PAGE ont été réalisés soit avec une coloration au bleu de Coomassie (Fig. 16) soit à l'imidazole-zinc. Les résultats obtenus sont similaires, quelle que soit la coloration, entre les différentes purifications. Les bandes associées au *flow-through* montrent la présence de nombreuses protéines qui ne possèdent pas le *tag* GST et qui ne s'accrochent donc pas à la colonne. La GST est bien clivée comme le montre la bande aux alentours de 25 kDa selon le marqueur de masse moléculaire, ce qui est attendu étant donné que la masse de la GST s'élève à 25.50 kDa. De plus, aucune bande ne correspond à la protéine de fusion DPF3b-GST qui aurait été potentiellement visible vers 68 kDa dans le cas d'une absence de clivage. Curieusement, aucune bande n'apparaît pour DPF3b FL peu importe le type de coloration. Bien que ce comportement reste singulier, il semble propre à la protéine étant donné sa reproductibilité. Nonobstant l'absence de bande associée à DPF3b, les gels montrent le bon déroulement de la purification et rendent compte d'une certaine pureté de par l'absence de bandes de protéines contaminantes à des masses moléculaires différentes de celle de DPF3b FL (~ 43 kDa).

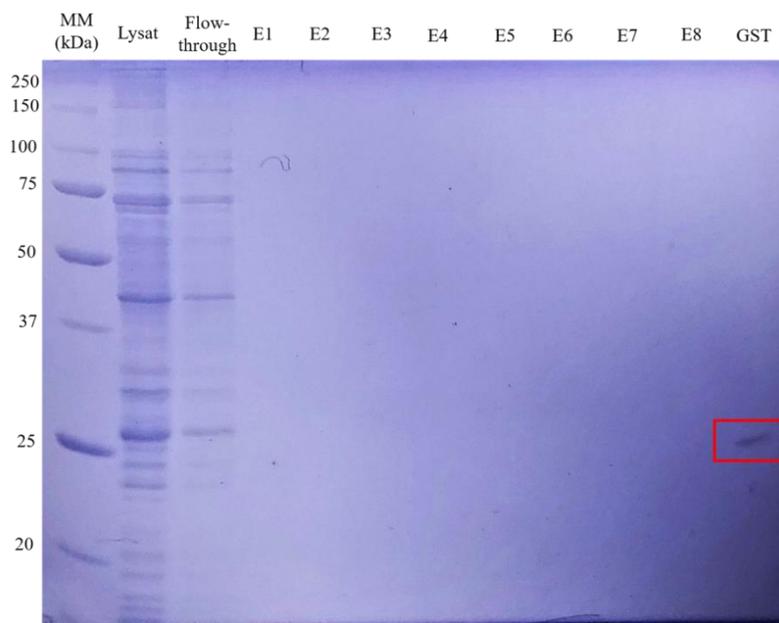


Figure 16 – Gel SDS-PAGE 12 % associé à la purification de DPF3b FL. Le gel correspond à une coloration au bleu de Coomassie. MM correspond à l'échelle du marqueur de masse moléculaire (en kDa) et E est associé aux différentes fractions d'éluion de DPF3b FL clivée. Le cadre rouge met en évidence la présence d'une bande associée à la GST clivée.

3.1.3. Identification protéique par spectrométrie de masse (LC-ESI-MS/MS)

La protéine DPF3b ne pouvant être visualisée sur gel SDS-PAGE, sa présence a été vérifiée par spectrométrie de masse (LC-ESI-MS/MS). DPF3b est bien identifiée avec une couverture de séquence de 45 %. Les dix peptides uniques, allant de 8 à 24 résidus, sont correctement répartis tout au long de la séquence (Fig. 17). Les liaisons peptidiques sont lysées par la trypsine après chaque résidu lysine et arginine. De ce fait, 47 sites de clivage potentiels sont identifiés dans la séquence de DPF3b et donc, 47 peptides devraient être formés après lyse par la trypsine. Toutefois, tous ne sont pas repérés par spectrométrie de masse car certains sont de trop petite ou de trop grande taille pour le détecteur.

MATVIHNPLK↓ALGDQFYK↓EAIEHCRSYNSRLCAERSVR↓LPFLDSQTGVAQNNCYIWMKRHR↓GPGLAPGQ
 LYTPAR↓CWRKKR↓RLHPPEDPK↓LRLEIKPEVELPLKK↓DGFTSESTLEALLR↓GEGVEKK↓VDAREEESIQE
 IQR↓VLENDENVEEGNEEEDLEEDIPK↓RKNRTRGRARGSAGGRRRHDAASQEDHDKPYVCDICGKRYKNRPGL
 SYHYAHTHLASEEGDEAQDQETRSPNHR↓NENHRPQK↓GPDGTVIPNNYCDFCLGGSNMNKK↓SGRPEELVSC
 ADCGR↓SGHPTCLQFTLNMTEAVKTYKWQCIECKSCILCGTSENDDQLLFCDDCDRGRYHMYCLNPPVAEPPEGS
 WSCHLCWELLKEKASAFGCQA

Figure 17 – Représentation schématique des peptides uniques exclusifs (surlignés en vert) identifiés dans la séquence de DPF3b FL clivée par LC-ESI-MS/MS. Les flèches rouges représentent les sites de clivage à la trypsine entre chaque peptide identifié.

3.1.4. Comportement singulier par spectroscopie d'absorption UV-visible

Afin de confirmer la présence de DPF3b FL et d'en déterminer la concentration, les fractions E1 à E6 ont été rassemblées et analysées par spectroscopie d'absorption UV-visible (Fig. 18). Alors qu'il est habituellement attendu un maximum d'absorbance à 280 nm de par la présence d'acides aminés aromatiques dans la séquence, en particulier des tryptophanes et des tyrosines, un déplacement hypsochrome de le bande d'absorption vers 258 nm est observé. A l'image du gel SDS-PAGE, ce comportement particulier a déjà été précédemment observé pour DPF3a FL et a été reproduit à chaque purification. L'absence de bande à 280 nm pour DPF3b FL prouve néanmoins que la GST a été efficacement clivée étant donné que la protéine de fusion DPF3b FL-GST présente un maximum d'absorbance à 280 nm.

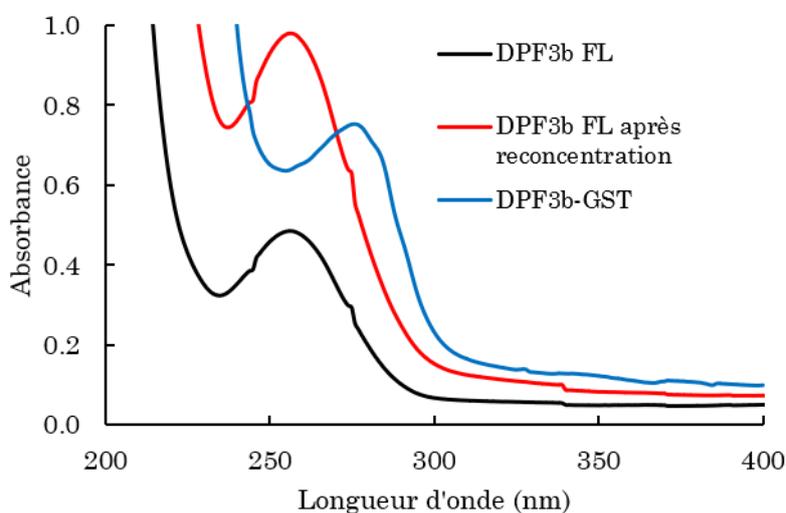


Figure 18 – Spectre d'absorption UV-visible de DPF3b FL avant (noir) et après (rouge) reconcentration par dialyse au PEG-20000 ainsi que de DPF3b-GST (bleu) dans le TPB à ~20 °C.

De plus, aucun déplacement du maximum d'absorption n'est observé à la suite de la reconcentration de DPF3b FL au PEG-20000. Seul un effet hyperchrome est visualisé. Etant donné l'absence de maximum à 280 nm, il n'est pas possible de calculer la concentration protéique à partir de la valeur d'absorbance à cette longueur d'onde. Le coefficient d'extinction molaire étant inconnu à 258 nm, une méthode alternative, consistant à calculer la concentration à partir de l'absorbance à 214 nm et décrite par B. J. H. Kuipers et H. Gruppen, a été utilisée.¹²³ En effet, le coefficient d'extinction molaire peut être déterminé à cette longueur d'onde à partir de la séquence de la protéine. Juste après purification, les 6 mL de DPF3b FL ont une concentration de 0.9 μ M (0.04 mg/mL). Après reconcentration, il reste environ 1 mL à une

concentration de 5.6 μM (0.24 mg/mL), montrant ainsi l'efficacité de cette méthode, indispensable pour les futures analyses biophysiques. Ces concentrations sont données à titre exemplatif mais elles sont représentatives de toutes les purifications réalisées.

Partie n°3.2

—

Caractérisation des isoformes de DPF3 : DPF3b FL et DPF3a FL

3.2.1. Prédiction de propriétés associées au désordre

Les résultats obtenus pour la caractérisation des isoformes de DPF3 (DPF3b FL et DPF3a FL) ainsi que des régions C-TER (C-TERb et C-TERa) ont fait l'objet de deux articles de recherche publiés dans des revues à comité de lecture au cours de ce mémoire.^{124,125}

Le caractère désordonné des isoformes de DPF3, et des IDPs en général, a une importance capitale dans la définition de leurs fonctions. Il est donc essentiel d'élucider cette propriété. Pour ce faire, il est dans un premier temps possible de prédire la propension au désordre d'une protéine sur base de sa structure primaire, autrement dit sa composition en acides aminés. La différence de composition entre les séquences des deux isoformes à l'extrémité C-terminale se reflète dans leur pourcentage de désordre. A partir d'un ensemble de 19 algorithmes, le pourcentage moyen de désordre intrinsèque prédit par résidu (PPID) de DPF3a FL s'élève à 63 % alors qu'il n'est que de 44 % pour l'isoforme DPF3b FL. Au regard des différents domaines (Fig. 2), ce résultat s'explique par les doigts de zinc (ZnFs) PHD-1 et PHD-2 qui sont ordonnés chez DPF3b FL. DPF3a FL ne contenant pas ces domaines, mais étant pourvue d'une partie C-terminale prédite désordonnée, il est cohérent que cette isoforme possède un PPID plus élevé. Nonobstant ces différences, les deux isoformes sont considérées comme des protéines hautement désordonnées et possèdent également des régions intrinsèquement désordonnées (IDRs) communes (Figs. 19A et B).

Les séquences étant identiques du premier au 292^{ème} résidu, les premières régions ordonnées et les deux premières IDRs sont identiques. En effet, à partir du côté N-terminal, les 90 premiers acides aminés, comprenant le domaine 2/3, sont partiellement ordonnés étant donné que le score de désordre varie entre 0.2 et 0.5, indiquant une certaine flexibilité surtout entre le 60^{ème} et le 90^{ème} résidu. Cette partie ordonnée est suivie de la première IDR (IDR-1) longue de 110 acides aminés et caractérisée par un score supérieur au seuil de 0.5. Cette valeur représente la limite au-dessus de laquelle la séquence est considérée désordonnée. Le deuxième domaine ordonné est le ZnF C₂H₂ comptant 22 résidus (199-220). La deuxième IDR (IDR-2), aussi commune aux isoformes, ne contient que 40 acides aminés et s'étend jusqu'au ZnF PHD-1 (221-260).

Pour DPF3b FL, le score de désordre reste sous la limite de 0.5 jusqu'à la fin de sa séquence (résidu 378) étant donné le caractère structuré des domaines PHD-1 et PHD-2 typique de la famille D4. En revanche, DPF3a FL montre une structure ordonnée, associée à son ZnF tronqué (PHD-1/2), du 261^{ème} résidu jusqu'à l'acide aminé 292 avant que le score réaugmente fortement au-dessus de 0.5. Cette dernière IDR (IDR-3), spécifique à DPF3a FL, est longue de

65 acides aminés (résidus 293-357). Au niveau des domaines en ZnFs (C₂H₂, PHD-1/2, PHD-1 et PHD-2), il était attendu qu'ils aient un faible score de désordre étant donné que les domaines PHD-1 et PHD-2 de DPF3b FL sont de structure connue et que tous les ZnFs sont ordonnés par la coordination d'ions Zn²⁺. L'ensemble de ces résultats montrent que les deux isoformes semblent faire partie de la classe des IDPs étant donné que le pourcentage de désordre est supérieur à 30 % et qu'au moins une IDR de plus de 30 résidus consécutifs est présente pour chacune d'entre elles.¹²⁴

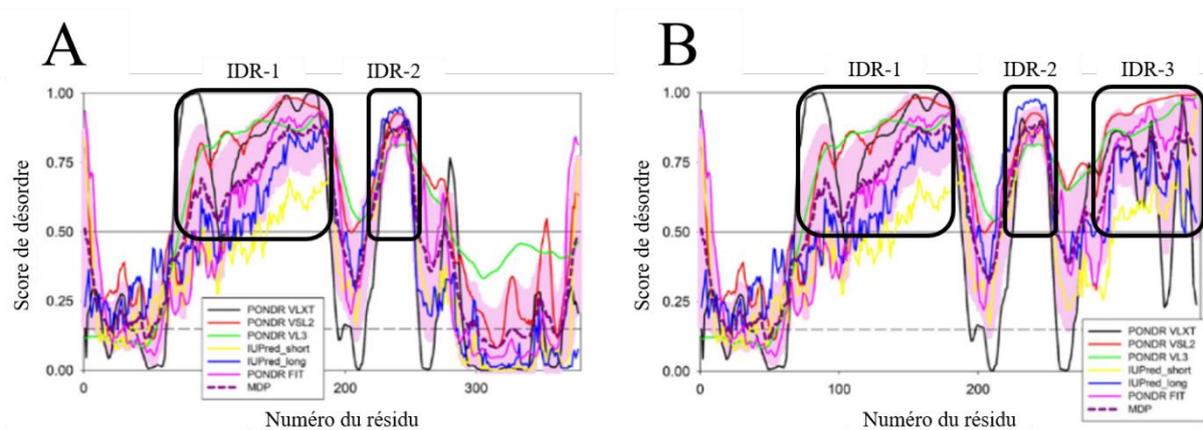


Figure 19 – Profil de désordre PONDRA et IUPred de (A) DPF3b FL et (B) DPF3a FL. MDP (mean disorder profile) correspond au profil de désordre moyen à partir des scores PONDRA et IUPred. Chaque résidu présentant un score supérieur à 0.5 est considéré désordonné.¹²⁴

Les IDPs étant enrichies en résidus promoteurs de désordre (DPRs), c'est aussi le cas pour les isoformes DPF3a FL (59 %) et DPF3b FL (54 %). Plus précisément, DPF3b FL est particulièrement enrichie en arginines (30 résidus équivalant à 8 %) et en glutamate (40 résidus comptant pour 11 %) et il en va de même pour DPF3a FL, composée de 36 arginines (10 %) et de 41 glutamates (11 %). En revanche, chaque isoforme est appauvrie en résidus promoteurs d'ordre (OPRs) puisque DPF3a FL n'en contient que 26 % contre 32 % pour DPF3b FL.

Les graphiques de fonction de distribution cumulative (CDF pour *cumulative distribution function*) et de charge-hydrophathie (CH) sont également utiles pour investiguer le caractère désordonné des isoformes de DPF3 (Figs. 20A et B). Concernant le graphe CDF, les courbes localisées sous la limite correspondent aux IDPs car elles ont un faible pourcentage de résidus avec un faible score de désordre au contraire des protéines globulaires, retrouvées au-dessus de la limite. Les courbes associées aux isoformes de DPF3 se trouvant en dessous de la limite, cela va dans le sens de leur caractère désordonné. En outre, le graphe CDF appuie la

nature plus désordonnée de DPF3a FL étant donné que sa courbe est sous celle relative à DPF3b FL.^{124,126}

En ce qui concerne la distribution CH, l'hydropathie est une indication du caractère hydrophobe de la protéine en se basant sur sa séquence. En effet, une valeur d'hydropathicité est attribuée à chaque résidu en fonction de sa chaîne latérale en sachant que plus l'hydropathie est élevée, plus l'acide aminé est hydrophobe. La charge nette correspond à la valeur absolue de la différence entre les résidus chargés positivement (lysine et arginine) et négativement (aspartate et glutamate) à pH 7.0. Le graphique CH est séparé en deux zones par une limite empirique définie par la droite d'équation $\langle R \rangle = 2.785 \langle H \rangle - 1.151$ où $\langle R \rangle$ est la charge nette moyenne et $\langle H \rangle$ est l'hydropathie moyenne de la protéine.¹²⁷ Ce paramètre $\langle H \rangle$ est calculé selon l'échelle définie par J. Kyte et R. F. Doolittle.¹²⁸ Les IDPs étant enrichies en résidus polaires et chargés, elles présentent généralement une faible hydropathie ainsi qu'une charge nette moyenne élevée. Les deux isoformes sont localisées au-dessus de la frontière ordonnée-désordonnée et se retrouvent, par conséquent, dans la zone associée aux IDPs, ce qui supporte une nouvelle fois leur nature désordonnée.¹²⁴

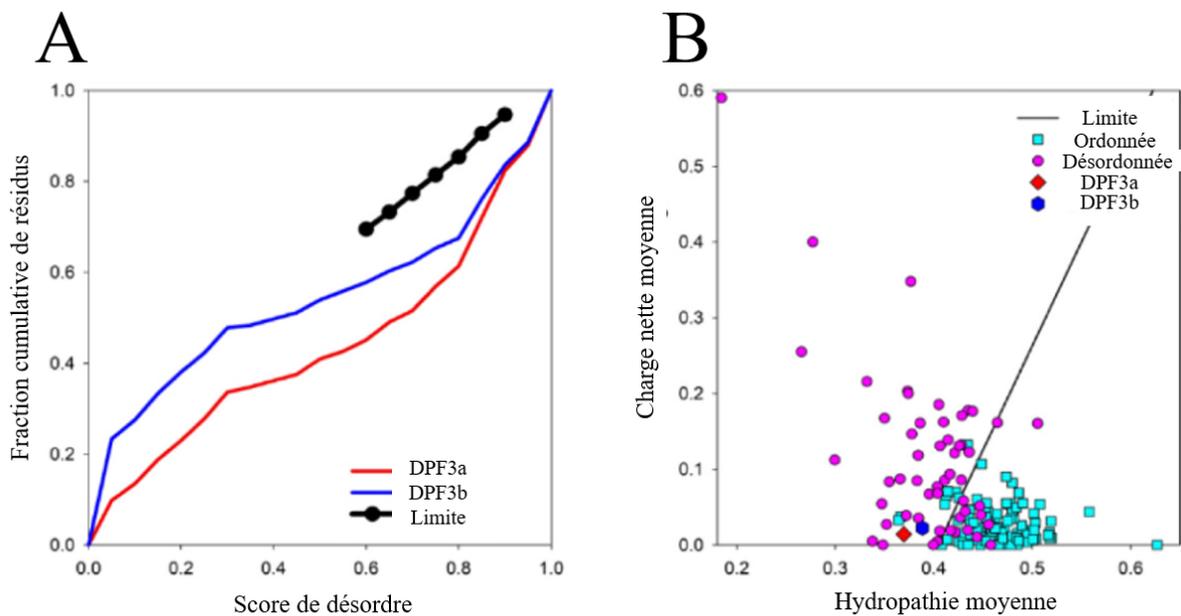


Figure 20 – (A) Graphique CDF et (B) distribution CH de DPF3b FL et DPF3a FL. Sur le graphique CDF, la courbe de DPF3a FL est représentée en rouge et celle de DPF3b FL en bleu. Sur la distribution CH, les protéines ordonnées sont représentées par des carrés bleu clair et les désordonnées par des cercles roses. DPF3a FL correspond au losange rouge et DPF3b FL à l'hexagone bleu foncé.¹²⁴

Le diagramme de phase de Das-Pappu des isoformes de DPF3 (Fig. 21) permet d'estimer leur ensemble conformationnel préférentiel. Comme 40 % des IDPs, les deux isoformes se trouvent dans la région R2. Il s'agit de la zone de transition entre les zones R1 et R3 respectivement associées aux polyampholytes faibles et forts. DPF3a est légèrement plus proche de cette dernière zone R3 qui regroupe 30 % des IDPs.¹²⁴

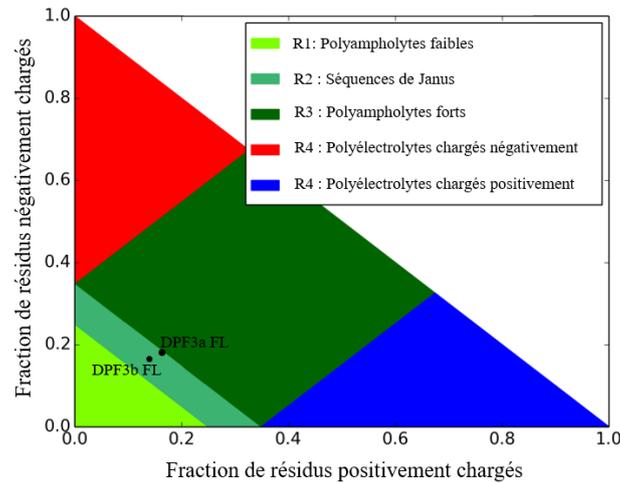


Figure 21 – Diagramme de phase de Das-Pappu de DPF3b FL et DPF3a FL.¹²⁴

Complémentairement, bien que la charge moyenne des isoformes soit faible, le facteur κ rend compte de la distribution des acides aminés chargés positivement et négativement le long de la séquence et permet d'estimer le repliement conformationnel des protéines. La valeur du facteur κ est définie sur une échelle allant de 0 à 1 où une faible valeur correspond à une conformation étendue de la protéine. DPF3b FL et DPF3a FL ont respectivement une valeur κ de 0.26 et 0.27, ce qui montre que les conformations étendues sont favorisées. En outre, le paramètre Ω rend compte, aussi défini sur une échelle de 0 à 1, de la disposition des prolines et des résidus chargés par rapport à tous les autres acides aminés. Une faible valeur correspond également à un conformère déplié. Les faibles valeurs Ω obtenues pour DPF3b FL et DPF3a FL, respectivement de 0.17 et 0.15, indiquent donc des conformations étendues.¹²⁴

Les domaines de faible complexité (LCDs), fréquemment retrouvés dans les IDRs et étant impliqués dans les fonctions des IDPs, sont caractérisés par une faible diversité de séquence. Les isoformes de DPF3 partagent deux LCDs prédits comme étant situés dans l'IDR-1. Le premier LCD (LCD-1) est riche en glutamate et fait 20 résidus de long (145-164) alors que le deuxième (LCD-2) est riche en arginine ainsi qu'en glycine et est long de 19 acides

aminés (171-189). DPF3a est caractérisée par deux LCDs supplémentaires se trouvant dans l'IDR-3 au niveau C-terminal. Le troisième LCD (LCD-3) est riche en alanine et est composé de 17 résidus (295-311), alors que le quatrième (LCD-4) est enrichi en arginine ainsi qu'en glycine et est long de seulement 12 acides aminés (337-348). Il est intéressant de noter que tous ces LCDs sont riches en DPRs, réaffirmant leur localisation dans des IDRs (Fig. 22).¹²⁴

Les caractéristiques de reconnaissance moléculaire (MoRFs) représentent de courts segments désordonnés qui peuvent se replier lors de la liaison avec des partenaires d'interaction spécifiques. Les deux isoformes partagent la première MoRF (MoRF-1), prédite comme faisant 14 résidus (69-82) et située à l'extrémité du domaine 2/3 (Fig. 22). A l'instar des LCDs, deux autres MoRFs sont retrouvées dans l'IDR-3 de DPF3a FL. La deuxième MoRF (MoRF-2) se trouve entre le LCD-3 et le LCD-4 et est longue de 19 acides aminés (321-339). La troisième (MoRF-3) chevauche le LCD-4 et fait 14 résidus de long (343-356). Ces deux MoRFs supplémentaires semblent être spécifiques au réseau d'interactions de DPF3a FL, comprenant des partenaires qui ne pourraient pas se lier à DPF3b FL. Bien qu'il s'agisse de prédictions, elles se corrélaient néanmoins toutes en montrant que les deux isoformes de DPF3 appartiennent à la classe des IDPs et que DPF3a FL est davantage désordonnée par rapport à DPF3b FL.¹²⁴

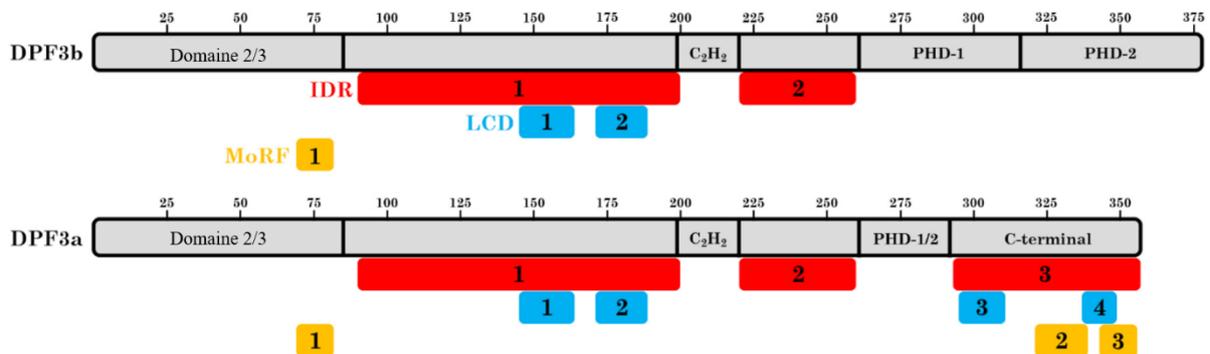


Figure 22 – Distribution et localisation des IDRs (rouge), LCDs (bleu) et MoRFs (jaune) prédits dans les séquences de DPF3b FL et DPF3a FL.¹²⁴

3.2.2. Analyses expérimentales de la nature désordonnée des isoformes de DPF3

Les propriétés structurales ainsi que le comportement *in vitro* des isoformes de DPF3 ont été investigués par diffusion dynamique de la lumière (DLS), dichroïsme circulaire dans l'UV lointain (FCD) et par spectroscopies de fluorescence et d'absorption. Chaque analyse a été réalisée juste après purification afin d'empêcher la protéine de s'agréger ou de se dégrader au cours du temps.

Des résultats similaires ont été obtenus par DLS pour les deux isoformes. En effet, DPF3b FL possède un diamètre hydrodynamique moyen (\bar{D}_h) de 24 ± 7 nm, tandis que le \bar{D}_h de DPF3a FL s'élève à 20 ± 4 nm. A la lueur de leur conformation étendue, ces valeurs sont dans la gamme des protéines désordonnées (de ~ 5 à ~ 22 nm) et sont plus grandes que les diamètres hydrodynamiques des protéines globulaires de séquence équivalente (de ~ 3 à ~ 10 nm).¹²⁹ Bien que ces valeurs apparaissent élevées, elles correspondent aux protéines sous leur forme monomérique. En effet, des études ont montré que le diamètre hydrodynamique d'une IDP est directement lié à sa longueur de séquence. Il en ressort que des protéines désordonnées de longueur de séquence de l'ordre de celles des isoformes de DPF3 (300 à 400 résidus de long) peuvent avoir un \bar{D}_h allant de 16 à 20 nm.^{124,130,131}

La spectroscopie FCD permet d'obtenir des informations sur les structures secondaires présentes au sein des protéines. Les spectres FCD des isoformes de DPF3 (Fig. 23) sont en accord avec les prédictions de désordre en révélant des signatures typiques d'une IDP hybride. Le spectre associé à DPF3b FL montre un minimum à 206 nm, relatif à du désordre ainsi qu'un maximum à 198 nm qui indique la présence de structures ordonnées étant donné qu'une protéine totalement désordonnée ne montre aucun maximum dans cette gamme de longueur d'onde. Un second minimum, plus large, entre 220 et 225 nm correspond à la présence de feuillets β antiparallèles. Il est également possible que cette dernière bande soit liée à des hélices α . En effet, une protéine uniquement composée de ce type de structure montre un minimum vers 222 nm ainsi qu'un autre vers 208 nm. Dans le cas de DPF3b FL, ce dernier pourrait être caché par le minimum intense à 206 nm. Des empreintes similaires sont observées sur le spectre de DPF3a FL, montrant un maximum déplacé à 200 nm associé à des structures ordonnées, un minimum à 206 nm lié à du désordre ainsi qu'un minimum davantage centré vers 225 nm associé à des feuillets β antiparallèles. Il est donc attendu que cette isoforme contienne moins d'hélices α mais autant de désordre et de feuillets β antiparallèles que DPF3b FL.

Selon l'outil BeStSel, qui permet d'estimer le pourcentage en structure secondaire de chacun de ces spectres, les deux isoformes sont majoritairement composées de désordre (69 % pour DPF3b et 72 % pour DPF3a). Il est à noter que le désordre résulte de l'addition des contributions de *turn* et de *coil*. Comme précédemment prédit, ces résultats confirment que DPF3a FL est plus désordonnée que son isoforme. Alors que DPF3b FL est composée de 24 % de feuillets β antiparallèles et 7 % d'hélices α , DPF3a FL contient 25 % de feuillets β antiparallèles et seulement 3 % d'hélices α . Ces estimations BeStSel sont en accord avec les

profils des spectres. Le pourcentage élevé en feuillets β antiparallèles s'explique par la présence des ZnFs C_2H_2 et PHD au sein des deux isoformes.¹²⁴

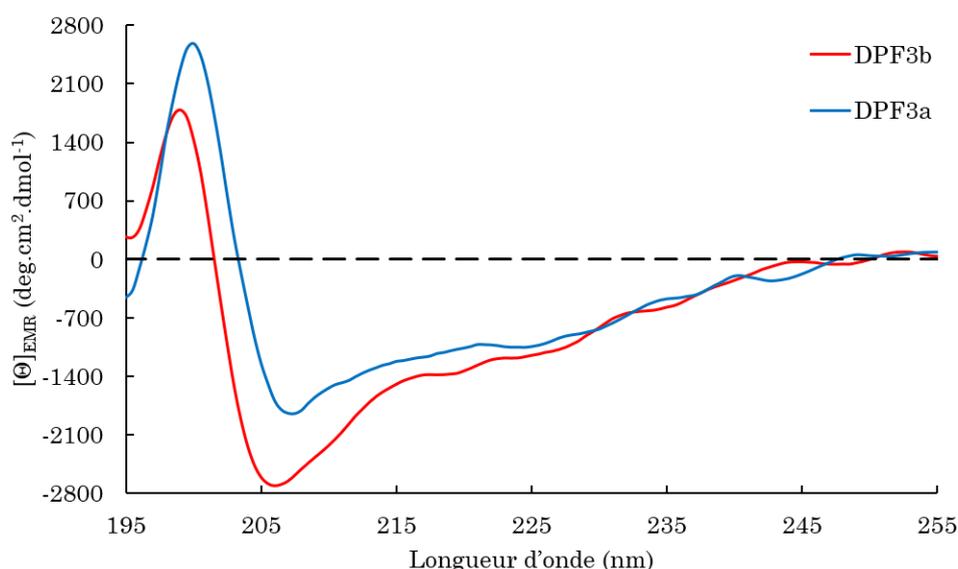


Figure 23 – Spectres FCD de DPF3b FL (rouge) et DPF3a FL (bleu) dans le TPB à ~ 20 °C.

Les spectroscopies de fluorescence, à savoir la fluorescence intrinsèque du tryptophane (ITF) et celle de la tyrosine (ITyrF), sont des techniques utiles pour appréhender localement l'état de repliement des protéines. En effet, après excitation à 295 nm, l'émission des tryptophanes (Trp) est solvatochromique et influencée par leur environnement : elle apparaît respectivement vers le rouge (*red shift*) ou vers le bleu (*blue shift*) dans un environnement polaire ou plus hydrophobe. Plus précisément, i) les résidus Trp totalement exposés à un solvant aqueux émettent à 350 nm, ii) ceux partiellement exposés à 330-340 nm, iii) ceux interagissant avec les groupements polaires voisins de 315 à 330 nm et iv) ceux situés dans un noyau apolaire vers 308-315 nm (Fig. 25).¹³²

DPF3b FL contient cinq résidus Trp dans sa séquence et montre une bande d'émission à 335 nm, supposant que ces résidus sont partiellement exposés au solvant et ne sont pas enfouis dans un noyau hydrophobe, comme cela serait le cas pour une protéine globulaire qui verrait sa bande d'émission déplacée vers les plus basses longueurs d'onde (Fig. 24A). Comparativement, DPF3a FL ne contient que deux résidus Trp et exhibe une bande d'émission à 340 nm ainsi qu'une large épaule à plus haute longueur d'onde, ce qui est mis en évidence par la différence entre les deux spectres. Ces résultats indiquent que les Trp de DPF3a FL sont plus exposés au solvant ou à des résidus polaires. Les résidus Trp⁵⁶ et Trp⁷⁹ sont communs aux

deux isoformes et se trouvent dans le domaine 2/3. Cette région, bien qu'ordonnée, possède un PPID variant entre 0.2 et 0.5, montrant une certaine flexibilité et pouvant expliquer l'état d'exposition partiel des Trp au solvant. Pour l'isoforme DPF3b FL, les trois Trp supplémentaires (Trp³¹¹, Trp³⁵⁶ et Trp³⁶⁴) sont localisés dans les domaines PHD-1 et PHD-2, particulièrement ordonnés, correspondant à la bande d'émission à une plus basse longueur d'onde.¹²⁴

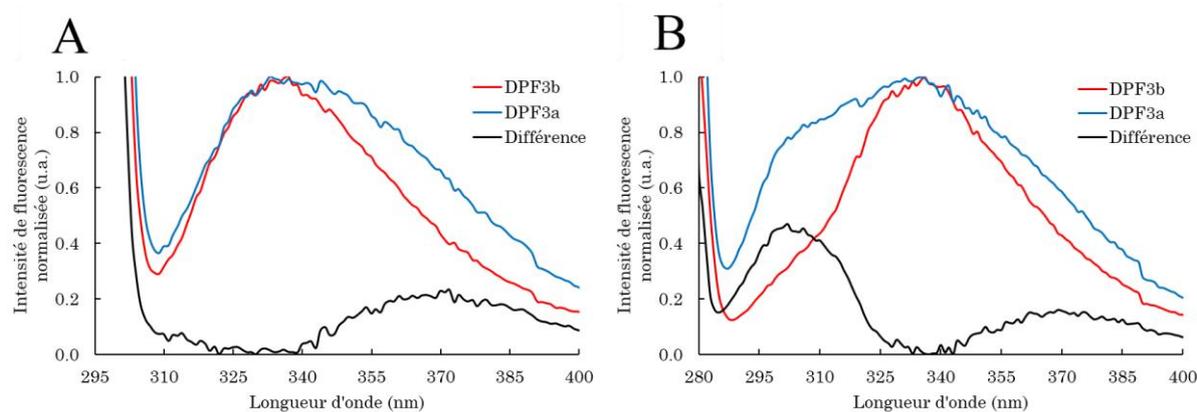


Figure 24 – Fluorescence intrinsèque de DPF3b FL et DPF3a FL dans le TPB à ~ 20 °C. (A) Spectres ITF normalisés ($\lambda_{ex} = 295 \text{ nm}$, $sw = 10 \text{ nm}$) de DPF3b FL (rouge) et DPF3a FL (bleu) ainsi que la différence entre les deux spectres (noir). (B) Spectres ITyrF normalisés ($\lambda_{ex} = 275 \text{ nm}$, $sw = 10 \text{ nm}$) de DPF3b FL (rouge) et DPF3a FL (bleu) ainsi que la différence entre les deux spectres (noir).

Complémentairement à l'ITF, les spectres ITyrF permettent d'étudier l'environnement des tyrosines (Tyr). En effet, l'excitation sélective des résidus Tyr à 275 nm résulte en une émission vers 305-310 nm. Cependant, l'ITyrF est moins sensible que l'ITF, ce qui engendre souvent des signatures se chevauchant lorsqu'une protéine contient à la fois des Trp et des Tyr. Lorsque les Tyr sont proches des Trp, cela amène même au phénomène de transfert d'énergie par résonance de fluorescence (FRET pour *fluorescence resonance energy transfer*). Alors que DPF3b FL contient 13 Tyr, DPF3a FL en possède seulement 10. Le spectre ITyrF de DPF3b FL dévoile une bande à 336 nm ainsi qu'une légère épaule à 305 nm associée à des Tyr exposées au solvant et donc indicatives de désordre (Fig. 24B). La bande principale, à 336 nm, trouve probablement son origine dans le FRET entre des résidus Tyr et les Trp proches dans l'espace et/ou la séquence. Le spectre ITyrF de DPF3a FL montre une épaule plus intense entre 300 et 305 nm, comme le souligne le spectre de différence entre les deux isoformes, prouvant que ses Tyr sont davantage exposées au solvant. La bande d'émission majoritaire est sensiblement la même, d'une part, que celle du spectre ITyrF de DPF3b FL montrant une bande

centrée vers 336 nm et, d'autre part, que celle du spectre ITF de DPF3a FL s'étendant jusqu'à 400 nm. Ces deux tendances suggèrent, à l'instar de DPF3b FL, un phénomène de FRET entre les résidus aromatiques de DPF3a FL grâce à leur proximité spatiale.¹²⁴

Afin de compléter ces résultats expérimentaux, il est possible de prédire l'accessibilité au solvant de chaque résidu aromatique au sein des isoformes. L'analyse de séquence par le serveur PaleAle donne à voir que les deux Trp communs aux isoformes (Trp⁵⁶ et Trp⁷⁹), autrement dit les deux seuls résidus Trp de DPF3a FL, sont estimés totalement ou partiellement exposés au solvant, ce qui est cohérent avec la bande d'émission à haute longueur d'onde ainsi que l'épaule observée entre 340 et 400 nm sur le spectre ITF de DPF3a. Les trois Trp supplémentaires dans la séquence de DPF3b FL sont prédits partiellement exposés au solvant, expliquant la bande d'émission centrée à 335 nm (Fig. 25).

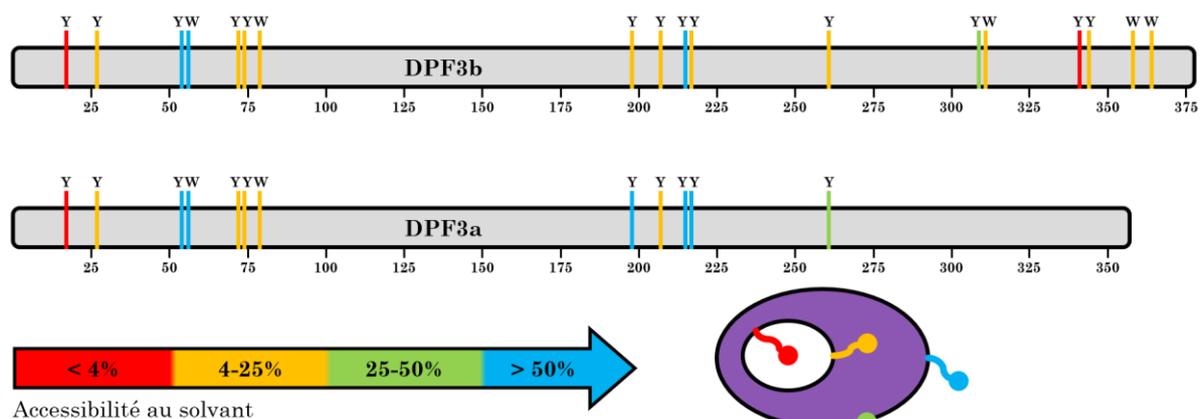


Figure 25 – Distribution et localisation des tryptophanes (W) et tyrosines (Y) dans les séquences de DPF3b et DPF3a. Les résidus sont classés en fonction de leur accessibilité au solvant : très enfoui (rouge), quelque peu enfoui (jaune), quelque peu exposé (vert) ou très exposé (bleu). Les états d'exposition des résidus aromatiques sont représentés schématiquement dans une protéine (mauve) en bas à droite : résidus aromatiques dans une poche hydrophobe (rouge), enfouis près de résidus polaires (jaune), une face (vert) et deux faces exposées au solvant (bleu).¹²⁴

En ce qui concerne les résidus Tyr, DPF3b FL possède 10 Tyr sur 13 (77 %) considérées comme enfouies contre 5 Tyr sur 10 (50 %) pour DPF3a FL. L'intensité supérieure de la bande à 305 nm sur le spectre ITyrF de DPF3a FL provient probablement de ces cinq autres Tyr exposées au solvant contre seulement trois pour DPF3b FL. Plus précisément, les résidus Tyr¹⁷ et Tyr²⁷, communs aux deux isoformes, sont prédits comme très et quelque peu enfouies et devraient par conséquent être impliqués dans le repliement du domaine 2/3. Le FRET Trp-Tyr impliquerait les Tyr⁵⁴, Tyr⁷² et Tyr⁷⁴ étant donné leur proximité avec les Trp⁵⁶ et Trp⁷⁹. Bien que les deux isoformes partagent les dix premiers résidus Tyr, une certaine divergence est observée. En effet, alors que la plupart des résidus Tyr du domaine C₂H₂ sont considérés

comme enfouis pour DPF3b FL, ces résidus sont prédits exposés pour DPF3a FL. Dans le cas de DPF3b FL, cela pourrait être une indication d'un repliement local, alors qu'inversement cela contribuerait à l'intensité d'émission plus élevée observée à 305 nm pour DPF3a FL. De même, du FRET se produit vraisemblablement au sein du tandem PHD de DPF3b entre les résidus Tyr³⁴¹, Tyr³⁴⁴, Trp³⁵⁸ et Trp³⁶⁴ ainsi qu'entre les Tyr³⁰⁹ et Trp³¹¹, ce qui favoriserait la réduction de la fluorescence des Tyr dans le spectre ITyrF de DPF3b FL.¹²⁴

Les propriétés d'absorption UV ont été employées pour sonder le repliement des protéines par le calcul de l'indice de repliement (FI pour *foldedness index*) à l'aide des rapports de valeurs d'absorbance : $FI = (A^{280}/A^{275}) + (A^{280}/A^{258})$. La plupart des protéines repliées ont un $FI > 2.5$. Lors de la dénaturation, et donc de la formation d'une conformation étendue, le FI diminue. Cela s'explique par le déplacement la longueur d'onde d'absorbance (λ_{abs}) de 280 à 258 nm, diminuant ainsi la contribution à 280 nm.¹³³ Comme expliqué précédemment, les deux isoformes de DPF3 FL possèdent des propriétés d'absorption singulières, présentant une absorption maximale à ~ 258 nm (Fig. 18). Il en résulte un faible FI de ~ 1.7 et ~ 1.4 pour DPF3b FL et DPF3a FL, ce qui prouve, une nouvelle fois, leur caractère désordonné. Le FI plus faible de DPF3a FL suggère un contenu plus élevé en désordre, ce qui est cohérent avec les observations précédentes.¹²⁴

3.2.3. Investigation de la propension à l'agrégation des isoformes de DPF3

Les IDPs étant impliquées dans les maladies associées à l'agrégation protéique et DPF3a FL ayant une tendance à se fibriller en feuillets β , le comportement d'agrégation de DPF3b FL a été étudié *in vitro* et comparé à celui de DPF3a FL. Pour ce faire, chaque échantillon, à une concentration de 10 μ M, a été incubé à ~ 20 °C pendant trois à quatre jours et subséquemment analysé par DLS, FCD et spectroscopies d'émission et d'absorption.

Afin de savoir si DPF3b FL s'agrège spontanément au cours du temps, des mesures DLS ont été réalisées après 120 h d'incubation et un \bar{D}_h de 245 ± 8 nm a été obtenu. Ce résultat prouve l'agrégation de la protéine car ce diamètre hydrodynamique moyen est environ dix fois supérieur au \bar{D}_h obtenu juste après purification ($\bar{D}_h = 24 \pm 7$ nm).

Les changements conformationnels et de composition en structure secondaire au cours de l'agrégation ont été suivis par FCD. Des spectres ont été enregistrés après 24 h, 48 h et 72 h pour DPF3a FL ainsi qu'après 96 h pour DPF3b FL (Fig. 26). Chaque spectre a été déconvolué à l'aide du calculateur de structure secondaire BeStSel afin d'évaluer l'évolution

en motifs structuraux au cours de l'agrégation (Tableau 1). En ce qui concerne DPF3b FL (Fig. 26A), des modifications structurales sont déjà visibles après 24 h de par l'importante diminution en intensité du minimum à 206 nm et le déplacement du maximum de 198 à 200 nm. Cela est associé à une perte en hélices α , qui passent de 7 à 2 %, et à un enrichissement en feuillets β antiparallèles de 24 à 29 %. Le pourcentage de désordre, incluant les *turns* et le *coil*, reste inchangé. Les spectres FCD après 48 et 96 h tendent progressivement vers une seule bande négative centrée à ~ 226 nm, montrant que la conformation de DPF3b FL continue de changer. Il a été rapporté que ce comportement, déjà observé pour DPF3a FL, est lié à la formation de feuillets β antiparallèles et « twistés » typiques des fibrilles amyloïdes.^{43,134} Plus précisément, après 48 h d'incubation, le pourcentage en feuillets β antiparallèles augmente jusqu'à 32 %, le pourcentage de désordre chute à 66%, tandis que le pourcentage en hélices α reste identique (2%). L'enrichissement en feuillets β antiparallèles se poursuit dans les 48 h suivantes, résultant en 62 % de désordre, toujours 2 % d'hélices α et 36 % de feuillets β antiparallèles. En seulement 96 h, la teneur en feuillets β a augmenté de 24 à 36 % au détriment, dans un premier temps, des hélices α et, dans un second temps, du désordre. Il s'agit de la première preuve de la tendance de DPF3b FL à s'agréger spécifiquement en feuillets β antiparallèles. En effet, le pourcentage associé aux feuillets β parallèles reste nul tout au long de l'agrégation.¹²⁴

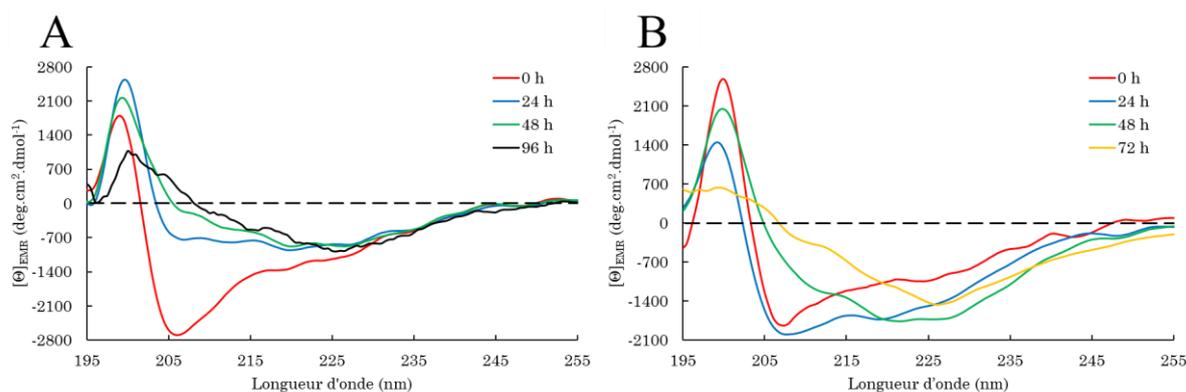


Figure 26 – Spectres FCD de (A) DPF3b FL et (B) DPF3a FL dans le TPB à ~ 20 °C après 0 h (rouge), 24 h (bleu), 48 h (vert), 72 h d'incubation pour DPF3a FL (jaune) et 96 h pour DPF3b FL (noir).

Quant à DPF3a FL (Fig. 26B), des changements conformationnels comparables sont observés au cours du temps. Après 24 h d'incubation, à l'instar de DPF3b FL, survient un enrichissement en feuillets β antiparallèles (de 25 à 31 %). Toutefois, une différence se manifeste par un gain en hélices α passant de 3 à 6 % et caractérisé par les minima plus marqués

à 207 et 220 nm. Contrairement à DPF3b FL, DPF3a FL voit son pourcentage en désordre diminuer de 9 % au total. Il peut être noté que la formation d'intermédiaires en hélices α a déjà été mise en évidence chez l' α -syn au cours de sa fibrillation.¹³⁵ Dans les 24 h suivantes, un comportement similaire à DPF3b FL est observé, s'accompagnant d'une perte en hélices α , un enrichissement en feuillets β antiparallèles ainsi qu'une faible diminution du désordre. Ces changements conformationnels sont caractérisés, sur le spectre, par un déplacement du minimum de 207 vers 224 nm. Après 48 h, il y a 35 % de feuillets β antiparallèles alors qu'il y en avait 36 % après 96 h pour DPF3b FL, suggérant que DPF3a FL se réarrange plus vite et que sa fibrillation est plus rapide. Après 72 h, les hélices α sont totalement perdues et le gain en feuillets β antiparallèles s'élève à 39 %.¹²⁴

Tableau 1 - Estimations (BeStSel) de la composition en structure secondaire et valeurs de RMSD associées de DPF3b FL et DPF3a FL dans le TPB à ~ 20 °C après 0 h, 24 h, 48 h, 96 h d'incubation pour DPF3b FL et 72 h pour DPF3a FL.

Protéine (Temps d'incubation)	RMSD	Composition en structure secondaire (%)			
		Hélice α	Feuillet β antiparallèle	Feuillet β parallèle	Désordre <i>Turn</i> <i>Coil</i>
DPF3b (0 h)	0.03	7	24	0	16 53
DPF3b (24 h)	0.03	2	29	0	16 53
DPF3b (48 h)	0.02	2	32	0	16 50
DPF3b (96 h)	0.03	2	36	0	16 46
DPF3a (0 h)	0.03	3	25	0	18 54
DPF3a (24 h)	0.02	6	31	0	16 47
DPF3a (48 h)	0.02	3	35	0	15 47
DPF3a (72 h)	0.03	0	39	0	16 45

Complémentairement à la spectroscopie CD, les fluorescences intrinsèques ITF et ITyrF permettent de suivre les modifications conformationnelles au cours de l'agrégation (Fig. 27). En ce qui concerne DPF3b FL, son spectre ITF (Fig. 27A) donne à voir, endéans 72 h, un déplacement de la bande d'émission de 335 vers 333 nm, accompagné d'une augmentation constante de l'intensité d'émission. Ces deux phénomènes sont associés à des Trp plus enfouis en sein de la protéine et, par conséquent, moins exposés au solvant. Ce résultat concorde avec la tendance en FCD et suggère que DPF3b FL se réarrange en formant des conformations plus compactes avant de se fibriller. En ITyrF, la bande d'émission associée au FRET Trp-Tyr est légèrement déplacée vers 333 nm et augmente en intensité au cours du temps, signifiant que les Tyr se trouvent dans un environnement moins polaire (Fig. 27C).¹³⁶

La légère épaule à 305 nm disparaît au cours de l'incubation, indiquant une augmentation du FRET Trp-Tyr due à des réarrangements de conformation.¹²⁴

Un comportement similaire est observé sur les spectres ITF de DPF3a FL (Fig. 27B). Après 24 h, la bande d'émission commence à être déplacée vers le bleu tandis que l'intensité de fluorescence n'augmente pas. L'épaule entre 340 et 400 nm disparaît également. Dans les 48 h suivantes, l'intensité d'émission augmente et la bande d'émission est déplacée vers 332 nm, montrant que les résidus Trp se trouvent dans un environnement plus apolaire. Concernant l'ITyrF (Fig. 27D), aucune augmentation de l'intensité d'émission n'est observée après 24 h mais la contribution à 305 nm diminue. De 24 à 72 h, l'émission de Tyr à 305 nm est définitivement perdue ainsi que l'élargissement de bande entre 340 et 400 nm. De plus, la bande associée au FRET Trp-Tyr s'accroît. Ces résultats, aussi bien pour DPF3b FL que pour DPF3a FL, sont en accord avec les spectres FCD. En effet, les réarrangements conformationnels et le gain en feuillets β au fil du temps s'accompagnent d'un enfouissement et d'un rapprochement des résidus Trp et Tyr au sein un noyau plus hydrophobe.¹²⁴

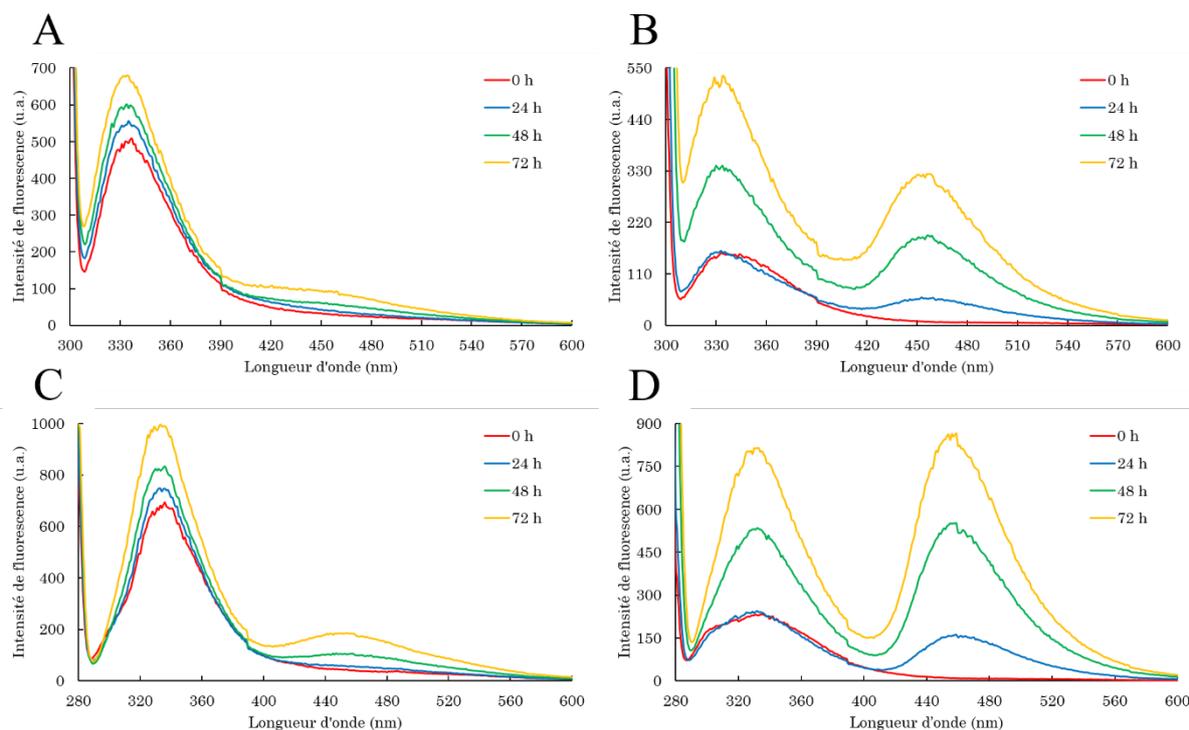


Figure 27 – Fluorescence intrinsèque de DPF3b FL et DPF3a FL dans le TPB à ~ 20 °C après 0 h (rouge), 24 h (bleu), 48 h (vert) et 72 h (jaune) d'incubation. Spectres ITF ($\lambda_{ex} = 295$ nm, $sw = 10$ nm) de (A) DPF3b FL et de (B) DPF3a FL. Spectres ITyrF ($\lambda_{ex} = 275$ nm, $sw = 10$ nm) de (C) DPF3b FL et de (D) DPF3a FL.

Sur l'ensemble des spectres ITF et ITyrF apparaît également une deuxième bande d'émission à 456 nm dont l'intensité augmente légèrement au cours du temps pour DPF3b FL et fortement pour DPF3a FL. Une telle bande a déjà été détectée chez les agrégats de la transthyrétine, une protéine impliquée dans la protéolyse et l'amyloïdose contenant une région N-terminale désordonnée, et de DPF3a.^{43,137} Elle s'apparente à un phénomène récemment décrit dans la littérature désignée par autofluorescence dans le bleu profond (dbAF). Cette autofluorescence a déjà été reportée chez des IDPs amyloïdogènes telles que l' α -syn et tau.¹³⁸ Bien que l'origine de la dbAF demeure débattue, de récentes études suggèrent que cette fluorescence viendrait de transitions électroniques de basse énergie causées par la délocalisation des électrons le long des ponts hydrogène présents dans les feuillettes β croisés.¹³⁹⁻¹⁴¹ C'est pourquoi, il semble que cette bande d'émission en ITF et ITyrF est un indicateur de l'agrégation et, en particulier, de la fibrillation de DPF3.¹²⁴

De sorte à mieux comprendre l'origine de la bande dbAF ainsi que les longueurs d'onde d'excitation qui y contribuent, les spectres d'excitation ($\lambda_{em} = 456$ nm) de DPF3b FL (Fig. 28A) et DPF3a FL (Fig. 28B) ont été enregistrés. Lors des 24 premières heures, le spectre associé à DPF3b FL révèle deux bandes dont les intensités augmentent avec le temps comme c'était le cas sur les spectres ITF et ITyrF. La première, assez large, est retrouvée entre 235 et 270 nm, expliquant la bande d'émission lors de l'excitation des Trp et des Tyr, et la deuxième est centrée à 355 nm. Au fil du temps, la première bande semble se séparer en deux. Cela est d'autant plus visible après 168 h, où le spectre présente clairement trois bandes d'excitation. La première vers 235 nm est attribuée à l'absorption des liaisons peptidiques^{140,142}, la seconde vers 270 nm correspond aux résidus aromatiques et la troisième déplacée à 400 nm présente une intensité plus importante et est la contribution dbAF principale. Ces trois bandes sont présentes dès 24 h d'incubation pour DPF3a FL (Fig. 28B). A l'instar de la bande dbAF sur les spectres ITF et ITyrF, l'intensité de ces bandes augmentent aussi avec le temps pour DPF3a FL. La différence majeure entre les deux isoformes est la troisième bande d'excitation qui est d'abord à 355 nm pour DPF3b FL et puis à 400 nm après 168 h, alors qu'elle est directement à 400 nm pour DPF3a FL. Comme cette dernière apparaît se fibriller plus rapidement que l'autre isoforme, l'excitation à 400 nm est considérée comme typique d'une fibrillation amyloïde directe. Au contraire, les signatures d'excitation dbAF de DPF3b FL semblent décrire des transitions structurales plus complexes, qui ne sont pas encore liées à un caractère amyloïde après 72 h mais qui le deviennent après 168 h. Par conséquent, le déplacement du maximum

d'excitation de dbAF de 355 à 400 nm serait un indicateur de l'agrégation en fibrilles amyloïdes.¹²⁴

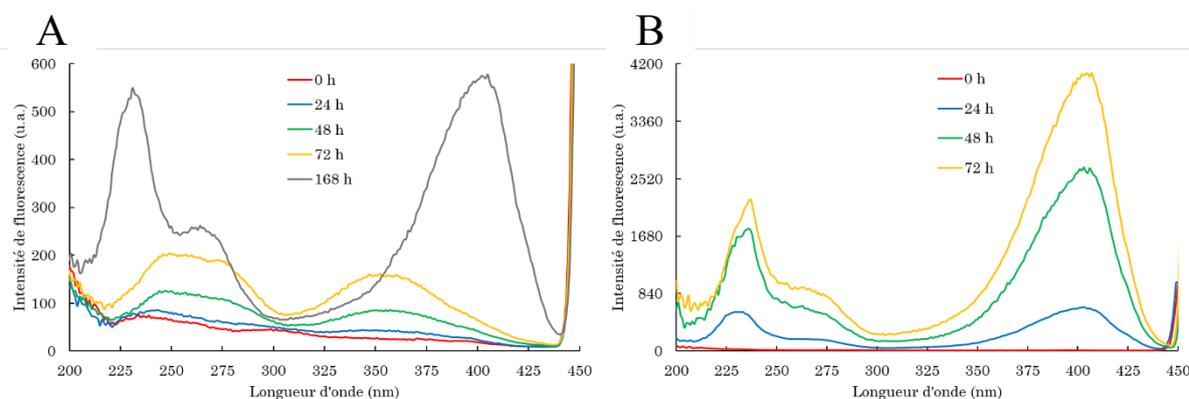


Figure 28 – Spectres d'excitation dbAF ($\lambda_{em} = 456 \text{ nm}$, $sw = 10 \text{ nm}$) de (A) DPF3b FL et (B) DPF3a FL dans le TPB à $\sim 20 \text{ }^\circ\text{C}$ après 0 h (rouge), 24 h (bleu), 48 h (vert) et 72 h (jaune) d'incubation ainsi qu'à 168 h (gris) pour DPF3b FL.

La liaison de colorants spécifiques aux fibrilles amyloïdes permet de suivre et de confirmer la fibrillation de la protéine. La thioflavine T (ThT) est un fluorophore extrinsèque qui, en se liant aux fibrilles amyloïdes et après excitation à 440 nm, exhibe une émission typique vers 485 nm (Fig. 29A). La spécificité de la liaison ThT découle de la structure des fibrilles, qui présentent des rainures au niveau de leur surface. Le mécanisme le plus accepté décrit que la rotation libre de la liaison carbone-carbone entre les deux cycles de la ThT est bloquée lorsque le fluorophore se lie aux feuillettes β .^{121,143}

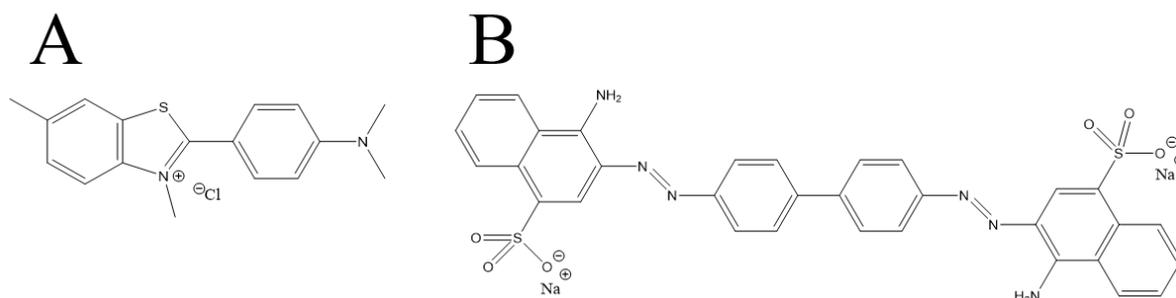


Figure 29 – Structures moléculaires de (A) la ThT et (B) du RC.

Le spectre d'émission de DPF3b FL en présence de ThT (Fig. 30A) est cohérent avec les résultats précédents. En effet, aucune bande d'émission n'apparaît durant les 72 premières heures, indiquant la formation d'agrégats non amyloïdes bien que riches en feuillettes β .

Cependant, après 168 h, une bande d'émission à 485 nm apparaît, validant la formation de fibrilles amyloïdes. Cette tendance concorde avec les émissions obtenues en dbAF. Comme l'on pouvait s'y attendre en raison de la propension plus élevée de DPF3a FL à la fibrillation, les spectres ThT associés (Fig. 30B) montrent une augmentation de l'intensité dès 24-48 h, confirmant ainsi la formation d'agrégats amyloïdes dès les premiers jours d'incubation.¹²⁴

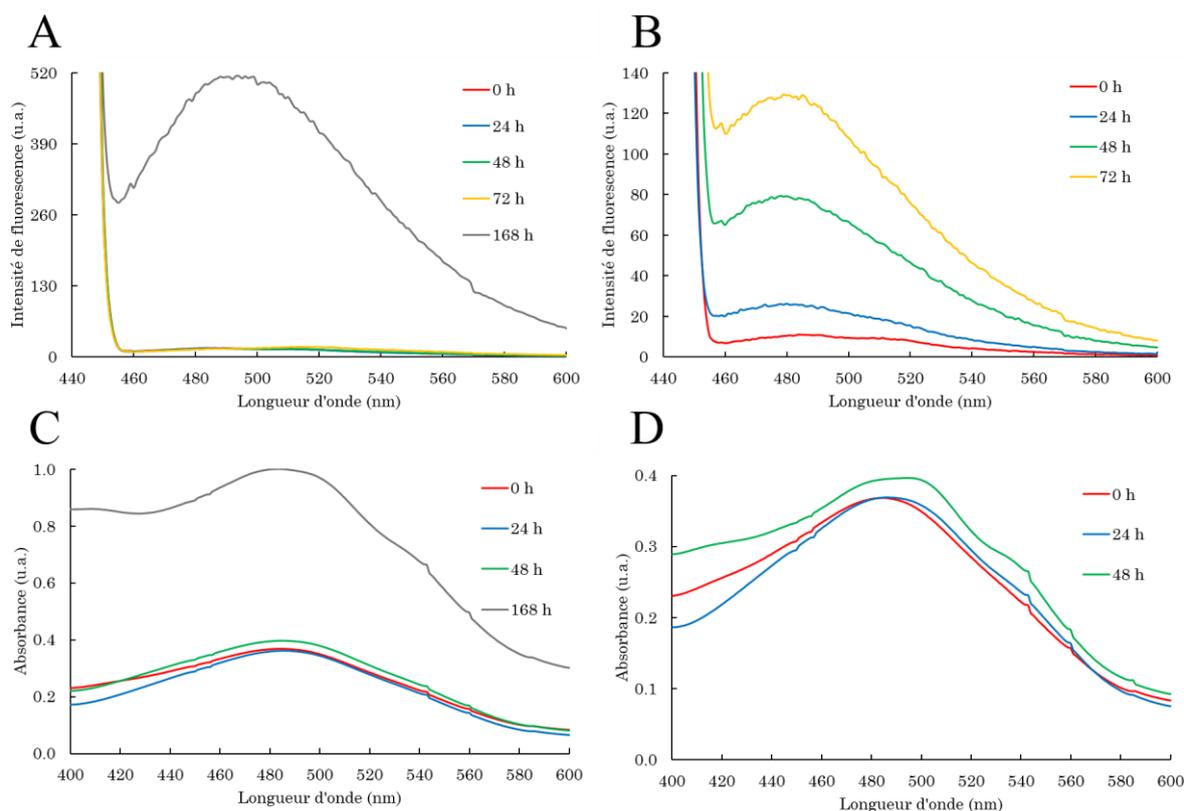


Figure 30 – Essais de liaison de colorants spécifiques aux fibrilles amyloïdes. Spectres d'émission de la ThT ($\lambda_{ex} = 440$ nm, $sw = 10$ nm, $C_{ThT} = 10 \mu M$) de (A) DPF3b FL et de (B) DPF3a FL dans le TPB à ~ 20 °C après 0 h (rouge), 24 h (bleu), 48 h (vert), 72 h (jaune) d'incubation ainsi que 168 h (gris) pour DPF3b FL. Spectres d'absorption UV-visible du RC ($C_{RC} = 10 \mu M$) de (C) DPF3b FL et (D) DPF3a FL dans le TPB à ~ 20 °C après 0 h (rouge), 24 h (bleu) et 48 h (vert) d'incubation ainsi que 168 h (gris) pour DPF3b FL.

Complémentairement, le chromophore rouge Congo (RC) est utilisé car il possède des propriétés similaires à la ThT (Fig. 29B). La liaison du RC aux fibrilles amyloïdes est connue pour entraîner des modifications de son spectre d'absorption, telles qu'un déplacement de la λ_{abs} de 490 à 500-510 nm, l'émergence d'un épaulement à 540 nm et un effet hyperchrome général.¹⁴⁴ Le spectre associé à DPF3b FL (Fig. 30C) ne montre qu'une légère augmentation d'intensité après 48 h. L'absence de déplacement de la bande d'absorption confirme l'absence de fibrilles amyloïdes à ce stade de l'agrégation. Une nouvelle fois, la formation de fibrilles

amyloïdes arrive plus tardivement pour DPF3b FL. En effet, après 168 h, la liaison de DPF3b FL au RC résulte en une augmentation de l'absorbance ainsi qu'en l'apparition de l'épaule à 540 nm. Comparativement, DPF3a FL présente toutes les signatures amyloïdes caractéristiques lors de l'ajout de RC après seulement 48 h (Fig. 30D). Le λ_{abs} est déplacé vers 508 nm, la bande présente un épaulement à 540 nm et l'intensité d'absorption est augmentée, ce qui appuie la formation d'un noyau amyloïde endéans 48 h.¹²⁴

En sachant que sur les spectres d'excitation dbAF ($\lambda_{\text{ém}} = 456 \text{ nm}$), la bande d'excitation la plus importante est à 400 nm, les spectres d'émission dbAF ($\lambda_{\text{ex}} = 400 \text{ nm}$) donnent lieu à des bandes d'émission d'une plus grande intensité. Dès lors, un profil cinétique peut être construit pour DPF3b FL et DPF3a FL en prenant la valeur de l'intensité dbAF à 456 nm après chaque temps d'incubation (Fig. 31A). Il est reconnu dans la littérature que des profils cinétiques sigmoïdaux en trois phases (latence, élongation et stationnaire) sont caractéristiques des protéines amyloïdogènes, ce qui est obtenu pour les deux isoformes.^{145,146} Concernant DPF3b FL, seule la première phase est visible sur une période de 168 h. En effet, la phase de latence, qui correspond aux transitions structurales, nucléations et assemblages d'oligomères, dure environ 120 h. En revanche, les trois phases sont visibles pour DPF3a FL. La phase de latence est plus courte (30 h), ce qui est cohérent avec les signatures dbAF augmentant plus rapidement pour cette isoforme. La phase d'élongation, correspondant aux regroupements des oligomères et à la formation de protofibrilles, est associée à l'augmentation rapide de l'intensité jusqu'à atteindre un plateau lié à la phase stationnaire où des fibrilles matures sont formées. Cette dernière est atteinte après 96 h pour DPF3a FL, ce qui correspond à un temps d'élongation de 66 h.

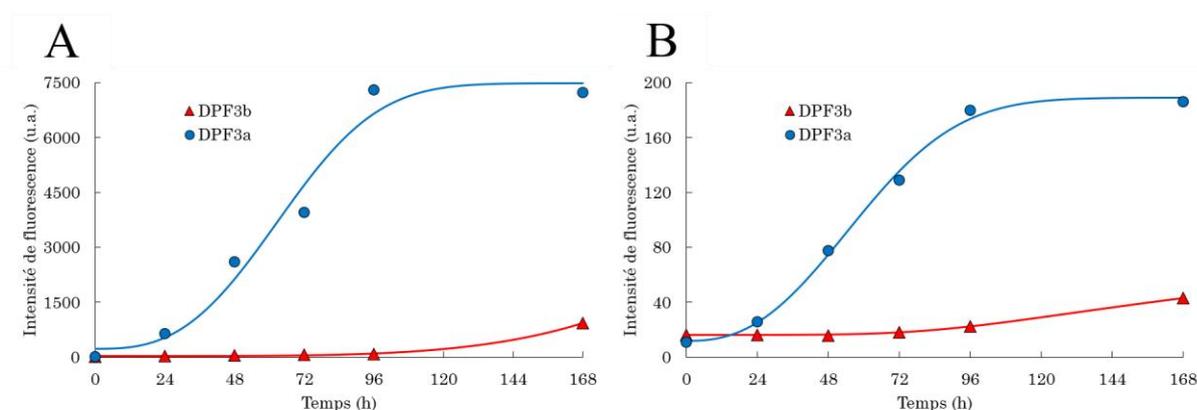


Figure 31 – Courbes cinétiques de l'agrégation de DPF3b FL (triangle et rouge) et de DPF3a FL (cercle et bleu) dans le TPB à $\sim 20 \text{ }^\circ\text{C}$ par suivi de (A) l'émission dbAF ($\lambda_{\text{ex}} = 400 \text{ nm}$, $\lambda_{\text{ém}} = 456 \text{ nm}$, $sw = 10 \text{ nm}$) et de (B) la fluorescence de la ThT ($\lambda_{\text{ex}} = 440 \text{ nm}$, $\lambda_{\text{ém}} = 485 \text{ nm}$, $sw = 10 \text{ nm}$).

La fluorescence de la ThT a été utilisée afin de corroborer les résultats obtenus par les cinétiques de dbAF. L'intensité à 485 nm est mise en relation avec le temps d'incubation de sorte à obtenir des cinétiques également en forme de sigmoïde (Fig. 31B). A nouveau, pour DPF3b FL, seule la phase d'élongation, estimée à 110 h, est visible endéans 168 h. Pour DPF3a FL, les trois phases sont derechef visibles et leurs durées sont proches de celles obtenues par les cinétiques d'autofluorescence. En effet, les phases de latence et d'élongation durent respectivement 28 et 68 h. De ce fait, 96 h sont nécessaires pour atteindre le plateau et la phase stationnaire. Bien qu'il soit difficile de comparer les deux isoformes entre elles, la différence de temps de latence confirme que DPF3a FL se fibrille plus rapidement que DPF3b FL. De plus, les deux techniques ayant révélé des résultats similaires, l'émission dbAF apparaît tout aussi efficace que la ThT pour étudier les cinétiques de fibrillation.

3.2.4. Morphologie des agrégats des isoformes de DPF3

Bien que les techniques spectroscopiques aient permis de décrire les changements conformationnels s'opérant au cours de l'agrégation, il est nécessaire d'utiliser la microscopie électronique à transmission (TEM) afin de différencier les formes d'agrégats et de comprendre le mécanisme de fibrillation. Les images TEM de DPF3b FL (Fig. 32) et DPF3a FL (Fig. 33) ont été prises après 168 h d'incubation à ~ 20 °C afin de s'assurer de la formation des fibrilles. Pour les deux isoformes, des unités de nucléation sphériques (SNU pour *spherical nucleation units*) se trouvent à proximité de protofibrilles qui semblent s'allonger à partir de l'assemblage de ces SNU de 25 à 40 nm de diamètre (Figs. 32A et 33A). Ces oligomères sphériques de haut ordre ont déjà été détectés chez tau.¹⁴⁷ Des fibrilles isolées, appelées filaments droits (SFs pour *straight filaments*), sont identifiées chez DPF3b FL (Fig. 32B) et DPF3a FL (Fig. 33D) et possèdent une largeur de 18 à 25 nm, ce qui correspond aux largeurs décrites dans la littérature pour les fibrilles amyloïdes.¹⁴⁸ Ces SFs s'auto-assemblent pour former des fibrilles striées et torsadées en forme de ruban (STRFs pour *striated twisted ribbon fibrils*) avec une certaine diversité de taille. Ces STRFs, retrouvées chez DPF3b FL (Fig. 32B) et DPF3a FL (Fig. 33B) sont caractérisées par une série de torsades (*twists*) avec une certaine périodicité, donnant un aspect hélicoïdal aux STRFs.¹⁴⁹ L'aspect strié de ces fibrilles semble provenir de l'alignement de plusieurs SFs (Figs. 32C et 33C). Chaque strie possède une largeur de 3 à 4 nm.

La différence entre les deux isoformes se trouve dans la taille des STRFs. Alors que DPF3b FL contient des torsades d'une largeur de 100 à 110 nm et des régions non torsadées de 200 à 230 nm de large, les zones torsadées et non torsadées de DPF3a FL ont respectivement

des largeurs de 30 à 40 nm et 80 à 90 nm. La distance entre deux torsades est comparable pour chaque isoforme et correspond à dix à quinze fois la largeur d'une torsade. Des rapports similaires ont été décrits pour la protéine tau.^{147,150} Ces STRFs sont retrouvées enchevêtrées au sein des réseaux fibrillaires (Figs. 32D et 33D), typiquement détectées dans les maladies neurodégénératives.¹⁵¹ Les deux isoformes peuvent aussi s'agréger en phase amorphe (Figs. 32E et 33E).¹²⁴

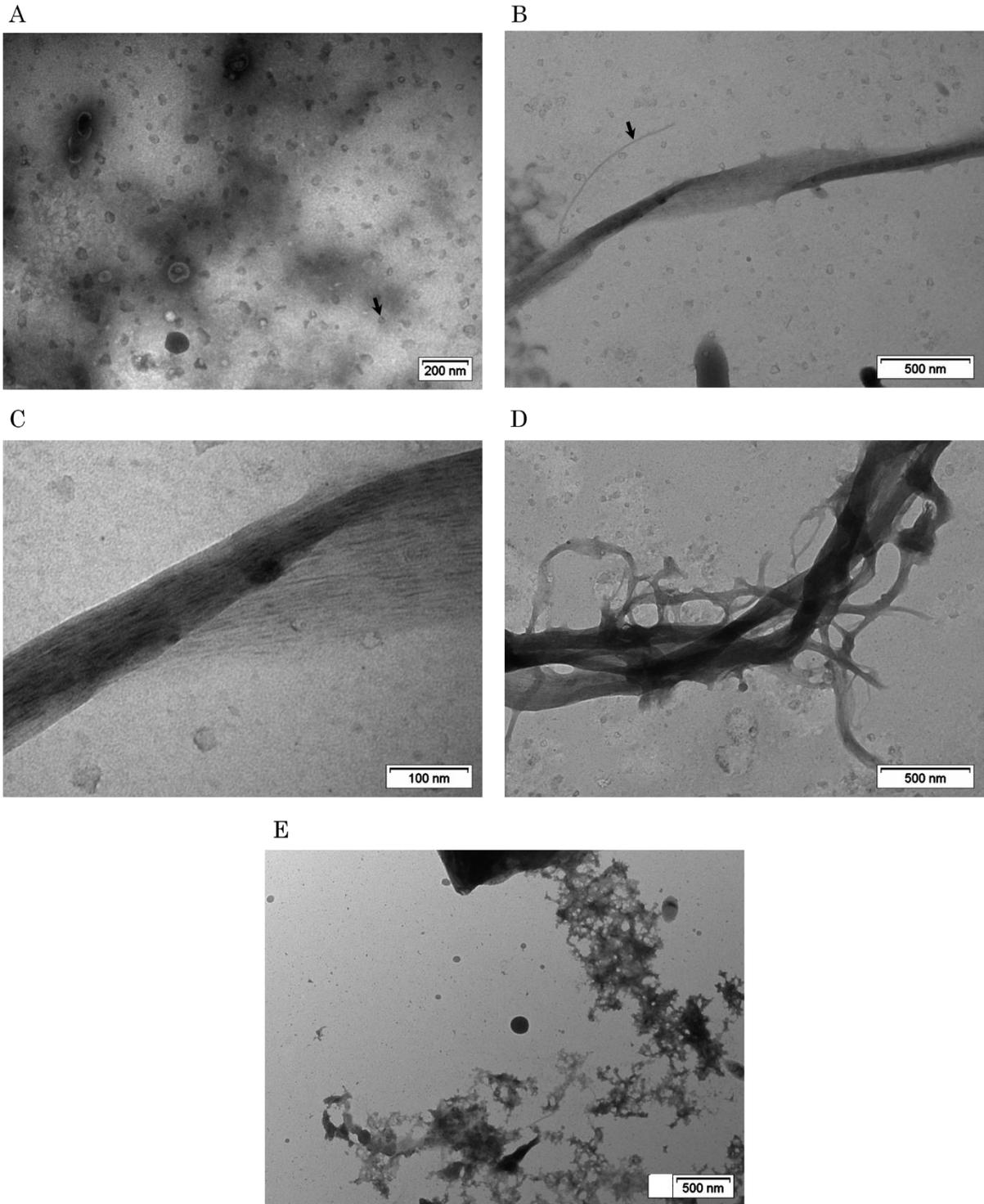


Figure 32 – Micrographies TEM par negative staining de DPF3b FL après 168 h d'incubation à ~ 20 °C dans le TPB. (A) SNUs (une unité est indiquée par une flèche noire) s'assemblant en protofibrilles. (B) STRF avec un SF indiqué par une flèche noire. (C) Aspect strié des STRFs. (D) Réseau de fibrilles enchevêtrées. (E) Agrégats amorphes. Sur chaque image, l'échelle est indiquée dans le coin inférieur droit.

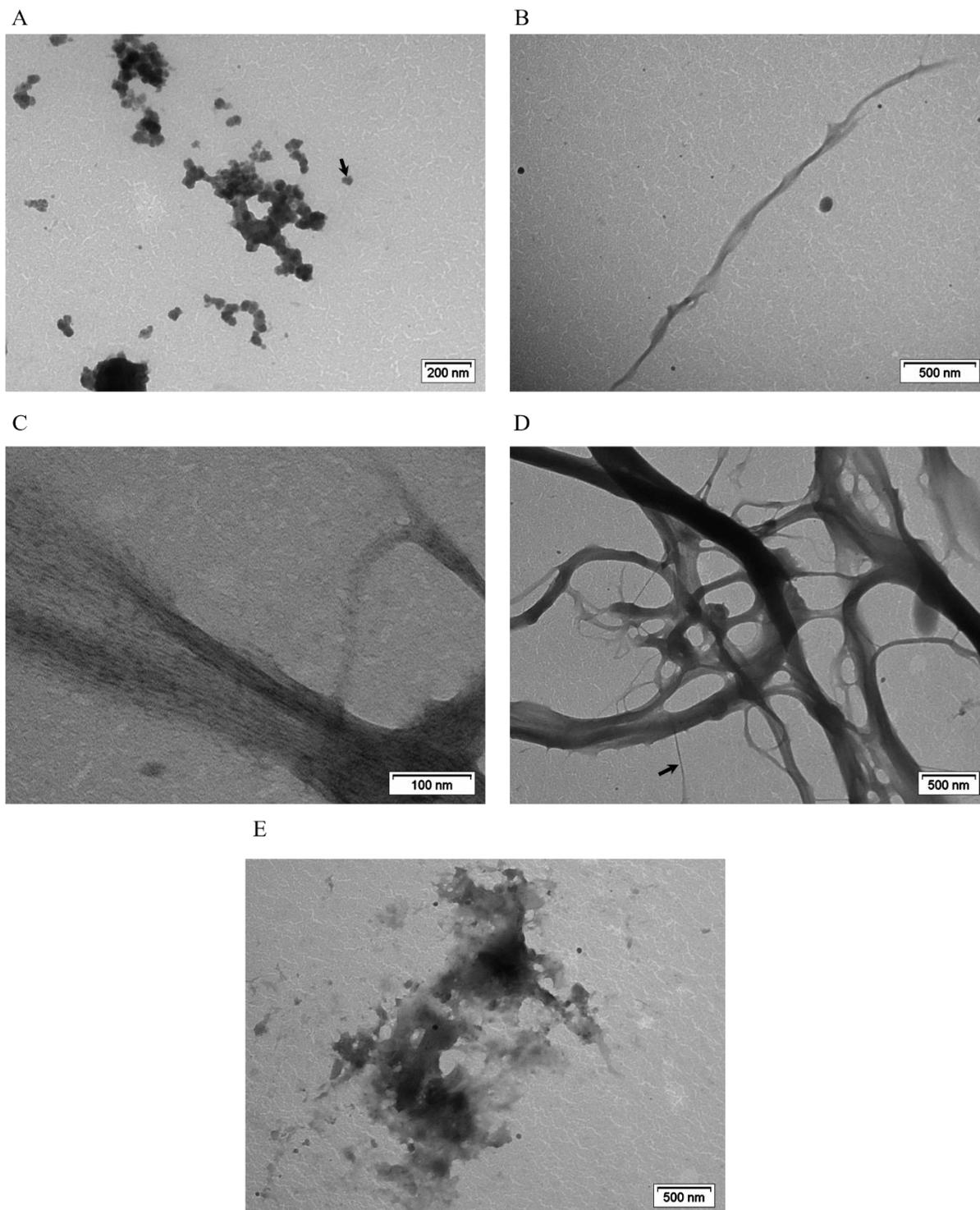


Figure 33 – Micrographies TEM par negative staining de DPF3a FL après 168 h d'incubation à ~ 20 °C dans le TPB. (A) SNUs (une unité est indiquée par une flèche noire) s'assemblant en protofibrilles. (B) STRF. (C) Aspect strié des STRFs. (D) Réseau de fibrilles enchevêtrées avec un SF indiqué par une flèche noire. (E) Agrégats amorphes. Sur chaque image, l'échelle est indiquée dans le coin inférieur droit.

Partie n°3.3

—

Caractérisation de la région C-terminale des isoformes de DPF3

3.3.1. Analyses expérimentales de la nature désordonnée des régions C-terminales

Les deux isoformes de DPF3 se différenciant par leur région C-terminale (C-TER), il est intéressant d'en étudier les propriétés. Les régions considérées vont du 200^{ème} au dernier résidu de chaque isoforme (Fig. 22). Le caractère désordonné de la partie C-terminale de DPF3b (C-TERb) et de DPF3a (C-TERa) a été investigué par une série de techniques spectroscopiques. C-TERb est caractérisé par le ZnF C₂H₂, une partie désordonnée (IDR-2) de 40 acides aminés suivie des ZnFs PHD-1 et PHD-2 considérés comme ordonnés. En revanche, C-TERa possède également le ZnF C₂H₂ et l'IDR-2 mais est composé du ZnF tronqué PHD-1/2 et d'une autre partie désordonnée (IDR-3) longue de 65 résidus.

Le contenu en structures secondaires des régions C-TERb et C-TERa a été déterminé par FCD (Fig. 34). Le spectre de C-TERb montre un maximum à ~ 202 nm, un léger minimum à ~ 219 nm et un autre à ~ 225 nm, qui sont associés à des feuillets β . Une épaule, indicatrice de structures en hélice α est également perceptible à ~ 208 nm. Un autre minimum, situé vers ~ 212 nm, provient probablement des contributions à la fois d'hélices α et de feuillets β . Le calculateur de structure secondaire BeStSel est en accord avec ces observations étant donné que 4 % d'hélices α et 38 % de feuillets β antiparallèles sont reconnus ainsi que 57 % de désordre (*turn* et *coil*). Le spectre FCD de C-TERa révèle des signatures similaires notamment avec un maximum à ~ 200 nm et un minimum à ~ 225 nm relatif à des feuillets β antiparallèles. Ce dernier peut également provenir du désordre, car il a été démontré que les polypeptides non ordonnés présentent une large bande négative dans la gamme 220-230 nm.¹⁵² Néanmoins, le large minimum observé entre ~ 206 et ~ 210 nm provient certainement d'un mélange d'hélices α et de désordre.¹⁵³ Il est donc attendu que les estimations donnent à voir plus de désordre pour C-TERa. En effet, selon BeStSel, C-TERa en contient 66 %, ce qui est cohérent avec le spectre et la présence de l'IDR-3. En regard de C-TERb, C-TERa ne possède que 2 % d'hélices α et 32 % de feuillets β antiparallèles, ce qui s'explique par l'absence du tandem PHD-1 et PHD-2 riche en ces structures. Alors que C-TERa possède globalement les mêmes proportions en structure secondaire que DPF3a FL, C-TERb contient moins de désordre et plus de feuillets β antiparallèles par rapport à DPF3b FL. Ces différentes proportions s'expliquent par l'absence de l'IDR-1 dans C-TERb alors que les feuillets β antiparallèles des ZnFs PHD-1 et PHD-2 prennent plus d'importance étant donné la séquence plus petite de C-TERb par rapport à DPF3b FL.¹²⁵

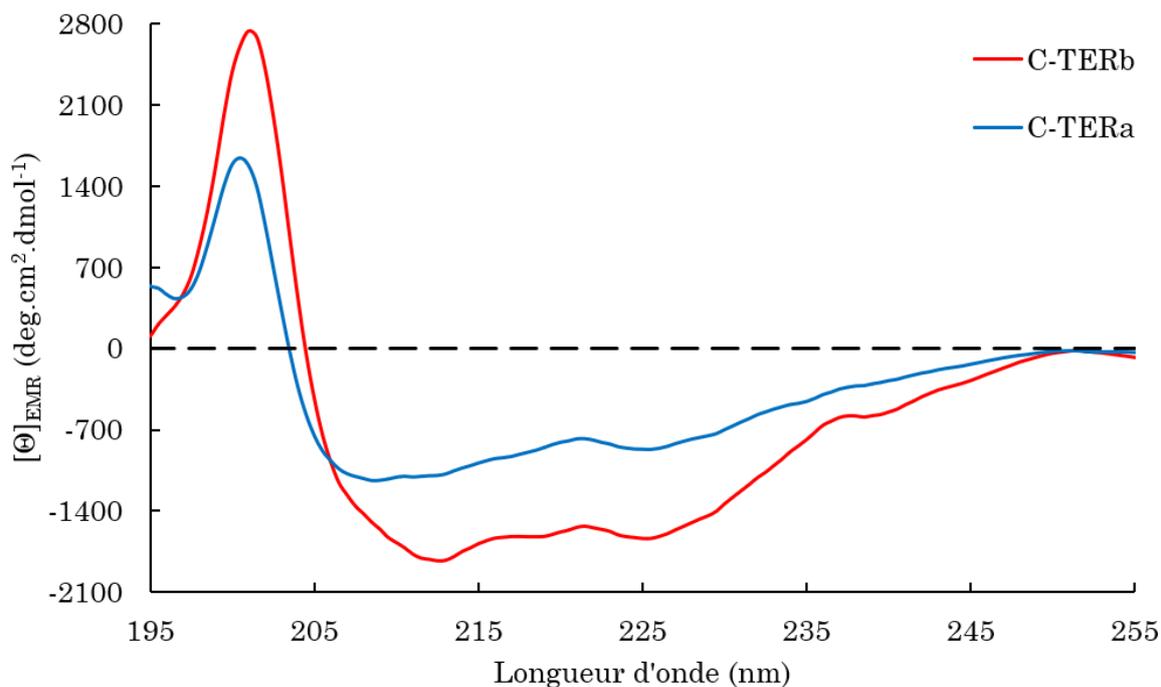


Figure 34 – Spectres FCD de C-TERb (rouge) et C-TERa (bleu) dans le TPB à ~ 20 °C.¹²⁵

L'environnement local et l'état d'exposition des Trp et Tyr au solvant ont été analysés respectivement par ITF et ITyrF. C-TERa ne possédant pas de Trp dans sa séquence, seul le spectre ITF de C-TERb, qui possède trois Trp, a été enregistré (Fig. 35A). A l'instar de DPF3b FL (Fig. 24A), une bande d'émission à 334 nm est observée, correspondant à des résidus partiellement exposés au solvant. Les Trp dans les ZnFs n'apparaissent donc pas totalement enfouis dans ces domaines structurés mais ont plutôt une face de leur cycle exposée. Le spectre ITyrF de C-TERb (Fig. 35B) montre également une bande d'émission à ~ 334 nm, ce qui indique, de nouveau, la présence de FRET entre les résidus Trp et Tyr. Ceci est cohérent avec la localisation des résidus Tyr dans la séquence de C-TERb. En effet, trois de ses sept résidus Tyr se trouvent à proximité dans la séquence de trois résidus Trp dans les domaines PHD-1 et PHD-2 : Tyr³⁰⁹, Tyr³⁴¹ et Tyr³⁴⁴ sont respectivement proches de Trp³¹¹, Trp³⁵⁸ et Trp³⁶⁴. Sur ce même spectre, une légère épaule entre 305 et 310 nm est visible et provient de la fluorescence des résidus Tyr contenu dans le domaine C₂H₂ qui ne sont pas impliqués dans le FRET. Le spectre ITyr de C-TERa (Fig. 35B) se différencie de celui de C-TERb de par la présence d'une bande d'émission à 328 nm, ainsi qu'une épaule plus marquée à 305 nm, comme cela est mis en évidence par la différence entre les deux spectres. Étant donné que C-TERa ne contient aucun Trp et de l'impossibilité de FRET Trp-Tyr, le déplacement à 328 nm provient probablement de la formation de tyrosinate due au transfert de proton entre le groupement

hydroxyle de la chaîne latérale tyrosile et un accepteur, soit un aspartate (Asp) soit un glutamate (Glu), dans la structure primaire.¹⁵⁴ Plus précisément, cela pourrait venir de la formation de tyrosinate à partir du résidu Tyr²⁶¹ qui se trouve à proximité du résidu Asp²⁶³. Bien que ce phénomène soit souvent observé dans des conditions de pH très basiques, il a déjà été observé, à pH physiologique, chez des protéines déplétées en résidus Trp mais dont les Tyr sont proches d'Asp et de Glu dans la structure tertiaire.¹⁵⁵ L'intensité plus importante de l'épaule à 305 nm peut également être expliquée par l'absence de résidus Trp, empêchant ainsi le FRET, et attribuée aux résidus Tyr dans le domaine C₂H₂.¹²⁵

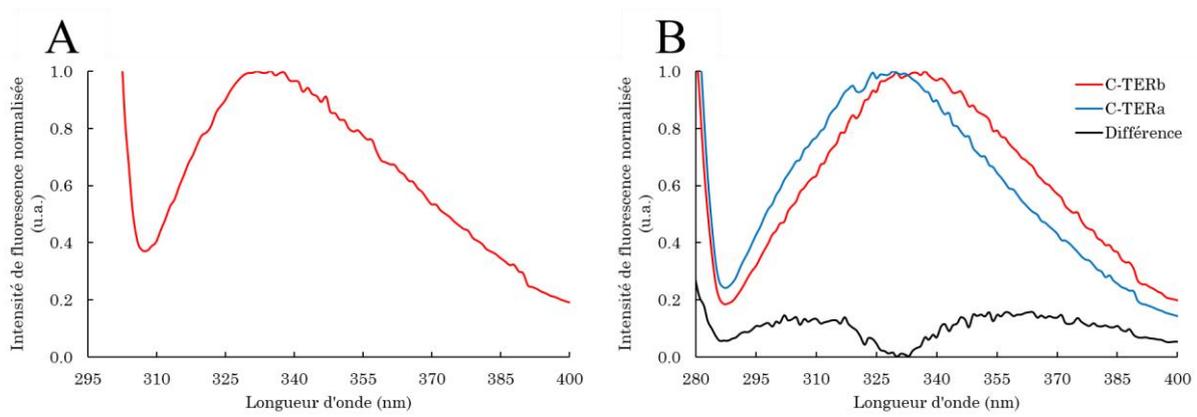


Figure 35 – Fluorescence intrinsèque de C-TERb et C-TERa dans le TPB à ~ 20 °C. (A) Spectre ITF normalisé ($\lambda_{ex} = 295$ nm, $sw = 10$ nm) de C-TERb. (B) Spectres ITyrF normalisés ($\lambda_{ex} = 275$ nm, $sw = 10$ nm) de C-TERb (rouge) et C-TERa (bleu) ainsi que la différence entre les deux spectres (noir).¹²⁵

3.3.2. Investigation de la propension à l'agrégation des régions C-terminales

Les deux isoformes de DPF3 FL ayant montré une tendance à s'agréger spontanément sous forme de feuillets β , le comportement d'agrégation des régions C-terminales, C-TERb et C-TERa, a été investigué *in vitro*. Pour ce faire, chaque échantillon, à une concentration de 10 μ M, a été incubé pendant quatre jours à ~ 20 °C dans le TPB et analysé par dichroïsme circulaire et différentes spectroscopies de fluorescence.

Les modifications au cours du temps en structures secondaires des domaines C-TERb (Fig. 36A) et C-TERa (Fig. 36B) ont été suivies par FCD. Les pourcentages en structures secondaires ont été estimés à l'aide du serveur BeStSel (Tableau 2). Après 24 h d'incubation, le spectre de C-TERb montre que le domaine subit déjà des transformations conformationnelles. La proportion en hélices α , associée à l'apparition d'un léger minimum à ~ 222 nm et une augmentation globale de l'intensité dans la région à 206-210 nm, passe de 4 à 5 %. La formation d'intermédiaires enrichis en hélices α a déjà été reportée chez l' α -syn et

DPF3a FL.¹³⁵ Aussi, les déconvolutions indiquent une perte en feuillets β antiparallèles (de 38 à 35 %) et de *turn* (de 17 à 15 %) ainsi qu'un gain de 6 % en *coil*. Après 48 h, le minimum est complètement déplacé à ~ 227 nm et une épaule vers 212 nm apparaît. A ce stade, C-TERb ne contient plus d'hélice α et une faible portion (3 %) en feuillets β parallèles se manifeste. Lors des deux jours suivants, l'intensité du minimum à 227 nm augmente significativement, ce qui est provoqué par la formation de feuillets β parallèles dont le pourcentage passe de 3 à 15 %. Cet enrichissement, qui n'a pas été observé pour DPF3b FL (Fig. 26A), se fait en partie au détriment des feuillets β antiparallèles dont le pourcentage diminue de 38 à 24 % en 96 h. La coexistence de feuillets β parallèles et antiparallèles peut expliquer l'élargissement de la bande à 227 nm après 96 h et n'est pas incohérent avec la formation de fibrilles. En effet, le réarrangement en feuillets β parallèles est aussi un marqueur de l'agrégation amyloïde, notamment chez l' α -syn.^{156,157} Contrairement à DPF3b FL, une augmentation du nombre de *turns* est observée endéans 96 h.¹²⁵

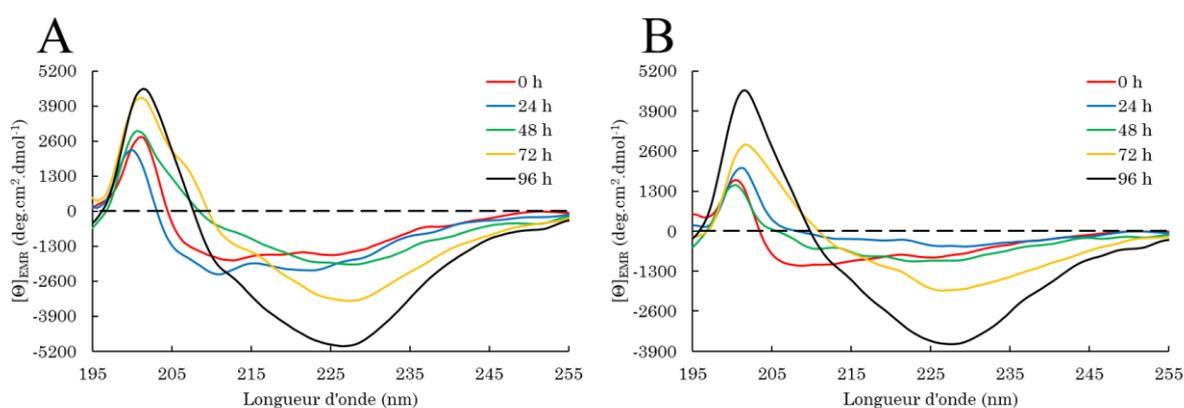


Figure 36 – Spectres FCD de (A) C-TERb et (B) C-TERa dans le TPB à ~ 20 °C après 0 h (rouge), 24 h (bleu), 48 h (vert), 72 h (jaune) et 96 h (noir) d'incubation.¹²⁵

A l'exception des premières 24 h, C-TERa présente des transformations conformationnelles comparables à C-TERb. En effet, au lieu d'un gain, le contenu en hélices α et en désordre diminue respectivement de 2 et 3 %, comme le montre la perte du minimum entre 206 et 210 nm. Un enrichissement de 5 % en feuillets β antiparallèles est également observé. Après 48 h, aucune modification majeure n'est observée. Dans les 48 h suivantes, similairement à C-TERb, il y a apparition d'un minimum à ~ 227 nm ainsi qu'une augmentation de l'intensité du maximum à 202 nm. Ces signatures s'accompagnent d'un enrichissement de 10 % en feuillets β parallèles et de 6 % en *turns* au détriment des feuillets β antiparallèles dont le pourcentage tombe à 25 %. Les deux C-TER présentent donc un

enrichissement en feuillets β parallèles, qui n'était pas présent pour les isoformes complètes, associé à l'apparition du minimum à 227 nm. Cependant, ces réarrangements se font plus rapidement pour C-TERb. En effet, après 72 h, 9 % de feuillets β parallèles sont formés alors qu'il faut 96 h à C-TERa pour en atteindre 10 %. Ce comportement est opposé à celui obtenu pour les isoformes complètes. Cela peut être expliqué par un mécanisme de fibrillation différent, des fibrilles plus longues ou la formation en plus grand nombre de fibrilles pour C-TERa.¹²⁵

Tableau 2 – Estimations (BeStSel) de la composition en structure secondaire et valeurs de RMSD associées de C-TERb et C-TERa dans le TPB à ~ 20 °C après 0 h, 24 h, 48 h, 72 h et 96 h d'incubation.¹²⁵

Protéine (Temps d'incubation)	RMSD	Composition en structure secondaire (%)				
		Hélice α	Feuillet β antiparallèle	Feuillet β parallèle	Désordre <i>Turn</i> <i>Coil</i>	
C-TERb (0 h)	0.04	4	38	0	17	40
C-TERb (24 h)	0.05	5	35	0	15	46
C-TERb (48 h)	0.05	0	38	3	16	43
C-TERb (72 h)	0.08	0	24	9	21	46
C-TERb (96 h)	0.08	0	24	15	21	40
C-TERa (0 h)	0.03	2	32	0	18	48
C-TERa (24 h)	0.02	0	37	0	18	45
C-TERa (48 h)	0.04	0	37	0	16	47
C-TERa (72 h)	0.06	0	34	5	19	42
C-TERa (96 h)	0.07	0	25	10	22	43

Les réarrangements de la structure tertiaire de C-TERb et C-TERa au cours du temps ont été étudiés par ITF et ITyrF. Le spectre ITF de C-TERb (Fig. 37A) présente, à l'instar de DPF3b FL, une bande d'émission à 334 nm avec une intensité augmentant au fur et à mesure du temps. Toutefois, aucun déplacement de bande n'est observé, ce qui signifie que certains Trp sont toujours partiellement exposés au solvant et que d'autres sont retrouvés dans un environnement plus hydrophobe. C'est pourquoi, les Trp, qui se trouvent dans les ZnFs PHD-1 et PHD-2, ne sont pas particulièrement affectés par les modifications en structures secondaires et, par conséquent, ne semblent pas impliqués dans la formation du noyau amyloïde. Similairement à DPF3b FL et DPF3a FL, une bande d'émission à 456 nm apparaît sur les spectres et dont l'intensité croît avec le temps. Cette bande d'autofluorescence dbAF est également retrouvée sur les spectres ITyrF de C-TERb (Fig. 37B), bien que légèrement déplacée à 460 nm. La bande d'émission principale reste à 334 nm avec un effet hyperchrome

au fil du temps, indiquant que les Tyr se trouvent dans un environnement plus hydrophobe. De plus, l'épaule à 305 nm disparaît complètement après 96 h, ce qui est associé à une augmentation du FRET Trp-Tyr liée au réarrangements conformationnels. Aussi, cela semble être un indicateur de l'implication des résidus Tyr dans le noyau amyloïde.

En ce qui concerne les spectres ITyrF de C-TERa (Fig. 37C), la bande d'émission principale est déplacée à 333 nm et l'épaule à 305 nm est progressivement perdue jusqu'à disparaître totalement après 96 h d'incubation. Ces deux tendances suggèrent des réarrangements structuraux menant à des résidus Tyr proches d'accepteurs de protons Asp ou Glu, favorisant la formation de tyrosinates étant donné que C-TERa est dépourvu de résidus Trp. Une seconde bande d'émission apparaît à 460 nm, indiquant que les Tyr participent également à l'autofluorescence. La bande dbAF devient même majoritaire après 48 h.¹²⁵

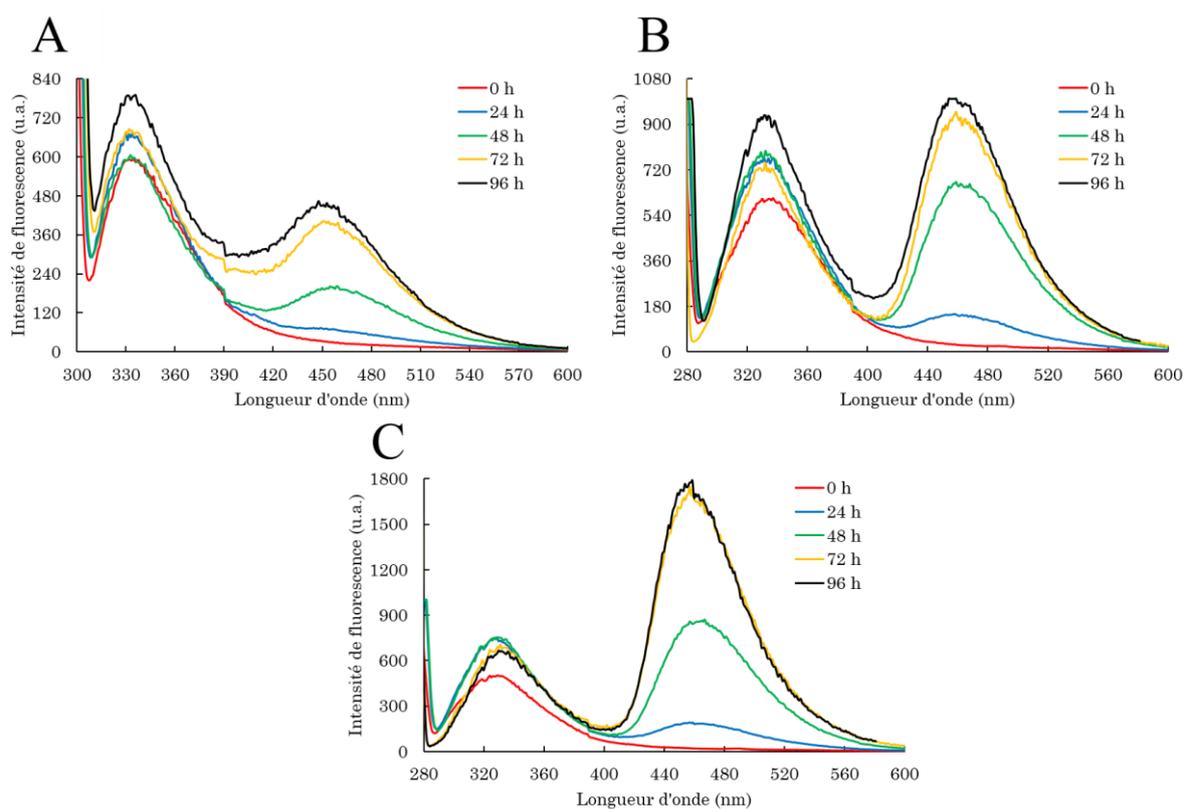


Figure 37 – Fluorescence intrinsèque de C-TERb et C-TERa dans le TPB à ~ 20 °C après 0 h (rouge), 24 h (bleu), 48 h (vert), 72 h (jaune) et 96 h (noir) d'incubation. (A) Spectre ITF ($\lambda_{ex} = 295$ nm, $sw = 10$ nm) de C-TERb. Spectres ITyrF ($\lambda_{ex} = 295$ nm, $sw = 10$ nm) de (B) C-TERb et (C) C-TERa.¹²⁵

Les spectres d'excitation ($\lambda_{ém} = 460$ nm) de C-TERb (Fig. 38A) et C-TERa (Fig. 38B) ont été enregistrés pour identifier les contributions dbAF. C-TERb présente quatre bandes

d'excitation dont trois étaient déjà présentes sur le spectre de DPF3b FL après 168 h (Fig. 28A). La première à 237 nm correspond à l'absorption des liaisons peptidiques, la seconde vers 270 nm est relative aux résidus aromatiques et explique la présence des bandes d'émission sur les spectres ITF et ITyrF, la troisième à 400 nm est la principale composante dbAF et la quatrième est une bande de plus faible intensité, située entre 325 et 330 nm, qui pourrait provenir d'interactions spécifiques au sein du noyau amyloïde de C-TERb et différentes de celles de DPF3b FL. C-TERa contient les trois mêmes bandes que DPF3a FL, c'est-à-dire à 237 nm, 270 nm et 400 nm. Comme pour les spectres ITF et ITyrF, un effet hyperchrome est observé au cours du temps.¹²⁵

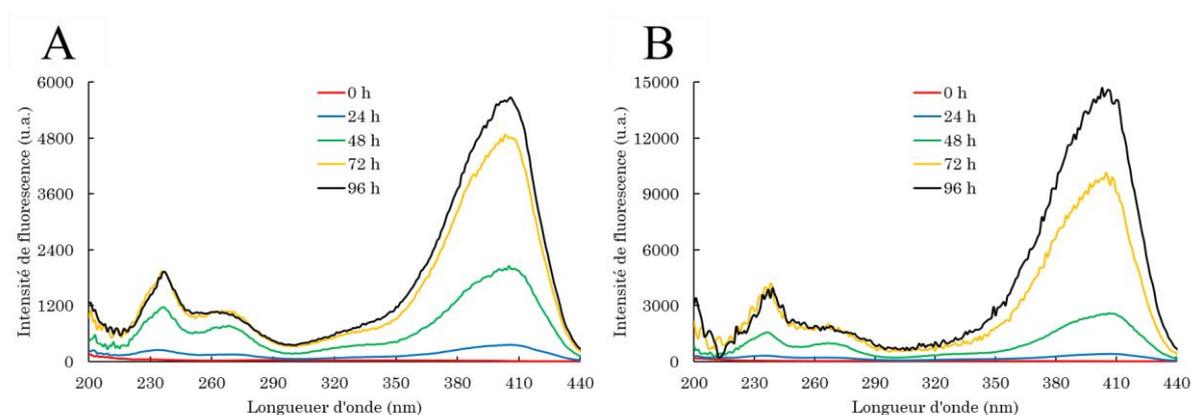


Figure 38 – Spectres d'excitation dbAF ($\lambda_{em} = 460$ nm, $sw = 10$ nm) de (A) C-TERb et (B) C-TERa dans le TPB à ~ 20 °C après 0 h (rouge), 24 h (bleu), 48 h (vert), 72 h (jaune) et 96 h (noir) d'incubation.¹²⁵

La bande d'excitation la plus importante étant à 400 nm, les spectres dbAF à $\lambda_{ex} = 400$ nm donnent lieu à des bandes d'émission d'intensité plus importante. En prenant la valeur de l'intensité dbAF à 456-460 nm après chaque temps d'incubation, il est possible de construire un profil cinétique pour chaque C-TER (Fig. 39A). Il en résulte une courbe sigmoïdale qui, comme pour les isoformes complètes, est reconnue comme caractéristique de protéines amyloïdogènes. La phase de latence est relativement la même pour les deux isoformes et dure respectivement 36 et 40 h pour C-TERb et C-TERa. La phase stationnaire est atteinte après 72 h pour C-TERb et 90 h pour C-TERa, ce qui signifie que l'étape d'élongation est plus courte (36 h) pour C-TERb que pour son isoforme (50 h). Cela est probablement dû au fait que le plateau de C-TERa a une intensité trois fois plus importante que celui associé à C-TERb. Cette différence peut être expliquée par la formation d'un plus grand nombre de fibrilles et/ou de

fibrilles plus longues. La phase d'élargissement plus courte de C-TERb trouverait son origine dans le réarrangement plus rapide en feuillets β parallèles observé en FCD.¹²⁵

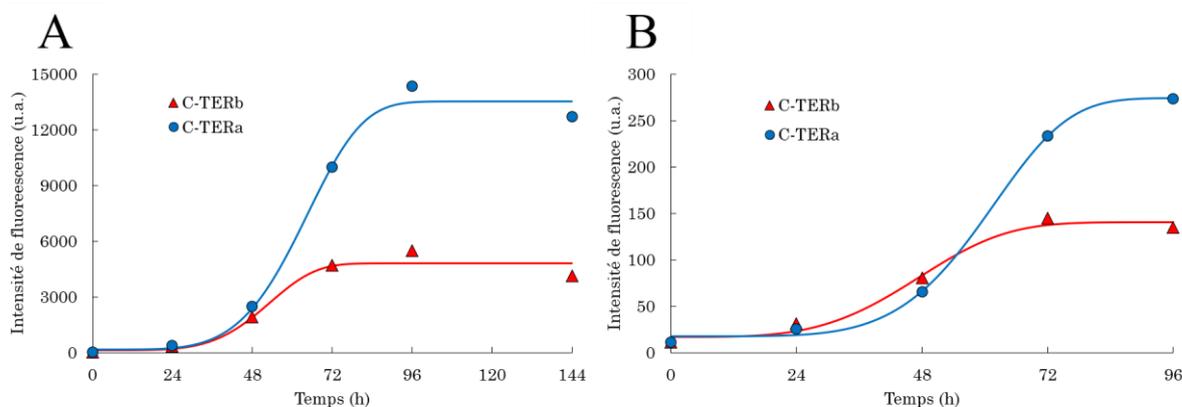


Figure 39 – Courbes cinétiques de l'agrégation de C-TERb (triangle et rouge) et de C-TERa (cercle et bleu) dans le TPB à ~ 20 °C par suivi de (A) l'émission dbAF ($\lambda_{ex} = 400$ nm, $\lambda_{ém} = 460$ nm, $sw = 10$ nm) et de (B) la fluorescence de la ThT ($\lambda_{ex} = 440$ nm, $\lambda_{ém} = 485$ nm, $sw = 10$ nm).¹²⁵

Les essais de liaison à la ThT permettent de confirmer la fibrillation amyloïde de la protéine en suivant l'émission à 485 nm après excitation à 440 nm. Afin de valider les résultats obtenus par les cinétiques en dbAF, l'intensité à 485 nm a été mise en graphique par rapport au temps d'incubation. Les cinétiques résultantes ont de nouveau un profil de sigmoïde (Fig. 39B). Pour chaque isoforme, l'intensité est très faible après 24 h, signifiant que, dans un premier temps, des agrégats non amyloïdes sont formés, ce qui correspond à la phase de latence. Pour C-TERb, la phase d'élargissement s'étend sur 48 h alors qu'elle prend 24 h de plus pour C-TERa. En effet, les plateaux sont respectivement atteints après 72 et 96 h pour C-TERb et C-TERa. Les profils cinétiques suivis par dbAF et par liaison à la ThT sont comparables pour les deux isoformes. En effet, les mêmes tendances sont observées pour C-TERa, à savoir une phase d'élargissement plus longue et une phase stationnaire dont le plateau a une intensité plus élevée.¹²⁵

3.3.3. Morphologie des agrégats des régions C-terminales

Afin de comparer la morphologie des agrégats des régions C-TERb et C-TERa, des images TEM ont été prises après 144 h d'incubation à ~ 20 °C afin de s'assurer de la formation de fibrilles. Malgré certaines différences spectrales, les deux isoformes présentent les mêmes types d'agrégat et de fibrille. Néanmoins, les fibrilles des régions C-TER sont quelque peu distinctes par rapport à celles obtenues pour les protéines complètes DPF3b et DPF3a.

La plupart des fibrilles retrouvées sont simples et droites (SSFs pour *single straight fibrils*) avec une largeur de 15 à 20 nm (Figs. 40A et 41A), correspondant à ce qui est retrouvé dans la littérature.¹⁴⁸ Ces SSFs peuvent s'assembler par paire pour former des fibrilles droites doubles (DSFs pour *double straight fibrils*). Les SSFs et DSFs peuvent aussi se rassembler davantage pour générer des fibrilles droites multiples (MSFs pour *multiple straight fibrils*) (Figs. 40B et 41B) sur lesquelles chaque SSF est distinguable. Les SSFs et DSFs sont également retrouvées dans des réseaux dans lesquels ces fibrilles s'entrecroisent, suggérant la présence de points de nucléation secondaires sur les fibrilles déjà formées (Figs. 40C et 41C). Aussi, comme pour DPF3b FL et DPF3a FL, des phases amorphes sont observées (Figs. 40D et 41D). L'absence de SNUs et de la formation de fibrilles striées et torsadées en forme de ruban (STRFs) suggère que les régions C-TER suivent un mécanisme d'agrégation différent de celui de DPF3b FL et DPF3a FL.¹²⁵

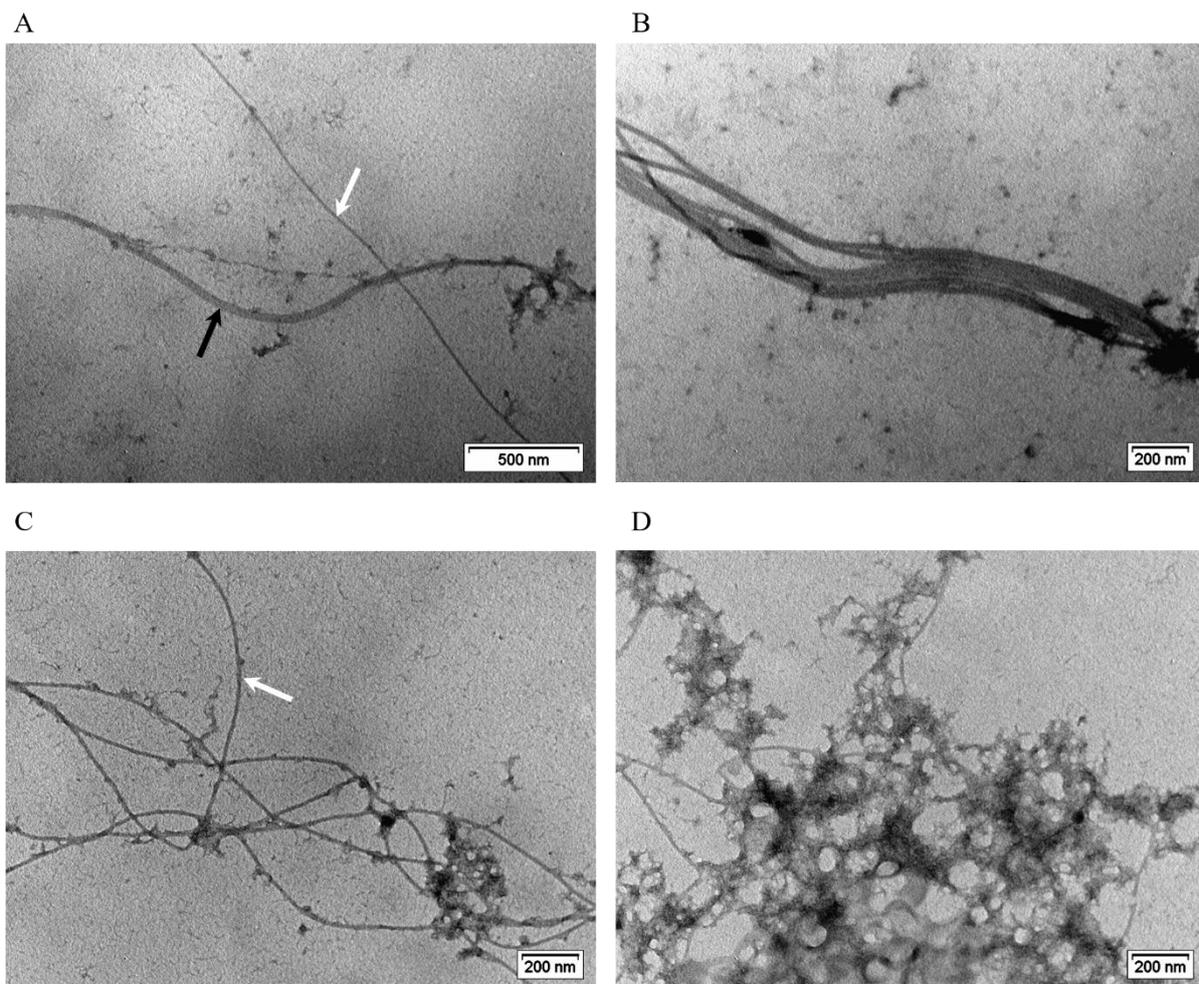


Figure 40 – Micrographies TEM en negative staining de C-TERb incubé pendant 144 h dans le TPB à ~ 20 °C. (A) Fibrilles droites simples (SSFs) (indiquées par une flèche blanche) et droites doubles (DSFs) (indiquée par une flèche noire). (B) Fibrilles droites multiples (MSFs). (C) Réseau composé de SSFs (une SSF est indiquée par une flèche blanche). (D) Agrégats amorphes. L'échelle est indiquée dans le coin inférieur droit de chaque image.¹²⁵

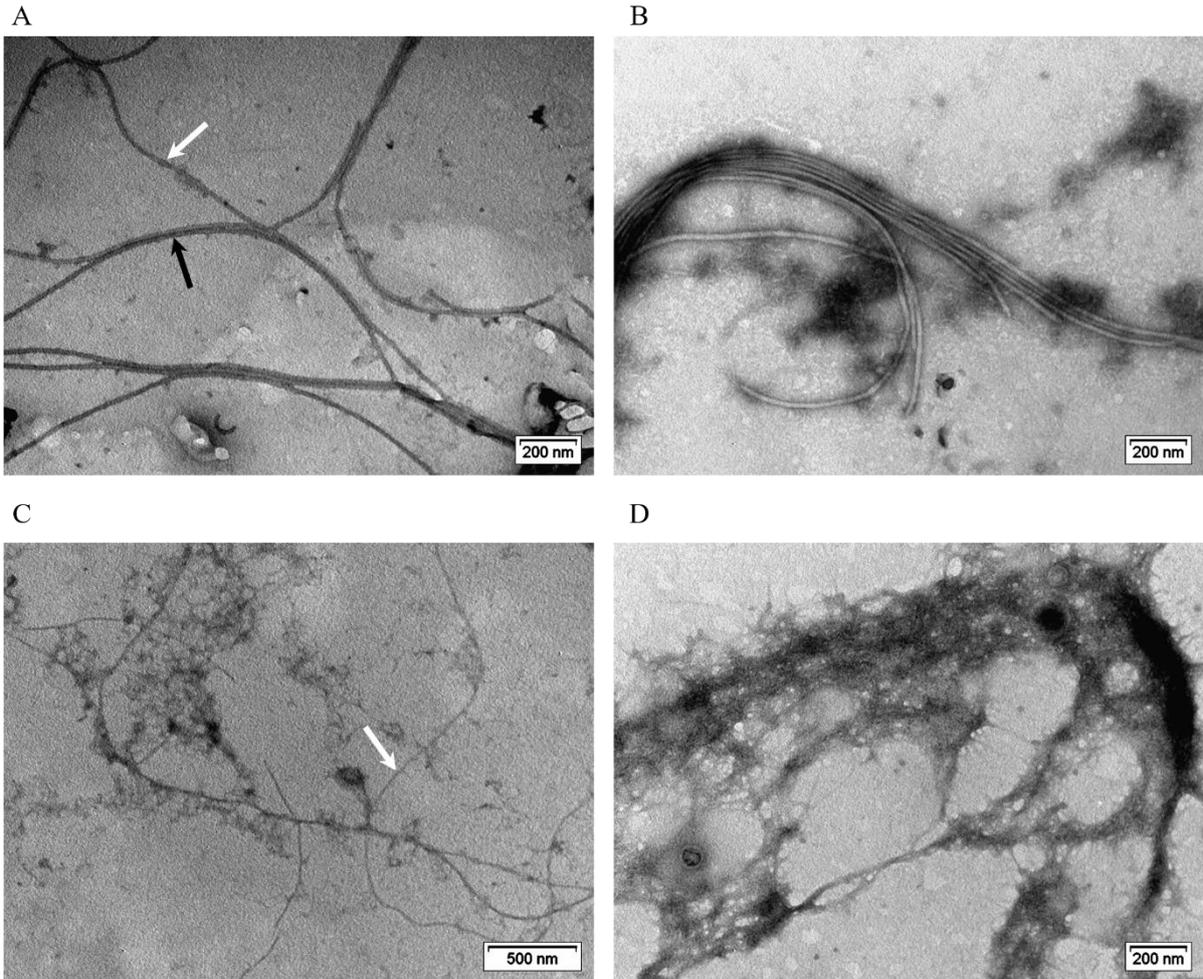


Figure 41 – Micrographies TEM en negative staining de C-TERA incubé pendant 144 h dans le TPB à ~ 20 °C. (A) Fibrilles droites simples (SSFs) (indiquées par une flèche blanche) et droites doubles (DSFs) (indiquée par une flèche noire). (B) Fibrilles droites multiples (MSFs). (C) Réseau composé de SSFs (une SSF est indiquée par une flèche blanche). (D) Agrégats amorphes. L'échelle est indiquée dans le coin inférieur droit de chaque image.¹²⁵

Partie n°3.4

—

Influence de cations divalents sur l'agrégation de la région C-terminale des isoformes de DPF3

Les métaux étant impliqués dans l'agrégation des protéines amyloïdogènes et DPF3 étant une protéine à doigt de zinc, l'influence de cations divalents (Cu^{2+} , Mg^{2+} , Ni^{2+} et Zn^{2+}) sur l'agrégation de C-TERb et C-TERa a été étudiée *in vitro* par une série de techniques spectroscopiques et microscopique. Pour ce faire, chaque isoforme, à une concentration de 5 μM , a été incubée à $\sim 20^\circ\text{C}$ dans le TPB pendant quatre jours en présence de métaux à une concentration de 100 μM . Les données présentées sont relatives à une période d'incubation de 72 h, représentative des tendances observées.

3.4.1. Influence des cations divalents sur la structure secondaire des régions C-terminales

L'effet des cations divalents sur la composition en structure secondaire a été analysé par FCD. Pour le C-TERb (Fig. 42A), les métaux n'impactent pas grandement son réarrangement conformationnel au regard des signatures spectrales comparables, en particulier en présence de Mg^{2+} et le Zn^{2+} . Néanmoins, le Cu^{2+} provoque le déplacement du minimum de 227 à 224 nm et du maximum de 202 à 200 nm. Cela est difficilement attribuable à un type précis de structure secondaire et pourrait provenir de motifs spécifiques dus à la liaison métallique. Les estimations BeStSel (Tableau 3) confirment l'absence d'impact majeur de la liaison aux métaux sur la composition en structure secondaire de C-TERb. Cependant, chaque métal promeut l'augmentation du pourcentage en feuillets β antiparallèles. Le Ni^{2+} possède l'effet le plus important. Pour rappel, le C-TERb sans métal se caractérise au cours de son agrégation par un enrichissement en feuillets β parallèles et de *turns*. Ces structures sont conservées en présence des métaux étant donné qu'un écart de maximum 3 % est observé.¹²⁵

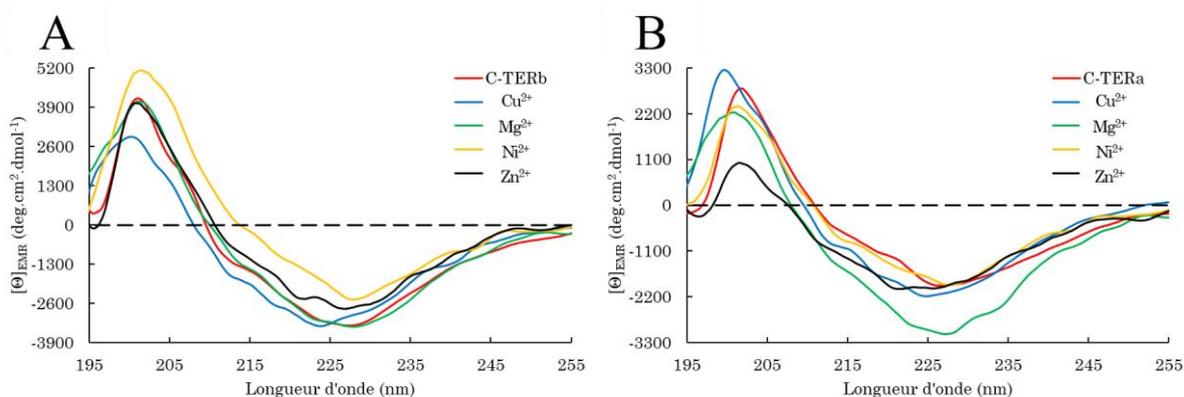


Figure 42 – Spectres FCD de (A) C-TERb et (B) C-TERa après 72 h d'incubation dans le TPB à $\sim 20^\circ\text{C}$ sans métal (rouge) et en présence de Cu^{2+} (bleu), Mg^{2+} (vert), Ni^{2+} (jaune) et Zn^{2+} (noir).¹²⁵

Quant au C-TERa, les métaux conduisent à plus de modifications visibles sur les spectres FCD (Fig. 42B). Alors que le minimum reste à 227 nm pour le Mg^{2+} et le Ni^{2+} , il est déplacé à 225 nm en présence de Cu^{2+} et à 221 nm avec le Zn^{2+} . Le maximum est également déplacé à de plus basses longueurs d'onde pour le Cu^{2+} (200 nm), le Mg^{2+} et le Ni^{2+} (201 nm). En outre, le Mg^{2+} provoque une nette augmentation de l'intensité du minimum à 227 nm. Etant donné les tendances observées pour les C-TER sans métaux (Fig. 36 et Tableau 2), ce comportement est associé à la formation de fibrilles enrichies en feuillet β parallèles. Selon les estimations BeStSel, un enrichissement de 6 % par rapport à C-TERa sans métal est obtenu. De ce fait, les ions Mg^{2+} pourraient faire office d'accélérateur de la fibrillation de C-TERa. Alors que ce métal engendre des conformations appauvries en feuillet β antiparallèles (21 %), le Ni^{2+} et le Zn^{2+} favorisent, à l'inverse, des conformations enrichies en ces structures (38 %). Similairement à C-TERb, les pourcentages de C-TERa associés au désordre (*turn* et *coil*) ne sont pas grandement influencés par les métaux.¹²⁵

Tableau 3 – Estimations (BeStSel) de la composition en structure secondaire et valeurs de RMSD associées de C-TERb et C-TERa dans le TPB à ~ 20 °C après 72 h d'incubation sans métal et en présence de Cu^{2+} , Mg^{2+} , Ni^{2+} et Zn^{2+} .¹²⁵

Protéine (Métal)	RMSD	Composition en structure secondaire (%)			
		Hélice α	Feuillet β antiparallèle	Feuillet β parallèle	Désordre <i>Turn</i> <i>Coil</i>
C-TERb	0.08	0	24	9	21 46
C-TERb (Cu^{2+})	0.06	0	31	8	18 43
C-TERb (Mg^{2+})	0.06	0	27	8	20 41
C-TERb (Ni^{2+})	0.07	0	33	6	20 41
C-TERb (Zn^{2+})	0.07	0	30	7	19 44
C-TERa	0.06	0	34	5	19 42
C-TERa (Cu^{2+})	0.05	0	33	5	17 45
C-TERa (Mg^{2+})	0.09	0	21	11	23 45
C-TERa (Ni^{2+})	0.05	0	38	3	18 41
C-TERa (Zn^{2+})	0.06	0	38	3	18 41

3.4.2. Influence des cations divalents sur la structure tertiaire des régions C-terminales

La composition en structure secondaire de C-TERb et C-TERa restant globalement inchangée par la liaison aux métaux, l'influence de ces derniers sur les modifications de la structure tertiaire au cours de l'agrégation a été analysée par ITF et ITyrF. La position de la première bande d'émission, à 334 nm, sur les spectres ITF de C-TERb (Fig. 43A) est identique

pour chaque condition, indiquant que les Trp restent partiellement exposés au solvant. Cependant, l'intensité de cette bande est diminuée en présence des cations, ce qui signifie que la liaison aux métaux engendre des réarrangements structuraux au sein desquels des groupements qui éteignent la fluorescence des Trp se trouvent dans leur environnement proche. Il a été reporté que le Cu^{2+} , Mg^{2+} et Ni^{2+} diminuent l'intensité de fluorescence des Trp.¹⁵⁸ Le même comportement étant détecté sur les spectres de C-TERb, cela suggère que ces métaux se lient directement aux résidus Trp.^{158,159} A l'inverse, le Zn^{2+} n'est pas un extincteur de l'émission des Trp. Cependant, l'intensité à 334 nm est quand même diminuée, ce qui montre que la liaison du Zn^{2+} provoque des réarrangements qui rapprochent des Trp de certains groupements éteignant la fluorescence. Les mêmes tendances sont retrouvées pour la bande d'autofluorescence dbAF. Il a été montré que certains paramètres, tels que les cations métalliques, peuvent modifier les propriétés dbAF des protéines agrégées.¹⁴¹ En effet, les métaux provoquent la réduction de l'intensité d'émission d'autofluorescence dans le même ordre que la bande à 334 nm : $\text{Ni}^{2+} > \text{Cu}^{2+} \geq \text{Zn}^{2+} > \text{Mg}^{2+}$. De plus, le signal dbAF disparaît presque totalement en présence de Ni^{2+} et est diminué de moitié avec le Mg^{2+} . Cela montrerait que les métaux freinent la fibrillation. Cette diminution d'intensité à 456 nm est également observée selon la même tendance sur les spectres ITyrF de C-TERb (Fig. 43B). Comme pour les spectres ITF, la bande à 334 nm n'est pas non plus déplacée. Cependant, seuls les ions Cu^{2+} provoquent un effet hypochrome significatif. Etant donné que le Cu^{2+} est connu pour éteindre la fluorescence des Tyr¹¹⁶, cela indique qu'il se lie à celles de C-TERb. Aussi, l'épaule à 305 nm n'est plus présente pour aucune condition.¹²⁵

Le domaine C-TERa étant dépourvu de Trp, seul les spectres ITyrF ont été enregistrés (Fig. 43C). Alors que la première bande à 333 nm ne présente aucun déplacement et plus aucune épaule, la bande d'autofluorescence est fortement influencée par l'addition des métaux. Tandis que l'intensité de la bande dbAF est fortement réduite en présence de Cu^{2+} et de Ni^{2+} , elle est doublée avec le Mg^{2+} . Quant au Zn^{2+} , il provoque non seulement une augmentation d'intensité mais aussi un déplacement bathochrome de 460 à 490 nm. A l'image de C-TERb, la fibrillation semble être affectée par les métaux. Aussi, le Zn^{2+} et le Cu^{2+} réduisent l'intensité de la première bande d'émission. Le Cu^{2+} étant connu pour diminuer l'intensité de fluorescence de la Tyr, ces cations sont vraisemblablement directement liés aux résidus Tyr de C-TERa. Au contraire, les ions Zn^{2+} ne contribuent pas à l'extinction de la fluorescence des Tyr, ce qui indique qu'ils sont responsables d'une réorganisation conformationnelle provoquant le rapprochement de groupements extincteurs de la fluorescence des résidus Tyr.¹¹⁶ Le Mg^{2+} et le

Ni^{2+} ne provoquant pas de diminution d'intensité de la bande à 333 nm, ces métaux induisent des changements conformationnels différents par rapport au Cu^{2+} et Zn^{2+} .¹²⁵

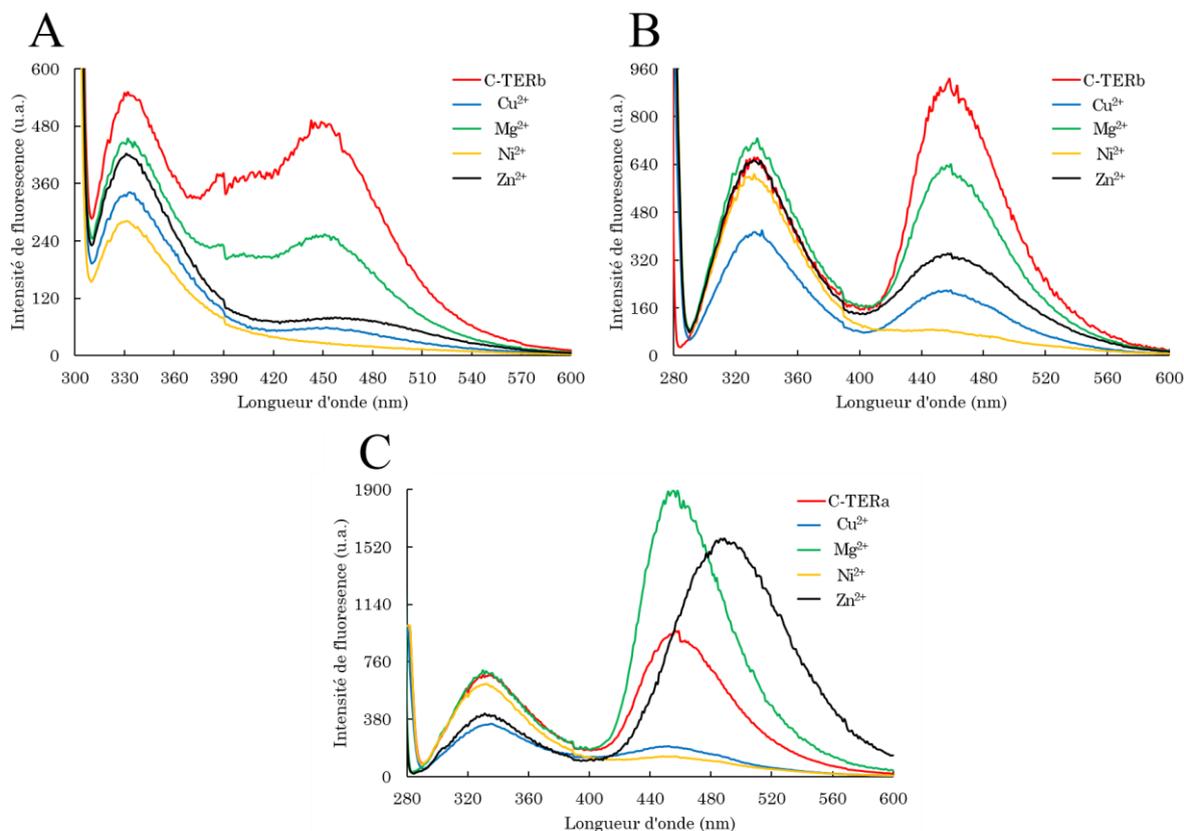


Figure 43 – Fluorescence intrinsèque de C-TERb et C-TERa après 72 h d'incubation dans le TPB à ~ 20 °C sans métal (rouge) et en présence de Cu^{2+} (bleu), Mg^{2+} (vert), Ni^{2+} (jaune) et Zn^{2+} (noir). (A) Spectres ITF ($\lambda_{ex} = 295$ nm, $sw = 10$ nm) de C-TERb. Spectres ITF ($\lambda_{ex} = 275$ nm, $sw = 10$ nm) de (B) C-TERb et (C) C-TERa.¹²⁵

Afin d'identifier les contributions dbAF ($\lambda_{ém} = 456$ nm), les spectres d'excitation de C-TERb (Fig. 44A) et de C-TERa (Fig. 44B) ont été réalisés. Pour C-TERb, aucune modification due à la liaison aux métaux n'est observée. Il existe toujours quatre contributions : une épaule entre 325 et 330 nm et trois bandes à respectivement 237, 270 et 400 nm. Ces bandes sont peu visibles pour le Ni^{2+} étant donnée la faible intensité d'autofluorescence à 456 nm. Les résultats sont similaires pour C-TERa, à l'exception des spectres associés au Cu^{2+} et au Zn^{2+} . En effet, la liaison au Cu^{2+} engendre une nouvelle bande d'excitation à 300 nm (Fig. 44C) et l'addition du Zn^{2+} déplace la quatrième contribution de 400 à 410 nm. De plus, étant donné que ce métal déplace la bande d'émission dbAF de 456 à 490 nm, un spectre d'excitation à cette dernière longueur d'onde a été enregistré (Fig. 44D). Alors que les deux premières bandes sont conservées à 237 et 270 nm, la bande d'excitation principale se retrouve déplacée à 415 nm et

une épaule vers 340 nm est également observée. Ces signatures impliquent que les propriétés dbAF de C-TERa sont grandement affectées par l'addition de Zn^{2+} . En accord avec les spectres ITyrF, des modifications structurales et spectrales de C-TERa et de ses résidus Tyr occurrent en présence de Cu^{2+} et Zn^{2+} .¹²⁵

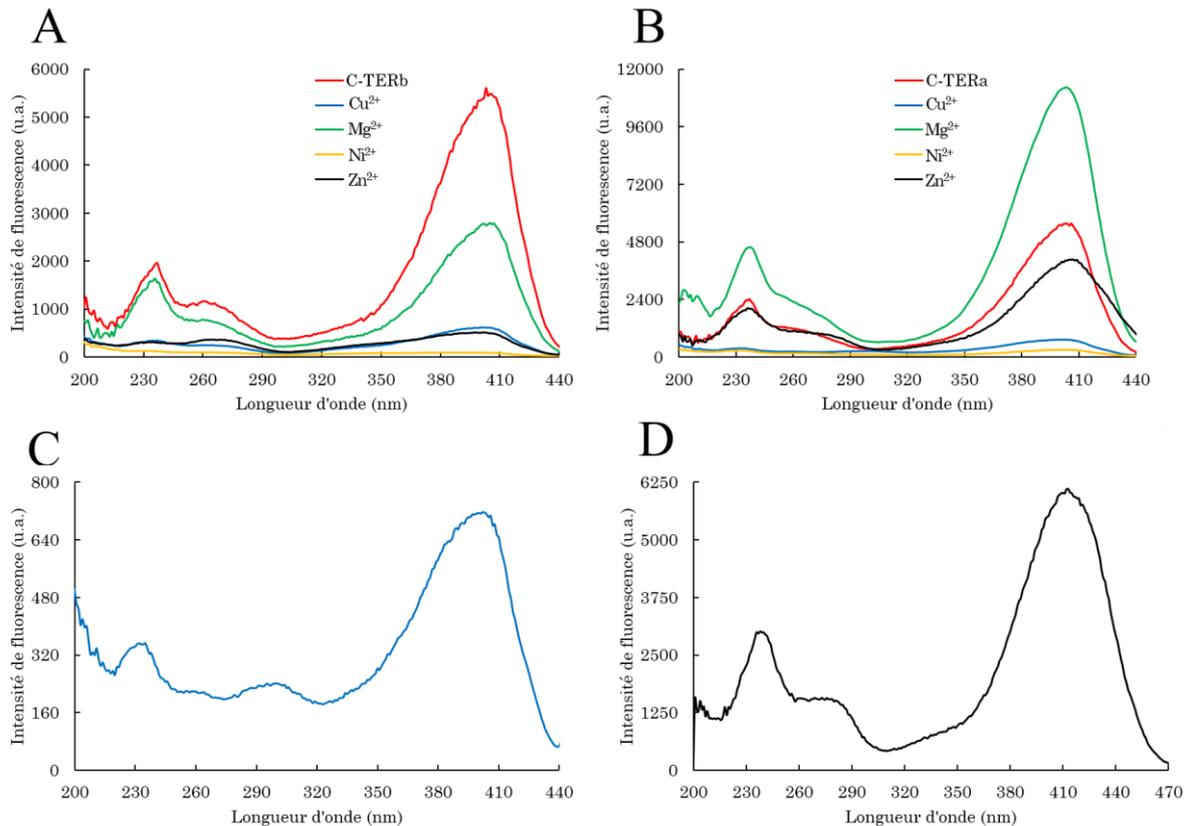


Figure 44 – Spectres d'excitation dbAF de C-TERb et C-TERa après 72 h d'incubation dans le TPB à ~ 20 °C sans métal (rouge) et en présence de Cu^{2+} (bleu), Mg^{2+} (vert), Ni^{2+} (jaune) et Zn^{2+} (noir). Spectres d'excitation dbAF ($\lambda_{em} = 456$ nm, $sw = 10$ nm) de (A) C-TERb et (B) C-TERa. Spectres d'excitation dbAF de C-TERa en présence de (C) Cu^{2+} et (D) Zn^{2+} ($\lambda_{em} = 490$ nm, $sw = 10$ nm).¹²⁵

L'état de fibrillation amyloïde a été vérifiée par test de liaison à la ThT. Pour le C-TERb (Fig. 45A), l'addition des métaux conduit à la diminution de l'intensité à 485 nm, indiquant que leur présence génère des espèces agrégées non amyloïdes et/ou fibrilles moins sensibles à la ThT. Alors que le Mg^{2+} diminue l'émission de moitié, le Ni^{2+} l'éteint presque totalement, signifiant qu'une faible proportion de fibrilles est formée, ce qui est en accord avec les résultats dbAF. C'est pourquoi, le Ni^{2+} favoriserait la formation d'oligomères qui autofluorescent moins et qui ne lient pas la ThT. Concernant le C-TERa (Fig. 45B), des résultats similaires sont obtenus pour le Ni^{2+} et le Cu^{2+} , présentant une forte diminution de la fluorescence de la ThT. Au contraire, le Mg^{2+} accélère la formation de fibrilles amyloïdes, ce qui est rendu visible par

l'effet hyperchrome observé sur la bande d'émission à 485 nm. Quant au Zn^{2+} , son addition s'accompagne d'une faible diminution de l'intensité ainsi qu'un déplacement de la bande d'émission vers 500 nm, suggérant, au regard des résultats dbAF, la formation de fibrilles de structure différente.¹²⁵

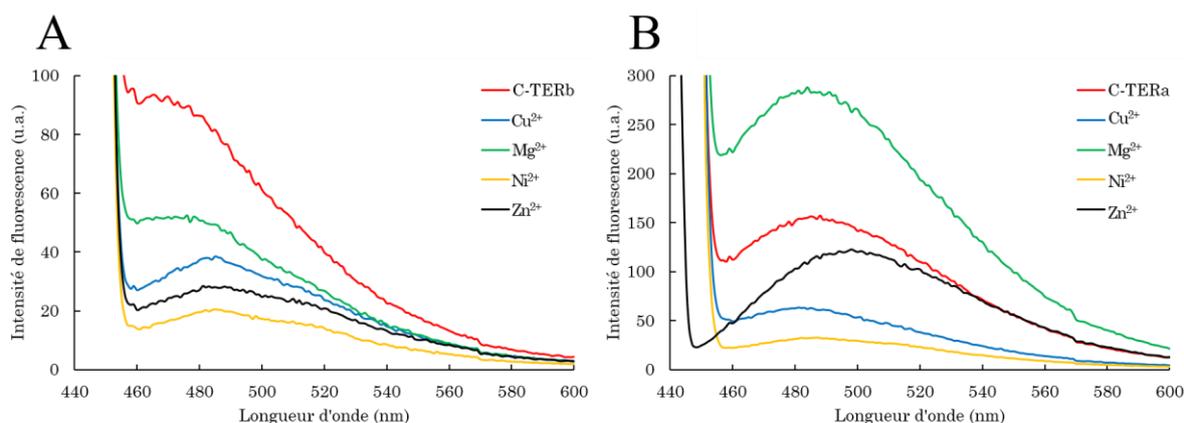


Figure 45 – Spectres d'émission de la ThT ($\lambda_{ex} = 440$ nm, $sw = 10$ nm) de (A) C-TERb et (B) C-TERa après 72 h d'incubation dans le TPB à ~ 20 °C sans métal (rouge) et en présence de Cu^{2+} (bleu), Mg^{2+} (vert), Ni^{2+} (jaune) et Zn^{2+} (noir).¹²⁵

3.4.3. Influence des cations divalents sur la cinétique d'agrégation des régions C-terminales

Etant donné qu'il a été montré à travers les essais précédents que la dbAF est aussi efficace et fiable que la ThT pour décrire les cinétiques d'agrégation, l'évolution de l'intensité dbAF à 456 nm, après excitation à 400 nm, a été mise en relation avec le temps d'incubation pour chaque condition. Concernant le C-TERb (Fig. 46A), la phase de latence ne dure que 24 h alors qu'elle prenait 36 h pour une concentration deux fois plus importante (Fig. 39A). Il est reconnu que les cinétiques d'agrégation dépendent de la concentration mais la phase de latence, correspondant à la nucléation, est souvent plus longue à plus faible concentration étant donné que moins d'interactions intermoléculaires se produisent.¹⁶⁰ Curieusement, l'effet inverse est ici obtenu. Nonobstant cette observation, les métaux ont des effets remarquables sur les profils cinétiques. Comme attendu à partir des différents spectres de fluorescence, l'intensité du plateau associé à la phase stationnaire est réduite en présence de chaque métal selon cet ordre : $Ni^{2+} > Zn^{2+} \geq Cu^{2+} > Mg^{2+}$. Le Ni^{2+} diminue fortement la vitesse de fibrillation, ce qui allonge la phase d'élongation de 26 à 66 h. Le Cu^{2+} et le Zn^{2+} voient leur temps de latence étendu à 36 h. De plus, la faible intensité de leur plateau est en accord avec leurs signaux en fluorescence de la ThT. En effet, il est probable qu'il y ait moins de fibrilles amyloïdes

formées par rapport à la quantité d'oligomères. Le Mg^{2+} impacte uniquement l'intensité de la phase stationnaire, suggérant également une plus faible proportion en fibrilles.¹²⁵

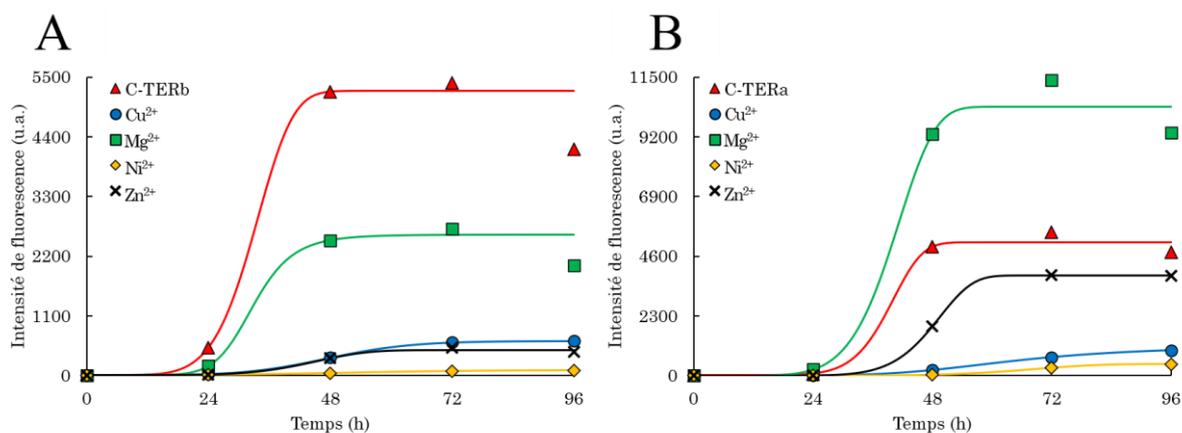


Figure 46 – Courbes cinétiques par suivi de l'émission dbAF ($\lambda_{ex} = 400 \text{ nm}$, $\lambda_{ém} = 456 \text{ nm}$, $sw = 10 \text{ nm}$) de l'agrégation de (A) C-TERb et de (B) C-TERa sans métal (triangle rouge) et en présence de Cu^{2+} (cercle bleu), Mg^{2+} (carré vert), Ni^{2+} (losange jaune) et Zn^{2+} (croix noire) dans le TPB à $\sim 20 \text{ }^\circ\text{C}$ après 72 h d'incubation.¹²⁵

A l'image de C-TERb, C-TERa (Fig. 46B) possède une phase de latence de 24 h en l'absence de métaux. Encore une fois, ce résultat est surprenant étant donné sa plus faible concentration. Concernant les métaux, le Ni^{2+} et le Cu^{2+} ont un effet similaire en rallongeant la phase de latence à 50 h et la phase stationnaire à plus de 96 h. Comparativement à C-TERb et comme attendu des différentes fluorescences, leur plateau se situe à une intensité bien plus faible qu'en l'absence de cations métalliques. Par conséquent, des oligomères ou des agrégats ne présentant qu'une faible autofluorescence et liaison à la ThT sont favorisés en présence de ces métaux. La présence de Mg^{2+} engendre, au contraire, un plateau avec une intensité plus élevée, alors que les phases de latence et d'élongation ne sont pas affectées. Cohéremment avec la fluorescence de la ThT, une plus grande proportion en fibrilles en est sans doute la cause. Le Zn^{2+} conduit à la phase d'élongation la plus courte ainsi qu'à une phase de latence plus longue (44 h). Ceci peut être expliqué par la formation d'intermédiaires et de fibrilles par des mécanismes différents des autres métaux.¹²⁵

3.4.4. Analyse de séquence des régions C-terminales

La diversité des effets due aux ions métalliques sur les isoformes C-TERb et C-TERa trouve vraisemblablement son origine dans la structure primaire de ces domaines, la distribution des résidus impliqués dans les liaisons aux métaux (Fig. 47) ainsi que dans la

spécificité que possède ces derniers vis-à-vis de certains résidus. En effet, il est connu que les cations Cu^{2+} et Ni^{2+} lient préférentiellement les résidus histidines (His) alors que le Mg^{2+} a une plus grande affinité pour les aspartates (Asp) et glutamates (Glu). Le Zn^{2+} quant à lui se lie aux cystéines (Cys) et aux His.^{161,162} Les comportements différents entre les C-TER induits par le Zn^{2+} sont explicables par le fait que le C-TERb est particulièrement enrichi en résidus Cys et en contient plus que le C-TERa, de par la présence, notamment, des domaines PHD-1 et PHD-2 qui en nécessitent pour coordonner au total quatre Zn^{2+} . Toutefois, un excès de Zn^{2+} peut induire des réarrangements conformationnels générant de nouveaux sites de liaison. L'effet similaire des cations Cu^{2+} et Ni^{2+} sur la cinétique d'agrégation de C-TERb et C-TERa peut être expliqué par le nombre comparable de résidus His dans la séquence et le fait que ces derniers sont répartis de manière comparable dans les deux séquences. Néanmoins, les différences spectrales entre ces deux métaux pourraient provenir de leur géométrie de coordination, à savoir trigonale plan pour Cu^{2+} et octaédrique pour Ni^{2+} .¹⁶¹ Les comportements distincts de C-TERb et C-TERa par rapport à la présence de Mg^{2+} peuvent venir de la localisation différente des résidus Asp et Glu.¹²⁵

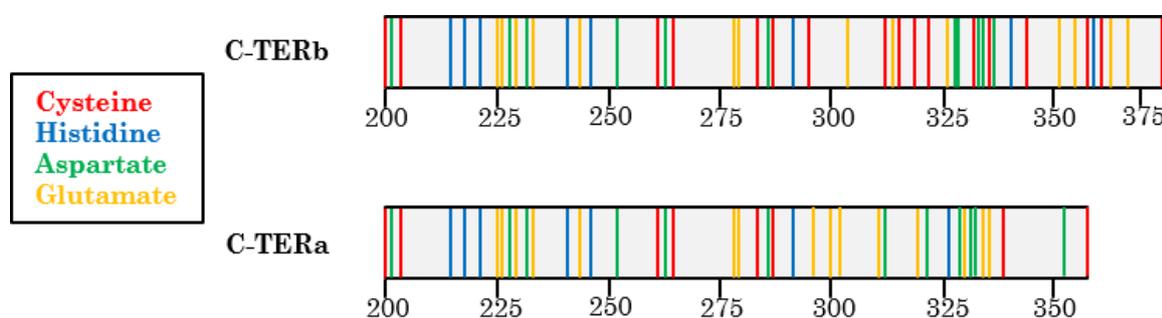


Figure 47 – Localisation des acides aminés impliqués dans les liaisons protéines-métal dans les séquences de C-TERb (en haut) et C-TERa (en bas). Les résidus représentés par des bâtonnets sont les cystéines (rouge), histidines (bleu), aspartates (vert) et glutamates (jaune).¹²⁵

3.4.5. Influence de cations divalents sur la morphologie des agrégats des régions C-terminales

À la lueur de l'influence des cations sur l'agrégation et plus particulièrement la fibrillation, la TEM a été employée afin de comparer les propriétés spectrales avec la morphologie des agrégats formés. Les images TEM ont été prises après 144 h d'incubation à ~ 20 °C dans le TPB.

En présence de Cu^{2+} , C-TERb forme des longues fibrilles courbées de 16 nm de large le long desquelles apparaissent des points de rupture (Fig. 48A). Cela fragilise la structure des fibrilles et peut amener à leur fragmentation comme observée pour C-TERa en présence de

Cu^{2+} (Fig. 49A). Des fragmentations sont également visualisées sur les fibrilles de 18 à 20 nm de large lors de la liaison du Mg^{2+} (Fig. 48B), Ni^{2+} (Fig. 48C) et Zn^{2+} (Fig. 48D) au domaine C-TERb. Pour le Ni^{2+} , des points de rupture sont aussi observés sur des fibrilles qui s'assemblent en fibrilles droites multiples (MSFs) provoquant leur dissociation. Le Ni^{2+} cause également des cassures sur les fibrilles de C-TERa (Fig. 49C). La génération de ces fragments ou de ces points de rupture vient de la fragilisation et de la rigidification des fibrilles induites par la liaison aux métaux. Ces fragments expliqueraient les émissions de faible intensité retrouvées en dbAF et fluorescence de la ThT. En effet, un comportement similaire a été décrit pour des protéines fibrillées et notamment pour la protéine tau en présence de Ni^{2+} .^{163,164}

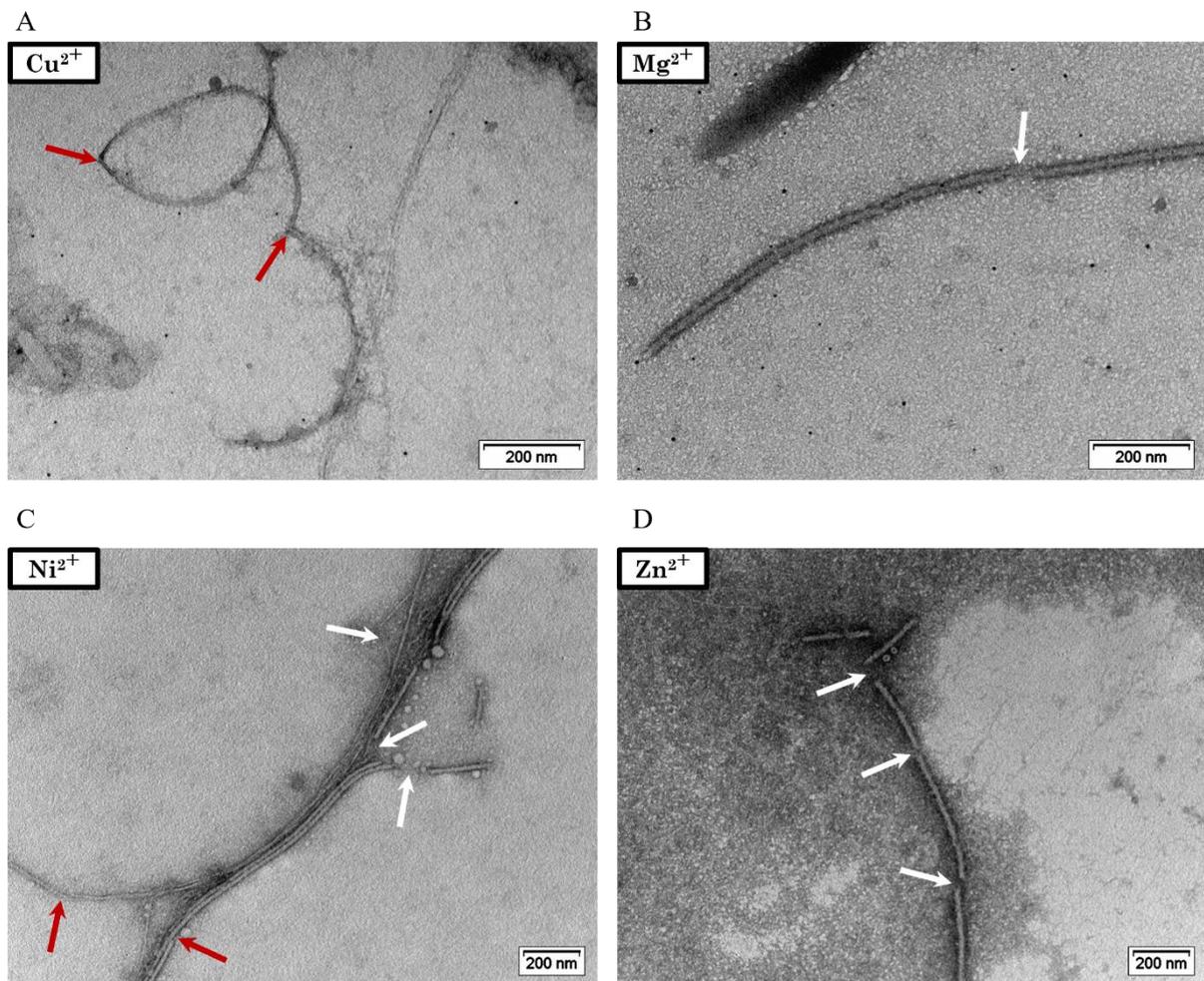


Figure 48 – Micrographies TEM par negative staining de C-TERb incubé pendant 144 h dans le TPB à ~ 20 °C en présence de (A) Cu^{2+} , (B) Mg^{2+} , (C) Ni^{2+} et (D) Zn^{2+} . Dans certaines conditions, les fibrilles formées présentent des points de rupture (indiqués par des flèches rouges) et peuvent être fragmentées (les ruptures sont indiquées par des flèches blanches). Sur chaque image, l'échelle est indiquée dans le coin inférieur droit.¹²⁵

En ce qui concerne le C-TERa en présence de Mg^{2+} (Fig. 49B), aucune fragmentation n'est observée. Au contraire, à l'image de C-TERa sans métal, des fibrilles droites simples (SSFs) de 18 nm de large ainsi que certaines droites doubles (DSFs) sont détectées. Leur propriété de flexibilité semble également conservée. La présence de Zn^{2+} sur C-TERa (Fig. 49D) n'influence pas la forme des fibrilles mais leur largeur. En effet, des SSFs plus fines, de 5 à 8 nm de large, sont retrouvées assemblées en MSFs. Cette différence de largeur peut notamment expliquer les signatures uniques retrouvées en autofluorescence et en FCD. Il est important de noter que des oligomères formés en présence des métaux ne sont probablement pas visualisés en TEM en raison de leur trop petite taille. Au regard des analyses spectroscopiques, cela doit être le cas pour C-TERb en présence de Ni^{2+} , Cu^{2+} et Zn^{2+} ainsi que pour C-TERa avec du Ni^{2+} et du Cu^{2+} .¹²⁵

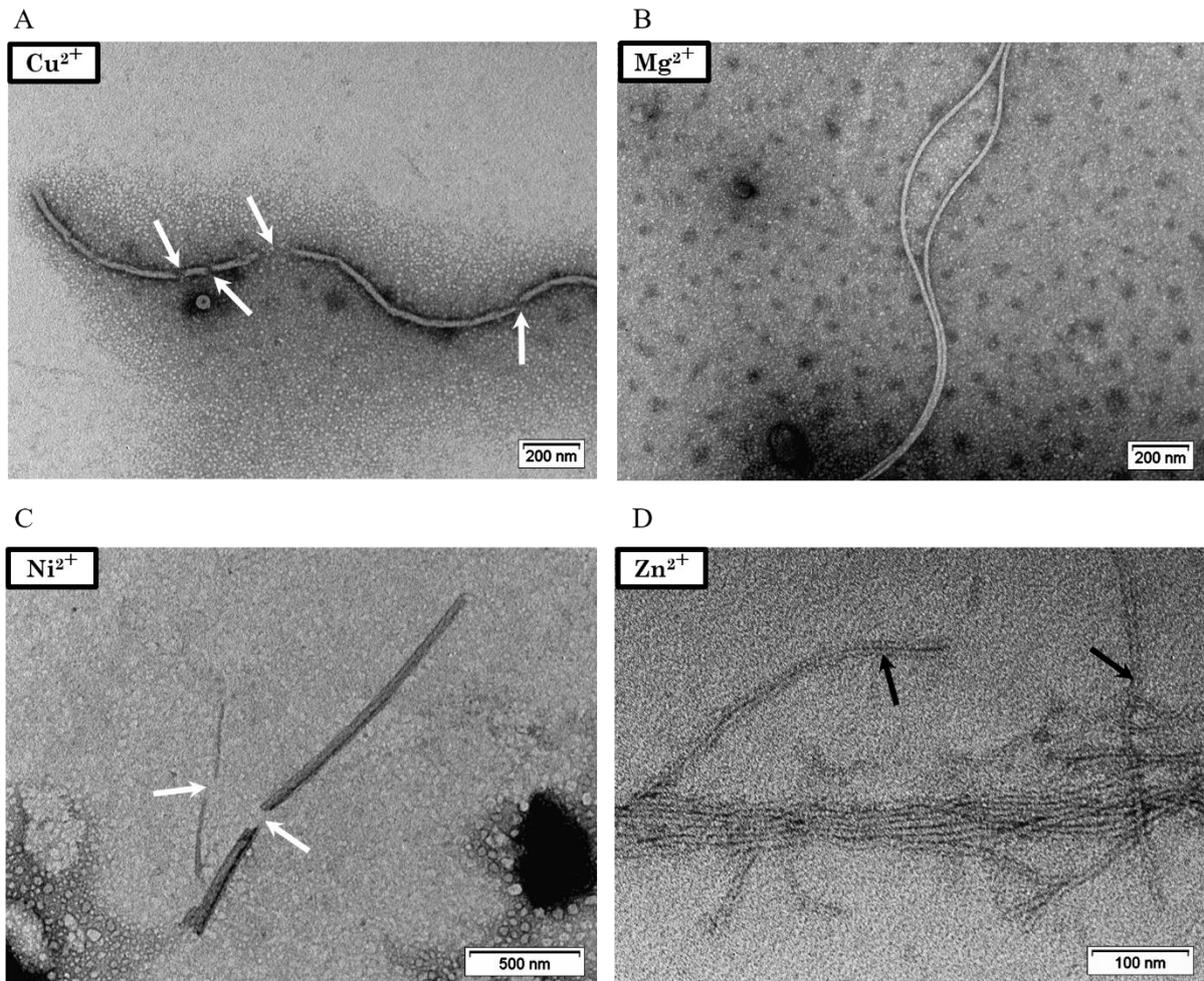


Figure 49 – Micrographies TEM par negative staining de C-TERa incubé pendant 144 h dans le TPB à ~ 20 °C en présence de (A) Cu^{2+} , (B) Mg^{2+} , (C) Ni^{2+} et (D) Zn^{2+} . Dans certaines conditions, les fibrilles formées peuvent être fragmentées (les ruptures sont indiquées par des flèches blanches). Sur le panneau D, les flèches noires indiquent les fibrilles droites simples (SSFs) qui se regroupent en fibrilles droites multiples (MSFs). Sur chaque image, l'échelle est indiquée dans le coin inférieur droit.¹²⁵

Partie n°4

—

Conclusions et perspectives

4.1. Conclusions

Ce travail avait pour objectif de recherche la caractérisation structurale de deux protéines amyloïdogènes intrinsèquement désordonnées, à savoir les isoformes DPF3b et DPF3a, par une combinaison de techniques prédictives et biophysiques. En effet, DPF3 est impliquée dans des différentes formes de cancer ainsi que dans des troubles neurodégénératifs, comme la maladie de Parkinson et d'Alzheimer. Dans un premier temps, la nature désordonnée et la propension à l'agrégation des protéines complètes (DPF3b FL et DPF3a FL) ont été étudiées. Dans un second temps, les caractéristiques structurales de leur région C-terminale (C-TERb et C-TERa), qui les différencient, ont été investiguées plus spécifiquement. Enfin, l'influence d'une série de cations métalliques divalents (Cu^{2+} , Mg^{2+} , Ni^{2+} et Zn^{2+}) sur les propriétés d'agrégation des domaines C-terminaux a été évaluée.

Avant les étapes de caractérisation, il a été montré que le procédé de purification par FPLC des protéines de fusion DPF3-GST est efficace. Nonobstant l'impossibilité de visualiser les protéines d'intérêt clivées sur gel SDS-PAGE, elles ont été identifiées par spectrométrie de masse.

Sur la base d'algorithmes de prédiction, les deux isoformes se sont révélées être de protéines désordonnées arborant un ensemble conformationnel étendu. De ce fait, le pourcentage de désordre moyen par résidu de DPF3a FL s'élève à 63 % alors qu'il n'est que de 44 % pour l'isoforme DPF3b FL. Cette dernière contient seulement deux régions intrinsèquement désordonnées (IDRs) contre trois pour DPF3a FL dont l'IDR supplémentaire est localisée dans la partie C-terminale.

La nature désordonnée des isoformes de DPF3 a été confirmée expérimentalement par DLS montrant des \bar{D}_h similaires pour les deux protéines et correspondant à des conformations étendues caractéristiques des IDPs. Des signatures majoritairement associées à du désordre ont par ailleurs été observées par FCD. Soutenus par les prédictions d'accessibilité au solvant, l'étude des spectres ITF et ITyrF a révélé que les résidus Trp et Tyr de DPF3b FL sont moins exposés au solvant que ceux de DPF3a FL. Ces spectres ont également dévoilé l'existence de FRET entre des résidus Tyr et Trp proches au niveau de la séquence. Complémentairement, l'analyse des spectres d'absorption UV-visible a montré que DPF3b FL et DPF3a FL possèdent un faible indice de repliement, appuyant le caractère étendu de leur état conformationnel. C'est pourquoi, bien que DPF3a FL soit davantage désordonnée de par l'absence des domaines ZnFs du côté C-terminal, les deux isoformes appartiennent à la classe des IDPs.

Concernant la propension à l'agrégation des isoformes complètes, la DLS a confirmé la formation d'espèces agrégées de grande taille pour DPF3b, ce qui avait déjà été prouvé pour DPF3a. Par FCD, une augmentation progressive du contenu en feuillets β antiparallèles a été observée au cours du temps pour les deux isoformes avec la formation d'intermédiaires enrichis en hélices α pour DPF3a FL. Ces réarrangements s'opèrent en seulement 48 h pour DPF3a FL contre 96 h pour DPF3b FL, suggérant que l'isoforme DPF3a FL se fibrille plus rapidement.

Le gain en feuillets β antiparallèles s'accompagne de modifications conformationnelles révélées par ITF et ITyrF. Après 72 h d'incubation, des déplacements hypsochromes et une augmentation d'intensité de la bande d'émission relative aux résidus Trp de DPF3b et DPF3a ont montré que les Trp sont plus enfouis. Les spectres ITyrF des deux isoformes voient l'épaule à 305 nm disparaître au cours de l'incubation ainsi que la bande d'émission associée au FRET Trp-Tyr augmenter en intensité au cours du temps. Les Tyr se trouvent donc dans un environnement moins polaire et des résidus Trp et Tyr sont d'autant plus rapprochés au sein d'un noyau hydrophobe, probablement amyloïde.

A la fois sur les spectres ITF et ITyrF, une deuxième bande d'émission correspondant au phénomène d'autofluorescence dans le bleu profond (dbAF) et indicatrice de fibrillation, a émergé à 456-460 nm pour chaque isoforme. L'étude des spectres d'excitation de cette autofluorescence a révélé trois contributions principales : une relative aux liaisons peptidiques vers 235 nm, une associée aux résidus aromatiques vers 270 nm ainsi qu'une dernière, majoritaire, à 400 nm.

La liaison des deux isoformes à deux colorants (ThT et RC) a validé la nature amyloïde des agrégats de DPF3 ainsi que la fibrillation plus rapide de DPF3a FL. En effet, l'agrégation de DPF3b FL semble plus lente en raison de transitions successives ordre-désordre-ordre, s'expliquant par le dépliage des ZnFs PHD-1 et PHD-2 suivi du repliage en un noyau amyloïde. Inversement, DPF3a FL se fibrille plus rapidement grâce à son domaine C-terminal dont les propriétés désordonnées et dynamiques permettent des transitions désordre-ordre énergétiquement moins coûteuses. Pour les deux isoformes, ces essais ont également montré que des agrégats non amyloïdes sont en premier lieu générés.

A partir des signatures spectroscopiques et des images TEM, il a été montré que les deux isoformes ont une cinétique d'agrégation différente mais suivent un mécanisme semblable, excepté lors des premières étapes de modification structurale (Fig. 50). En effet, DPF3b FL se réarrange en premier lieu en perdant des hélices α , tandis que DPF3a FL forme

des intermédiaires enrichis en hélices α . Par la suite, la proportion en feuillets β antiparallèles des deux isoformes augmente progressivement afin de former un noyau amyloïde, conduisant à l'assemblage en oligomères. Ces oligomères se rassemblent en unités de nucléation sphériques (SNUs), agissant comme des points de nucléation pour la formation de protofibrilles granulaires. Ces dernières s'allongent, s'assemblent et s'étirent pour former des filaments droits (SFs), se regroupant en fibrilles striées et torsadées en forme de ruban (STRFs) dont les stries correspondent à des SFs individuels. Les STRFs de tailles diverses s'enchevêtrent dans de denses réseaux fibrillaires.

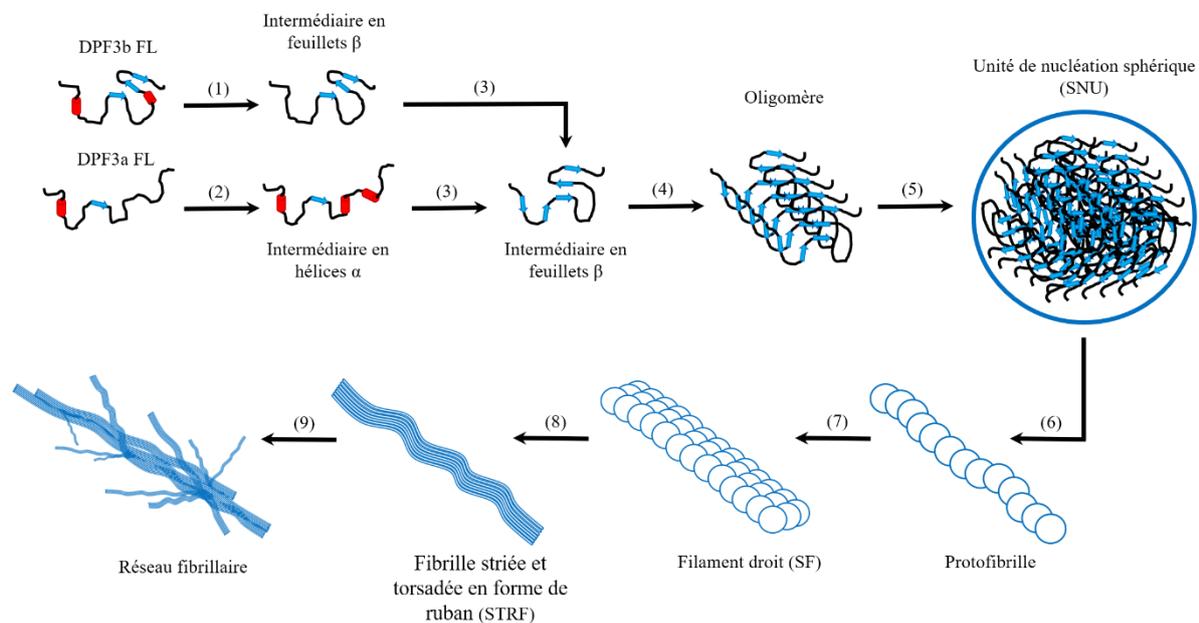


Figure 50 – Représentation schématique du mécanisme de fibrillation proposé des isoformes de DPF3. (1) Réarrangement de DPF3b FL avec une perte en hélices α . (2) Intermédiaire de DPF3a FL enrichis en hélices α . (3) Noyau amyloïde des deux isoformes riche en feuillets β antiparallèles. (4) Assemblage d'oligomères. (5) Formation de SNUs. (6) Formation de protofibrilles granulaires. (7) Assemblage en SFs et (8) en STRFs. (9) Réseau fibrillaire. Les hélices α sont représentées par des cylindres rouges et les feuillets β antiparallèles par des flèches bleues.¹²⁴

Les isoformes se différenciant par leur région C-terminale, le caractère désordonné de C-TERb et C-TERa a été investigué par différentes techniques spectroscopiques. Les spectres FCD ont révélé 57 % de désordre pour C-TERb contre 66 % pour C-TERa, ce qui est cohérent au vu de l'IDR-3 présente pour C-TERa. C-TERb contient également plus de feuillets β antiparallèles et d'hélices α de par ses domaines ZnFs PHD-1 et PHD-2 riches en ces structures. Les Trp de C-TERb, qui se trouvent dans les ZnFs PHD-1 et PHD-2, sont partiellement exposés au solvant et possèdent donc une face exposée. La présence de FRET entre les résidus Trp et Tyr de C-TERb a été supportée par les spectres ITyrF. Pour C-TERa, le spectre ITyrF présente

une bande d'émission déplacée à 328 nm, ce qui en l'absence de résidus Trp, provient vraisemblablement de la formation de tyrosinates de par la proximité avec des Asp et/ou des Glu.

A l'instar des protéines complètes, C-TERb et C-TERa ont montré, par combinaison de spectroscopies et de microscopie, une tendance à s'agréger spontanément. Les spectres FCD ont révélé que les parties C-terminales s'enrichissent au cours du temps en feuillets β parallèles au détriment des feuillets β antiparallèles initialement présents. Bien que C-TERb et C-TERa gagnent progressivement des feuillets β parallèles, les deux isoformes passent par des intermédiaires différents. Alors que C-TERb s'enrichit en hélices α avant de les perdre après 48 h, C-TERa forme directement des intermédiaires riches en feuillets β antiparallèles.

Comme pour les isoformes complètes, une bande d'autofluorescence dbAF à 456-460 nm est apparue au fil du temps sur les spectres ITF et ITyrF. Le spectre ITF de C-TERb a dévoilé que les résidus Trp ne semblent pas impliqués dans la formation d'un noyau amyloïde de par l'absence de déplacement de la bande d'émission. A l'inverse, l'environnement des Tyr est modifié au cours du temps, avec une augmentation du FRET Trp-Tyr et un enfouissement des résidus Tyr. Concernant C-TERa, ses résidus Tyr sont retrouvés dans un environnement qui vraisemblablement favorise davantage la génération de tyrosinates.

L'évolution de l'intensité dbAF au cours du temps a indiqué une agrégation en fibrilles amyloïdes plus rapide pour le C-TERb par rapport au C-TERa. Ce comportement, notamment opposé aux isoformes complètes, peut s'expliquer par des mécanismes d'agrégation différents, un nombre et une taille plus importants des fibrilles formées pour le C-TERa. Au-delà de confirmer ces tendances, les tests de liaison à la ThT ont également montré que des agrégats non amyloïdes sont en premier lieu formés pour les deux régions C-terminales.

A partir des données spectroscopiques et des images obtenues par TEM, un mécanisme de fibrillation semblable, à l'exception des premières étapes, entre les deux C-TER a été proposé (Fig. 51). En effet, C-TERb se réarrange en formant des intermédiaires enrichis en hélices α avant de les perdre et de s'enrichir en feuillets β parallèles. C-TERa forme des intermédiaires riches en feuillets β antiparallèles avant de gagner en feuillets β parallèles comme C-TERb. Les monomères restructurés s'assemblent en oligomères composés de feuillets β parallèles et antiparallèles. Par addition de monomères, les oligomères s'allongent en fibrilles droites simples (SSFs) qui peuvent se regrouper par paire en fibrilles droites doubles

(DSFs). Les SSFs et DSFs ont également tendance à se rassembler en fibrilles droites multiples (MSFs). Toutes ces fibrilles sont retrouvées enchevêtrées dans des réseaux fibrillaires, à l'instar des isoformes complètes.

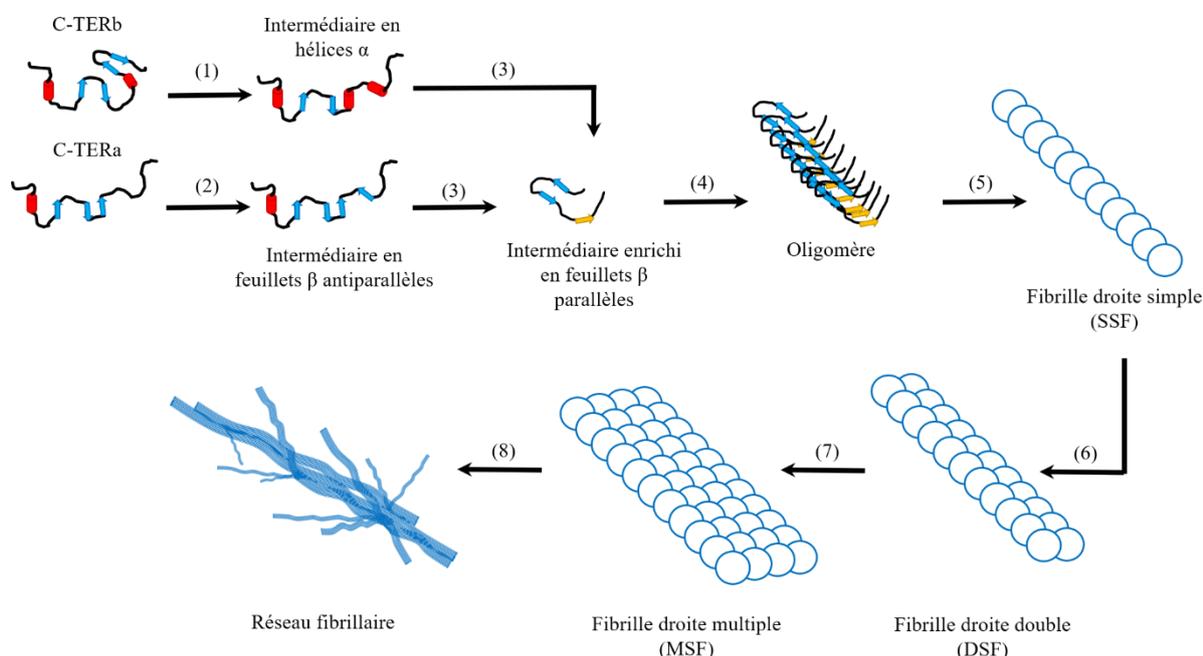


Figure 51 – Représentation schématique du mécanisme de fibrillation proposé de C-TERb et C-TERa. (1) Intermédiaires de C-TERb enrichis en hélices α . (2) Intermédiaires de C-TERa enrichis en feuillets β antiparallèles. (3) Intermédiaires enrichis en feuillets β parallèles. (4) Assemblage d'oligomères. (5) Formation de fibrilles droites simples (SSFs). (6) Formation de fibrilles droites doubles (DSFs). (7) Formation de fibrilles droites multiples (MSFs). (8) Réseau fibrillaire. Les hélices α sont représentées par des cylindres rouges, les feuillets β antiparallèles et parallèles respectivement par des flèches bleues et jaunes.

Etant donné l'implication des métaux dans les processus de neurodégénérescence, l'influence des cations métalliques divalents, à savoir Cu^{2+} , Mg^{2+} , Ni^{2+} et Zn^{2+} , sur les propriétés d'agrégation des régions C-TER a également été investiguée. Les spectres FCD ont montré que les métaux n'impactent pas grandement le réarrangement conformationnel de C-TERb. Même si chaque métal promeut la formation de feuillets β antiparallèles, la formation de feuillets β parallèles n'est pas altérée. Plus de modifications sont visibles pour C-TERa. Ainsi, alors que le Cu^{2+} n'influence que très peu la composition en structure secondaire, le Mg^{2+} favorise le réarrangement en feuillets β parallèles et appauvrit les conformations en feuillets β antiparallèles. Le Ni^{2+} et le Zn^{2+} favorisent au contraire la formation de ces derniers.

En ce qui concerne la structure tertiaire de C-TERb, au-delà de ne pas affecter la position de la bande d'émission des Trp, les métaux provoquent un effet hypochrome de cette dernière. Alors que le Cu^{2+} , Mg^{2+} et Ni^{2+} sont des extincteurs de la fluorescence des Trp

capables de s'y lier, le Zn^{2+} conduit à des réarrangements rapprochant des Trp de C-TERb à certains groupements diminuant l'intensité d'émission. Sur les spectres ITyrF de C-TERb, seul le Cu^{2+} provoque un effet hypochrome de la bande d'émission de FRET, suggérant que les résidus Tyr sont liés au Cu^{2+} et sont retrouvés dans un environnement plus polaire. Comme pour C-TERb, le Cu^{2+} diminue aussi l'intensité de la bande d'émission des Tyr de C-TERa par liaison à ces résidus. Quant au Zn^{2+} , il ne se lie pas directement aux Tyr mais est responsable d'un changement conformationnel engendrant le rapprochement, dans l'environnement des Tyr, de groupes extincteurs de fluorescence.

L'étude des cinétiques d'agrégation dbAF a révélé pour C-TERb que l'intensité du plateau associé à la phase stationnaire est réduite en présence de chaque métal selon cet ordre : $Ni^{2+} > Zn^{2+} \geq Cu^{2+} > Mg^{2+}$. L'ensemble des métaux diminuerait donc la cinétique de formation des fibrilles. Le Ni^{2+} et le Cu^{2+} ont un effet comparable sur C-TERa, en allongeant à la fois la phase de latence et la phase stationnaire. Le déplacement bathochrome de la bande dbAF de C-TERa en présence de Zn^{2+} s'expliquerait par la formation de fibrilles de structure différente. De ce fait, un temps de latence plus long mais une phase d'élongation très courte sont observés en présence de ce métal. Le Mg^{2+} influence uniquement l'intensité du plateau en l'augmentant par rapport à C-TERa sans métal de par le plus grand nombre de fibrilles formées.

Il a été montré par la fluorescence de la ThT que la présence de cations métalliques, en particulier le Ni^{2+} , favorise la formation d'oligomères plutôt que de fibrilles pour C-TERb. En regard des spectres de C-TERa, des résultats identiques ont été obtenus pour le Ni^{2+} et le Cu^{2+} . Cependant, la liaison de Mg^{2+} au C-TERa augmente l'émission de la ThT, ce qui confirme une proportion plus élevée en fibrilles amyloïdes. Pour le Zn^{2+} , un déplacement bathochrome de la bande d'émission de la ThT supporte la formation de fibrilles structurellement distinctes.

Alors que des oligomères n'ont pas pu être détectés par TEM, probablement à cause de leur trop petite taille, des fibrilles présentant des points de rupture et des cassures, dus à leur fragilisation ainsi qu'à leur rigidification, ont été observées pour C-TERb en présence des métaux (Fig. 52). Le même phénomène a été identifié pour C-TERa avec du Cu^{2+} et du Ni^{2+} . Cette fragmentation est à l'origine des faibles intensités d'autofluorescence et d'émission de la ThT. A l'inverse, le Mg^{2+} , à l'instar de C-TERa sans métal, induit des SSFs ainsi que des DSFs. Enfin, l'addition de Zn^{2+} génère des fibrilles plus fines assemblées en MSFs, ce qui expliquerait les signatures spectrales (FCD, dbAF et ThT) uniques observées pour C-TERa en présence de Zn^{2+} .

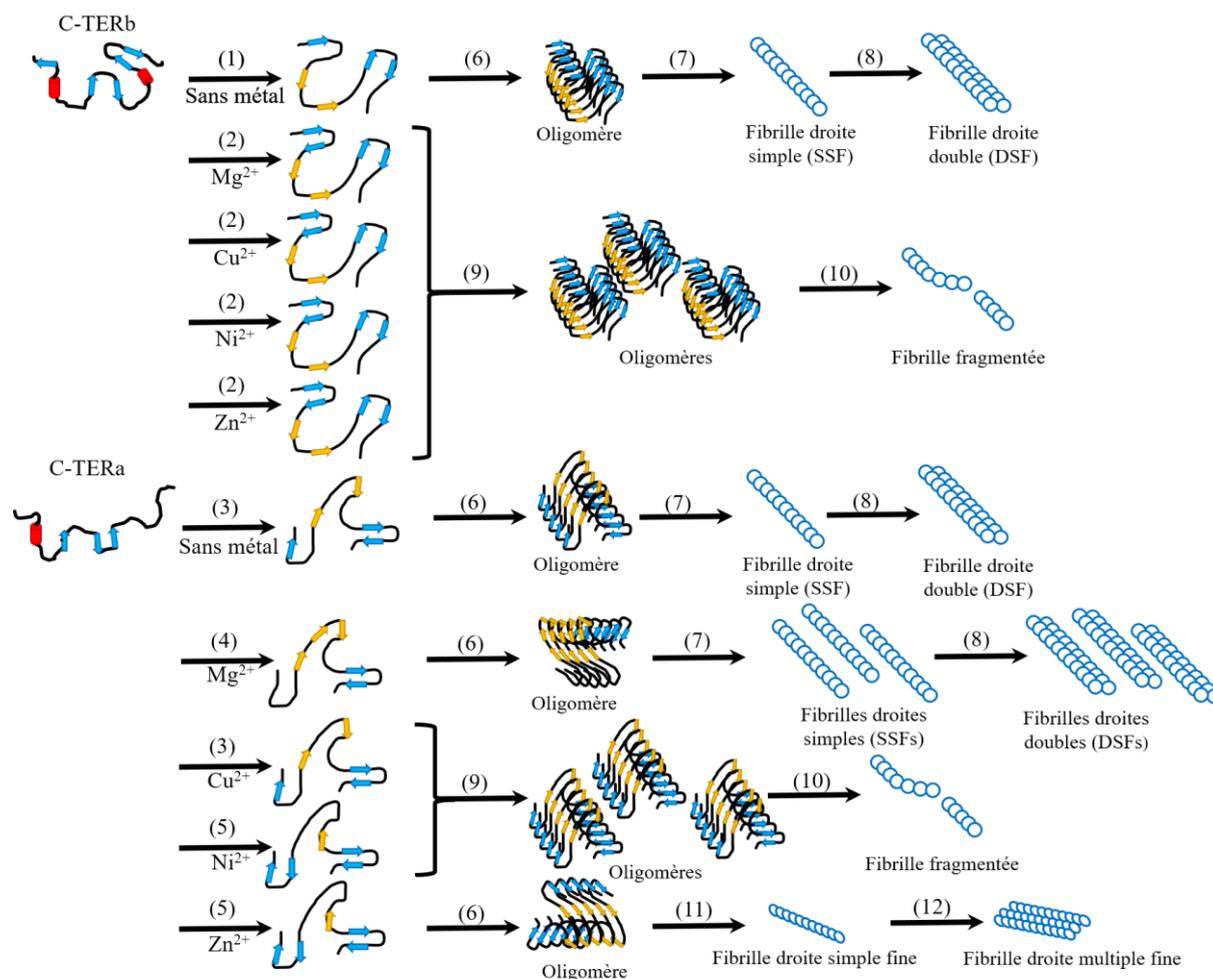


Figure 52 – Représentation schématique du mécanisme de fibrillation proposé de C-TERb et C-TERa en présence de Cu^{2+} , Mg^{2+} , Ni^{2+} et Zn^{2+} . (1) Intermédiaires de C-TERb sans métal. (2) Intermédiaires de C-TERb en présence des cations. (3) Intermédiaires de C-TERa sans métal et en présence de Cu^{2+} . (4) Intermédiaires de C-TERa en présence de Mg^{2+} . (5) Intermédiaires de C-TERa en présence de Ni^{2+} et de Zn^{2+} . (6) Formation d'une petite fraction en oligomères. (7) Formation de fibrilles droites simples (SSFs). (8) Formation de fibrilles droites doubles (DSFs). (9) Formation d'une grande fraction en oligomères. (10) Formation de fibrilles droites simples plus fines. (11) Formation de fibrilles droites simples plus fines. (12) Formation de fibrilles droites multiples plus fines. Les hélices α sont représentées par des cylindres rouges, les feuilletts β antiparallèles et parallèles respectivement par des flèches bleues et jaunes.

4.2. Perspectives

Ce travail de recherche s'est porté sur la caractérisation structurale de deux isoformes amyloïdogènes intrinsèquement désordonnées. La nature désordonnée et la propension à la fibrillation de DPF3b et DPF3a ainsi que de leur régions C-terminales ont été démontrées. Il serait intéressant d'appliquer les études réalisées dans le cadre de mémoire à la région N-terminale, commune aux deux isoformes. En effet, les divergences spectrales et morphologiques entre les isoformes complètes et leur régions C-terminales suggèrent l'implication de la partie N-terminale dans le mécanisme de fibrillation de DPF3b FL et DPF3a FL.

Complémentaire à ce qui a déjà été réalisé, la RMN (résonance magnétique nucléaire) apporterait une autre approche structurale afin de résoudre la structure des isoformes de DPF3 à l'échelle atomique et d'obtenir leur ensemble conformationnel respectif. Des simulations par dynamique moléculaire y seraient également complémentaires pour évaluer le repliement de DPF3 dans différentes conditions.

Afin d'investiguer plus en profondeur les propriétés amyloïdogènes des isoformes, l'influence de plusieurs paramètres physico-chimiques, tels que le pH, la température, la concentration en protéine ou encore la force ionique, sur l'agrégation et la fibrillation pourra être testée à la fois sur les protéines complètes ainsi que sur les domaines C et N-terminaux. Bien que, dans le présent travail, seul l'effet de certains cations métalliques divalents sur C-TERb et C-TERa ait été analysé, il est possible d'étudier d'autres métaux divalents ainsi que des cations monovalents et trivalents, connus également pour leur implication dans des mécanismes neurodégénératifs.

Aussi, pour mieux comprendre les mécanismes d'auto-assemblage des isoformes, leur tendance à la séparation de phase, autrement dit à former des LLPS (pour *liquid-liquid phase separation*), pourra être investiguée. En plus d'impliquer les domaines de faible complexité (LCDs), les LLPS contribuent à l'agrégation protéique et au développement de maladies neurodégénératives.^{165,166} La génération de mutants, par mutagenèse dirigée, permettra d'identifier les régions promouvant les LLPS ainsi que les régions promotrices d'agrégation (APRs). Il est également possible d'étudier l'agrégation d'autres domaines spécifiques (2/3, C₂H₂, PHD) afin de distinguer les APRs. Typiquement, il peut être considéré de retirer les régions désordonnées afin de déterminer leur implication sur les processus d'agrégation et de fibrillation de DPF3.

Au-delà des techniques qui ont été employées dans le cadre de ce mémoire, la fluorescence extrinsèque de l'acide 8-anilino-1-naphtalènesulfonique (ANS) et la microscopie à force atomique couplée à la spectroscopie infrarouge (AFM-IR pour *atomic force microscopy-infrared spectroscopy*) apporterait des informations structurales complémentaires vis-à-vis de l'agrégation de DPF3. L'ANS assure le suivi des changements conformationnels par modification de son émission, tandis que l'AFM-IR est une méthode puissante pour l'obtention d'images de morphologie fibrillaire ainsi que pour la détermination de la composition en structure secondaire d'espèces agrégées. Complémentaire à la TEM, cela permettrait l'élaboration d'un mécanisme de fibrillation plus précis.

D'un point de vue physiopathologique et complémentaiement aux analyses *in vitro*, il serait intéressant, par des approches *in* et *ex vivo*, de déterminer si les isoformes de DPF3 ont également tendance à s'agréger et à se fibriller dans le cerveau des patients atteints des maladies de Parkinson et d'Alzheimer.

Ce mémoire s'étant inscrit dans le cadre de l'étude des propriétés désordre-agrégation et s'étant focalisé sur ce point, l'implication de DPF3 dans la mitose reste à étudier par l'élucidation des relations désordre-fonction et des interactions protéine-protéine (PPIs) qu'entretiennent les isoformes DPF3b et DPF3a ainsi que leurs domaines N- et C-terminaux avec différents partenaires. L'ensemble de ces études sont indispensables à l'appréhension du réseau d'interactions qu'occupe DPF3 au cours de la mitose, ce qui permettrait la conception et le développement de nouveaux antimitotiques et anticancéreux spécifiques à DPF3 et ses propriétés désordonnées. Etant donné l'implication de DPF3 dans les maladies neurodégénératives, la compréhension approfondie de ses mécanismes et voies d'agrégation et de fibrillation pourrait conduire à la conception d'anti-agrégants et d'agents anti-amyloïdes.

Partie n°5

—

Matériels et méthodes

5.1. Référence et composition des tampons de purification

- Tampon de lyse (LB pour *lysis buffer*) : tampon phosphate (PBS pour *phosphate buffered saline*) pH 7.3, 0.5 % TritonX-100, 200 mM KCl.
- Tampon de fixation (BB pour *binding buffer*) : PBS pH 7.3, 200 mM KCl.
- Tampon d'éluion associé à la protéase TEV (TPB pour *TEV protease buffer*) : 50 mM Tris-HCl pH 8.0, 150 mM NaCl.
- Tampon d'éluion de la GST (EB pour *elution buffer*) : 50 mM Tris-HCl pH 8.0, 10 mM GSH.

5.2. Transformation de souches compétentes de *E. coli* pour l'expression d'une protéine de fusion marquée à la GST

Les plasmides utilisés, à savoir DPF3a FL, DPF3a C-TER, DPF3b FL et DPF3b C-TER, ont été construits avec les vecteurs pGEX-3X de GE Healthcare et pET. La transformation desdits plasmides a été réalisée dans des souches de *E. coli* BL21 Rosetta 2 et stockée dans du glycérol 60 %. Le vecteur pET code pour une protéine de fusion recombinante marquée à la GST à l'extrémité N-terminale et inclue un site de clivage, inséré entre le *tag* et la protéine d'intérêt, spécifique à la protéase TEV. Ce facteur contient un gène de résistance à l'ampicilline. La protéase TEV, ou protéase du virus de la gravure du tabac, est une protéase à cystéine hautement spécifique. La séquence de reconnaissance de TEV ayant la plus grande efficacité est E-N-L-Y-F-Q↓-S. Cependant, l'acide aminé juste après le site de clivage peut également être G, A, M, C ou H. Evidemment, les protéines d'intérêt ne possèdent pas, dans leur séquence primaire, ces motifs de reconnaissance.

Pour les transformations, 100 µL de tampon Tris-EDTA (Tris-HCl 10 mM pH 8.0, EDTA 1 mM) sont transférés dans un Eppendorf de 1.5 mL contenant le papier avec le plasmide. Après passage au vortex, l'Eppendorf est incubé à température ambiante pendant 5 minutes et centrifugé à 21000 x g pendant 60 secondes à l'aide d'une microcentrifugeuse Eppendorf 5420 GLP. Sur glace, 3 µL du plasmide resolubilisé sont transférés dans 50 µL de la souche BL21 (DE3) dans un Eppendorf avant d'être incubés sur glace pendant 30 minutes et de subir un choc thermique dans un bain thermostaté à 42 °C pendant 45 secondes. Ce choc est suivi d'une nouvelle incubation sur glace de 3 minutes. Ensuite, 450 µL de milieu LB (*lysogeny broth*) Lennox stérilisé sont ajoutés et les bactéries sont incubées dans un thermomixeur Eppendorf Comfort AP-895 à 800 rpm et 37 °C pendant 60 minutes. Pour deux boîtes de Petri, 50 µL d'ampicilline 100 mg/mL sont additionnés à 50 mL de LB Lennox agar

(gélose) stérilisé. La solution est séparée entre les deux boîtes. Après incubation, 100 µL de transformation sont déposés sur une boîte de Petri et étalés sur toute la surface de la boîte à l'aide d'une dizaine de billes de verre stériles. Le restant de la transformation est centrifugé à 8000 rpm quelques secondes à l'aide d'une microcentrifugeuse Eppendorf 5420 GLP et 360 µL de surnageant sont éliminés de sorte à reconcentrer dix fois. Le culot est resuspendu dans le milieu et une nouvelle dispersion est réalisée dans la deuxième boîte de Petri. Une nouvelle fois, la transformation est étalée à l'aide d'une dizaine de billes de verre stériles. Les boîtes sont refermées et incubées pendant ~ 16-18 h à 37 °C avant de récupérer les colonies transformées le lendemain matin. Toutes les manipulations sont réalisées en condition stérile sous une hotte d'aspiration ou autour d'une flamme.

En ce qui concerne le stockage à long terme des bactéries transformées, seules les colonies correctement transformées, c'est-à-dire celles apparaissant sous la forme de taches blanches sur la gélose, sont utilisables. Il est nécessaire de lancer une préculture à partir d'une colonie bactérienne dont le détail est donné ci-après. Après une nuit d'incubation, les précultures sont récupérées et 500 µL sont transférées dans une dizaine d'Eppendorf stérile de 1.5 mL. Dans chacun des Eppendorf, 500 µL d'une solution de glycérol 60 % stérilisée sont ajoutés. L'ensemble des Eppendorf sont vortexés quelques secondes et stockés à -80 °C. Ces stocks bactériens peuvent être ultérieurement utilisés pour l'étape de surexpression.

5.3. Surexpression hétérologue dans *E. coli* d'une protéine de fusion marquée à la GST

La surexpression se déroule en quatre étapes : préculture, culture, induction à l'IPTG et récupération des culots bactériens. Les trois premières étapes sont réalisées sous conditions stériles. Pour la préculture, 100 mL de LB Lennox stérilisé sont transférés dans un erlenmeyer stérile de 500 mL avec chicanes. Dans le milieu, 125 µL d'ampicilline 100 mg/mL sont ajoutés ainsi qu'une fraction du stock glycérol de la construction d'intérêt. Celle-ci est ajoutée à l'aide d'une boucle d'inoculation. La préculture est incubée à 37 °C pendant ~ 16-18 h sous agitation. Pour la culture, quatre erlenmeyers stériles de 1000 mL avec chicanes sont préparés en ajoutant 250 mL de LB Lennox stérile, 125 µL d'ampicilline 100 mg/mL et 5.0 mL de préculture avec une pipette graduée stérile. Les erlenmeyers sont incubés à 37 °C sous agitation jusqu'à ce qu'une densité optique, mesurée à 600 nm à l'aide d'un spectrophotomètre Thermo Spectronic Genesys 8, comprise entre 0.5 et 0.6 soit obtenue. Les cultures ayant atteint une densité optique acceptable, 125 µL d'IPTG 1 M sont ajoutés dans chaque erlenmeyer avant induction à 37 °C sous agitation pendant 4 heures. Pour la récupération des culots bactériens induits, le contenu

de chaque erlenmeyer est transféré dans quatre biberons de 250 mL ou deux de 500 mL. Ces biberons sont placés à centrifuger pendant 30 minutes à 4000 rpm à 4 °C à l'aide d'une centrifugeuse Eppendorf 5804 R. Les surnageants sont éliminés avant de resuspendre les culots avec de l'eau déminéralisée dans deux Falcon de 50 mL. Après centrifugation de 15 minutes à 4000 rpm à 4 °C et élimination des surnageants, les culots sont conservés à -20 °C.

5.4. Lyse par sonication de souches *E. coli* induites à l'IPTG

Le culot bactérien est décongelé à température ambiante, décollé et suspendu dans environ 15 mL de tampon de lyse (LB pour *lysis buffer*) de composition suivante : tampon phosphate (PBS pour *phosphate-buffered saline*) pH 7.3, 0.5 % TritonX-100 et 200 mM KCl. La solution stock de PBS pH 7.3 contient : 1.76 mM KH₂PO₄, 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄. Le fluorure de phénylméthylsulfonyl (PMSF) 25 mM (dans l'isopropanol), un inhibiteur non spécifique de sérine protéase, est ajouté à hauteur de 120 µL. Il est nécessaire d'ajouter 120 µL après chaque heure car il est instable dans l'eau et perd son activité inhibitrice après un certain temps. Son temps de demi-vie dans l'eau est de 55 minutes à pH 7.5 et 25 °C.¹⁶⁷ Pour la sonication, six cycles sur glace de 30 secondes espacés par 30 secondes de pause, pour éviter la surchauffe, sont réalisés. Le sonicateur est calibré sur une puissance de 6. La suspension soniquée est centrifugée pendant 30 minutes à 12000 x g à 4 °C à l'aide d'une centrifugeuse VWR Mega Star 600. Le surnageant, aussi appelé le lysat, est récupéré et le culot contenant les débris cellulaires est jeté. Pour la deuxième fois, 120 µL de PMSF 25 mM sont ajoutés au lysat.

5.5. Purification et clivage par chromatographie d'affinité (FPLC) d'une protéine de fusion marquée à la GST sur colonnes pré-packées

Les purifications par FPLC sur colonnes pré-packées sont réalisées sur le système FPLC Äkta Purifier et avec une colonne de glutathion pré-packée GSTrap FF de 5 mL de chez GE Healthcare. Les chromatogrammes sont générés avec le programme Unicorn. Le système FPLC est couplé à une boucle d'injection de 5 mL et le tout est rincé à l'eau distillée (débit de 3.0 à 5.0 mL/min) jusqu'à stabilisation des signaux de conductivité et d'absorbance à 280 nm. La colonne GSTrap est connectée au système et sa solution de stockage (éthanol 20 %) est drainée avec de l'eau distillée (débit de 3.0 mL/min) jusqu'à nouvelle stabilisation des signaux. Le système et la colonne sont équilibrés avec le tampon de liaison (BB pour *binding buffer*),

composé de PBS pH 7.3, 200 mM KCl, à un débit de 1.0 mL/min jusqu'à observer un plateau de conductivité dû à la présence du KCl.

Pour la fixation de la protéine de fusion, le lysat centrifugé est chargé en trois fois 5 mL dans la boucle d'injection et injecté sur la colonne à un débit de 1 mL/min. Une fois les trois *flow-through* passés et rassemblés, la colonne est lavée par passage du tampon BB (débit de 3.0 mL/min) jusqu'à trouver des signaux similaires à ceux d'avant chargement du lysat.

L'étape de clivage commence par un lavage au tampon d'élution associé à la protéase TEV (TPB pour *TEV protease buffer*), composé de 50 mM Tris-HCl pH 8.0 et 150 mM NaCl, à un débit de 3.0 mL/min jusqu'à stabilisation des signaux d'absorbance à 280 nm et de conductivité. Il est également nécessaire d'injecter préalablement 5 mL de TPB dans la boucle pour la conditionner. La solution de clivage est préparée par transfert de 20 µL de la protéase TEV 1 mg/mL dans 4980 µL de TPB. Cette quantité de protéase permet de cliver environ 1 mg de protéine de fusion. La solution de clivage est chargée dans la boucle d'injection et 6 mL sont injectés sur colonne afin de s'assurer du passage de toute la solution de protéase. Ce volume correspond à 6 minutes d'injection à un débit de 1 mL/min. La colonne GSTrap est déconnectée et placée à incuber à 30 °C pendant 2 heures. Après l'incubation, la colonne GSTrap est reconnectée et la protéine clivée est éluee par fraction de 1.0 mL (débit de 1.0 mL/min). La colonne est lavée avec le tampon TPB (débit de 3.0 mL/min) jusqu'à stabilisation des signaux.

L'élution de la GST clivée de la GSTrap est réalisée à l'aide du tampon d'élution de la GST (EB pour *elution buffer*), composé de 50 mM Tris-HCl pH 8.0 et 10 mM GSH, à un débit de 1.0 mL/min. Le tampon EB doit être préparé juste avant utilisation étant donné que le GSH s'oxyde assez rapidement en GSSG une fois en solution. Son temps de demi-vie dans l'eau est d'environ 9 heures à pH 7.5 à 20 °C.¹⁶⁸ La colonne est lavée au EB (débit de 3.0 mL/min) jusqu'à stabilisation des signaux. Le système et la colonne sont rincés à l'eau distillée puis à l'éthanol 20 % (débit de 3.0 à 5.0 mL/min) jusqu'à stabilisation des signaux. Ces solutions sont également passées dans la fractionneuse. La colonne GSTrap est fermée et stockée à 4 °C.

5.6. Analyse par SDS-PAGE d'échantillons protéiques

Les gels de migration sont préparés dans des cassettes Novex Life Technologies de 1.0 mm. Pour ce faire, le *running gel* est d'abord préparé par l'addition d'eau déminéralisée, d'une solution d'acrylamide-bisacrylamide, de Tris-HCl 1.5 M à pH 8.8, de dodécylsulfate de

sodium (SDS pour *sodium dodecyl sulphate*) 10 %, de persulfate d'ammonium (APS pour *ammonium persulfate*) 10 % et de tétraméthyléthylènediamine (TEMED). Ce gel est coulé et laissé polymériser pendant 30 minutes. Les proportions des composants sont déterminées en fonction du degré de réticulation souhaité. Le *stacking gel* est ensuite préparé par addition d'eau déminéralisée, d'une solution d'acrylamide-bisacrylamide, de Tris-HCl 0.5 M à pH 6.8, de SDS 10 %, d'APS 10 % et de TEMED. Le gel est coulé et un peigne contenant le nombre de puits souhaités est ajouté avant 30 minutes de polymérisation. Les gels sont emballés dans du papier humidifié et stockés à 4 °C.

La migration sur gel est réalisée dans une cuve d'électrophorèse Novex Life Technologies XCell SureLock. La cassette du gel est placée dans la cuve avant de remplir cette dernière avec le tampon de migration de SDS. Les échantillons protéiques sont préparés en ajoutant, à 10 µL d'échantillon d'intérêt, 10 µL de tampon de charge contenant du SDS 10 % pour la dénaturation, du β-mercaptoéthanol pour la réduction des ponts disulfure et du bleu de bromophénol 0.5 % pour la visualisation de la migration. Après avoir retiré le peigne de la cassette, 10 µL de marqueur de masse moléculaire ainsi que 10 µL de mélange protéine-tampon de charge sont déposés dans les différents puits. La migration est lancée après avoir réglé le générateur sur 200 V.

Une fois la migration terminée, le gel est récupéré et rincé à l'eau distillée avant d'être coloré dans une solution de Bleu de Coomassie R-250 sur un agitateur orbital pendant 30 à 45 minutes. Le gel est placé dans une solution de décoloration, contenant du méthanol 30 % et de l'acide acétique glacial 10 % dans l'eau distillée pendant ~ 16-18 h. Il est également possible d'utiliser une solution de coloration de Quick Blue Coomassie, pendant environ 30 minutes, qui ne nécessite pas d'étape de décoloration. La méthode de révélation à l'imidazole-zinc consiste en l'immersion du gel dans une solution d'imidazole 0.2 M pendant dix minutes sur un agitateur suivi d'une seconde immersion dans une solution de chlorure de zinc 0.3 M pendant environ une minute. Le gel est photographié sur une plaque blanche en PMMA rétroéclairée.

5.7. Reconcentration d'un échantillon protéique à l'aide du PEG-20000

Les échantillons clivés sont reconcentrés à l'aide de la méthode de dialyse dont une membrane semi-perméable d'un *cut-off* de 6-8 kDa est plongée dans l'eau distillée et puis dans le tampon dans lequel se trouve la protéine d'intérêt pendant quelques minutes. La solution

protéique est placée dans le boudin formé par la membrane de dialyse après avoir fermé ce dernier par le dessous. Après avoir versé l'entièreté de la solution, une deuxième pince est placée à environ 2 cm au-dessus du niveau de solution dans le boudin. Ce dernier est placé sur une feuille d'aluminium sur laquelle du PEG-20000 (pour 20000 g/mol), qui va adsorber l'eau, a été déposé. Il est nécessaire de vérifier régulièrement le volume de solution à l'intérieur du boudin de dialyse. Typiquement, il faut environ 2 heures pour passer d'un volume de 6 mL à 1 mL.

5.8. Préparation des solutions protéiques contenant des cations divalents

Les solutions de travail de cations métalliques divalents (20 mM) sont préparées dans du TPB à partir de sulfate de cuivre (II) anhydre (CuSO_4) pour le Cu^{2+} , de sulfate de magnésium heptahydraté ($\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$) pour le Mg^{2+} , de sulfate de nickel (II) hexahydraté ($\text{NiSO}_4 \cdot 6 \text{H}_2\text{O}$) pour le Ni^{2+} et de sulfate de zinc heptahydraté ($\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$) pour le Zn^{2+} . Les échantillons de protéine sont incubés en présence de 100 μM de cations métalliques, en mélangeant 1 μL de solution de travail de cations métalliques avec 100 μL de matériel protéique et 99 μL de TPB (concentration finale en protéine $\sim 0.1 \text{ mg/mL}$ ou $\sim 5 \mu\text{M}$). Le rapport protéine:métal est d'environ 1:20.

5.9. Prédiction des propriétés structurales et analyses bioinformatiques

Des analyses basées sur la séquence sont réalisées à partir de la structure primaire connue de DPF3b (Uniprot ID : Q92784-1) et de DPF3a (Uniprot ID : Q92784-2). Le pourcentage moyen prédit de désordre intrinsèque par résidu (PPID) est déterminé en utilisant un ensemble de 18 prédicteurs de désordre disponibles gratuitement en ligne, à savoir les algorithmes VL-XT, XL1-XT, CaN-XT, VL3 et VSL2 de *Predictors of Natural Disordered Regions* (PONDR) ;¹⁶⁹⁻¹⁷¹ l'ensemble de données ESpritz entraînées par RMN, DRX et Disprot ;¹⁷² le système de prédiction de l'ordre des protéines (PrDOS pour *Protein DisOrder prediction System*) ;¹⁷³ les régions désordonnées longues et courtes à partir de l'outil *Prediction of Intrinsically Unstructured Proteins* (IUPred3) ;¹⁷⁴ l'outil *metapredict* (v2.1), un prédicteur basé sur l'apprentissage profond (*deep-learning*) ;¹⁷⁵ les algorithmes NORsnet, Ucon et MetaDisorder à partir du serveur web PredictProtein ;¹⁷⁶ les prédicteurs basé sur *Prediction of Order and Disorder by evaluation of NMR data* (ODiNPred).¹⁷⁷ Le désordre d'une protéine dépend de la nature des acides aminés qui la composent dont le score PPID est placé sur une

échelle allant de 0 (résidus totalement ordonnés) à 1 (résidus totalement désordonnés). Les résidus dont le PPID est supérieur à la valeur seuil de 0.5 participent au désordre.

A partir de l'outil PONDR, des graphiques de fonction de distribution cumulative (CDF) et de charge-hydrophobie (CH) sont générés.¹⁷⁸ Les paramètres de configuration associés aux séquences désordonnées (κ et Ω) et les groupes conformationnels prédits sont extraits des diagrammes de phase de Das-Pappu, générés avec le serveur *Classification of Intrinsically Disordered Ensemble Regions* (CIDER).¹⁷⁹ L'occurrence et la localisation des LCDs et des MoRFs dans les séquences de DPF3 sont déterminées à l'aide des outils de recherche SMART (*Simple modular architecture research tool*)¹⁸⁰ et *PLATform ot Tools Complexity* (PlaToLoCo) pour les LCDs¹⁸¹ ainsi que *MoRFchibi* pour les MoRFs.¹⁸² L'état d'enfouissement des résidus Trp et Tyr est prédit pour chaque isoforme à partir de leur structure primaire en utilisant le serveur de prédiction de l'accessibilité relative aux solvants des protéines (PaleAle 5.0).¹⁸³ Les résidus sont classés en quatre classes d'accessibilité au solvant : très enfouis (moins de 4 % exposés), quelque peu enfouis (entre 4 et 25 % exposés), quelque peu exposés (entre 25 et 50 % exposés) et très exposés (plus de 50 % exposés).

5.10. Détermination de la concentration protéique

Les concentrations protéiques sont déterminées en mesurant l'absorbance à 214 nm avec un spectrophotomètre UV-63000PC (VWR), en utilisant une cellule QS en quartz de 10 mm (Hellma) et en connaissant le coefficient d'extinction molaire de chaque isoforme à 214 nm ($\epsilon_{\text{DPF3b}} = 763132 \text{ M}^{-1} \cdot \text{cm}^{-1}$, $\epsilon_{\text{DPF3a}} = 622915 \text{ M}^{-1} \cdot \text{cm}^{-1}$, $\epsilon_{\text{C-TERb}} = 397435 \text{ M}^{-1} \cdot \text{cm}^{-1}$, $\epsilon_{\text{C-TERa}} = 257218 \text{ M}^{-1} \cdot \text{cm}^{-1}$) calculé à partir de la méthode décrite par B. H. J. Kuipers et H. Gruppen.¹²³ La concentration des échantillons protéiques après reconcentration est de ~ 0.1 ou 0.2 mg/mL (~ 5 ou 10 μM) dans chaque expérience. A partir des spectres d'absorption, l'indice de repliement (FI) est calculé à l'aide de l'équation suivante : $\text{FI} = (A^{280}/A^{275}) + (A^{280}/A^{275})$, où A correspond à la valeur d'absorbance à une longueur d'onde donnée, indiquée par l'exposant correspondant.

5.11. Spectrométrie de masse (LC-ESI-MS/MS)

Avant l'analyse par spectrométrie de masse en phase liquide, les échantillons sont digérés par la procédure FASP (*filter-aided sample preparation*) et stockés à -20 °C. L'analyse consiste en une injection de 15 μL d'échantillon digéré sur un spectromètre de masse Bruker

nano-LC-ESI-MS /MS maXis Impact UHR-TOF couplé à un Thermo 2D-LC Dionex Ultimate 3000. Les échantillons sont séparés par chromatographie liquide en phase inverse sur une colonne Thermo Acclaim PepMap 100 C18 72 μm x 250 mm. Les phases mobiles sont composées d'eau, d'acétonitrile et d'acide formique en différentes proportions afin d'appliquer un gradient d'éluion linéaire. Les éluats sont connectés à un ionisateur Bruker Captive Spray et les diagrammes, générés à l'aide du logiciel Bruker DataAnalysis 4.2, sont confrontés à la base de données UniProt avec le logiciel Mascot 2.4. Les protéines d'intérêt sont identifiées grâce au logiciel Scaffold 4.8. Les analyses sont réalisées par l'équipe de Marc Dieu de la Plateforme Technologique de Spectrométrie de Masse de l'Université de Namur (MaSUN).

5.12. Dichroïsme circulaire dans l'UV lointain (FCD)

Les spectres FCD (gamme 190-260 nm) sont enregistrés avec un spectropolarimètre MOS-500 à ~ 20 °C dans du TPB, en utilisant une cellule Suprasil en quartz de 1 mm de chemin optique (Hellma). Quatre balayages sont moyennés, les lignes de base du tampon sont soustraites et les spectres corrigés sont lissés. Les paramètres suivants sont utilisés : vitesse de balayage de 30 nm/min, largeur de bande de 2 nm, incrément de pas de 0.5 nm, temps d'intégration numérique de 1 s. Pour chaque spectre, les données sont présentées en ellipticité moyenne par résidu ($[\Theta]_{\text{EMR}}$), calculée comme suit : $[\Theta]_{\text{EMR}} = (M\theta)/(n-1)(10\gamma l)$, où M est la masse moléculaire de la protéine (Da), θ est l'ellipticité mesurée (mdeg), n le nombre de résidus dans la séquence de la protéine, γ la concentration massique de la protéine (mg/mL) et l est la longueur du chemin optique de la cellule (cm). A partir des spectres FCD, le contenu en structure secondaire est estimé en utilisant le serveur web de déconvolution en ligne *Beta Structure Selection* (BeStSel).¹⁸⁴

5.13. Spectroscopies de fluorescence intrinsèque (ITF, ITyrF et dbAF)

Les spectres d'émission ITF et ITyrF ainsi que les spectres dbAF sont enregistrés à partir de leur longueur d'onde d'excitation respective jusqu'à 600 nm avec un spectrophotomètre de fluorescence Agilent Cary Eclipse à ~ 20 °C dans du TPB, en utilisant une cellule GS en quartz de 10 mm de chemin optique (Hellma). L'incrément de pas est réglé sur 1.0 nm et les largeurs des fentes d'excitation et d'émission (sw pour *slit width*) sur 10 nm chacune. Une longueur d'onde d'excitation (λ_{ex}) de 295 nm est utilisée pour ITF, 275 nm pour ITyrF, et 400 nm pour dbAF. Les spectres d'excitation en dbAF sont enregistrés de 200 nm jusqu'à la longueur d'onde d'émission d'autofluorescence avec les paramètres susmentionnés.

5.14. Fluorescence extrinsèque de la thioflavine T (ThT)

La solution de travail de ThT (20 μM) est préparée dans du TPB et filtrée sur du polyéther sulfone 0.2 μm . Juste avant la mesure, 75 μL de la solution de travail de ThT sont ajoutés à 75 μL de solution protéique (concentration finale en ThT de 10 μM). Les spectres de fluorescence de la ThT sont enregistrés avec un spectrophotomètre de fluorescence Agilent Cary Eclipse à ~ 20 °C dans du TPB, en utilisant une cellule GS en quartz de 10 mm de chemin optique (Hellma), un incrément de pas de 1.0 nm et des sw d'excitation et d'émission de 10 nm.

5.15. Test de liaison au rouge Congo (RC)

Les spectres d'absorption du RC sont enregistrés sur la gamme de 400 à 600 nm par incrément de 1.0 nm avec un spectrophotomètre UV-63000PC (VWR) à ~ 20 °C dans le TPB, en utilisant une cellule QS en quartz de 10 mm de longueur de parcours (Hellma). Les échantillons sont préparés en mélangeant 75 μL de solution protéique avec 75 μL de RC 20 μM dans du TPB (concentration finale en RC de 10 μM).

5.16. Diffusion dynamique de la lumière (DLS)

Les mesures DLS sont effectuées à 20 °C à l'aide d'un analyseur de nanoparticules Horiba Zetasizer SZ-100 avec le détecteur à 90°. La fonction d'autocorrélation est ajustée avec succès 20 fois par analyse et les résultats sont exprimés sous la forme du diamètre hydrodynamique moyen \bar{D}_h . Chaque échantillon de protéine a été passé à travers un filtre de polyéther sulfone de 0.2 μm avant analyse.

5.17. Microscopie électronique à transmission (TEM)

Les agrégats et fibrilles de protéine sont visualisés par coloration négative (*negative staining*) par microscopie électronique à transmission (TEM) sur un microscope PHILIPS/FEI Tecnai 10 fonctionnant à une tension de 100 kV. Des gouttes de 5 μL de matériel protéique sont déposées pendant 3 minutes sur des grilles de cuivre recouvertes d'un film de formvar/carbone et préalablement hydrophilisées par décharge lumineuse. Tout excès est absorbé avec un morceau de papier buvard et la grille est placée sur une gouttelette de 5 μL d'acétate d'uranyle à 0.5% pendant 1 min avant d'être séchée à l'air pendant 5 min.

Partie n°6

—

Bibliographie

1. Huber, F. M. *et al.* Histone-binding of DPF2 mediates its repressive role in myeloid differentiation. *Proc Natl Acad Sci U S A* **114**, 6016–6021 (2017).
2. Chestkov, A. *et al.* *The d4 Gene Family in the Human Genome*. (1996).
3. Ninkina, N. N. *et al.* Cerd4, third member of the d4 gene family: Expression and organization of genomic locus. *Mammalian Genome* **12**, 862–866 (2001).
4. Buchman, V. L. *et al.* Differential splicing creates a diversity of transcripts from a neurospecific developmental regulated gene encoding a protein with new zinc-finger motifs. *Nucleic Acids Research* **20**, (1992).
5. Cai, C. *et al.* The RING finger protein family in health and disease. *Signal Transduction and Targeted Therapy* **7**, (2022).
6. Sanchez, R. *et al.* The PHD finger: A versatile epigenome reader. *Trends in Biochemical Sciences* **36**, (2011).
7. Viola, I. L. *et al.* Structure and Evolution of Plant Homeobox Genes. in *Plant Transcription Factors: Evolutionary, Structural and Functional Aspects* 101–112 (Elsevier Inc., 2016).
8. Kulikova, D. A. *et al.* d4 family genes: Genomic organization and expression. *Russian Journal of Developmental Biology* **44**, (2013).
9. Mertsalov, I. B. *et al.* Generation of mutant mice with targeted disruption of two members of the d4 gene family: Neuro-d4 and cer-d4. *Dokl Biochem Biophys* **419**, 65–68 (2008).
10. Lange, M. *et al.* Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev* **22**, 2370–2384 (2008).
11. Fedotova, A. A. *et al.* C2H2 Zinc Finger Proteins: The Largest but Poorly Explored Family of Higher Eukaryotic Transcription Factors. *Acta Naturae* **9**, 47 (2017).
12. Li, W. *et al.* Crystal structure of DPF3b in complex with an acetylated histone peptide. *J Struct Biol* **195**, 365–372 (2016).
13. Panamarova, M. *et al.* The BAF chromatin remodelling complex is an epigenetic regulator of lineage specification in the early mouse embryo. *Development (Cambridge)* **143**, 1271–1283 (2016).
14. Hohmann, A. F. *et al.* A rationale to target the SWI/SNF complex for cancer therapy. *Trends in Genetics* **30**, 356–363 (2014).
15. Alfert, A. *et al.* The BAF complex in development and disease. *Epigenetics and Chromatin* **12**, (2019).
16. Son, E. Y. *et al.* The role of BAF (mSWI/SNF) complexes in mammalian neural development. *Am J Med Genet C Semin Med Genet* **166**, 333–349 (2014).
17. Hadidy, N. *et al.* Intrinsic Disorder of the BAF Complex: Roles in Chromatin Remodeling and Disease Development. *Int J Mol Sci* **20**, (2019).
18. Chen, R. *et al.* Release and activity of histone in diseases. *Cell Death and Disease* **5**, (2014).

19. Nelissen, H. *et al.* Impact of core histone modifications on transcriptional regulation and plant growth. *Critical Reviews in Plant Sciences* **26**, 243–263 (2007).
20. Jenuwein, T. *et al.* *Proc. Natl. Acad. Sci. U.S.A* **11**, (2001).
21. Luger, K. *et al.* New insights into nucleosome and chromatin structure: An ordered state or a disordered affair? *Nature Reviews Molecular Cell Biology* **13**, 436–447 (2012).
22. Zeng, L. *et al.* Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* **466**, 258–262 (2010).
23. Richard Mcintosh, J. Mitosis. *Cold Spring Harb Perspect Biol* **8**, (2016).
24. Hoyal, C. R. *et al.* Genetic polymorphisms in DPF3 associated with risk of breast cancer and lymph node metastases. *J Carcinog* **4**, (2005).
25. Lin, W. hao *et al.* Downregulation of DPF3 promotes the proliferation and motility of breast cancer cells through activating JAK2/STAT3 signaling. *Biochem Biophys Res Commun* **514**, 639–644 (2019).
26. Theodorou, M. *et al.* Identification of a STAT5 Target Gene, Dpf3, Provides Novel Insights in Chronic Lymphocytic Leukemia. *PLoS One* **8**, (2013).
27. Hiramatsu, H. *et al.* The role of the SWI/SNF chromatin remodeling complex in maintaining the stemness of glioma initiating cells. *Sci Rep* **7**, (2017).
28. Banaganapalli, B. *et al.* Integrative weighted molecular network construction from transcriptomics and genome wide association data to identify shared genetic biomarkers for COPD and lung cancer. *PLoS One* **17**, (2022).
29. Protze, J. *et al.* The renal cancer risk allele at 14q24.2 activates a novel hypoxia-inducible transcription factor-binding enhancer of DPF3 expression. *Journal of Biological Chemistry* **298**, (2022).
30. Colli, L. M. *et al.* Altered regulation of DPF3, a member of the SWI/SNF complexes, underlies the 14q24 renal cancer susceptibility locus. *Am J Hum Genet* **108**, 1590–1610 (2021).
31. Cui, H. *et al.* The SWI/SNF chromatin remodeling factor DPF3 regulates metastasis of ccRCC by modulating TGF- β signaling. *Nat Commun* **13**, (2022).
32. Zhu, X. *et al.* HRP2-DPF3a-BAF complex coordinates histone modification and chromatin remodeling to regulate myogenic gene transcription. *Nucleic Acids Res* **48**, 6563–6582 (2020).
33. Cui, H. *et al.* Phosphorylation of the chromatin remodeling factor DPF3a induces cardiac hypertrophy through releasing HEY repressors from DNA. *Nucleic Acids Res* **44**, 2538–2553 (2015).
34. Wang, G. *et al.* Epigenetics in Congenital Heart Disease. *J Am Heart Assoc* **11**, (2022).
35. Bailliard, F. *et al.* Tetralogy of Fallot. *Orphanet J Rare Dis* **4**, (2009).
36. Sato, Y. *et al.* Association of TUSC1 and DPF3 gene polymorphisms with male infertility. *J Assist Reprod Genet* **35**, 257–263 (2018).

37. Liu, S. Y. *et al.* Strong association of SLC1A1 and DPF3 gene variants with idiopathic male infertility in Han Chinese. *Asian J Androl* **18**, (2016).
38. Vieira, A. R. *et al.* Candidate gene/loci studies in cleft lip/palate and dental anomalies finds novel susceptibility genes for clefts. *Genetics in Medicine* **10**, 668–674 (2008).
39. Liu, H. *et al.* Expression profiles of HA117 and its neighboring gene DPF3 in different colon segments of Hirschsprung's disease. *Int J Clin Exp Pathol* **7**, (2014).
40. Rudnitskaya, E. A. *et al.* Changes in Glial Support of the Hippocampus during the Development of an Alzheimer's Disease-like Pathology and Their Correction by Mitochondria-Targeted Antioxidant SkQ1. *Int J Mol Sci* **23**, (2022).
41. Satoh, J. *Molecular network analysis of human microRNA targetome: from cancers to Alzheimer's disease Jun-ichi Satoh.* (2012).
42. Huang, M. *et al.* Cell–Cell Communication Alterations via Intercellular Signaling Pathways in Substantia Nigra of Parkinson's Disease. *Front Aging Neurosci* **14**, (2022).
43. Mignon, J. *et al.* Revealing Intrinsic Disorder and Aggregation Properties of the DPF3a Zinc Finger Protein. *ACS Omega* **6**, 18793–18801 (2021).
44. Hong, S. *et al.* Mechanisms of macromolecular interactions mediated by protein intrinsic disorder. *Molecules and Cells* **43**, 899–908 (2020).
45. Morris, O. M. *et al.* Intrinsically disordered proteins: Modes of binding with emphasis on disordered domains. *Open Biology* **11**, (2021).
46. Cumberworth, A. *et al.* Promiscuity as a functional trait: Intrinsically disordered regions as central players of interactomes. *Biochemical Journal* **454**, 361–369 (2013).
47. Burger, V. M. *et al.* Intrinsically disordered proteins: Where computation meets experiment. *Polymers* **6**, 2684–2719 (2014).
48. Kumar, N. *et al.* Comprehensive Intrinsic Disorder Analysis of 6108 Viral Proteomes: From the Extent of Intrinsic Disorder Penetrance to Functional Annotation of Disordered Viral Proteins. *J Proteome Res* **20**, 2704–2713 (2021).
49. Wright, P. E. *et al.* Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology* **16**, 18–29 (2015).
50. Chong, S. *et al.* Towards Decoding the Sequence-Based Grammar Governing the Functions of Intrinsically Disordered Protein Regions. *Journal of Molecular Biology* **433**, (2021).
51. Basile, W. *et al.* Why do eukaryotic proteins contain more intrinsically disordered regions? *PLoS Comput Biol* **15**, (2019).
52. Uversky, V. N. Intrinsically disordered proteins and their 'Mysterious' (meta)physics. *Frontiers in Physics* **7**, (2019).
53. Deiana, A. *et al.* Intrinsically disordered proteins and structured proteins with intrinsically disordered regions have different functional roles in the cell. *PLoS One* **14**, (2019).

54. Das, R. K. *et al.* Relating sequence encoded information to form and function of intrinsically disordered proteins. *Current Opinion in Structural Biology* **32**, (2015).
55. Holehouse, A. S. *et al.* CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys J* **112**, 16–21 (2017).
56. Martin, E. W. *et al.* Sequence Determinants of the Conformational Properties of an Intrinsically Disordered Protein Prior to and upon Multisite Phosphorylation. *J Am Chem Soc* **138**, 15323–15335 (2016).
57. Singh, G. P. *et al.* Role of intrinsic disorder in transient interactions of hub proteins. *Proteins: Structure, Function and Genetics* **66**, 761–765 (2007).
58. Haynes, C. *et al.* Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput Biol* **2**, 0890–0901 (2006).
59. Wright, P. E. *et al.* Linking folding and binding. *Current Opinion in Structural Biology* **19**, 31–38 (2009).
60. Patil, A. *et al.* Domain distribution and intrinsic disorder in hubs in the human protein-protein interaction network. *Protein Science* **19**, 1461–1468 (2010).
61. Huntley, M. A. *et al.* Simple sequences are rare in the Protein Data Bank. *Proteins: Structure, Function and Genetics* **48**, 134–140 (2002).
62. Dyson, H. J. *et al.* Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **6**, 197–208 (2005).
63. Watson, M. *et al.* Disordered domains in chromatin-binding proteins. *Essays in Biochemistry* **63**, 147–156 (2019).
64. Iakoucheva, L. M. *et al.* Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol* **323**, 573–584 (2002).
65. Cheng, Y. *et al.* Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry* **45**, 10448–10460 (2006).
66. Fuertes, G. *et al.* Perspectives on drug discovery strategies based on IDPs. in *Intrinsically Disordered Proteins* 275–327 (2019).
67. Mignon, J. *et al.* Propriétés associées au désordre et mécanisme d'agrégation d'une nouvelle IDP amyloïdogène: la protéine à doigts de zinc DPF3a Propriétés associées au désordre et mécanisme d'agrégation d'une nouvelle IDP amyloïdogène: la protéine à doigts de zinc DPF3a. *Chimie Nouvelle* **138**, (2021)
68. Wang, S. *et al.* Targeting the MDM2-p53 protein-protein interaction for new cancer therapy: Progress and challenges. *Cold Spring Harbor Perspectives in Medicine* **7**, (2017).
69. Kruiswijk, F. *et al.* P53 in survival, death, and metabolic health: A lifeguard with a licence to kill. *Nature Reviews Molecular Cell Biology* **16**, 393–405 (2015).
70. Yin, X. *et al.* Low molecular weight inhibitors of Myc-Max interaction and function. *Oncogene* **22**, 6151–6159 (2003).

71. Carabet, L. A. *et al.* Computer-aided drug discovery of Myc-Max inhibitors as potential therapeutics for prostate cancer. *Eur J Med Chem* **160**, 108–119 (2018).
72. Ayyadevara, S. *et al.* Intrinsically disordered proteins identified in the aggregate proteome serve as biomarkers of neurodegeneration. *Metabolic Brain Disease* **37**, 147–152 (2022).
73. Almeida, Z. L. *et al.* Structure and aggregation mechanisms in amyloids. *Molecules* **25**, (2020).
74. Chatani, E. *et al.* Current Understanding of the Structure, Stability and Dynamic Properties of Amyloid Fibrils. *International Journal of Molecular Sciences* **22**, 4349 (2021).
75. Ke, P. C. *et al.* Half a century of amyloids: Past, present and future. *Chem Soc Rev* **49**, 5473–5509 (2020).
76. Eisenberg, D. *et al.* The amyloid state of proteins in human diseases. *Cell* **148**, 1188–1203 (2012).
77. Verma, M. *et al.* Toxic species in amyloid disorders: Oligomers or mature fibrils. *Annals of Indian Academy of Neurology* **18**, 138–145 (2015).
78. Yoshimura, Y. *et al.* Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation.
79. Kosolapova, A. O. *et al.* Biological functions of prokaryotic amyloids in interspecies interactions: Facts and assumptions. *International Journal of Molecular Sciences* **21**, 1–26 (2020).
80. Gebbink, M. F. B. G. *et al.* Amyloids - A functional coat for microorganisms. *Nature Reviews Microbiology* **3**, 333–341 (2005).
81. Rideout, H. J. *et al.* Regulation of α -synuclein by bFGF in cultured ventral midbrain dopaminergic neurons. *J. Neurochem* **84**, 803–813 (2003).
82. Stefanis, L. α -Synuclein in Parkinson's disease. *Cold Spring Harb Perspect Med* **2**, (2012).
83. Singleton, A. B. *et al.* Synuclein Locus Triplication Causes Parkinson's Disease, (2003).
84. Kalia, L. v. *et al.* Parkinson's disease. *The Lancet* **386**, 896–912 (2015).
85. Wong, Y. C. *et al.* α -synuclein toxicity in neurodegeneration: Mechanism and therapeutic strategies. *Nature Medicine* **23**, 1–13 (2017).
86. Barghorn, S. *et al.* Tau Paired Helical Filaments from Alzheimer's Disease Brain and Assembled in Vitro Are Based on β -Structure in the Core Domain. *Biochemistry* **43**, 1694–1703 (2004).
87. von Bergen, M. *et al.* Tau aggregation is driven by a transition from random coil to beta sheet structure. *Biochimica et Biophysica Acta - Molecular Basis of Disease* **1739**, 158–166 (2005).
88. Tepper, K. *et al.* Oligomer formation of tau protein hyperphosphorylated in cells. *Journal of Biological Chemistry* **289**, 34389–34407 (2014).
89. Olsson, F. *et al.* Characterization of intermediate steps in amyloid beta (A β) production under near-native conditions. *Journal of Biological Chemistry* **289**, 1540–1550 (2014).

90. Takami, M. *et al.* γ -Secretase: Successive tripeptide and tetrapeptide release from the transmembrane domain of β -carboxyl terminal fragment. *Journal of Neuroscience* **29**, 13042–13052 (2009).
91. Joshi, G. *et al.* Golgi defects enhance APP amyloidogenic processing in Alzheimer's disease. *BioEssays* **37**, 240–247 (2015).
92. Chen, G. F. *et al.* Amyloid beta: Structure, biology and structure-based therapeutic development. *Acta Pharmacologica Sinica* **38**, 1205–1235 (2017).
93. Yu, L. *et al.* Structural characterization of a soluble amyloid β -peptide oligomer. *Biochemistry* **48**, 1870–1877 (2009).
94. Ehrnhoefer, D. E. *et al.* EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat Struct Mol Biol* **15**, 558–566 (2008).
95. Lamberto, G. R. *et al.* Structural and mechanistic basis behind the inhibitory interaction of PcTS on-synuclein amyloid fibril formation, (2009).
96. Akoury, E. *et al.* Inhibition of tau filament formation by conformational modulation. *J Am Chem Soc* **135**, 2853–2862 (2013).
97. Mignon, J. *et al.* Short Communication The intrinsically disordered DPF3 zinc finger protein: a promising new target in cancer therapy. *J Cancer Biol* **3**, (2022).
98. Palanikumar, L. *et al.* Protein mimetic amyloid inhibitor potently abrogates cancer-associated mutant p53 aggregation and restores tumor suppressor function. *Nat Commun* **12**, (2021).
99. Li, J. *et al.* p53 amyloid aggregation in cancer: function, mechanism, and therapy. *Exp Hematol Oncol* **11**, (2022).
100. Silva, J. L. *et al.* Prion-like aggregation of mutant p53 in cancer. *Trends in Biochemical Sciences* **39**, 260–267 (2014).
101. Silva, J. L. *et al.* Expanding the prion concept to cancer biology: Dominant-negative effect of aggregates of mutant p53 tumour suppressor. *Biosci Rep* **33**, 593–603 (2013).
102. Kirkland, A. E. *et al.* The role of magnesium in neurological disorders. *Nutrients* **10**, (2018).
103. Mezzaroba, L. *et al.* The role of zinc, copper, manganese and iron in neurodegenerative diseases. *NeuroToxicology* **74**, 230–241 (2019).
104. Anyachor, C. P. *et al.* Mechanistic considerations and biomarkers level in nickel-induced neurodegenerative diseases: An updated systematic review. *IBRO Neuroscience Reports* **13**, 136–146 (2022).
105. Breydo, L. *et al.* Role of metal ions in aggregation of intrinsically disordered proteins in neurodegenerative diseases. *Metallomics* **3**, 1163–1180 (2011).
106. Leal, S. S. *et al.* Metal ions as modulators of protein conformation and misfolding in neurodegeneration. *Coordination Chemistry Reviews* **256**, 2253–2270 (2012).
107. Faller, P. *et al.* Metal ions and intrinsically disordered proteins and peptides: From Cu/Zn amyloid- β to general principles. *Acc Chem Res* **47**, 2252–2259 (2014).

108. Poulson, B. G. *et al.* Aggregation of biologically important peptides and proteins: Inhibition or acceleration depending on protein and metal ion concentrations. *RSC Advances* **10**, 215–227 (2019).
109. Tōgu, V. *et al.* Interactions of Zn(ii) and Cu(ii) ions with Alzheimer’s amyloid-beta peptide. Metal ion binding, contribution to fibrillization and toxicity. *Metallomics* **3**, 250–261 (2011).
110. Arena, G. *et al.* Zn²⁺ Interaction with Amyloid-B: Affinity and Speciation. *Molecules (Basel, Switzerland)* **24**, (2019).
111. Rana, M. *et al.* Cu and Zn interactions with A β peptides: Consequence of coordination on aggregation and formation of neurotoxic soluble A β oligomers. *Metallomics* **11**, 64–84 (2019).
112. Gorantla, N. V. *et al.* Transition metal nickel prevents Tau aggregation in Alzheimer’s disease. *Int J Biol Macromol* **156**, 1359–1365 (2020).
113. Ahmadi, S. *et al.* Aggregation of Microtubule Binding Repeats of Tau Protein is Promoted by Cu²⁺. *ACS Omega* **4**, 5356–5366 (2019).
114. Cristóvão, J. S. *et al.* Metals and Neuronal Metal Binding Proteins Implicated in Alzheimer’s Disease. *Oxidative Medicine and Cellular Longevity* **2016**, (2016).
115. Kumar, G. A simple method for detecting phosphorylation of proteins by using Zn²⁺-Phos-Tag SDS-PAGE at Neutral pH. in *Methods in Molecular Biology* **1853**, 223–229 (2018).
116. Uversky, V. N. *et al.* Metal-triggered structural transformations, aggregation, and fibrillation of human α -synuclein: A possible molecular link between parkinson’s disease and heavy metal exposure. *Journal of Biological Chemistry* **276**, 44284–44296 (2001).
117. Wright, J. A. *et al.* Unique copper-induced oligomers mediate alpha-synuclein toxicity. *The FASEB Journal* **23**, 2384–2393 (2009).
118. Golts, N. *et al.* Magnesium inhibits spontaneous and iron-induced aggregation of α -synuclein. *Journal of Biological Chemistry* **277**, 16116–16123 (2002).
119. Valiente-Gabioud, A. A. *et al.* Structural basis behind the interaction of Zn²⁺ with the protein α -synuclein and the A β peptide: A comparative analysis. in *Journal of Inorganic Biochemistry* **117**, 334–341 (2012).
120. Al-Harhi, S. *et al.* Zinc ions prevent α -synuclein aggregation by enhancing chaperone function of human serum albumin. *Int J Biol Macromol* (2022).
121. Biancalana, M. *et al.* Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1804**, 1405–1412 (2010).
122. Efficient, rapid protein purification and on-column cleavage using GStrap FF columns.
123. Kuipers, B. J. H. *et al.* Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. *J Agric Food Chem* **55**, 5445–5451 (2007).
124. Mignon, J. *et al.* Structural characterisation of amyloidogenic intrinsically disordered zinc finger protein isoforms DPF3b and DPF3a. *Int J Biol Macromol* **218**, 57–71 (2022).

125. Leyder, T. *et al.* Unveiling the Metal-Dependent Aggregation Properties of the C-terminal Region of Amyloidogenic Intrinsically Disordered Protein Isoforms DPF3b and DPF3a. *Int J Mol Sci* **23**, 15291 (2022).
126. Xue, B. *et al.* CDF it all: Consensus prediction of intrinsically disordered proteins based on various cumulative distribution functions. *FEBS Lett* **583**, 1469–1474 (2009).
127. Uversky, V. N. *et al.* Why Are ‘Natively Unfolded’ Proteins Unstructured Under Physiologic Conditions? *Proteins* **41**, (2000).
128. Kyte, J. *et al.* A Simple Method for Displaying the Hydrophobic Character of a Protein. *J. Mol. Biol* **157**, (1982).
129. Tcherkasskaya, O. *et al.* Denatured Collapsed States in Protein Folding: Example of Apomyoglobin. (2001).
130. Nygaard, M. *et al.* An Efficient Method for Estimating the Hydrodynamic Radius of Disordered Protein Conformations. *Biophys J* **113**, 550–557 (2017).
131. Tomasso, M. E. *et al.* Hydrodynamic Radii of Intrinsically Disordered Proteins Determined from Experimental Polyproline II Propensities. *PLoS Comput Biol* **12**, (2016).
132. Ghisaidoobe, A. B. T. *et al.* Intrinsic tryptophan fluorescence in the detection and analysis of proteins: A focus on Förster resonance energy transfer techniques. *International Journal of Molecular Sciences* **15**, 22518–22538 (2014).
133. Biter, A. B. *et al.* A method to probe protein structure from UV absorbance spectra. *Anal Biochem* **587**, (2019).
134. Manning, M. C. *et al.* Circular dichroism studies of distorted α -helices, twisted β -sheets, and 3_1 -turns. *Biophysical Chemistry* **31**, (1988).
135. Ghosh, D. *et al.* Structure based aggregation studies reveal the presence of helix-rich intermediate during α -Synuclein aggregation. *Sci Rep* **5**, (2015).
136. Noronha, M. *et al.* Intramolecular Fluorescence Quenching of Tyrosine by the Peptide α -Carbonyl Group Revisited. *Journal of Physical Chemistry A* **108**, 2155–2166 (2004).
137. Wiczorek, E. *et al.* Deep blue autofluorescence reveals the instability of human transthyretin. *Int J Biol Macromol* **191**, 492–499 (2021).
138. Ziaunys, M. *et al.* Exploring the potential of deep-blue autofluorescence for monitoring amyloid fibril formation and dissociation. *PeerJ* **2019**, (2019).
139. Chan, F. T. S. *et al.* Protein amyloids develop an intrinsic fluorescence signature during aggregation. *Analyst* **138**, 2156–2162 (2013).
140. Niyangoda, C. *et al.* Carbonyl-based blue autofluorescence of proteins and amino acids. *PLoS One* **12**, (2017).
141. Tikhonova, T. N. *et al.* Dissection of the deep-blue autofluorescence changes accompanying amyloid fibrillation. *Arch Biochem Biophys* **651**, 13–20 (2018).

142. Saraiva, M. A. Interpretation of α -synuclein UV absorption spectra in the peptide bond and the aromatic regions. *J Photochem Photobiol B* **212**, (2020).
143. Xue, C. *et al.* Thioflavin T as an amyloid dye: Fibril quantification, optimal concentration and effect on aggregation. *R Soc Open Sci* **4**, (2017).
144. Yakupova, E. I. *et al.* Congo Red and amyloids: History and relationship. *Bioscience Reports* **39**, (2019).
145. Robustelli, P. *et al.* Molecular Basis of Small-Molecule Binding to α -Synuclein. *J Am Chem Soc* **144**, 2501–2510 (2022).
146. Alali, S. *et al.* Cannabidiol inhibits tau aggregation in vitro. *Cells* **10**, (2021).
147. Xu, S. *et al.* Characterization of tau fibrillization in vitro. *Alzheimer's and Dementia* **6**, 110–117 (2010).
148. Close, W. *et al.* Physical basis of amyloid fibril polymorphism. *Nat Commun* **9**, (2018).
149. Willbold, D. *et al.* Amyloid-type Protein Aggregation and Prion-like Properties of Amyloids. *Chemical Reviews* **121**, 8285–8307 (2021).
150. Wegmann, S. *et al.* Human tau isoforms assemble into ribbon-like fibrils that display polymorphic structure and stability. *Journal of Biological Chemistry* **285**, 27302–27313 (2010).
151. Kurouski, D. Supramolecular Organization of Amyloid Fibrils. in *Exploring New Findings on Amyloidosis* (2016).
152. Schramm, A. *et al.* An arsenal of methods for the experimental characterization of intrinsically disordered proteins – How to choose and combine them? *Archives of Biochemistry and Biophysics* **676**, (2019).
153. Gong, Z. *et al.* Secondary structure of cell-penetrating peptides during interaction with fungal cells. *Protein Science* **27**, 702–713 (2018).
154. Szabo, A. G. *et al.* Tyrosinate fluorescence maxima at 345 nm in proteins lacking tryptophan at pH 7. *FEBS Lett* **94**, 249–252 (1978).
155. Mulvihill, S. *et al.* Characteristics of Tyrosinate Fluorescence Emission in cy-and P-Purothioninst. *Proc. Natl. Acad. Sci. U.S.A* **23**, (1984).
156. Camino, J. D. *et al.* The extent of protein hydration dictates the preference for heterogeneous or homogeneous nucleation generating either parallel or antiparallel β -sheet α -synuclein aggregates. *Chem Sci* **11**, 11902–11914 (2020).
157. Gracia, P. *et al.* Multiplicity of α -synuclein aggregated species and their possible roles in disease. *International Journal of Molecular Sciences* **21**, 1–27 (2020).
158. Wang, G. *et al.* Tryptophan fluorescence quenching by alkaline earth metal cations in deionized bacteriorhodopsin. *Journal of Photochemistry and Photobiology B: Biology* **59**, (2000).
159. Chen, R. F. Fluorescence Quenching As A Parameter For Measuring Complex Formation Between Metal Ions And Aromatic Amino Acids And Peptides. *Anal Lett* **19**, 963–977 (1986).

160. Meisl, G. *et al.* Molecular mechanisms of protein aggregation from global fitting of kinetic models. *Nat Protoc* **11**, 252–272 (2016).
161. Barber-Zucker, S. *et al.* Transition metal binding selectivity in proteins and its correlation with the phylogenomic classification of the cation diffusion facilitator protein family. *Sci Rep* **7**, (2017).
162. Cao, X. *et al.* Identification of metal ion binding sites based on amino acid sequences. *PLoS One* **12**, (2017).
163. Gorantla, N. V. *et al.* Transition metal nickel prevents Tau aggregation in Alzheimer’s disease. *Int J Biol Macromol* **156**, 1359–1365 (2020).
164. Ozawa, D. *et al.* Destruction of amyloid fibrils of keratoepithelin peptides by laser irradiation coupled with amyloid-specific thioflavin T. *Journal of Biological Chemistry* **286**, 10856–10863 (2011).
165. Brocca, S. *et al.* Liquid–liquid phase separation by intrinsically disordered protein regions of viruses: Roles in viral life cycle and control of virus–host interactions. *International Journal of Molecular Sciences* **21**, 1–31 (2020).
166. Ahmad, A. *et al.* Aberrant liquid-liquid phase separation and amyloid aggregation of proteins related to neurodegenerative diseases. *International Journal of Biological Macromolecules* **220**, 703–720 (2022).
167. James, G. T. Inactivation of the Protease inhibitor Phenylmethylsulfonyl Fluoride in Buffers. *ANALYTICAL BIOCHEMISTRY* **86**, (1978).
168. Stevens, R. *et al.* The stabilities of various thiol compounds used in protein purifications. *Biochem Educ* **11**, 70 (1983).
169. Dunker, A. K. *et al.* The unfoldomics decade: An update on intrinsically disordered proteins. in *BMC Genomics* **9**, (2008).
170. Peng, K. *et al.* OPTIMIZING LONG INTRINSIC DISORDER PREDICTORS WITH PROTEIN EVOLUTIONARY INFORMATION. *Journal of Bioinformatics and Computational Biology* **3**, (2005).
171. Williams, R. M. *et al.* THE PROTEIN NON-FOLDING PROBLEM: AMINO ACID DETERMINANTS OF INTRINSIC ORDER AND DISORDER. (2001).
172. Walsh, I. *et al.* Espritz: Accurate and fast prediction of protein disorder. *Bioinformatics* **28**, 503–509 (2012).
173. Ishida, T. *et al.* PrDOS: Prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res* **35**, (2007).
174. Erdos, G. *et al.* IUPred3: Prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. *Nucleic Acids Res* **49**, W297–W303 (2021).
175. Emenecker, R. J. *et al.* Metapredict: a fast, accurate, and easy-to-use predictor of consensus disorder and structure. *Biophys J* **120**, 4312–4319 (2021).

176. Bernhofer, M. *et al.* PredictProtein - Predicting protein structure and function for 29 years. *Nucleic Acids Res* **49**, W535–W540 (2021).
177. Dass, R. *et al.* ODinPred: comprehensive prediction of protein order and disorder. *Sci Rep* **10**, (2020).
178. Xue, B. *et al.* PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochim Biophys Acta Proteins Proteom* **1804**, 996–1010 (2010).
179. Holehouse, A. S. *et al.* CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys J* **112**, 16–21 (2017).
180. Letunic, I. *et al.* SMART: Recent updates, new developments and status in 2020. *Nucleic Acids Res* **49**, D458–D460 (2021).
181. Jarnot, P. *et al.* PlaToLoCo: The first web meta-server for visualization and annotation of low complexity regions in proteins. *Nucleic Acids Res* **48**, W77–W84 (2020).
182. Malhis, N. *et al.* MoRFchibi SYSTEM: software tools for the identification of MoRFs in protein sequences. *Nucleic Acids Res* **44**, W488–W493 (2016).
183. Kaleel, M. *et al.* PaleAle 5.0: prediction of protein relative solvent accessibility by deep learning. *Amino Acids* **51**, 1289–1296 (2019).
184. Micsonai, A. *et al.* BeStSel: A web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**, W315–W322 (2018).