Université de Montréal

Études des premières étapes d'oligomérisation de la région Non-A β Component de l' α -synucléine et de A β_{1-40}

par Eugene Cindie

Département de Chimie Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures et postdoctorales en vue de l'obtention du grade de Maître ès sciences (M.Sc.) en chimie

Avril, 2014

© Eugene Cindie, 2014.

RÉSUMÉ

Les protéines amyloïdes sont retrouvées sous forme de fibres dans de nombreuses maladies neurodégénératives. En tentant d'élucider le mécanisme de fibrillation, les chercheurs ont découvert que cette réaction se fait par un phénomène de nucléation passant par des oligomères. Il semblerait que ces espèces soient la principale cause de la toxicité observée dans les cellules des patients atteints d'amyloïdose. C'est pourquoi un intérêt particulier est donc porté aux premières étapes d'oligomérisation. Dans ce mémoire, nous nous intéressons à une séquence d'acide aminé fortement hydrophobe de l' α -synucléine appelée composante non β -amyloïde (Non-Amyloid β Component ou NAC). Cette dernière est retrouvée sous forme de fibres dans les corps et les neurites de Lewy des patients atteints de la maladie de Parkinson. De plus, elle constitue une composante minoritaire des fibres impliquées dans la maladie d'Alzheimer. Nous avons observé les changements structuraux qui ont lieu pour le monomère, le dimère et le trimère de la séquence NAC de l'α-synucléine. Nous nous sommes aussi intéressés aux conséquences structurelles observées dans des oligomères hétérogènes qui impliqueraient, $A\beta_{1-40}$. Pour cela nous utilisons des dynamiques moléculaires, d'échange de répliques couplées au potentiel gros-grain, OPEP. Nous constatons une disparition des hélices α au profit des feuillets β , ainsi que le polymorphisme caractéristique des fibres amyloïdes. Certaines régions se sont démarquées par leurs capacités à former des feuillets β . La disparition de ces régions lorsque NAC est combinée à A β laisse entrevoir l'importance de l'emplacement des résidus hydrophobes dans des structures susceptibles de former des fibres amyloïdes.

Mots clés: protéines amyloïdes, maladies neurodégénératives, dynamique moléculaire, potentiel gros-grain .

ABSTRACT

Amyloid proteins are found in fiber form in many neurodegenerative diseases. In attempting to elucidate the mechanism of fibrillation, researchers have found that fibril formation occurs by a nucleation mechanisms involving oligomers. It seems, in particular, that the latter species are responsible for the toxicity observed in the cells of patients suffering from amyloidosis. That is why special interest is focused in the early stages of oligomerization. In this this work, we focus on a highly hydrophobic amino acid sequence of the α -synuclein called Non-Amyloid β Component (NAC). The NAC is recovered in the form of fibers in the body and Lewy neurites in patients with Parkinson's disease. Moreover, it is a minority component of the fibers involved in Alzheimer's disease. In particular, we observe the structural changes taking place for the monomer, dimer and trimer of the NAC region of α -synuclein. We are also interested in the structural consequences observed in heterogeneous oligomers which involve $A\beta_{1-40}$. We use Hamiltonian and temperature replica exchange molecular dynamics (HT-REMD) simulations combined with the coarse-grained OPEP potential. We observe a loss of α -helices in favor of β -strands and the characteristic polymorphism of amyloid fibers. We also find that some regions are distinguished by their ability to form β -strands. The disappearance of these regions when combined A β with NAC suggests the importance of the location of hydrophobic residues in amyloid fibers structures.

Keywords: amyloid proteins, neurodegenerative diseases, molecular dynamic, coarse grained potential.

TABLE DES MATIÈRES

RÉSUM	1É		ii
ABSTR	ACT.		iv
TABLE	DES M	IATIÈRES	١
LISTE	DES TA	ABLEAUX	vii
LISTE	DES FI	GURES	ix
DÉDIC	ACE .		X
REME	RCIEM	ENTS	xi
INTRO	DUCTI	ON	1
CHAPI	TRE 1	: INTRODUCTION BIOCHIMIQUE	3
1.1	Les pr	otéines	3
1.2	Les pr	otéines amyloïdes	7
	1.2.1	Structures	7
	1.2.2	Mécanismes	8
	1.2.3	Toxicité et oligomères	ç
1.3	Choix	du système : La région NAC de l' α -synucléine et A β_{1-40}	10
	1.3.1	La séquence NAC de l' α -synucléine	10
	1.3.2	la séquence $A\beta_{1-40}$ issue du clivage de la glycoprotéine Amy-	
		loid precursor protein (APP)	11
	1.3.3	Problématiques	12
CHAPI	TRE 2	: MÉTHODOLOGIE	14
2.1	OPEP-	Optimized Potential for efficient protein prediction	14
	2 1 1	Description du potentiel	15

	•
T 7	1
v	ı

	2.1.2	Quelques champs de force	21
	2.1.3	Avantages du potentiel OPEP	22
	2.1.4	Limites du potentiel OPEP	22
2.2	La dyı	namique moléculaire	23
	2.2.1	Description de la dynamique moléculaire	24
	2.2.2	Avantages et limites de la dynamique moléculaire	25
	2.2.3	Échange de répliques en dynamiques moléculaires :HT-REMD	26
2.3	Les ou	itils d'analyses	28
	2.3.1	Le domaine de convergence	28
	2.3.2	Suivi de l'évolution des structures secondaires	30
	2.3.3	Les contacts	31
	2.3.4	Rassemblement de structures semblables	32
2.4	Conclu	usion	32
CHAPI	TDF 2	: EARLY OLIGOMERIZATION STAGES FOR THE NON-	
СПАГІ	I KE 3		
		AMYLOID COMPONENT OF α -SYNUCLEIN AMYLOID	
3.1		act	33
3.2	Introd	Introduction	
3.3	Metho	ds	36
3.4	Result	s	38
	3.4.1	Monomer	39
	3.4.2	Dimer	40
	3.4.3	Trimer	45
3.5	Discus	ssion	48
	3.5.1	Evolution of the secondary structure	48
	3.5.2	The polymorphism	50
	3.5.3	Comparison with other amyloid proteins	51
3.6	Conclu	Conclusion	
3.7	Ackno	owledgments	54

CHAPI'	TRE 4: AGGREGATION PROCESS OF $A\beta_{1-40}$ WITH NON- $A\beta$	
	AMYLOID COMPONENT OF α -SYNUCLEIN	55
4.1	Methods	55
4.2	Results	58
	4.2.1 The dimer of $A\beta_{1-40}$ with non- $A\beta$ amyloid component of α -	
	synuclein	59
4.3	Discussion	63
4.4	Conclusion	65
4.5	Acknowledgments	66
CHAPI	TRE 5: CONCLUSION	67
RIRLIC	OGRAPHIE	71

LISTE DES TABLEAUX

3.I	Percentage of secondary structure taken at 300 K over the full	
	NAC peptide	40
3.II	The region implicated in secondary structure for monomer, dimer,	
	trimer	41
4.I	Percentage of secondary structure taken at 300 K over the full ma-	
	jor cluster and for all conformations	59

LISTE DES FIGURES

1.1	Les 20 acides aminés	3
1.2	Les niveaux de structure	4
1.3	Diagramme de Ramachandran	6
1.4	Structure cross beta	7
1.5	Mécanisme de fibrillation	9
2.1	Représentation de la leucine en modèle tout-atome et gros grain	
	OPEP (Optimized Potential for efficient protein prediction)	16
2.2	Paramètres pour $\Delta(ijkl)$ =1	19
2.3	Paramètres pour $\Delta'(ijkl)$ =1	20
2.4	Schématisation de l'échange de répliques en dynamiques molécu-	
	laire	27
3.1	Initial configurations from monomer to trimer	37
3.2	Entropy as a function of temperature for the NAC monomer, dimer	
	and trimer	39
3.3	Evolution of α -helix and β -strands as a function of residue num-	
	ber at 300K from monomer to trimer	42
3.4	The clusters' center for NAC of monomer, dimer and trimer	43
3.5	The contact maps display the contact propensity of total, intra-	
	chain, interchain contacts for NAC of monomer, dimer and trimer	44
3.6	The contact maps display the contact propensity of total contacts	
	for the dominant morphology of monomer	45
3.7	The contact maps display the contact propensity of total, intra-	
	chain and interchain contacts for the dominant morphology of di-	
	mer	46
3.8	The contact maps display the contact propensity of total, intra-	
	chain and interchain contacts for the dominant morphology of trimer	47

3.9	Comparaison of the sequence of NAC of α -synuclein with hIAPP	
	and A eta_{1-40}	52
4.1	Entropy as a function of temperature for the NAC and $Aeta_{1-40}$	57
4.2	Specific heat as a function of temperature for the NAC and $Aeta_{1-40}$	
	from 400 to 600 ns	58
4.3	lpha-helix and eta -strand propensity as a function of residue number	
	at 300K for the NAC/A eta dimer	60
4.4	The contact propensity of total, intrachain and interchain contacts	
	for NAC/A eta_{1-40} dimer	61
4.5	The clusters' center of the dimer NAC with A eta_{1-40}	62
4.6	The contact maps display the contact propensity of total, intra-	
	chain and interchain contacts for the major clusters of the dimer of	
	NAC with A β_{1-40}	63



REMERCIEMENTS

Tout d'abord j'aimerais remercier mes co-directeurs de recherche de m'avoir donner l'opportunité de faire une maîtrise de recherche en bio-informatique. Leurs marques de confiance et leur foi en mes capacités d'adaptation ont été pour moi de véritables encouragements à toujours me dépasser. La patience et la pédagogie de Monsieur Radu Ion Ifitimie, que j'avais déjà pu apprécier durant mon baccalauréat en chimie ont été un atout indéniable durant mon cheminement académique. Ma gratitude va aussi à Monsieur Normand Mousseau. Sa supervision, trouvant le juste équilibre entre exigences élevées et liberté d'exécution, m'ont fourni un terreau fertil pour m'épanouir en tant qu'étudiante et en tant qu'individu. J'ai eu dans son groupe de recherche, l'espace et le lieu pour devenir pleinement moi-même en toute confiance et apprendre à exploiter au mieux mes compétences. J'y ai rencontré des gens qui resteront longtemps dans ma tête et mon coeur : Jean-François Joly, Mickaël Trochet, Laurent Karim Béland, Gawonou Kokou Ntsouaglo, Vincent Binette, Jean-Francois St-Pierre, Peter Brommer. Une tendre pensée pour Jessica Nassica-Labouze et Lilianne Dupuis, 2 femmes extraordinaires, qui ont toujours été présentes pour partager mes joies et mes peines durant ces deux années et avec qui je garde un lien empli d'affection et de respect. À leur contact, j'ai pu constater qu'avec organisation et persévérance, une femme peut allier vie personnelle et académique avec harmonie. Je tiens aussi à remercier Sébastien Côté pour sa disponibilité et ses suggestions. Nos discussions et petits défis m'ont confirmé l'importance d'un travail accompli dans le respect, la bonne humeur et l'entraide. Je salue Rozita Laghaei, pour sa rigueur et sa gentillesse durant notre collaboration qui a donné lieu à l'écriture d'un article.

Je voudrai remercier des organismes comme le GEPROM, dont les conférences et les séminaires m'ont permis d'apprendre plus sur la recherche au niveau des protéines mais aussi de m'exprimer sur l'avancée de mes recherches dans un cadre motivant. Cet organisme et d'autres comme le Conseil de Recherches en Sciences Naturelles et en Génie du Canada (CRSNG), la Fondation des Chaires de Recherche du Canada, les Fonds Québécois de Recherche en Santé et la Faculté des Études Supérieures et Postdocto-

rales (FESP) qui fournissent des fonds permettant de financer les travaux des étudiants désireux de poursuivre leurs parcours universitaires.

De façon plus personnelle, j'aimerais souligner la présence dans ma vie de la famille Nséké-Pérutin, dont chacun des membres a su m'apporter douceur et légèreté dans mes moments de doute. Je les remercie pour la confiance qu'ils ont mise en moi et les beaux moments de vie partagés qui m'ont rendue plus responsable et mature. Je remercie mes parents, car sans eux, je n'aurais jamais pu suivre mon rêve de partir effectuer mes études à l'étranger. Ils me donnent un soutien de tous les instants dans lequel je puise mon courage et ma force pour continuer à avancer vers mes objectifs. Une douce pensée pour ma grand-mère dont les prières et bougies allumées ont souvent fourni une étincelle particulière à mes réussites. Enfin, un grand merci à Walter Frankem Wakem pour ces bons mots d'encouragement et son soutien à toutes épreuves.

INTRODUCTION

Les séquences d'acides aminés qui composent une protéine sont responsables de son repliement et lui permettent d'adopter sa forme native et d'assurer son rôle dans le métabolisme cellulaire. Il arrive parfois que ce repliement se fasse mal et provoque une suite d'événements ayant des conséquences diverses pouvant mener à la mort cellulaire [1]. Ces dernières années, les recherches expérimentales et numériques s'allient dans le but de proposer des solutions thérapeutiques qui permettraient de prévenir les maladies neurodégénératives provoquées par les protéines amyloïdes [2]. Il a été démontré que leurs mépliements mènent à des fibres que l'on retrouve dans les cellules nerveuses du cerveau de patients atteints d'amiloïdoses telles que la maladie de Parkinson, la maladie d'Alzheimer et les maladies à prions. Les différentes maladies neurodégénératives dans lesquelles la formation de fibres amyloïdes est impliquée permettent de constater que les protéines à leurs origines peuvent être très différentes [3]. De plus, on peut voir qu'une même protéine peut mener à la formation de fibres ayant des structures très différentes, démontrant ainsi un certain polymorphisme [4]. Cependant, la structure cross- β , fortement impliquée dans la stabilisation de la fibre, sera présente dans chacune de ces structures [5–7].

Les détails concernant le mécanisme de formation des fibres amyloïdes sont peu connus. Toutefois, celui-ci semble dirigé par un phénomène de croissance par nucléation dont la première étape a pour but de former un noyau stable, le noyau critique. Cette étape est suivie d'une phase de croissance, lors de laquelle d'autres protéines viennent se fixer sur le noyau en adoptant la structure de feuillets β . Enfin, après avoir formé des oligomères qui se regroupent en protofibriles, il y a formation de fibre [8]. Un intérêt particulier est porté à l'étape d'oligomérisation car des études faites dans les dix dernières années démontrent et confirment la forte implication des petits oligomères dans la toxicité [9–13].

Dans ce mémoire, nous avons fait l'étude numérique des premières étapes d'agrégation de l' α -synucléine dont la séquence d'acides aminés *non-Amyloid* β *component* (NAC) impliquée dans les neurites et corps de Lewy est retrouvée dans le cerveau des

malades atteints de la maladie de Parkinson [14–16] et combinée avec $A\beta_{1-40}$ dans la maladie d'Alzheimer [17, 18]. Nos objectifs de recherche sont de caractériser les premières étapes d'oligomérisation de la région NAC de l' α -synucléine pour des structures allant du monomère au trimère (chapitre 3) et d'observer les changements qui ont lieu lorsque cette région est mise en présence de la protéine $A\beta_{1-40}$ (chapitre 4). De cette façon, nous pourrons donner quelques hypothèses sur les natures et fonctions de certains résidus ou séquences de résidus, de la région NAC, favorisant la formation de feuillets β et donc de fibres amyloïdes.

Nous avons utilisé une méthode de dynamique moléculaire avec répliques optimisée dans notre groupe couplée au potentiel gros-grain, OPEP. Cette méthode a permis de simuler des systèmes allant jusqu'au trimère pendant des temps de l'ordre de la microseconde, beaucoup plus longs que dans d'autres études numériques et plus pertinents dans la comparaison avec des études expérimentales. Un grand nombre d'études numériques traitent de la protéine $A\beta_{1-40}$ [19–23], mais ces interactions avec NAC sont encore très peu étudiées. Ainsi, ce mémoire permet de faire le parallèle entre la région NAC et d'autres protéines amyloïdes plus étudiées, en mettant l'emphase sur les acides aminés de la structure primaire qui favorisent certaines structures tertiaires propres aux protéines amyloïdes. Finalement, nous tentons de fournir quelques hypothèses sur le rôle que pourrait jouer NAC sur la stabilisation de certaines structures dans les premières étapes d'oligomérisation de la protéine $A\beta_{1-40}$.

CHAPITRE 1

INTRODUCTION BIOCHIMIQUE

1.1 Les protéines

Dans ce chapitre, nous allons éclaircir le vocabulaire et les notions se rattachant à l'étude des protéines en général (section 1.1) et des amyloïdes en particulier (section 1.2). Enfin, nous tenterons d'expliquer le choix de l' α -synucléine et $A\beta_{1-40}$ dans l'élucidation de la problématique.

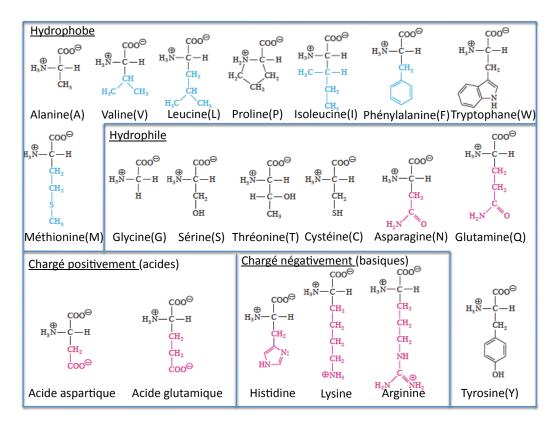


Figure 1.1 – Les 20 acides aminés qui composent une protéine, séparés en 4 catégories selon les propriétés physico-chimiques de leurs chaînes latérales. Les images ont été modifiées et proviennent du livre de Horton $et\ al\ [24]$.

Les protéines sont le fondement chimique de la vie, elles sont à la base de nombreuses fonctions indispensables au métabolisme d'une cellule. Elles peuvent jouer un rôle au niveau de la forme et de la résistance de la structure d'une cellule, du transport de substances indispensables à la survie du corps humain (oxygène par hémoglobine), de la catalyse sous forme d'enzyme, de la protection et la reconnaissance sous forme d'anticorps, de messager sous forme d'hormones et permettre la mobilité (flagelle du spermatozoïde) et bien d'autres fonctions encore, primordiales dans le métabolisme cellulaire.

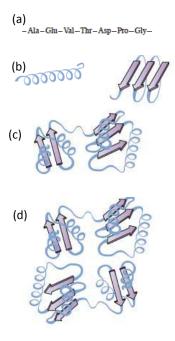


Figure 1.2 – Les niveaux de structure du niveau primaire au niveau quaternaire.

- a) structure primaire : la séquence d'acides aminés et les liens covalents qui les lient ainsi que les ponts disulfures qui se forment entre deux cystéines.
- b) structure secondaire : les structures régulières adoptées par la chaîne polypeptidique : hélice α et feuillet β .
- c) structure tertiaire : l'organisation spatiale des différentes structures secondaires et l'orientation des chaînes latérales.
- d) structure quaternaire : organisation spatiale des différentes chaînes les unes par rapport aux autres.

L'image a été modifiée et provient du livre de Horton et al. [24]

Pour qu'une protéine soit fonctionnelle, elle doit être sous sa forme tridimensionnelle la plus stable, appelée aussi état natif. Ce sont les interactions entre les résidus des acides aminés et l'eau qui vont diriger et chaperonner le repliement et permettre d'adopter une conformation adéquate à la réalisation de la fonction de la protéine :

- Les interaction de van der Waals entre deux atomes adjacents qui agissent à courte distance (de l'ordre de quelques Å).
- L'effet hydrophobe s'applique aux résidus avec une chaîne hydrophobe. La protéine adopte une conformation dans laquelle un coeur hydrophobe est formé afin de minimiser le contact avec le solvant.
- La force électrostatique qui traduit l'attraction ou la répulsion qu'il peut y avoir entre deux résidus chargés et s'applique donc sur les résidus acides et basiques.
 Les structures formées par deux résidus de charge opposée s'appellent des ponts salins.
- Les ponts hydrogène à l'origine de la formation des hélices α et feuillets β qui relient les atomes d'oxygène d'un groupe carbonyle et les atomes hydrogène d'un groupe amine.

20 acides aminés sont indispensables afin de constituer la chaîne polypeptidique d'une protéine. Ces acides aminés possèdent tous une fonction amine et une fonction carbonyle sur un carbone commun qu'on appelle $C\alpha$. (Fig 1.1)

Il existe 4 niveaux structuraux qui permettent de décrire de façon détaillée cette conformation (Fig 1.2) :

- 1. La structure primaire est la séquence d'acides aminés et les liens covalents qui les lient ainsi que les ponts disulfures qui se forment entre deux cystéines.
- 2. La structure secondaire désigne les structures régulières adoptées par la chaîne polypeptidique : hélice α et feuillet β .
 - Le diagramme de Ramachadran (Fig 1.3) permet de prédire quelles structures secondaires seront préférentiellement choisies par un acide aminé. En effet, les

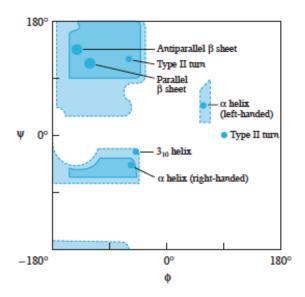


Figure 1.3 – Diagramme de Ramachandran. L'axe des abscisses correspond à l'angle ϕ (C_{n-1} -N- C_{α} -C) et l'axe des ordonnées à l'angle ψ (N- C_{α} -C- N_{n+1}). Ce diagramme prend en compte les encombrement stériques et les répulsions électrostatiques qui vont contraindre les angles ϕ et ψ à adopter uniquement certaines valeurs qui vont favoriser certaines structures secondaire. Les structures décrites dans ce diagramme sont adoptées par une séquence lorsque plusieurs de ces résidus appartiennent à une région identifiée du diagramme. L'image a été modifiée et provient du livre de Horton *et al.* [24]

angles ψ et ϕ ne peuvent pas prendre n'importe quelle valeur à cause de l'encombrement stérique et des répulsions électrostatiques. On peut appliquer ce diagramme sur tous les acides aminés sauf la glycine et la proline. On sait qu'une structure décrite par l'une des régions va être adoptée lorsque plusieurs résidus sont dans une région particulière du diagramme.

3. La structure tertiaire décrit l'organisation spatiale des différentes structures secondaires et l'orientation des chaînes latérales. Elle correspond à la structure tridimensionnelle que va prendre la protéine sous la contrainte imposée par le regroupement des régions hydrophobes. D'autres interactions peuvent s'ajouter pour stabiliser la structure pour qu'elle adopte sa forme native.

4. La structure quaternaire s'applique pour les protéines composées de plusieurs chaînes polypeptidiques et donne des informations sur l'organisation spatiale des différentes chaînes les unes par rapport aux autres.

1.2 Les protéines amyloïdes

Les protéines amyloïdes regroupent une classe de protéines mépliées qui seraient impliquées dans de nombreuses maladies neurodégénératives comme les maladies d'Alzheimer, de Parkinson, de Huntington. Nous tenterons ici de donner les informations connues aux niveaux structurel et mécanistique, pour lesquelles les protéines amyloïdes impliquées dans les amiloïdoses semblent partager des similitudes.

1.2.1 Structures

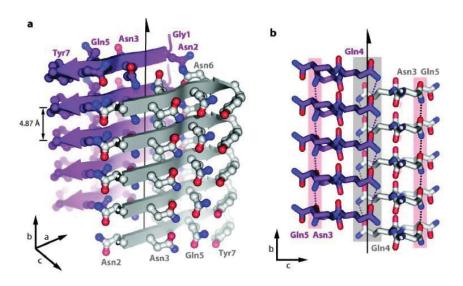


Figure 1.4 – Structure cross- β de GNNQQNY. a) figure montrant des brins parralèles dans chaque feuillet et des feuillets antiparralèles les uns par rapport aux autres.(les chaînes latérales sont représentées par des bâtons et des sphères). b) interdigitation des chaînes latérales vu de côté (l'axe a). Cette figure provient de Nelson *et al.* [25]

Les protéines amyloïdes sont des protéines mépliées qui, sous certaines conditions physiologiques, forment des agrégats non solubles appelés fibres amyloïdes [26]. De nombreuses protéines, très différentes les unes des autres, rentrent dans cette catégorie et semblent avoir les mêmes effets au niveau cellulaire. Pourtant une conformation commune, la structure cross- β semble se démarquer et laisse présager que des structures primaires très différentes peuvent mener à la formation de fibres [5, 27–30]. Ceci est confirmé par des expériences dans lesquelles des oligomères provenant de différentes protéines auraient tous une affinité qui leur permet de se lier à un même anticorps [25].

Il existe peu d'information sur les détails de la structure des fibres amyloïdes. Toutefois, on sait qu'elles sont des structures filamenteuses d'une dizaine de nanomètres de diamètre et de 0,1 à 10 micromètres de long. La structure cross- β , riche en feuillets β semble être primordiale dans la structure fibrillaire et se définit par des paires de feuillets β dont les chaînes latérales se croisent par interdigitation dont l'axe de symétrie sera perpendiculaire à l'axe de la fibre (Fig 1.4). La nature même des fibres est un défi, car on ne peut pas utiliser les techniques habituelles d'élucidation des structures, comme la résonnance magnétique nucléaire (RMN) ou la cristallographie. Cependant, d'autres techniques en passant par des mesures indirectes permettent de donner des informations sur les fibres [3, 31, 32]. L'observation des structures secondaires, comme les feuillets β caractéristiques de la formation des fibres amyloïdes, est donc devenue une bonne alternative et a permis d'avoir certaines informations structurelles.

1.2.2 Mécanismes

Le mécanisme de formation des fibres amyloïdes se fait par un phénomène de croissance par nucléation qui peut être séparé en 3 phases. La première étape correspond à la nucléation, qui est caractérisée par la formation de petits oligomères riches en feuillets β qui vont se rassembler jusqu'à la formation d'un noyau critique. La durée de cette phase peut varier selon la nature de la protéine. La deuxième phase permet la croissance du noyau et la formation de protofobrile. Enfin, la réaction s'arrête lorqu'un équilibre est atteint et qu'il ne reste plus assez de monomères pour se fixer sur le noyau [8](Fig 1.5).

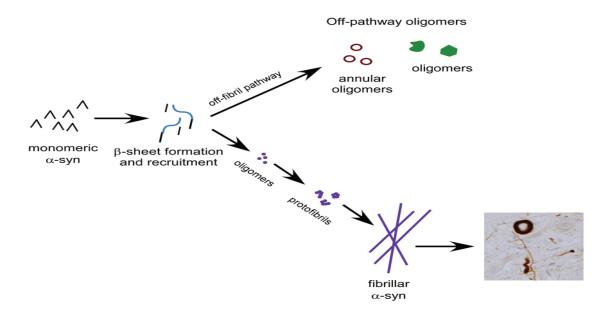


Figure 1.5 – Mécanisme de fibrillation. Croissance par nucléation en 3 étapes : la nucléation, la croissance du noyau menant à la formation de protofibrile. Enfin, arrêt de la réaction lorqu'un équilibre est atteint. L'image provient de Waxman et al(2009) [8].

1.2.3 Toxicité et oligomères

L'étude des petits oligomères n'a pas toujours été centrale dans la course vers la recherche de l'espèce responsable de la toxicité observée dans les amyloïdoses. Pourtant, depuis une dizaine d'années les chercheurs portent l'emphase sur l'étude des petits oligomères, car ils semblent être responsables de la toxicité observée [9–13]. De nombreuses études semblent confirmer l'hypothèse selon laquelle les oligomères formeraient des pores sur les parois cellulaires provoquant un déséquilibre ionique qui serait responsable de la mort cellulaire [33–35]. Les fibres pourraient donc avoir un rôle bénéfique car elle d'empêcheraient les petits oligomères solubles et toxiques de se propager dans la cellule [13, 31].

1.3 Choix du système : La région NAC de l' α -synucléine et A β_{1-40}

Dans cette étude, nous avons choisi un fragment particulier de l' α -synucléine qui semble être important dans la formation des fibres amyloïdes présentes dans les cellules nerveuses de patients atteints de la maladie de Parkinson [14–16]. Cette séquence est aussi le deuxième élément constituant dans les fibres amyloïdes impliquées dans la maladie d'Alzheimer. Dans cette partie, nous présenterons l' α -synucléine et l' $A\beta_{1-40}$, en justifiant les raisons qui nous ont motivé à choisir ces protéines amyloïdes pour tenter de répondre à certains points de la problématique liée aux premières étapes d'oligomérisation.

1.3.1 La séquence NAC de l' α -synucléine

L'α-synucléine est une protéine présynaptique retrouvée dans la substance noire du cerveau. On connait très peu sa fonction, mais les effets observés dans les maladies neurodégénératives, comme la maladie de Parkinson permet d'émettre l'hypothèse selon laquelle elle serait impliquée dans le transport de la dopamine. Lorsque cette protéine est mépliée, elle subit des phénomènes de fibrillation et d'agrégation à la suite desquels on peut la retrouver sous forme de neurites et corps de Lewy dans le cerveau des patients. Cette protéine est constituée de 140 acides aminés que l'on peut séparer en 3 régions. La première est la région du N-terminal (1-60) dans laquelle on retrouve la séquence répétée KTEKEGV. Elle a une forte propension à former des hélices amphiphatiques qui, par définition, présentent tous leurs résidus hydrophobes d'un même côté de la structure et sont caractéristiques des lipides liant les apolipoprotéines [36, 37]. La deuxième région est la région du C-terminal (96-140), hautement chargée négativement avec plusieurs résidus acides. D'une part, l'observation de la β -synucléine et de la γ -synucléine, dont les régions C-terminales diffèrent de l'a-synucléine au niveau du nombre de résidus hydrophobes et aussi de la charge, permettent de pressentir un lien entre ces deux facteurs et la régulation de l'agrégation. D'autre part, la comparaison des taux d'agrégation de l'α-synucléine et deux de ses mutations (A30P et A53T) [38–41] permet de constater que les mutations réduisent l'hydrophobicité de la région et aussi sa capacité à former des hélices α et augmentent celle des feuillets β . Ainsi, même si ces mutations n'altèrent pas la structure du monomère formé, elles vont affecter l'agrégation à cause de leur implication dans la stabilisation des feuillets β [42]. Ainsi, nous pouvons constater que les régions du N-terminal et du C-terminal servent de modulateurs et qu'il existe une séquence minimale nécessaire pour la formation de fibre qui se situe entre ces deux régions.

Cette région est appelée la région NAC, *Non-Amyloid* β *Component*, et semble avoir une importance cruciale dans le mécanisme d'agrégation et de fibrillation de l' α -synucléine à l'origine de dépôts retrouvés dans le cerveau des malades atteints de la maladie de Parkinson et aussi de façon minoritaire dans la maladie d'Alzheimer. Thompson *et al.* vont dans le sens de mécanismes dans lesquels des régions d'au moins sept acides aminés issuent de protéines amyloïdes pourraient suffire à la fibrillation[43]. Des critères comme la biréfringence verte observée par le microscope polarisé lorsque mis en contact avec le rouge Congo, des structures fibrillaires, rigides et fines sous microscope électronique et la non-solubilité en milieu aqueux ont été choisis et vérifiés par Iwai *et al.* pour confirmer la nature amyloïde de NAC [18]. Ils ont aussi observé que lorsque mis en contact avec $A\beta$, le NAC n'intervient que dans la première partie du mécanisme sous forme de noyau (Fig. 1.5).

Depuis de nombreuses années, l'emphase est mise sur la recherche de la plus petite séquence qui pourrait être responsable de la toxicité observée pour les oligomères des protéines amyloïdes. Bodles *et al.* [9] se sont arrêtés sur la région 8-16 de NAC et nuancent leurs résultats avec ceux d'autres chercheurs comme Giasson et *et al.* [44] qui trouvent une région de douze acides aminés indispensables (71-82 correspondants à NAC(10-21)) lorsqu'ils s'intéressent à $1^{\circ}\alpha$ -synucléine au complet.

1.3.2 la séquence $A\beta_{1-40}$ issue du clivage de la glycoprotéine Amyloid precursor protein (APP)

Lors du clivage de la glycoprotéine APP par les α et β -secrétase, plusieurs produits de tailles différentes se forment. La protéine $A\beta_{1-40}$ est le produit majoritaire de cette réaction [13] et est plus étudiée expérimentalement que d'autres fragments comme

 $A\beta_{1-42}$ qui est toxique. Il existe de nombreuses études numériques utilisant des méthodes diverses afin de parfaire la connaissance sur les petits oligomères de cette protéine qui est impliquée de façon majoritaire dans les fibres amyloïdes de la maladie d'Alzheimer. Même si tous leurs résultats ne s'accordent pas on peut extraire quelques informations. Par exemple, des études menées sur la région 21-30 de A β_{1-40} par RMN [45] permettent de penser que la structure du monomère aura une forte tendance à former une boucle entre les résidus 24 et 28. Au niveau expérimental, l'étude du monomère est devenue centrale depuis que l'on sait que l'on retrouve des oligomères aussi petits que les dimères dans les fibres amyloïdes de la maladie d'Alzheimer [46]. De plus, il semblerait que les monomères réussissent à coexister avec les autres oligomères lors de la polymérisation. Il semblerait que la région N-terminal soit très importante dans le phénomène de fibrillation. En effet, des mutations de cette région semblent affecter la structure des monomères et le processus d'oligomérisation [47]. De nombreuses études se sont aussi intéressées à différents fragments de A β , mais il en existe très peu qui explorent des éventuelles interactions entre NAC et $A\beta_{1-40}$ alors que ces protéines sont deux composantes importantes des fibres amyloïdes retrouvées chez les patients atteints de la maladie d'Alzheimer.

1.3.3 Problématiques

Nous venons de souligner l'importance de l' α -synucléine et d' $A\beta_{1-40}$ dans deux grandes maladies neurodégénératives très répandues. La recherche numérique qui s'est pourtant intéressée de près à d'autres protéines amyloïdes semble s'être moins intéressée à la région NAC de l' α -synucléine. De plus, malgré le fait que NAC et $A\beta_{1-40}$ soient retrouvées dans la maladie d'Alzheimer, il est étonnant que très peu d'intérêt soit porté sur les particularités structurelles et mécanistiques d'oligomères hétérogènes combinant ces deux protéines . En faisant des simulations numériques mettant en jeu des petits oligomères de NAC et de $A\beta_{1-40}$ nous tenterons de répondre à certaines questions : Peut-on identifier des structures primaires ou séquences d'intérêt de NAC, favorisant la formation de structures secondaires caractéristiques de la fibrillation ? Y a-t-il alors des particularités observées sur les structures adoptées au niveau tertiaire et secondaire ?

Peut-on généraliser et vérifier les hypothèses mécanistiques avancées pour NAC dans des oligomères homogènes en l'introduisant dans des oligomères hétérogènes dans lesquels elle est mise en contact avec $A\beta_{1-40}$?

CHAPITRE 2

MÉTHODOLOGIE

Au cours des dernières années, l'attention des chercheurs s'est portée vers les premières étapes d'agrégation et la formation d'oligomères semblant être encore plus toxiques que les fibres. Les mécanismes d'assemblage de ces protéines sont complexes et leur nature hors-équilibre rend les études expérimentales très difficiles. Dans ces conditions, il est utile de se tourner vers la simulation numérique pour essayer de comprendre la dynamique des premières étapes d'oligomérisation.

Ici, nous nous concentrons sur le monomère, le dimère et le trimère de l' α -synucléine, une protéine associée à la maladie de Parkinson, et les interactions qui ont lieu lors-qu'on la met en présence des premiers oligomères de la protéine $A\beta_{1-40}$. Combinant des méthodes de dynamique moléculaire avec répliques, optimisées dans notre groupe, à un potentiel gros-grain, OPEP, nous nous intéressons au rôle des divers acides aminés caractérisant les structures secondaires et regroupons l'ensemble des configurations selon leurs ressemblances à ce niveau. Nous décrivons en quelques mots les principes des outils numériques utilisés ainsi que certains des outils d'analyse, en soulignant leurs avantages mais aussi leurs limites.

2.1 OPEP- Optimized Potential for efficient protein prediction

Un champ de force est un ensemble de fonctions d'énergie potentielle, complété par des paramètres numériques expérimentaux, qui permet de décrire la structure de l'énergie potentielle d'un système. Habituellement, un champ de force est construit et utilisé dans le cadre de l'approximation de Born Oppenheimer. Elle permet de séparer la dynamique des noyaux et des électrons, ce qui permet de traiter de façon classique la dynamique des noyaux et de supposer que les électrons sont en constante équilibre avec le mouvement des noyaux. On peut obtenir la surface d'énergie potentielle des noyaux en résolvant l'équation de Schrodinger électronique seulement. Cette surface d'énergie

potentielle est traitée de façon effective dans un champ de force. Ainsi, des calculs quantiques permettent d'établir des paramètres et fonctions d'énergie potentielles, qui seront ajustées grâce aux mesures expérimentales effectuées dans les zones thermodynamiques d'intérêt.Une alternative consiste à considérer que les termes d'énergie potentielle intramoléculaire peuvent etre modélisés en utilisant l'approximation harmonique : la structure moléculaire est composé de billes et de ressort associés à des fonctions potentielles. La somme de ces fonctions est exprimée sous la forme d'un champ de force. Dans ce cas, le champs de force est constitué de fonctions d'énergie potentielle qui décrivent les interactions entre atomes liés et non-liés (équation 2.1)(Méthode empirique). C'est le cas du champ de force utilisé dans ce mémoire, OPEP, qui nous permet d'observer l'évolution des structures secondaires pour des oligomères allant jusqu'au trimère de la région NAC de l' α -synucléine et de la protéine A β_{1-40} . Ce potentiel gros-grain, basé sur un modèle simplifié dans lequel les résidus sont représentés par une bille, entraine une perte de la résolution. Cependant, une telle modélisation permet de diminuer le temps computationnel lié aux calculs de potentiel faisant intervenir les termes de paires du système (équation 2.8). De plus, il a déjà été beaucoup utilisé dans les dernières années couplé à différents outils numériques [48–51] sur diverses protéines amyloïdes [52– 61]. Ces études ont donc prouvé la pertinence de l'utilisation d'OPEP dans l'étude des phénomènes liés au repliement des protéines amyloïdes.

2.1.1 Description du potentiel

Dans le potentiel gros-grain OPEP, les acides aminés qui composent la structure primaire d'une protéine sont représentés sous forme de 6 billes : 5 pour les atomes de la chaîne principale, N, H, C, O, C α communs à tous les acides aminés et une dernière bille qui correspond à la chaîne latérale, sauf pour la proline. La géométrie et le rayon de van der Waals de la bille qui est identifié à la chaîne latérale sont fixés grâce à un échantillonnage représentatif de diverses structures (Fig.2.1). Globalement, la façon dont ces billes seront placées les unes par rapport aux autres prendra en compte la longueur des liens, l'angle entre les atomes ainsi que les angles dièdres [62]. Comme de nombreuses réactions chimiques et biologiques, celles que nous avons étudiées ont lieu

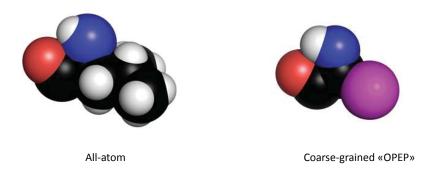


Figure 2.1 – Représentation de la leucine en modèle tout-atome et gros grain OPEP (Optimized Potential for efficient protein prediction). Avec le potentiel OPEP, les acides aminés composant la structure primaire d'une protéine sont représentés sous forme de 6 billes : 5 pour les atomes N, H, C, O, $C\alpha$ qui constituent la chaîne principale et une dernière bille correspondant à la chaîne latérale, sauf pour la proline. Les carbones, azotes, hydrogènes et oxygènes sont représentés respectivement par des billes, noires, vertes, blanches et rouges dans les deux modèles. Une bille rose représente la chaîne latérale de la structure obtenue par OPEP.

dans un solvant aqueux. Il faut donc prendre en compte les effets dynamiques de l'eau qui contribue à la stabilisation des protéines. En effet, pour des molécules amphiphiles qui possèdent à la fois des groupes polaires et apolaires comme les protéines, il y a solubilité lorsque la structure tertiaire permet de mettre les résidus polaires à l'extérieur et de former des poches hydrophobes, inaccessibles au solvant, à l'intérieur. Il existe deux façons de représenter l'eau : de manière explicite, par des boites périodiques entourant la molécule d'intérêt et en y rajoutant des molécules d'eau afin de représenter la densité du solvant. De manière implicite, avec un solvant considéré comme une perturbation de nature diélectrique caractérisée par sa constante diélectrique. La représentation implicite du solvant permet de réduire la taille du système et d'accélérer l'échantillonnage des différentes structures en négligeant les frottements qui ont lieu avec le solvant. Par exemple, un phénomène qui prend 1 μ s avec un solvant explicite peut ne prendre que 10 ns avec un solvant implicite [63].

L'énergie du système sera donnée par la somme des interactions covalentes, non liées et des ponts hydrogène :

$$E = E_{local} + E_{non-liant} + E_{HB} \tag{2.1}$$

Les énergies locales, E_{local} , prennent en compte les forces associées aux déformations des longueurs et des angles des liaisons covalentes ainsi que les forces reliées aux torsions et torsions impropres des chaînes principales et latérales :

$$E_{local} = w_b \sum_{liens} k_b (r - r_{eq})^2 + w_a \sum_{angles} k_\alpha (\alpha - \alpha_{eq})^2 + w_\Omega \sum_{imp-tor} k_\Omega (\Omega_{eq})^2 + E_{\phi,\psi}$$
 (2.2)

où r, α et Ω sont les valeurs de longueurs de liaisons, d'angles et de torsions hors équilibre. L'indice "eq" correspond aux même paramètre pris à l'équilibre. k_b , k_α et k_Ω sont les constantes de forces qui traduisent la rigidité des liaisons, angles et torsions. Les valeurs à l'équilibre et les constantes de forces de la chaînes principale sont tirées du potentiel AMBER [63] et celles associées à la chaîne latérale sont très similaires à celles d'AMBER. w_b , w_a et w_Ω sont des poids attribués aux composantes énergétiques afin d'assurer un bon équilibre entre les forces de courte et de longue portée[48].

$$E_{\phi,\psi} = w_{\phi,\psi} \left(\sum_{\phi} E_{\phi} + \sum_{\psi} E_{\psi} \right) \tag{2.3}$$

Les termes E_{ϕ} et E_{ψ} permettent de s'assurer que les angles dièdres que vont adopter les structures générées par OPEP soient représentatifs des structures favorisées dans le diagramme de Ramachandran (Fig. 1.3) :

$$E_{\phi} = k_{\phi \psi} (\phi - \phi_0)^2 \tag{2.4}$$

$$E_{\psi} = k_{\phi\psi}(\psi - \psi_0)^2 \tag{2.5}$$

où $k_{\phi\psi}$ est une constante de force. $\phi_0 = \phi$ si ϕ appartient à l'interval [ϕ_{inf} : ϕ_{sup}] et $\phi_0 = \min(\phi - \phi_{inf}; \phi - \phi_{sup})$ autrement, avec $\phi_{inf} = -160^\circ$ et $\phi_{sup} = -60^\circ$. Les mêmes critères sont appliqués pour les angles ψ , avec $\psi_{inf} = -60^\circ$ et $\psi_{sup} = 160^\circ$.

Les énergies non-liantes, $E_{non-liant}$, sont données par :

$$E_{non-liant} = w_{1,4} \sum_{1,4} E_{VdW} + w_{C\alpha - C\alpha} \sum_{C\alpha - C\alpha} E_{VdW} + w_{1>4} \sum_{S',S'} E_{VdW}$$
 (2.6)

$$+w_{1>4} \sum_{S',C\alpha} E_{VdW} + w_{1>4} \sum_{S',C\alpha} E_{VdW} + w_{1>4} \sum_{R,R} E_{VdW}$$
 (2.7)

L'indice 1,4 représente les interactions 1-4 à courte portée, qui font le lien entre les atomes aux extrémités d'un angle dièdre, "S'" et "R", représentent respectivement les fonctions amine et carboxyle de la chaîne principale et la chaîne latérale des acides aminés. Les atomes pour lesquels j > i+4 sont considérés comme des interactions à longue portée et identifiés par l'indice 1 > 4. Le terme E_{VdW} est donné par :

$$E_{wdw} = \varepsilon_{ij} \left(\left(\frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^0}{r_{ij}} \right)^6 \right) H(\varepsilon_{ij}) - \left(\frac{r_{ij}^0}{r_{ij}} \right)^6 H(-\varepsilon_{ij})$$
 (2.8)

H(x)=1 si $x \ge 0$ et H(x)=0 si x < 0, r_{ij} est la distance entre les particules i et j, r_{ij}^0 est la moyenne des rayons de Van der Waals de i et j. Pour les particules de la chaîne principale $r_N^0=1,75$ Å, $r_H^0=1,00$ Å, $r_{C\alpha}^0=2,385$ Å, $r_C^0=1,85$ Ået $r_O^0=1,60$ Å. Les valeurs des rayons de la chaîne latérale peuvent être retrouvées dans une table (ref article OPEP). Pour des interactions qui n'ont pas lieu entre deux chaînes latérales, H(x)=1 et un potentiel 12-6 sont utilisés. Pour les chaînes latérales hydrophobes ou de charges contraires, H(x)=1 et on garde le potentiel 12-6. Dans tout les autres cas, H(x)=0 et un potentiel -6 répulsif sont utilisés. Les coefficients ε_{ij} prennent différentes valeurs selon la nature des deux particules impliquées.

Enfin, le potentiel pour les liaisons hydrogène comprend un terme pour les intégrations à deux corps E_{HB1} et un terme pour les interactions à 4 corps qui prend en compte la coopérativité des liaisons hydrogène E_{HB2} :

$$E_{HB} = E_{HB1} + E_{HB2} \tag{2.9}$$

$$E_{HB1} = w_{hb1-4} \sum_{bonds} \varepsilon_{hb1-4} \mu(r_{ij}) v(\alpha_{ij}) + w_{hb1>4} \sum_{bonds} \varepsilon_{hbl>4} \mu(r_{ij}) v(\alpha_{ij})$$
(2.10)

où

$$\mu(r_{ij}) = 5(\frac{\sigma}{r_{ij}})^{12} - 6(\frac{\sigma}{r_{ij}})^{10}$$
(2.11)

$$v(\alpha_{ij}) = \cos^2 \alpha_{ij} \text{ si } \alpha_{ij} > 90^o, 0 \text{ sinon}$$
 (2.12)

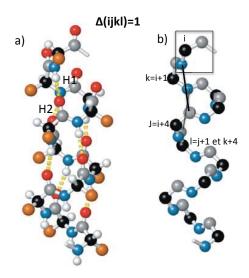


Figure 2.2 – Paramètres pour $\Delta(ijkl)$ =1. a) Représentation d'une hélice α et de deux ponts hydrogène H1 et H2. b) Le pont hydrogène H1 relie l'hydrogène de l'amide de l'acide aminé i et l'oxygène du carbonyle de l'acide amine j = i+4. Le pont hydrogène H2 relie l'hydrogène de l'amide de l'acide amine k = i+1 et l'oxygène du carbonyle de l'acide amine l = k+4 ce qui correspond aussi à l = j+1. Les carbonyle sont en gris, les C_{α} en noir et les amides en bleu. Le carré noir délimite ce qui compose un acide aminé et les traits noirs les deux ponts hydrogène qui sont en pointillés en jaune sur la figure a). Les images ont été modifiées et proviennent du livre de Horton *et al* [24]

La somme se fait sur tout les résidus i et j pour lesquels $j \ge i+4$, r_{ij} est la distance

$\Delta'(ijkl)=1$

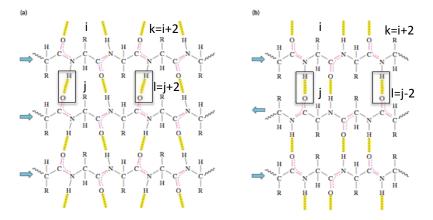


Figure 2.3 – Paramètres pour $\Delta'(ijkl)$ =1 . a) Les paramètres (k,l)=(i+2,j+2) décrivent les feuillets β -parallèles. b) Les paramètres (k,l)=(i+2,j-2) décrivent les feuillets β anti-parallèles. Les images ont été modifiées et proviennent du livre de Horton $et\ al\ [24]$

O...H du pont hydrogène qui se fait entre l'oxygène du carbonyle et l'hydrogène de l'amide, α_{ij} est l'angle NHO et σ la valeur à l'équilibre de la distance O...H (1,8 Å). les paramètres ε_{hb1-4} et $\varepsilon_{hbl>4}$ représentent les liaisons hydrogène de courte (helices α) et de longue portées (feuillet β).

Le calcul de coopérativité fait état de la formation des hélices α et des feuillets β à travers deux termes :

$$E_{HB2} = E_{HB2\alpha} + E_{HB2\beta} \tag{2.13}$$

$$E_{HB2\alpha} = \sum_{bonds} \varepsilon_{\alpha}^{coop} \exp\left(-\frac{1}{2}(r_{ij} - \sigma)^2\right) \exp\left(\left(-\frac{1}{2}(r_{kl} - \sigma)^2\right) \times \Delta(ijkl)$$
 (2.14)

$$E_{HB2\beta} = \sum_{bonds} \varepsilon_{\beta}^{coop} \exp\left(-\frac{1}{2}(r_{ij} - \sigma)^2\right) \exp\left(-\frac{1}{2}(r_{kl} - \sigma)^2\right) \times \Delta'(ijkl)$$
 (2.15)

Le paramètre $\Delta(ijkl)$ prendra comme valeur 1 si (k, l) = (i+1,j+1) et (j=i+4, l=k+4)

(Fig. 2.2), dans les autres cas ce paramètre sera égal à zéro. Ainsi, les hélices π ne sont pas stabilisées. Le paramètre $\Delta'(ijkl)$ est égal à 1 si (k, l) = (i+2, j-2) ou (i+2, j+2) (Fig. 2.3) et sera nul dans les autres cas. Ces conditions permettent d'aider à la stabilisation des hélices α et des feuillets β parallèles et anti-parrallèles, sans prendre en compte l'angle dièdre entre ϕ et ψ . Il s'applique aussi à tout segment respectant les conditions imposées pour i, j, k et l.

2.1.2 Quelques champs de force

Il existe de nombreux champs de forces et chacun on leur spécificité. la façon dont les différentes interactions vont être représentées va déterminer sur quelle genre de molécule il va être appliqué.

- Le champ ab initio, MM2, permet de traiter séparément les système avec des électrons π et ceux avec les électrons σ. La méthode SCF (Self-Consistent Field), semi-empirique, est utilisé pour les système π et la mécanique moléculaire sur les système σ. En général, ce sont plutôt les chimistes organiciens qui utilise ce champs de force, car il a été développé pour les hydrocarbures [64]. Les versions améliorées MM3 [65] et MM4 [66] permettent de simuler des molécules organiques de plus en plus complexes. MMFF94 est un champs de force de classe 2, il prend en compte des termes croisés afin de tenir compte d'interactions mixtes. Il est utilisé dans l'étude de conformations de petites molécules organiques en phase gaz. Très mauvais pour la phase condensée ou pour calculer les interactions intermoléculaires. Il est très utilisé dans l'industrie pharmaceutique pour effectuer des analyses conformationnelles. [71].
- Gromos (Groningen Molecular Simulation Program Package) permet d'observer les interactions entre les molécules d'eau et les groupements polaires des protéines et est donc souvent utilisé pour l'étude des biomolécules en milieux aqueux [67].
- AMBER (Asociated Model Building with Energy Refinement). Ce champ de force utilise les coordonnées cartésiennes des atomes de la molécule et les paramètres

qui y sont utilisés sont issues de calculs *ab initio* sur des molécules modèles regroupés dans une bibliothèque définie dans le programme. Il possède un mode "united atoms" dans lequel on ne considère pas les hydrogènes non-polaire. L'autre mode est le tout-atome et utilise des paramètres détaillés prenant en compte les charges pour chacun des acides aminés [70]. OPLS (Optimized Potentials for liquid simulations) optimise le potentiel qui décrit les propriété de solvatation et est compatible avec les TIP4P, TIP3P et SPC (modèle de l'eau). L'étirement des liens, les angles de flexions et les torsions sont modélisés d'après le champ de force AMBER dans son mode "united-atom" [68]. CHARMM (Chemistry at HARvard Macromolecular Mechanics) se spécialise dans les calculs concernant les biomolécules, en particulier, les aminoacides et et protéines et est semblable à AMBER [69].

2.1.3 Avantages du potentiel OPEP

Le champs de force est choisi en fonction de la nature de la molécule mais aussi des paramètres qui seront importants lors de l'analyse des résultats. En réduisant le nombre d'interactions à calculer par diminution du nombre de centres, le potentiel OPEP permet de diminuer le coût numérique lié au calcul des forces. Ceci permet de simuler des systèmes plus gros, mais aussi d'atteindre des temps de simulation plus représentatifs au niveau expérimental. De même, il a été démontré que la formation des fibres amyloïdes était souvent précédée et sûrement favorisée par l'augmentation des feuillets β . Ainsi, malgré les simplifications au niveau des chaînes latérales, le fait que le potentiel prenne en compte la capacité des résidus à former des hélices α et des feuillets β est un atout majeur dans l'étude des protéines amyloïdes. De plus, il a déjà beaucoup été utilisé pour diverses protéines amyloïdes. [52–61]

2.1.4 Limites du potentiel OPEP

La modélisation OPEP est un frein dans l'étude de plus gros oligomères, dans laquelle on tenterait d'élucider le mécanisme de grosses structures faisant intervenir les chaînes latérales, comme les motifs d'interdigitation stabilisant les structures cross- β dans les fibres amyloïdes (Fig.1.4). Il peut arriver que la viscosité de l'eau ou son effet sur certains résidus chargés jouent un rôle central dans le mécanisme de repliement d'une protéine; dans ces cas-là, la représentation explicite du solvant permettrait de reproduire les effets de telles interactions. Notre étude, concernant uniquement les premières étapes d'agrégation, n'est pas affectée par ces limitations et permet donc de former une structure primaire que nous pourrons soumettre à la dynamique moléculaire et pour laquelle nous pourrons examiner les structures secondaires, tertiaires et quaternaires avec divers outils d'analyse.

2.2 La dynamique moléculaire

La dynamique moléculaire est une méthode numérique qui permet de suivre l'évolution d'une structure en fonction du temps. En appliquant les principes fondamentaux de la dynamique à chaque élément d'un système étudié, nous obtenons des équations de trajectoire découlant de l'intégration de la deuxième loi de Newton :

$$\sum_{i \neq j} \overrightarrow{F_i}(\overrightarrow{r_j}) = m_j \overrightarrow{a_j} \tag{2.16}$$

où $\overrightarrow{F_i}$ sont les forces extérieurs exercées sur le système j par la particule i lorsque $i \neq j$, r_j la position de j, m_j la masse et $\overrightarrow{a_j}$ l'accélération du centre d'inertie de la particule j.

La trajectoire globale de la molécule ainsi obtenue donne des informations sur les propriétés dynamiques et statiques du système étudié. Dans l'étude des première étapes d'agrégation des protéines amyloïdes, nous cherchons d'une part à connaître les déplacements des acides aminés d'une protéine, mais aussi les déplacements de chacun des oligomères les uns par rapport aux autres. Ainsi la dynamique moléculaire semble appropriée pour l'étude des déplacements des petits oligomères de la région NAC de l' α -synucléine et la protéine A β_{1-40} sur les échelles de temps qui nous intéressent.

2.2.1 Description de la dynamique moléculaire

Après avoir transformé la structure initiale en format OPEP, on place celle-ci dans une sphère. Cette dernière à une taille différente selon la taille de l'oligomère. Cette sphère impose des conditions aux frontières réfléchissantes : un monomère qui sortant de la sphère sera repoussé vers le centre par l'inversion de la composante axiale de sa vitesse. Une telle approche est approprié ici par qu'il n'y a pas de solvant explicite et que nous nous sommes intéressés aux propriétés d'oligomères dont la taille est bien en dessous de celle de la sphère.

Au début d'une dynamique moléculaire, une protéine est d'abord soumise à une minimisation : cette étape permet de rectifier et relaxer les liens, angles et contacts qui auraient pu être affectés par les contraintes imposées par OPEP lors de la simplification de la structure. La dynamique moléculaire amortie permet de faire des minimisations tout au long de la simulation. La structure minimisée est ensuite thermalisée. Cela se fait en 10 étapes de 10000 pas pendant lesquels la température doit atteindre une valeur cible déterminée par la distribution de vitesse qui suit la distribution de Maxwell-Boltzmann et respecte la condition :

$$\langle v_i \rangle^2 = \frac{k_B T}{m_i}$$
 (2.17)

où v_i est la vitesse moyenne de la particule i, k_B la constante de Boltzmann, T la température et m_i la masse de la particule i.

Cette simulation se fait, comme on l'a dit, en intégrant les équations de Newton. Nous utilisons ici l'algorithme "Verlet Vitesse" [72], qui exige peu de stockage tout en assurant une bonne préservation de l'énergie total du système. En tout temps t lorsqu'on avance dans le temps d'un pas Δt :

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{1}{2}\Delta t^2 \frac{F_i(t)}{m_i}$$
 (2.18)

$$v_i(t + \Delta t) = v_i(t) + \frac{F_i(t) + F_i(t+T)}{2m_i} \Delta t$$
 (2.19)

 $F_i(t)$ et $F_i(t + \Delta t)$ sont respectivement l'ancienne et la nouvelle force qui s'appliquent sur l'élément i suivant le potentiel OPEP. $r_i(t + \Delta t)$ est la nouvelle position après temps $t + \Delta t$ et $r_i(t)$ l'ancienne. De même pour les vitesse $v_i(t + \Delta t)$ et $v_i(t)$.

Afin de faciliter la stabilité de la trajectoire et de permettre de plus grands pas d'intégration, nous avons combiné "Verlet vitesse" à RATTLE[73]. À chaque itération, l'algorhitme applique une correction à chaques paire d'atomes contraints, permettant d'imposer la distance à l'équilibre pour les forces harmoniques. Une contrainte sur les vitesses des atomes d'hydrogène, qui vise à permettre des pas d'intégration plus grands, est aussi imposée et n'affecte pas sensiblement la dynamique du système. D'ailleurs le choix du temps d'intégration est primordial : avec un pas trop court, la simulation ne converge pas et l'espace des phases n'est pas bien exploré. Avec un pas trop grand, le système est instable. Pour s'assurer que la trajectoire soit bien physique et que l'intégration soit stable, il faut prendre un pas de temps plus petit que la fréquence de vibration la plus haute du sytème. Ce qui donne, avec la correction sur l'hydrogène discutée ci-haut, un pas de temps de l'ordre de quelques fs. Le pas de temps de nos simulations est de 1,5 fs et respecte donc l'ordre de grandeur imposé pour avoir un système sur lequel ont peut appliquer les intégrateurs. Le couplage à un bain de Berendsen permet de garder une température constante lors des changements de vitesse, en ajoutant un terme supplémentaire qui contrôle le taux de relaxation du couplage du bain aux équations de mouvement [74]

.

2.2.2 Avantages et limites de la dynamique moléculaire

Cette technique permet d'avoir un suivi temporel de l'évolution d'un système. De plus, elle suit un échantillonage selon Boltzmann. On peut déplacer temporairement vers le haut une barrière de potentiel afin de trouver dans l'espace une meilleur région énergetique. Toutefois, on observer les limites de la dynamique moléculaire lorsque des barrières de l'ordre de quelques k_BT doivent être franchies. Les avantages de la méthode utilisée ici résident aussi dans le choix d'algorithme approprié à l'étude des protéines amyloïdes. Ainsi la méthode d'intégration "Verlet vitesse" a été choisie du fait de sa stabilité et de sa réversibilité au niveau temporel. De plus, le fait qu'elle n'ait besoin que

d'une seule génération de positions, vitesses et forces pour passer au pas suivant permet un gain de mémoire. De son côté, le bain thermique de Berendsen est une technique facile à implémenter et permet de maintenir de façon efficace une température voulue jusqu'a l'équilibre.

Les algorithmes utilisés peuvent aussi avoir des limitations qui viennent nuancer les effets positifs de leur utilisation sur le système. Par exemple, le bain de Berendsen [74] impose des corrections sur la température et limite les fluctuations d'énergie cinétique qui devraient avoir lieu dans un système canonique, correspondant à un système maintenant à température constante. La trajectoire résultante n'est donc pas, formellement, physique. Toutefois, la stabilité de l'algorithme est un avantage important qui l'emporte sur cette limite dont l'effet réel sur des systèmes tels qu'étudiés ici est négligeable. C'est pourquoi le couplage au bain Berendsen est encore beaucoup utilisé dans la communauté, même avec des codes tout-atomes disposant d'algorithmes alternatifs [75–77].

2.2.3 Échange de répliques en dynamiques moléculaires :HT-REMD

L'échange de répliques permet d'atteindre des minima plus bas mais aussi de débloquer les conformations qui n'arrivent pas à franchir certaines barrières énergetiques en dynamique moléculaire simple. Ainsi, le fait que les trajectoires puissent visiter des hautes et des basses température autorise une certaine flexibilité et permet de mettre l'emphase sur les proprietés thermodynamiques du système. Afin de faire l'échange de réplique on utilise la méthode Monte Carlo. En dépit de ses avantages favorisant un meilleur échantillonage, cette approche détruit les trajectoires temporelles. La HT-REMD est un hybride de *Temperature replica exchange* (T-REMD) et *Hamiltonian replica exchange* (H-REMD) [52] qui va chercher les conformations bloquées dans des minima locaux de la fonction d'énergie générées lors du repliement. Pour la T-REMD, plusieurs simulations sont lancées en parallèles à des températures différentes. A intervalles de temps réguliers, tous les 5000 pas, il peut y avoir un échange de température entre deux trajectoires adjacentes thermiquement avec une probabilité décrite par [52] (Fig. 2.4):

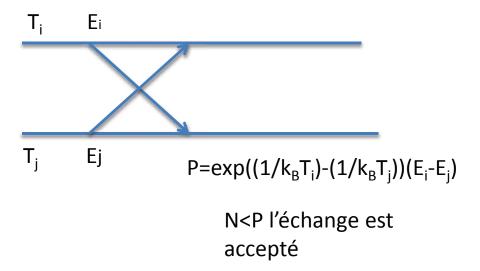


Figure 2.4 – Schématisation de l'échange de répliques en dynamique moléculaire : Afin de déterminer si l'échange aura lieu, on compare les probabilités de Boltzmann d'avoir la configuration i à T_i et la configuration j à T_j avec celle où les températures sont inversées. Si le rapport de la probabilité du deuxième est plus grand que 1, on accepte toujours l'échange. Sinon, on accepte avec une probabilité p, en comparant ce rapport à un nombre aléatoire p entre 0 et 1. Si le rapport est plus petit que p alors l'échange est accepté.

$$p(i,j) = \min\left\{1.0, \exp\left[\left(\frac{1}{k_B T_i} - \frac{1}{k_B T_j}\right) (E_i - E_j)\right]\right\}$$
 (2.20)

Ensuite en restant à la température la plus haute, la H-REMD consiste à réduire progressivement les forces d'attractions non-liées est appliquée. On dénature la protéines afin d'aller explorer des niveaux d'énergie encore plus bas. C'est un peu comme si on déplacait les atomes qui composent la molécule en la dépliant afin de voir si elle adopte des nouvelles conformations lors de son repliement.

Avec les dynamiques moléculaires d'échange de réplique, La distribution et le nombre de températures est très important. Au point de fusion, où l'énergie potentielle varie subitement, il faut que les températures soit rapprochées afin de favoriser les échanges de

réplique et s'assurer que l'échantillonage reste uniforme. De telles précautions, assurent un certain recouvrement des énergies, et empêchent que certaines configurations restent bloquées dans un minimum local.

2.3 Les outils d'analyses

Notre but en faisant l'étude des premières étapes d'agrégation des protéines amyloïdes était de trouver des structures et des régions d'acides aminés qui favorisent leur formation. Ainsi, lors de l'analyse, nous cherchons tout d'abord le domaine de convergence grâce à une analyse thermodynamique du système (PTWHAM) [78]. À la température correspondant aux conditions physiologiques (298 K), nous nous focalisons sur l'observation de la distribution des structures secondaires (STRIDE) [79] par rapport aux résidus et les différents contacts entre acide aminés responsable de leurs cohésions. Enfin les structures en 3D sont rassemblées selon leur ressemblance au niveau structurel sous forme de regroupement qui permettront d'extraire des informations visuelles indiquant quelles sont les structures tertiaires et quaternaires les plus fréquentes lors du rassemblement des oligomères.

2.3.1 Le domaine de convergence

Tout d'abord, nous cherchons le domaine de convergence. Ceci nous permet de réduire l'intervalle d'étude sur une portion de temps ou le système est équilibré. En effet avant la convergence les changements structuraux ne sont pas représentatifs du système. Afin de redéfinir l'intervalle nous observons l'entropie. Cette fonction dépendante de la température traduit le degré d'organisation d'un système. La convergence du système correspond au moment à partir duquel l'entropie sera indépendante du temps. Pour cela on trace différentes courbes de l'entropie en fonction de la température sur différents intervalles de temps que l'on met sur le même graphique. On sait que le système a convergé quand les courbes sont superposables. Afin d'avoir accès aux données thermodynamiques d'un système simulé par dynamique moléculaire d'échange de répliques, nous devons utiliser des méthodes d'analyse statistique. Ces dernières sont indispen-

sables à une analyse de données provenant de systèmes indépendants et certaines ont été adaptées pour des systèmes qui ne sont pas totalement indépendants car ils échangent des configurations comme en T-REMD. Notre choix s'est porté sur l'analyse des histogrammes par assignation de poids pour un système de répliques (PTWHAM). PTW-HAM [78] est une variante de WHAM [80], qui avait été prévue au départ pour combiner les données provenant d'échantillons indépendants formant un ensemble canonique. PTWHAM conserve les hypothèses utilisées dans WHAM mais peut être appliquée sur des ensembles non indépendants comme des séries réalisés par échanges de répliques. Lorsque les simulations sont indépendantes, on peut tracer des histogrammes de densité de probabilité des conformations pour chacune d'elles, dans lesquelles on aura la densité de probabilité en ordonnées et la donnée thermodynamiques par rapport à laquelle ont observe le système (entropie, rayon de gyration, etc) en abscisse. On normalise les densités provenant des différentes simulations, puis on les somme pour construire un autre diagramme qui prend en compte l'ensemble des simulations. Afin de simuler nos systèmes certains biais sont introduits et cette méthode nous donne l'occasion d'enlever la contribution des contraintes dans la répartition des données. Les calculs de WHAM permettent en premier lieu d'avoir l'énergie libre grâce à sa relation avec la fonction de partition : en physique statistique, une fonction de partition réuni toutes les propriétés statistiques d'un système à l'équilibre thermodynamique.

$$A(x) = -k_b T \ln(Z(x)) \tag{2.21}$$

où A est l'énergie libre du sytème non-biaisé, k_b la constante de Boltzman et Z la fonction de partition. Ainsi pour une simulation :

$$A(x) = -k_b T \ln P'(x) - U'(x) + F$$
 (2.22)

où P'(x) est la meilleure estimation de la probabilité de distribution non biaisée, U'(x) l'énergie libre biaisée et F l'énergie libre lorque le biais est enlevé. Concrètement deux équations sont résolues par itération dans WHAM.

$$P(x) = \frac{\sum_{i=1, N sim} n_i(x)}{\sum_{i=1, N sim} N_i(x) exp([F_i - U_{biais, i}(x)]/k_B T)}$$
(2.23)

$$F_i = -k_B T \ln\left(\sum_{xbins} P(x) exp[-U_{biais,i}(x)/k_B T]\right)$$
 (2.24)

N est le nombre de simulations, $n_i(x)$ est le nombre de divisions observe dans l'histogramme associée au paramètre x, $U_{biais,i}$ est le potentiel biaisé, F_i le déplacement en énergie libre pour la simulation i, P(x) la meilleure estimation de la probabilité de distribution sans le biais.

Par la suite on obtient l'entropie et la chaleur spécifique (énergie nécessaire pour faire augmenter la température de 1K par unite massique) grâce aux relations.

$$F = U - TS \tag{2.25}$$

et

$$C_{v} = \frac{1}{m} \frac{\partial U}{\partial T} \tag{2.26}$$

où F est l'énergie libre, U est l'énergie potentielle du sytème, T la température et S l'entropie.

2.3.2 Suivi de l'évolution des structures secondaires

Afin de suivre l'évolution des structures secondaires entre les oligomères, nous avons utilisé STRIDE [79]. C'est un algorithme qui définis les structures secondaires. Il part des coordonnées atomiques et combine l'utilisation de l'énergie des liens hydrogènes aux informations statistiques concernant les angles de torsions. L'énergie des pont hydrogènes E_{hb} est calculée en utilisant des fonctions empiriques provenant d'analyse des données experimentales concernant la géométrie des liens hydrogénes dans la structure cristalline de polypeptides, peptides amino-acides et de petits composés organiques :

$$E_{hb} = E_r \times E_t \times E_p \tag{2.27}$$

 E_r dépend de la distance du lien hydrogène et E_t et E_p sont des propriétés directionnelles. Les détails des paramètres de ces fonctions d'énergie sont donnés par Frishmann et al. Ils ont aussi définis les conditions que doivent remplir les résidus pour être identifiés comme une helice α ou un feuillet β . Ces conditions prennent en compte les critères déterminés pour l'énergie des lien hydrogène mais surtout les probabilités d'un résidu d'être identifié comme une hélice α ou un feuillet β selon ses contraintes sur les angles ϕ et ψ qui correspondent à ces structures sur le diagrammes de Ramachadran.

Grâce a STRIDE, nous pouvons tracer des graphiques sur lesquels apparaissent le pourcentage de structure secondaire en fonction des résidus. Ce graphique donne une moyenne des régions d'acides aminés pour lesquelles on observe des feuillets β et des hélices α . Lorsque l'on met sur un même graphique les courbes des pourcentages en hélices α ou en feuillets β des trois oligomères, on a une bonne idée de comment évoluent ces structures secondaires en moyenne lors des étapes d'oligomérisation.

2.3.3 Les contacts

Pour avoir plus de détails sur les contacts exacts responsables de la stabilisation et de l'apparition des structutures secondaires, on trace des cartes de contacts. On considère qu'il y a des contacts entre deux acides aminés, lorsque les billes représentant les chaînes latérales sont séparées d'une distance qui est plus petite que la somme de leurs rayons de Van Der Walls à laquelle on rajoute 0,5 Å. Les cartes possèdent en abcsisse et en ordonnée les numéros des résidus et le pourcentage avec lequel un contact est présent sur l'ensemble des structures de l'intervalle est donné par des couleurs. Elles vont du noir lorqu'aucune structures ne présentent le contact au blanc lorsque le contact est dans toute les conformations. De plus, pour le dimère et le trimère, on distingue les contacts interchaînes des contacts intrachaînes. Les cartes de contacts interchaînes donnent des informations sur l'organisation des structures secondaires les unes par rapport aux autres et leurs orientation (structures tertiaire). Les cartes de contacts intrachaînes donnent accès à l'organisation tridimensionnelle des chaînes et des structures secondaires qui les composent.

2.3.4 Rassemblement de structures semblables

Le rassemblement de structures semblables ou *clustering* se fait en utilisant l'écart quadratique moyen, RMSD (root-mean-square deviation) et en suivant la procédure de Daura [81] qui consiste à trouver le groupe de structure similaires le plus peuplé avec une RMSD de 2,5 Å et de déplacer les configurations qui répondent à ce critère dans un groupe. Les structures de ce groupe sont séparées du reste des configurations et la même procédure est appliquée de nouveau aux configurations restantes. Il faut faire cela tant qu'il reste des configurations. Les centres ou structures les plus représentatives d'un groupe sont choisis afin d'avoir une idée de l'organisation spatiale des structures secondaires en 3D. Afin de compléter ces informations, des cartes de contacts ont aussi été prises sur ces centres et donnent une idée de la distribution des structures secondaires par rapport aux structures des centres.

2.4 Conclusion

Nous venons de faire une description de la méthode de simulation et des outils d'analyses qui ont été utilisés pour extraire les informations thermodynamiques et structurelles de systèmes complexes comme les protéines amyloïdes. Dans les prochains chapitres, nous utiliserons ces techniques sur le monomère, le dimère et le trimère de la région NAC de l' α -synucléine et aussi sur des oligomères où cette région a été combinée à la protéine $A\beta_{1-40}$.

CHAPITRE 3

EARLY OLIGOMERIZATION STAGES FOR THE NON-AMYLOID COMPONENT OF α -SYNUCLEIN AMYLOID

Cet article soumis au *Journal of Chemical Physics*, présente l'étude numérique des premières étapes d'agrégation de la composante non-amyloïde de α -synucléine. Ce fragment est très intéressant car il est retrouvé dans les fibres amyloïdes associées à la maladie de Parkinson et la maladie d'Alzheimer associées à $A\beta$. Malgrés son implication dans ces deux maladies neurodegénératives, il existe peu d'études concernant ce fragment. Des simulations sur le monomère, le dimère et le trimère nous permettent d'extraire des motifs et régions d'intérêt très impliquées dans la formation de structures secondaires reconnus pour favoriser la fibrillation. Cette étude permet de vérifier que cette région partage bien les caractéristiques observées dans les protéines amyloïdes. La comparaison à d'autres simulations numériques concernant d'autres protéines amyloïdes permet d'extraire les différences qui caractérisent la région de l' α -synucléine que nous avons choisie.

Rozita Laghei et Normand Mousseau ont décidé de la nature et de la taille du système ainsi que de la méthode numérique qui serait utilisée. Rozita Laghei a effectué les simulations des monomères et dimères. Ma contribution a consisté à lancer la simulation du trimère, faire l'extraction et l'analyse des données afin d'écrire cet article révisé par Normand Mousseau.

3.1 Abstract

In recent years, much effort has focused on the early stages of aggregation and the formation of amyloid oligomers. Aggregation processes for these proteins are complex and their non-equilibrium nature makes any experimental study very difficult. Under these conditions, simulations provide a useful alternative for understanding the dynamics of the early stages of oligomerization. Here, we focus on the non-A β amyloid

component (NAC) of the monomer, dimer and trimer of α -synuclein, an important 35-residue sequence involved in the aggregation and fibrillation of this protein associated with Parkinson's disease. Using Hamiltonian and temperature replica exchange molecular dynamics (HT-REMD) simulations combined with the coarse grained OPEP potential, we identify the role of the various regions and the secondary structures for the onset of oligomerization. For this sequence, we clearly observe the passage from α -helix to β -sheet, a characteristic transition of amyloid proteins. More precisely, we find that the NAC monomer is highly structured with two α -helical regions, between residues 2-13 and 19-25. As the dimer and trimer form, β -sheet structures between residue 2-14 et 26-34 appear and rapidly structure the system. The resulting conformations are much more structured than similar dimers and trimers of β -amyloid and amylin proteins and yet display a strong polymorphism at these early stages of aggregation. In addition to its inherent experimental interest, comparison with other sequences show that NAC could be a very useful numerical model for understanding the onset of aggregation.

3.2 Introduction

Amyloid aggregates and insoluble fibrils are associated with many neurodegenerative diseases such as Alzheimers's and Parkinson's diseases. [82, 83] The latter is characterized by the presence of Lewy bodies and neurites composed, in large parts, of the α -synuclein protein. [14] The presence of α -synuclein has also been confirmed for a number of other important neurodegenerative disorders such as Down's syndrome and familial Alzheimer's disease, making this protein a major focus in the study of amyloid diseases. [84–87] α -syn is a 140-residue protein characterized by three major regions: an N-terminal region with KTEKEGV imperfect repeats, a hydrophobic center region named non-A β component (NAC) and a highly negatively charged C-terminal region. [8] The NAC is also the second major component found, after β -amyloid, in the amyloid components associated with Alzheimer's disease(AD). [88]

The polymerization of α -synuclein takes place through conformational changes that bring the protein from disordered conformations to β -sheet structures as it assembles

into toxic oligomers. [9, 10] Yet, little direct information is known about the first steps of oligomerization as its dynamical nature makes it difficult to characterize experimentally. [89, 90] To overcome this limitation, it is useful to turn to computer simulations that can provide detailed information about these small assemblies.

For α -synuclein, most simulations focused on the full-length sequence. A combination of paramagnetic relaxation enhancement NMR spectroscopy and ensemble molecular dynamics (MD) simulations were used to probe the topology of native α -synuclein. [91] More recently, molecular simulations combined with NMR were used to characterize the structure of full-length monomeric α -synuclein at low and neutral pH. It was shown that, while the protein is more structured at low pH, [92] it is far from random coil, even though it is disordered, at neutral pH, in agreement with a Bayesian ensemble analysis.[93] Molecular dynamic was also used to vary the stability and structural properties of several α -synuclein oligomers and mutated sequences. [94–98]

In this article, we are interested in characterizing the first steps of α -synuclein oligomerization. More precisely, we focus on the 35-residue NAC fragment. As mentioned above, this fragment represents the second major component of amyloid deposits associated with Alzheimer's diseases, with a ratio of up to 10 %. [17] The NAC forms α -synuclein's most hydrophobic region. It was early discovered that it self-aggregates into insoluble amyloid β -pleated structures. [18] The NAC also plays a central role in α -synuclein aggregation as is demonstrated by the behavior of β -synuclein. This highly homologous protein to α -synuclein lacks an 11-residue segment located in the NAC domain at position 73-83, [99] which overlaps with the minimal toxic aggregate-prone segment, residues 68-78. [100] It does not aggregate [42] and even inhibits α -synuclein aggregation. [101] NAC's *in vivo* function is still unknown. The neurotoxicity of extracellular NAC is well established, however. [102–104] A recent study suggests that NAC could activate p53, Cdk5 and Bas-dependent apoptotic signaling pathways, causing cell death. [105]

Previous studies tried to identify a signature motif and a particular part of NAC responsible for the fibrillation propensity [42, 106] and the interaction and the conformational state when NAC is linked with a membrane. [107] Others have focused on the

characteristic mutations present in this region. [96, 97, 108] Yet, there is still very little structural information regarding the onset of oligomerization for the NAC peptide with only rare simulations done to serve as a reference model for experiments. [109] The simulations presented here form therefore a first atomistic description of the onset of oligomerization for the NAC peptide.

Here, we use REMD coupled with the OPEP coarse-grained potential [54, 110, 111], to characterize the fragment's structural changes as the system goes from monomer to dimer and trimer. These unbiased simulations allow us to follow aggregation and generate a rich set of structures that provide a first picture as to the assembly mechanism for this sequence. Comparison with dimers of other amyloid peptides simulated in the same conditions also offers an indication of generic and specific properties for early aggregates.

3.3 Methods

In this study, we investigate the structure and the thermodynamics of monomeric, dimeric and trimeric non-A β amyloid component (NAC) of α -synuclein using Hamiltonian-temperature replica exchange molecular dynamics (HT-REMD) coupled with the coarse-grained OPEP potential. [53, 54, 60] The 35 amino acid sequence of the non-A β amyloid component of α -synuclein is given by EQVTNVGGAVVTGVTAVAQKTVEGAGSIAA-ATGFV. The GXXX motif is associated with amyloidogenic propensity in many amyloidogenic proteins where X might be the uncharged residues, GLY, ALA, VAL, ILE, LEU, PHE, TYR, TRP, THR, SER or MET. For the NAC, we find three GXXX motifs 8 GAVV 11 . 13 GVTA 16 and 24 GAGS 27 corresponding to positions 68-71, 73-76, 84-87, respectively, in the α -synuclein protein. [42, 112]

To describe intra and interpeptide interactions, we use the coarse-grained force field OPEP (Optimized Potential for Efficient peptide structure Prediction) version 3.2. OPEP reduces most amino acids to six beads, focusing on the heavy backbone atoms: $C\alpha$, N, H, C, and O and one bead for the side chain except for the proline amino acid which is represented by all heavy atoms. [62]It takes into account the propensities of each

residue to adopt α and β conformations, a crucial aspect for prediction of the aggregation rate of amyloid-forming proteins like synuclein. This potential has been applied with success to study the aggregation of many amyloid peptides such polyglutamine [52], amylin [53, 54] and different segments of $A\beta$. [55–61]

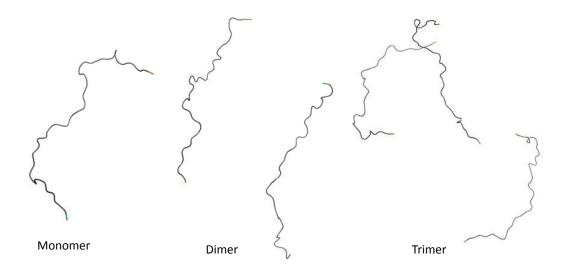


Figure 3.1 – Initial configurations from monomer to trimer. The N-terminal and the C-terminal are shown in green and orange.

The OPEP force field is combined with Hamiltonian-temperature replica exchange molecular dynamics, HT-REMD, a hybrid of temperature replica exchange (T-REMD) and Hamiltonian replica exchange (H-REMD), which was shown to accelerate sampling compared to T-REMD. [53] In the α -synuclein simulation, a set of 35 replicas are launched at logarithmically-distributed temperatures ranging from 240 to 555 K. Five other replicas are run at the highest temperature with progressively reduced non-bonded attractive forces scaled by 0.8, 0.7, 0.6, 0.4 and 0.2 of the initial interaction, respectively. Following the standard REMD procedure, at regular intervals, exchanges are attempted between adjacent replicas with a Metropolis probability. [53, 54]

The simulated temperature is controlled using the Berendsen thermostat. [74] The integration timestep is set at 1.5 fs with an external coupling constant of 100 fs. We use the RATTLE algorithm for covalent constraints [73] and replica exchanges are attempted

every 7.5 ps. The monomer is simulated in a 40 Å-radius sphere and the dimer and trimer are placed in a 140 Å-radius sphere, all with reflecting boundary conditions. The monomer is simulated for 500 ns per replica (total 20 μ s), the dimer for 700 ns (total 28 μ s) and the trimer for 1200 ns (total 48 μ s). The three systems are started from an extended configuration with chains, for the dimer and the trimer, placed at random in the box. Initial configurations, after a first zero-temperature energy minimization, are shown in Fig. 3.1.

The reconstruction of the thermodynamical properties of the NAC monomer, dimer and trimer is done using the replica-exchanged adapted weighted histogram method adapted to Hamiltonian, PTWHAM. [78] Secondary and tertiary structures are predicted using STRIDE, a protein secondary structure assignment based on the combined use of hydrogen bond energy and statistically derived backbone torsional angle informations. [79] Clustering is performed using the $C\alpha$ root-mean square deviations (rmsd) and following Daura's procedure [81]: the largest cluster with a rmsd of 2.5 Å is first identified and its member configurations removed; this procedure is repeated as long as there are remaining configurations.

We verify the convergence of the simulations by ensuring that the entropy as a function of temperature is constant. Entropy is obtained from the first law of thermodynamics, F = E - TS and the free energy calculated by WHAM as implemented by Chodera *et al.* [78] We compute the entropy over non-overlapping time intervals and identify the minimum simulation time over which S(T) is time independent. As shown in Figure 3.2, the monomer reaches equilibrium after 100 ns/replica, the dimer 200 ns/replica and the trimer, 1000 ns/replica. Error bars are estimated with the bootstrapping algorithm. [113]

3.4 Results

We are interested in the characterization of the structural and thermodynamical properties in the early step of aggregation the α -synuclein non-amyloid component (NAC). For this, we follow the evolution of its monomer, dimer and trimer at 300 K. We first present results for each system separately and compare those in the next section.

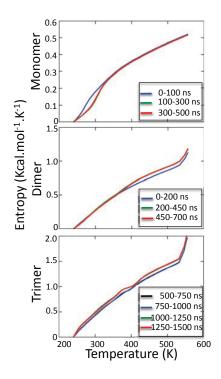


Figure 3.2 – Entropy as a function of temperature taken over various time intervals for the NAC monomer, dimer and trimer. From top to bottom, we see the monomer converged after 100 ns/replica, the dimer after 200 ns/replica and the trimer after 700 ns/replica. Statistical error (not shown for legibility) by bootstrapping is ± 0.01 kcal/(mol*K).

3.4.1 Monomer

We simulated the monomer for 500 ns/replica, collecting statistics over the last 400 ns. Table 3.I gives the average secondary structure propensity over all residues. α -helices dominate, with a propensity of 40%, while β -strands are present at less than 1%. The secondary structure per residue (Fig. 3.3) reveals that α -helices are dominant in two regions: between residues 2-13 and residues 19-25 with a respective probability of \sim 75% and \sim 90%.

As expected with the high proportion of secondary structure, clustering analysis for the monomer (Fig. 3.4) finds relatively well-defined conformations. The first cluster (M1), with a weight of 43 %, shows the two dominant α -helical structures at residues

Secondary structure(%)	Monomer	Dimer	Trimer
α-helix	40	8	7
β -strand	<1	24	23
Turn	39	48	49
Random coil	21	20	21

Tableau 3.I – Percentage of secondary structure taken at 300 K over the full NAC peptide. Results for the monomer, dimer and trimer are averaged over their respective interval of convergence (see Methodology). Statistical error is estimated at $\pm 1\%$.

2-13 and residues 19-25 (Table 3.II) positioned in a perpendicular conformation, with a disordered C-terminal. The same helices (at positions 3-13 and 19-24), in an antiparallel organization, form the second largest clusters, M2, at 16 % probability with the loop at residues 14-18 being present in both clusters. No secondary structure is observed in the third cluster, M3, already much smaller at 8.5 %. Yet, the contact map for M3 overlaps significantly with that of M2 (Fig. 3.6). This suggests that the M3 cluster is simply a slightly destabilized version of M2 and that the NAC monomer, while unstable against any specific native state, samples a relatively small number of well-defined conformations.

3.4.2 Dimer

We simulated the NAC dimer for 700 ns/replica, collecting statistic over the last 500 ns. Table 3.I shows that the presence of a second peptide affects strongly the secondary structure. While the monomer is α -helical at 40% with almost no β -sheet, the dimer is dominated by β -strands, with a propensity of 24%, and only 8% for α -helices. The first β -strand, at residues 2 to 14, overlaps completely with the corresponding α -helix found for the monomer while the second β -strand is found at residues 26 to 34 and does not overlap with the monomer's second α -helix. As shown in Fig. 3.3, maximum stability is observed in the 6-9 (\sim 90%) and 30-33 (\sim 30%) regions while α -helices form mostly between residues 15 and 25, a wider region than for the monomer, with a relatively small probability of 20 to 30%.

The dimer displays a much more diverse set of structures than the monomer. In the

other β						25-29(B)/6-11(A)		13-15(A)/5-7(A)	12-14(B)/27-29(B)		5-8(B)/26-29(C)	6-8(B)/28-30(A)	11-13(C)/32-34(A)	27-30(A)/31-34(C)	9-11(B)/23-25(B)		24-26(A)/29-31(C)		
β4												22-24(B)/31-33(B)						22-24(A)/31-33(A)	31-33(C)/22-24(C)
β3						6-8(A)/6-8(B)	6-8(A)/6-8(B)	4-8(A)/6-10(B)			6-8(C)/6-8(B)	6-8(C)/6-8(A)	2-8(A)/6-12(C)		6-8(A)/6-8(C)	6-9(C)/5-8(B)	6-9(A)/5-8(B)	6-9(C)/5-8(A)	
β2	Monomer			Dimer	30-34(A)/30-34(B)		31-33(A)/31-33(B)			Trimer									
β 1					6-8(B)/31-33(A) 31-33(B)/6-8(A)							30-33(B)/6-9(C)					30-32(B)/7-9(C)	30-33(A)/6-9(C)	6-9(A)/30-33(C)
$\alpha 3$						15-19(A) 14-19(B)	14-19(B)				14-18(C)	15-19(A)	15-19(A)		14-19(B)		15-19(C)		
α 2		19-25	19-24				19-25(A)												
$\alpha 1$		2-13	3-13								2-12(A)								
		MI	M2		DI	D2	D3	D4			TI	T2	T3		T4		T5	9L	

Tableau 3.II – This table give the region implicated in secondary structure for monomer, dimer, trimer see in 3.4 and identified them to 7 regions. The letter A, B and C allow to distinguish the different chains and extract some information about 3D organization in the secondary structure.

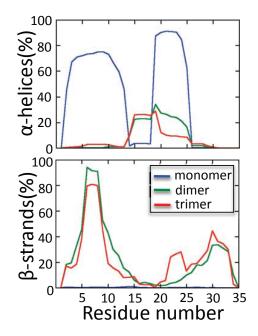


Figure 3.3 – Evolution of α -helix and β -strands as a function of residue number at 300K from monomer to trimer. Statistical error (not shown for legibility) obtained by bootstrapping is estimated at ± 0.5 %.

second row of Fig. 3.4, we show the four clusters present with a probability of more than 5 % (D1 to D4) defined with the 2.5 Å-RMSD cut-off as for the monomer. Together, they represent only 30 % of all sampled conformations. Another 30 % of the conformations are member of clusters present with a probability between 1 and 5 %. While presenting some structure, they typically show less β -sheet secondary structure than clusters D1 to D4.

As can be seen in Fig. 3.4 and Table 3.II, clusters D1 and D4 display purely β -strand conformations. In the first case, the strands are fully interdigitated while D4 shows rather the two NAC peptides side by side. For their part, clusters D2 and D4 show a mixture of α -helical and β -strand structures. Here again, one of the clusters is intertwined (D2), leading to quaternary structure, while the other offers rather the image of two tertiary structures stabilized once against the other (D3).

Two α -helical regions are identified in clusters D2 and D4 : residues 14 (or 15) to

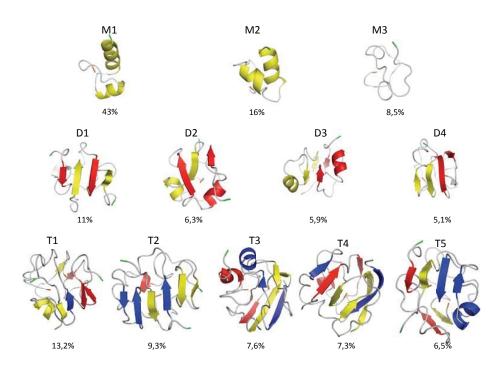


Figure 3.4 – The clusters' center are shown for NAC of monomer (first column), dimer (second column) and trimer (third column). The N-terminal and the C-terminal are shown in green and orange. The first chain is shown in yellow, the second in red and the third in blue.

19 is present on both chains in D2 and on one chain in D3. The second helix, in the latter cluster, is shifted to residues 19 to 25. Antiparallel β -sheets are also well defined: residues 6 to 8 form a sheet on both chains in all four clusters with, in some case, extensions (4-8 and 6-10 in D4 and 6-11 in D2). The region 30-34 also forms β -strands in D1 and D3 while 25-29 and 27-29 are seen in D2 and D4, respectively. D4 also displays a β -strand at 13-15 and 12-14, each on one chain.

Table 3.I gives some additional informations about the region involved in the secondary structure and its spatial organization. The highest contact density region is associated with interchain hydrophobic interactions between the region 6-8 with itself (D2

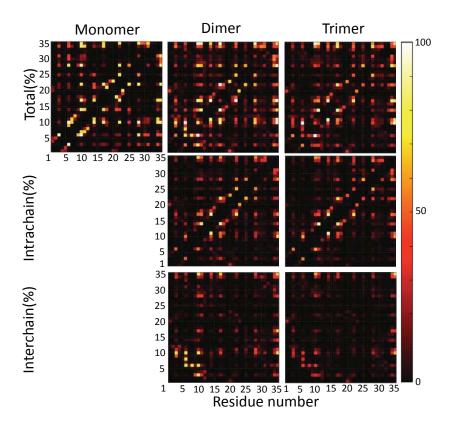


Figure 3.5 – The contact maps display the contact propensity of total (first row), intrachain (second row) and interchain (third row) contacts for NAC of monomer (first column), dimer (second column) and trimer (third column). The total contact propensity is calculated from the intrachain and interchain contacts.

and D3). This region can undergo additions of residue at the beginning or the end as in D4 where regions 4-8 and 6-10 are linked across chains. To a lesser extent, as for D1, 6-8 region may also be linked to the 31-33 region although this region (31-33 or 30-34) prefers to bind with its copy on the second monomer in D1 and D3. Intrachains β -sheets are only observed in D4, where the region 13-15 (12-14) binds to both 5-7 in chain A and 27-29 in chain B.

While interchains and intrachains contacts vary from cluster to cluster, we observe a striking similarity in total map contacts (Fig. 3.7), particularly between D1, D2 and D4. The variations between these intrachain and interchain contact maps reflect a rich-

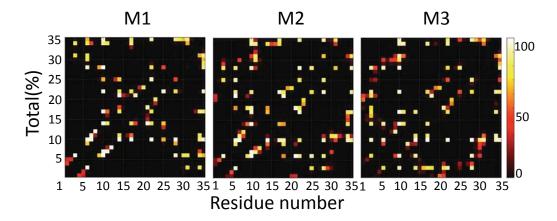


Figure 3.6 – The contact maps display the contact propensity of total (first row) contacts for the dominant morphology of monomer see in the first column of Fig. 3.5.

ness in the organizations at the tertiary and quaternary structure level observed in Table 3.II. Yet, these are constrained by the stability of the secondary structures that enforce a specific set of contacts that can be satisfied either through intra or interchain contacts. This explains the similar total contact map for very different clusters, when projected on the monomeric sequence. In terms of salt-bridges, while the intramolecular bridge at LYS20-GLU23 is present on both chains for all dominant clusters, we also observe a contact between GLU1 and LYS20 on either chain 1 or 2 for all clusters. While these few charged residues play an important role for determining structure of this sequence, their behavior is almost the same for all clusters, however.

3.4.3 Trimer

The NAC trimer is run for 1500 ns/replica, with equilibrium being reached in the last 500 ns. As seen in Table 3.I, the trimer not more structured than the dimer : β -strands are found with an average propensity of 23% and α -helices, with a propensity of 7%. These values are, within the error margin, equivalent to those found for the dimer. However, β -sheets are found mostly in region 2-10, with a peak at 6-8 reaching a \sim 80% probability, and region 22-33, with a two peaks at 22-24 (\sim 25%) and 28-33 (\sim 45%) as

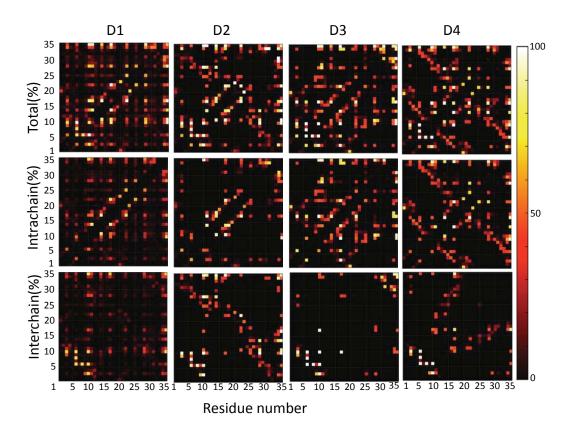


Figure 3.7 – The contact maps display the contact propensity of total (first row), intrachain (second row) and interchain (third row) contacts for the dominant morphology of dimer see in the second column of 3.5. The total contact propensity is calculated from the intrachain and interchain contacts.

shown in Figure 3.3. Interestingly, the first peak in this region is absent in the dimer suggesting that more regions can be stabilized into β -strands as monomers are being added. α -helices, for their part, display a weak but non-zero formation probability for almost the whole sequence with a peak reaching 30 % between residues 15 and 19.

The third row of Fig. 3.4 shows the center of the five clusters, T1 to T5, accounting for at least 5% of the total population. Together, they represent 44% of all sampled structures at 300 K. Like for the dimer, clusters counting at least 1% of all configurations represent 30% of the total set and add diversity mostly in the tertiary structures determined by the β -sheet organization. Globally, as seen in Table 3.II, α -helices spanning

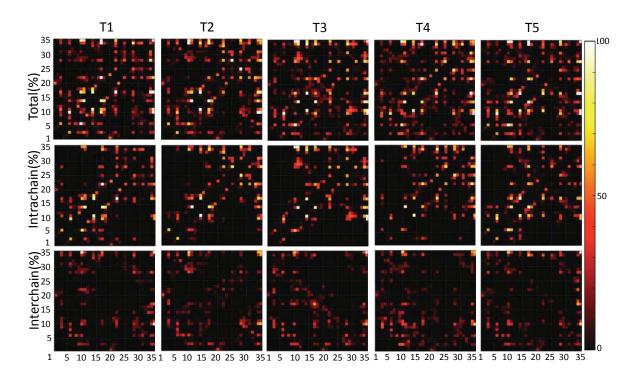


Figure 3.8 – The contact maps display the contact propensity of total (first row), intrachain (second row) and interchain (third row) contacts for the dominant morphology of trimer see in the third column of 3.5. The total contact propensity is calculated from the intrachain and interchain contacts

residues 14-19 are present in T1, T2, T3 and T5. α -helices spanning residue 2-12, also observed in the monomer, appear only in T3.

As in the dimer, we observe that a few β -strand prone regions adopt a rich diversity of contacts for forming exclusively anti-parallel β -sheets. For example, region 6-8 is the region with the highest contact density. In all five dominant clusters (T1 to T5), we observe at least one antiparallel β -sheet formed by the 6-8 region of two chains. The sheets extend to four residues (6-9/6-9) in T5, to seven residues (2-8/6-12) in T3 and nine residues (2-10/4-12) in T4. This region also forms β -sheet with other NAC segments. In one case (T2), we observe a six-strand β -sheet formed by the 6-8, 30-33 and 22-24 regions of two chains. As with the dimer, the region 6-8 also interacts with the 31-33 region (T1,T2, T4 and T5) and the 26-29 region (T5), always in anti-

parallel orientation. We also note that a region of high contact density not observed for the dimer appears at residues 22-24 and binds with the 31-33 region (T2,T4 and T5) into intrachain antiparallel β -sheets. As for the dimer, interchain β -sheets are found in all clusters, showing the importance of these interactions for breaking α -helices.

Figure 3.8 shows the contact maps for the intrachain, interchain and total contacts for all five clusters. While the contact density is smaller than for the dimer, explaining the reduced structural stability observed in the clustering, for example, we can still observe the same interplay between intrachain and interchain contacts that lead to very similar total contact maps for the the six clusters. The role of salt-bridges for the trimer is similar to the dimer: they occur with a high probability both intrachain and interchain, but with no specificity attached to particular clusters, beyond what is described above.

3.5 Discussion

3.5.1 Evolution of the secondary structure

We first look at the structural changes that occur as NAC goes from monomer to dimer to trimer. Previous studies have shown that the formation of β -sheet was an important factor promoting the initiation of nucleation causing fibrillation [38, 114, 115]. Following these studies, we expect that the ability to form β -sheets increases with the number of available chains.

In the case of the NAC fragment, the monomer shows no β -strand at all but two predominantly α -helical regions at residues 2-13 and 19-25 (positions 62-73 and 79-85 with respect to the full α -synuclein). Full-length α -synuclein simulations and experiments indicate that the monomer displays a β -sheet propensity of 11 % and an α -helical propensity of only 2-3 %, with little helices in the NAC region. [93] Yet, an NMR study of an extended NAC fragment (57-102 as compared to 61-95), with both the N-terminal and the C-terminal dominated by charged residues, finds no β -strand and suggests three-helix structures at 58-63, 70-80 and at 88-92, with the latter one, rather unstable. [107] The first helix overlaps with a relatively high α -helical propensity region for the full length α -synuclein, and the third one is in a region that is found to form helices with a

non-zero but low probability (around 5 %). [93] Results on the extended NAC suggests therefore that, for the monomer, β -strands require contacts outside of the NAC region to be stabilized. This is consistent with one of the β -rich conformation shown in Figure 6 of Ullman *et al.* [93]

The first helix observed in our simulation of the NAC monomer is shifted compared to both experiments. Although heteronuclear single quantum coherence measurements would indicate the NAC structure is not affected by those [107, 116], the shift on this first helix is to be expected since the α -helices observed in both experiments are centered on residue 60 and extend only to residues 63 or 64 (residue 3 or 4 of the NAC region). Interestingly, the second helix, while shifted with respect to Bisaglia *et al.* overlaps with a region found to display a low but non-zero helical propensity in full-length α -synuclein. [93] While our structure for the monomer is compatible with available experiments, direct measurement on the structure of the NAC monomer, an important fragment, is needed to confirm these conclusions and provide more detailed information.

We observe a clear evolution from α -helical to β -strand secondary going from monomer to trimer (see Table 3.I). In total, the α -helix propensity goes from 40 to 8%, between the monomer and dimer, and remains almost constant, at 7%, for the trimer, while the β -strand propensity moves from less than 1 % in the monomer to plateau around 24 and 23 % for the dimer and trimer, respectively, a transition characteristic of amyloid aggregation. [117–119] As we see in Fig. 3.2, however, the entropy for the trimer stills shown some small but visible fluctuations when computed over 250 ns time intervals. This suggests that the relevant phase space for this system is significantly larger than for the dimer. Longer simulations, which are beyond our current computational capabilities, could therefore lead to further increase in secondary structure at the trimer level. Already, however, the addition of a third monomer increases the regions where β -sheets is significant.

The first α -helix at residues 2-13, present at 75 % in the monomer, disappears and is replaced instead by a β -strand most stable at residues 6-8 that form exclusively interchain sheets. The second helix, at residues 19-25, more stable in the monomer (90 %

propensity) is shifted mostly to residues 14-19, and is present in most dominant clusters except D1, D4 and T4, with an overall 25 % propensity.

The β -strands are mostly found between residue 6-9 (90%) and 30-33 (30%) in both dimers and trimers although a third important region, between 22-24 is also found in the trimer, extending the β -strand prone regions with respect to the dimer (25 %). Previous studies define five important regions of β -strands appear in oligomerization of α -synuclein, with three in the NAC region corresponding to residues 2-6, 8-17 and 30-35. [35, 114, 115]The last region is in agreement with our findings. The first two are shifted by a few residues. This could be due to the fact that the experimental results are for the full protein and allow for interactions with other segments. [42]

A number of α -synuclein studies emphasize the role of the GXXX motif, generally associated with amyloidogenic propensity in many amyloidogenic proteins. [42, 112] As discussed in the Methods section, the NAC region contains three GXXX motifs at positions 8, 13 and 24. These motifs do not overlap significantly with any of the dominant β -strands observed for the dimer and trimer, in agreement with various experimental observations that find that this motif is not involved in the fibrillation process of α -synuclein. [42, 106]

3.5.2 The polymorphism

Polymorphism appears to be a hallmark for a majority of amyloid proteins and applies to the existence of distinct morphologies with well-defined structures. [120, 121] It can be observed during the formation of amyloid and affect the final form of fibrils. [122] Studies have been conducted to link polymorphism and multiple misfolding pathway [123] and the neurodegenerative diseases in which they appear. [124] Experimentally, polymorphism means that the final structure observed experimentally does not only depend on the sequence but also on the concentration and the preparation details. It reflects the possibility that there are several mechanisms of oligomerization leading to aggregation. [125]

Polymorphism should not be confused with the diversity of partially disordered conformations that are generally observed for the small oligomers of amyloid peptides such as those studied here. Beyond this rich structural diversity, however, polymorphism is clearly observable in the NAC dimer and trimer. Indeed, already at the dimer level, the NAC peptide presents a very defined secondary structure, with a few dominant β -strand and α -helical regions. These regions, however, interact to generate a rich set of tertiary and quaternary structures while maintaining a distribution of interchain and intrachain contacts that project an almost constant overall contact map and salt-bridges (Figs. 3.7 and 3.8). For example, the 30-34 region has great flexibility and can bind to very different regions: 6-8 (D1, T1, T2, T4 and T5), 30-34 (D1 and D3) and 22-24 (T2, T4 and T6). For the monomer, we observe that the relative position of the helices then change it to be composed of the same residues. In the dimer, the antiparallel β -strands can alternate between chains (D1 and D2) or show a strong symmetry (D3 and D4). For the trimer, T2 and T5 have the same regions involved in their secondary structure. However T5 structure use the three chain in interchain contact, while T2 use only two chain for it β -sheets and the third adopt an α -helix conformation.

These results show that polymorphism can appear at the very early stages of aggregation. Even when structures are still partially disordered, we observe that some regions show very stable secondary structure elements that can assemble into a rich set of well-defined tertiary structures. This suggests that the onset of oligomeric growth can already play a defining role in determining the final amyloid structure in agreement with experiments. [120, 121]

3.5.3 Comparison with other amyloid proteins

It is interesting to compare the structural changes taking place from the monomeric to the dimeric state for various amyloid sequences. Oligomers share a number of structural signatures [25] but very little is known, at the moment, about the microscopic similarities and differences in the growth of these sequences. Beyond the fundamental interest, there are also two more pragmatic advantages: (1) comparing the behavior of various peptides simulated with the same conditions, it is possible to separate, at least partially, the technical limitations, associated with the forcefield and the simulation approach, from the sequence-related behavior, and to extract, from there global and sequence-specific pro-

Figure 3.9 – We compare the sequence of NAC of α -synuclein with hIAPP and $A\beta_{1-40}$. We colored the hydrophobic residue in red and add the charge of charged residue which are implicated in the hydrophobic interaction and ionic bond respectively.

perties; (2) the comparison also allows to identify which sequence is more accessible from a numerical point of view, given the computational timescale and size limitations that we are still facing. This is particular useful when using a coarse-grained potential, such as OPEP, which allows better equilibrium on large systems such as those presented here, but at the costs of more uncertainty with respect to the final predictions. In spite of the advantage of comparing results over many sequences, most comparisons have been limited to very close sequences such as $A\beta_{1-40}$ and $A\beta_{1-42}$ [126–128] or mutations. [111]Here, we compare our results on the NAC region of α -synuclein to results obtained to two other amyloid proteins studied with the same potential and computational techniques: $A\beta$ and hIAPP, giving us a first insight on these questions. [52, 54, 60, 111]

 $A\beta$ is a 40 or 42 residue protein and the major component of the fibrils involved in Alzheimer disease. It contains many charged residues, but the structure is globally neutral and has the same number of hydrophobic residues as the NAC region. Pancreaticamyloid deposits in humans consist mainly of β -sheet fibrillar aggregates of the 37-residue polypeptide human islet amyloid polypeptide (hIAPP). This protein is negatively charged and counts fewer hydrophobic residues than the NAC (Fig. 3.9).

REMD-OPEP computational studies of $A\beta_{1-40}$ and $A\beta_{1-42}$ show, in agreement with most other simulations of these systems, that the monomer is mostly disordered with very low α -helical propensity (between 5 and 6 %) and 11 to 12 % β -strand propensity due, in large part, to a relatively stable β -sheet between residues 2-5 and 10-13. With little secondary structure, $A\beta$ monomer samples a much more diverse set of structures than NAC, with its top 3 clusters representing between 28 and 42 % of all visiting conforma-

tions, compared with 68 % for NAC. [60] This changes as a second chain is added: The top four clusters for A β represent 38 (A β_{1-40}) to 53 % (A β_{1-42}) of all conformations while they account only for 28 % of all visited conformations for NAC.

Looking at β -strand propensity, we find that it increases only slightly for $A\beta_{1-40}$, from 11.8 to 12.6 %, and triples for A β_{1-42} , reaching 30.8 % while, for NAC's dimer, it goes from less than 1% to 24 %, a much higher increase. A β dimerization for both sequences is characterized by a global increase in intermolecular hydrophobic contacts, mostly in the hydrophobic core (17-21) and the C-terminal (30-40) regions. Nevertheless, as observed also, for example, by Barz and Urbanc, in MD simulations with OPLS-AA forcefield and explicit solvent, intrachain contacts are more important than interchain contacts and the dimerization can be in large part described as a docking process [111, 127] We observe a similar behavior with hIAPP. The presence of a disulfide bridge at Cys2-Cys7 stabilizes an α -helix at residues 5-19 in the N-terminal, a structure that is maintained during the dimerization (with propensity above 75 %). This helix is combined with a β -strand (17-26 and 30-35) that sees its propensity going from 20 to 40 % as the monomer combines into a dimer, for an overall β propensity going from about 5 % to around 12% [52, 54]. In spite of this co-stabilization due to a higher interchain contact probability for the β -strand regions, the long N-terminal α -helix seems to limit the interdigitation between the two chains. The final dimeric structure is therefore akin to stacked monomers with structural elements stabilized by the presence of a nearby chain. At 300 K, the four dominant clusters account, here, for only 18 % of all structures, in large part because of the high flexibility of α -helical N-terminal, that shows very limited interchain interactions but remains very stable. [54]

This comparison with other sequences simulated with the same potential and method helps identify the specificities of the NAC sequence. While all three proteins do form more β -sheets as they go from monomer to dimer, the transformation is most important for the NAC with the clearest α to β transition. This transformation is accompanied by a much more significant interchain contact map than is observed for $A\beta$ and hIAPP. If, for the latter two sequences, interchain contacts occur at a few points to stabilize secondary structure, NAC shows significant interdigitation, associated with quaternary

structures, that is not compatible with a simple docking of one peptide over another one and an oligomerization process based on well-defined secondary structures than can form a richer set of tertiary structures than the two other sequences as can be seen while maintaining a very stable overall contact map. This suggests that, from a numerical point of view, the onset of aggregation for NAC might be much easier to simulate and characterize than for other sequences, allowing for much better complement between simulation and experiments.

3.6 Conclusion

In this study, simulated the onset of aggregation of the non-amyloid component of the α -synuclein protein (NAC). Our simulations show a clear α to β transition from the NAC monomer to the dimer as well as the presence of well-defined secondary structures, in agreement with the general consensus. Long simulations (up to 1,2 μ s per temperature for the trimer) allow us to reach equilibrium and to identify a rich polymorphism characterized by stable secondary structures that assemble to form a wide set of tertiary organizations with very similar contact maps. Overall, the NAC presents a much clearer aggregation pathway, at the dimer and trimer level, than other amyloid sequences, such as $A\beta$ and hIAPP, that have been much more studied recently. This suggests that more attention, numerically, but also experimentally, should be given to this sequence to develop a first route to aggregation.

3.7 Acknowledgments

This work is supported in part by the Natural Science and Engineering Research Council of Canada and the Canada Research Chair Foundation. This work was made possible through generous allocation of computer time by Calcul Québec/Compute Canada.

CHAPITRE 4

AGGREGATION PROCESS OF $A\beta_{1-40}$ WITH NON- $A\beta$ AMYLOID COMPONENT OF α -SYNUCLEIN

Cet article en préparation et qui sera soumis au Journal of Chemical Physics, présente l'étude numérique d'oligomères combinant les protéines amiloïdes $A\beta_{1-40}$ et la composante non- $A\beta$ amiloïde de α -synucléine que l'on retrouve dans des proportions différentes dans les fibres amiloïdes observées dans la maladie d'Alzheimer. Il s'inscrit dans la continuité de l'article précédent qui faisait l'étude numérique de petits oligomères composés uniquement de la région NAC. Ici, les simulations numériques sont faites sur des oligomères hétérogènes : des dimères et de trimères formés de $A\beta_{1-40}$ et NAC. En premier lieu, un dimère avec $A\beta_{1-40}$ et NAC dépliées. Puis un trimère avec deux $A\beta_{1-40}$ et un NAC afin de respecter le fait que dans les plaques amiloïdes la région NAC est minoritaire. Enfin, un trimère combinant les *clusters* majoritaires d'un dimère de $A\beta_{1-40}$ et d'un monomère de NAC obtenu d'études précédentes faites dans le groupe du professeur Mousseau. Les trimères étant toujours en cours de simulation, seule l'étude du dimère est présentée ici. Cette étude à pour but d'étudier les intéractions qui ont lieu entre les deux protéines et leurs implications dans la stabilisation de structures favorisant la fibrillation.

Normand Mousseau et moi même avons décidé des systèmes à étudier. J'ai fait les simulations, l'extraction et l'analyse de données afin d'écrire l'article qui a été révisé par Normand Mousseau. Ce dernier a aussi effectué la modification de certains outils d'analyse qui n'étaient pas adaptés à des oligomères composés de protéines de tailles différentes.

4.1 Methods

In this study, we investigate the structure and the thermodynamics of α -synuclein non-A β amyloid component coupled with monomeric or dimeric A β_{1-40} using Hamiltonian-

temperature replica exchange molecular dynamics (HT-REMD) coupled with the coarse-grained OPEP potential. [52, 54, 60]. The amino acid sequence of the non-A β amyloid component of α -synuclein and A β_{1-40} is given by EQVTNVGGAVVTGVTAVAQKT-VEGAGSIAAATGFV and DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMV-GGVV respectively.

Intra and interpeptide interactions are described with the coarse-grained force field OPEP (Optimized Potential for Efficient peptide structure Prediction) version 3.2. This potential focuses on the heavy backbone atoms, reducing the amino acid representation to six beads: $C\alpha$, N, H, C, and O and one bead for the side chain except for the proline amino acid which is represented by all heavy atoms [62]. The propensities of each residues to adopt α and β conformations are taken into account and are very important in the observation of secondary structures involved in the aggregation. In fact, this potential was already used successfully, for prediction of the aggregation rate in the amyloid-forming proteins like polyglutamine [52], amylin [53, 54] and different segments of $A\beta$ [55–61].

The Hamiltonian-temperature replica exchange molecular dynamics, HT-REMD, a hybrid of temperature replica exchange (T-REMD) and Hamiltonian replica exchange (H-REMD) is combined with OPEP. In the simulation, 35 replicas are launched at temperatures ranging from 240 to 555 K and distributed on a logarithmic scale. Five additional replicas are run at the highest temperature with progressively reduced non-bonded attractives forces to facilitate chain rearrangements (0.8, 0.7, 0.6, 0.4 and 0.2 of the initial interaction).

An external bath coupling, the Berendsen thermostat, controlled the temperature with an external coupling constant of 100 fs and a integration timestep setting at 1.5 fs [74]. The Rattle algorithm is applied to allow this large timesteps [73]. Replica exchanges are attempted every 7.5 ps. The dimeric system composed of the monomer NAC of α -synuclein and $A\beta_{1-40}$ is simulated for 700 ns per replica (total 28 μ s) in a 40 Å-radius sphere with reflecting boundary conditions.

The replica-exchanged adapted weighted histogram method adapted to Hamiltonian, PTWHAM [78] is used for the reconstruction of the thermodynamical properties of the

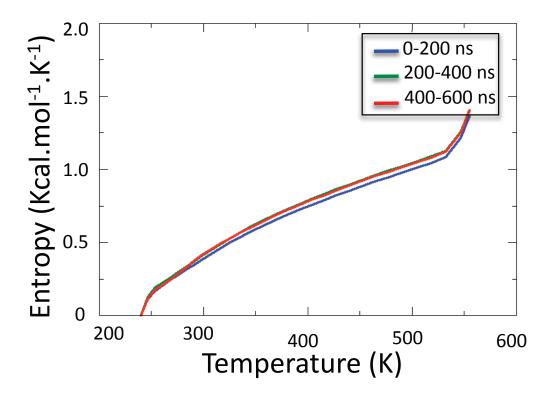


Figure 4.1 – Entropy as a function of temperature taken over three subsequent 200-ns time intervals for the NAC and $A\beta_{1-40}$. Equilibrium is reached after the first 200 ns.

oligomer. STRIDE, a protein secondary structure assignment based on the combined use of hydrogen bond energy and statistically derived backbone tortional angle informations [79], is selected to predict the secondary and tertiary structures. For the clustering we use the $C\alpha$ root-mean square deviations (rmsd) and following Daura's procedure [81]: we first identify the largest cluster with a rmsd of 2.5 Å; all configurations belonging to this cluster and then removed and the procedure is repeated to identify the second and following clusters.

The entropy, obtained from PTWHAM, is used to determine when the system has converged as it is a direct measure of the number of visited configurations. To estimate convergence, we compute the entropy over a sequence of time intervals and consider that

equilibrium is reached when the entropy, computed over the whole temperature range, becomes constant. As seen in Fig. 4.1, the entropy converges after 200 ns/replica for the NAC+A β dimer and it remains constant between the 200-400 ns/replica and 400-600 ns/replica time intervals.

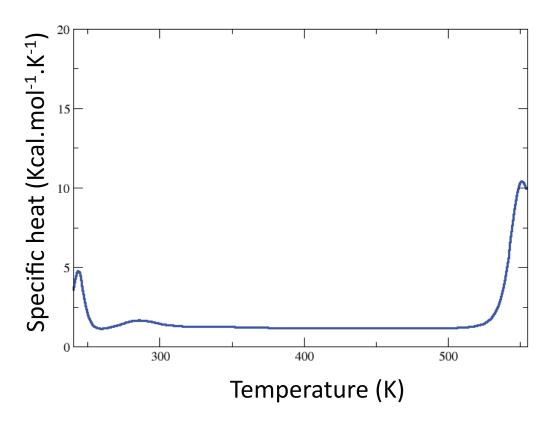


Figure 4.2 – Specific heat as a function of temperature for the NAC and $A\beta_{1-40}$ from 400 to 600 ns.

4.2 Results

In this section, we are interested in the structural and thermodynamic characterization of the dimer compound $A\beta_{1-40}$ with NAC starting from a totally unfolded configuration generated from OPEP File Generator and with the two chains are 15Å apart. Here, we analyze data obtained in the equilibrium phase focusing on the results at 300 K. We chose

this temperature because there are not transition temperature as we can see in Figure 4.2. The starting and ending peaks are not transition but side effects. These results will be compared to the dimers of $A\beta_{1-40}$ and of NAC in the following section.

4.2.1 The dimer of $A\beta_{1-40}$ with non- $A\beta$ amyloid component of α -synuclein

Secondary structure(%)	Total	D1	D2	D3	D4	D5	D6
α-helix	$11(NAC) - < 1(A\beta_{1-40})$	<1	12	3	7	5	<1
β -strand	$9(NAC) - 9(A\beta_{1-40})$	3	6	16	4	13	14
Turn	$56(NAC) - 53(A\beta_{1-40})$	67	59	57	57	49	51
Random coil	$23(NAC) - 38(A\beta_{1-40})$	30	22	25	32	34	35

Tableau 4.I – Percentage of secondary structure taken at 300 K over the total run and the six major clusters (D1 to D6). Results for all the clusters are averaged in the 200-700 ns interval of convergence (see Section 4.1). Statistical error is estimated to less than 1% with the bootstrapping method.

We simulated the dimer for 700 ns per replica and collected the last 500 ns for analysis. In this interval, we find relatively little overall secondary structure: α -helix propensity is about 6 % while β -stand propensity is 50 % higher, at 9% (Table 4.2.1). As shown in Fig. 4.3, α -helices are concentrated in the NAC chain, with two peaks between residues Q2-T12 and G13-A29 with a \sim 22% probability and \sim 30%. We also observe a small helical tendency at residues 56-63 (A21-G28 in A β ₁₋₄₀) with a propensity of \sim 6%.

As could be expected, β -strands are more common and appear with a propensity of at most 25 %, indicating that no region dominates in the formation of secondary structure. In NAC, two regions dominate: residues Q2-T12 and K20-F34 with probabilities of about 25 % and 14 %, respectively. A β shows, for its part, three regions with non-zero probability of forming β -sheets: residues 37-42, 43-53 and 58-75, corresponding to residues A2-D7, S8-H18 and D23-V40 of A β 40, with a maximum propensity per residue of \sim 25%, \sim 17% and 15%, respectively. For the total run, we can see in (Table 4.2.1) the α -helix have a better propensity with \sim 11% to be in NAC than A β 40 with a propensity under 1%. For β -strands, NAC and A β 40 have the same propensity to be in both.

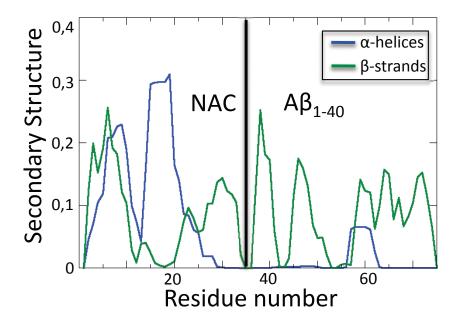


Figure 4.3 – α -helix and β -strands as a function of residue number at 300K for the NAC/A β dimer calculated over the 200-700 ns time interval.

Maps showing total, intrachain and interchain contacts are presented in Fig. 4.4. Here again, we observe a wide distribution of contacts, both interchain and interchain. While most of these contacts occur with a probability below 50 %, a number of long-range interactions are present in a majority of structures. We note, in particular, 10-66 and 17-53 and 35-54.(respectively with \sim 65% for the first and second and \sim 80% for the third interaction.)

The relatively small probability of forming well-defined contacts and secondary structure elements is explained by the diversity of tertiary structures that are visited by dimer. Figure 3.4 shows the six dominant clusters that represent, together, only about 30 % of all configurations, as defined with a 2.5 Å-RMSD cut-off. These six clusters show a very diverse set of secondary and tertiary elements that we describe below in some details.

D1, the first cluster, with an occurrence of 8.7 %, shows no β -strand or α -helical secondary structure. However, it displays many interchain contacts responsible for main-

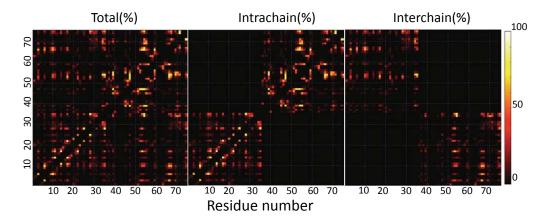


Figure 4.4 – Map of the contact propensity of total, intrachain and interchain contacts for NAC/A β_{1-40} dimer. The NAC runs from residues 1 to 35 and A β from residues 36 to 75. The total contact propensity is calculated from the intrachain and interchain contacts.

taining the tertiary structure in a mostly antiparallel arrangement (see Fig. 4.6). D2, that represents 4.7 % of all visited structures, is more structure. It forms a short anti-parallel interchain β -sheet, composed of residues N5-V6 of NAC and A2-E3 of A β_{1-40} . We also see an α -helix in the region 14-20. Looking at contacts (Fig. 4.6), we observe that the few interchains contacts are dominated, beside the β -sheet, by residues V6 and F39 (F4 of A β) that form contacts, with a high probability, with a number of residues from the other chain.

D3, with probability of 4.6 %, shows a 3-stranded intrachain antiparallel β -sheets at residues E3-R5/E11-H14/V24-D26 in A β_{1-40} , with few interchain contacts. D4, representing 3.9 % of all visited configurations, form also a short antiparallel interchain β -sheet, like D2, but between NAC's Q2-V3 and A β 's D23-V24. It also forms an α -helix in NAC between residues V6 and A9. Looking at contacts (Fig. 4.6),the residues I66 and I67 in A β_{1-40} have a high propensity to form interchain contacts with some regions of NAC: the region V3-V6 and V10-V14 and F34-V35. The residues 66 and 67 are important in the interchain stabilization of the structure.

D5, that occurs 3.5 % of the time, is the only structure with both an intrachain and an interchain β -sheet. The antiparallel interchain β -sheet is found between residues T4-G8 of NAC and K28-I32 of A β . A short intrachain antiparallel sheet is also observed at

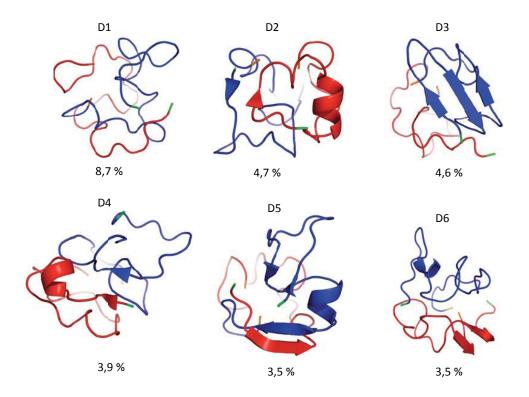


Figure 4.5 – The clusters' center of the dimer NAC with $A\beta_{1-40}$. The N-terminal and the C-terminal are shown in green and orange, respectively. The NAC chain is shown in red, the $A\beta_{1-40}$ in blue.

residues A2-E3/Q16-K17 of A β . D6, finally, with an occurrence probability of 3.5 %, shows, like for D3 and D4, relatively few interchain contacts, with the exception of an antiparallel region between residues V17-V22 of NAC and L17-E22 of A β . D6 is also rich in secondary structure with an intrachain antiparallel β -sheet formed by NAC's residues G7-V10 and A29-T32, as well as a short α -helix at residues E11-H13 of A β .

In terms of salt-bridges, while the intramolecular bridge at K20-E23 is present on NAC for all dominant clusters, we also observe a contact between E11-K16, K16-E22, E22-K28 and D23-K28 of $A\beta_{1-40}$) with a percentage close to 80 in a majority of the six clusters, suggesting that these few charged residues, particularly K16 and E22 (51 and

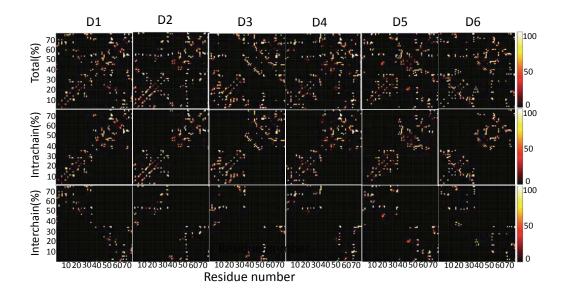


Figure 4.6 – The contact maps display the contact propensity of total, intrachain and interchain contacts for the major clusters of the NAC/A β_{1-40} dimer. The NAC runs from residues 1 to 35 and A β from residues 36 to 75. The total contact propensity is calculated from the intrachain and interchain contacts.

57 in the total sequence) play an important role for stabilize intrachain structures. There are no interchain salt-bridges.

4.3 Discussion

As discussed above, the NAC is present at about 10 % in amyloid fibrils associated with Alzheimer's disease. [17] Its role remains to be understood and it is not clear whether it is simply a contaminant of whether it impacts the fibril growth. This impact could be multiple. It could help stabilize the $A\beta$ monomers or oligomers as they grow into fibril or play a more active role into the oligomeric structures.

There exists very few experimental papers that attempt to characterize the interaction between the NAC and $A\beta$. Indeed, amyloid oligomers are particularly difficult to control and, even for simply monospecies oligomers, very little precise structural information exists on the structure of the dimer or small assemblies.

In in vitro growth studies, Han et al. explored the nucleation dependent polymeriza-

tion mechanism between NAC and A β_{1-40} . [129] They show, first, that the NAC peptide does aggregate into amyloid fibrils. Han et al., moreover, demonstrate that NAC fibrils can see A β fibrils' formation and *vice-versa*, that A β fibrils can seen NAC amyloid aggregation. While this heterogeneous seeding is possible, it remains less efficient than homogeneous seeding as determined by the lag time before fibril growth takes place, suggesting that the structural match between these two sequences is less amyloid prone that with a pure system. [129] While this work supports a positive interaction between NAC and A β leading to amyloid formation, it is not clear whether or not the two chains can interact at the monomeric level or only through preformed fibril structures. Indeed, Han et al. find no evidence in increase in the formation of insoluble aggregates in mixed NAC and A β solutions compared with pure ones. In some cases, the rate is even slowed down, although this result says nothing about the formation of soluble oligomers. This question is partially answered by Paik et al. in an experiment on the self-oligomerization of the full-length α -synuclein with A β_{25-35} . Paik and collaborators confirmed that the 25-35 region of the A β is sufficient for promoting α -synuclein aggregation. They also show that, although A β_{25-35} could interact with other regions of α -synuclein, the NAC region seems to be a preferred binding site. Repeating their experiments with a reverse 35-25 sequence, Paik et al. show that the orientation was essential for promoting oligomerization [89].

Our simulations, that show binding of both peptides, are incompatible with these observations. For example, Fig. 4.4 shows that contacts between A β and NAC are dominated by residues 17 to 53 and 35 to 54, showing that, indeed, A β 's N-terminal is relatively inactive here. Moreover, interchain β -sheets involve residues in this region: D3(24-26), D4(23-24), D5(28-32) and D6(26-32). These results confirm the interaction and the costabilization of secondary structure between region 25-35 of A β and NAC.

Comparison between our results and the formation of pure NAC and $A\beta$ dimers allow us to link our simulations to Han *et al.* results. Côté *et al.* studied the assembly of $A\beta_{1-40}$, $A\beta_{1-40}$ (D23N) and $A\beta_{1-42}$ dimers using the same OPEP and HT-REMD methodology, allowing a direct comparison with results presented here. [111] They show that while β -strand propensity increase comparing with the monomer, particularly at the

fibril-loop region(22-28) and C-terminal(29-40), the overall change in secondary structure remains small for $A\beta_{1-40}$: α -helical propensity goes from 6 to 1.3 % while β -strand propensity moves form 12 to 13 %, similar to what is observed here. For the NAC peptide, however, the difference is more important: while the β -sheet content reaches 24 % for the NAC dimer, it is only 9 % here.

Moreover, if the propensity is lower, β -sheets observed in the heterogeneous dimer are similar to those observed in the homogenous dimers, with a relatively high β -sheet propensity in the N-terminal for A β and at the N-terminal and C-terminal for NAC, even though some of these β -sheets are formed by the two chains. We also observe an important presence for the salt-bridge between D23 and K28 of A β ₁₋₄₀, and K20-E23 is present on NAC, in agreement with Côté *et al.* results. [60, 111] and our study of NAC dimer.

4.4 Conclusion

In this study we simulated the dimer of the non-amyloid component of α -synuclein (NAC) with $A\beta_{1-40}$. We are interested in this system because of the presence of NAC in a minority of the amyloid fibrils associated with Alzheimer's disease. Our simulations confirm experimental observations by Paik *et al.*; the 25-35 region is a significant area of secondary structure stabilization with NAC. When compared to NAC and $A\beta_{1-40}$ dimers using the same simulation procedure and analysis, we find that the propensity to form beta-sheet remains the same for $A\beta_{1-40}$ when there was a significant decrease for NAC. The ionic interactions between salt-bridge D23 and K28 in $A\beta_{1-40}$ and K20 and E23 in NAC of the heterogeneous dimer are consistent with the salt-bridges found in NAC and $A\beta_{1-40}$ homogenous dimers and allow to see that these interactions don't contribute in the interchain stabilizations. We have shown by our numerical simulation that there are many interactions between the two monomers of NAC and $A\beta_{1-40}$. Continuing our simulations on heterogeneous trimers we can propose some mechanisms which should be confirmed by experimental studies.

4.5 Acknowledgments

This work is supported in part by the Natural Science and Engineering Research Council of Canada and the Canada Research Chair Foundation. This work was made possible through generous allocation of computer time by Calcul Québec/Compute Canada.

CHAPITRE 5

CONCLUSION

Au début de ce mémoire, nous voulions identifier des structures primaires ou sequences d'intérêt de NAC qui favorisent la formation de structures secondaires caractéristiques de la fibrillation. Nous nous demandions aussi si ces structures possédaient des particularités au niveau de leurs structures secondaire et tertiaire. Enfin nous voulions savoir si les résultats observés à partir de l'étude de NAC étaient extrapolable à d'autre protéines amyloïdes.

Tout d'abord, nos résultats ont permis de reconfirmer des propriétés partagées par les protéines amyloïdes : la transition de l'hélice α au feuillet β , une proportion grandissante à former des feuillet β anti-parrallèles, en conformité avec les conformations retrouvées dans les structures cross- β . On retrouve aussi un certain polymorphisme. Ces quelques observations, et la rapidité particulière avec laquelle des feuillets β anti-parrallèles se forment, confirme le choix des séquences de protéines amyloïdes comme point de départ de simulations ayant pour but la collecte d'information sur les structures observées lors des premières étapes d'oligomérisation de protéines amyloïdes.

Lors du suivi de la réaction d'oligomérisation allant du monomère au trimère de NAC , les régions 6-9 et 30-33 se sont demarquées . Dans la majorités des résultats obtenus, ces régions participent aux contacts des feuillet β antiparralèles caractéristiques des fibres amyloïdes. De plus, la grande diversité observée aux niveaux tertiaire et quaternaire pour des structures qui impliquent toujours ces deux régions en font des zones d'intérêt à explorer plus en détails. Toutefois, ces régions ne sont pas retrouvées dans les structures secondaires des oligomères hétérogènes de NAC avec $A\beta_{1-40}$. Une telle observation va dans le sens des travaux de Han et al [129], qui font état de deux mécanismes qui ne font pas intervenir les même regions de $A\beta_{1-40}$ lorsque NAC est présent. En effet, après observations expériementales, ils constatent que lors de la formation de dimères homogènes et hétérogènes, la durée du mécanisme hétérogène est plus longue. Ils attribuent cela à une plus grande affinité entre deux séquences identiques. Notre inté-

rêt s'est aussi porté sur les ponts salins formés entre résidus de charges opposées. On a pu observer que même si ces interactions aident à la stabilisation intermoléculaire responsable de conformations adoptés aux niveau tertiaires des structures, elles sont rarement à l'origine d'interactions reliant les chaînes dans la structure quaternaire.

La particularité des systèmes choisis pouvant avoir comme noyau de nucléation tantôt A β_{1-40} ou NAC permet d'extraire certaines informations sur le mécanisme de nucléation et prend un intérêt particulier lorqu'on sait que des oligomères aussi petits que les dimères sont retrouvés dans les fibres amyloïdes de certaines protéines amyloïdes [46]. Ce genre d'étude s'inscrit aussi dans la recherche des espèces toxiques, depuis qu'on sait que les petits oligomères seraient responsables de la toxicité en formant des pores dans les membranes cellulaires. [33–35]. Dans un tel contexte, les fibres semblent avoir un rôle protecteur, empêchant les petits oligomères solubles et toxiques de se propager dans la cellule[13, 31]. Dans nos simulations on a observe dans le cas de l'oligomérisation hétérogène, on peut avoir deux types de structures. 1)Une des protéines a un rôle de soutien, elle ne forme pas de structures secondaires, mais vient consolider par certaines interactions les structures secondaires adoptées. 2)on observe aussi des structures tertiaires dans lesquelles des feuillets β sont formés en utilisant des résidus de chacunes des protéines. Des études d'oligomères plus gros permettraient de savoir si un de ces deux chemins est privilégié ou s'ils coexistent dans certaines proportions durant l'oligomérisation.

On ne peut savoir clairement si les résultats obtenus ici sont généralisables sur toutes les protéines amyloïdes. Cependant, les informations acquises sur certaines régions et l'importance de l'étude des interactions hydrophobes et ioniques dans la prédiction de régions susceptibles de former des feuillets β restent centrales.

La modélisation OPEP a déjà fait ses preuves sur de nombreuses protéines amyloïdes donnant des résultats relativement proches de ce qui est observé expérimentalement. Cependant, la modélisation du solvant de façon implicite et la représentation des chaînes latérales sous forme de billes limitent et nuancent les observations que nous pouvons faire. En effet, un telle modélisation surestime certaines interactions qui pourraient être moins fortes si on prenait en compte tous les paramètres des chaînes latérales et du

solvant. De plus, si nous voulions simuler des oligomères plus gros, afin d'observer l'interdigitation des chaînes latérales qui est importante pour la stabilisation de la structure cross- β [5, 27–30], nous serions limités par l'absence des détails des chaîne latérale. Il faudrait alors utiliser un potentiel tout-atome afin de ne plus avoir cette limitation. La charge et l'hydrophobicité des résidus sont des paramètres qui peuvent réagir en fonction du solvant dans lequel les protéines sont simulées. Étudier l'effet de l'eau sur certains résidus déjà identifiés, grâce à un solvant explicite, pourrait être un bon complément à nos études.

La meilleure façon d'avoir des précisions par rapport à certaines surestimations d'interaction par le potentiel pourrait être de comparer nos résultats à des études expérimentales. Toutefois, cela reste encore difficile pour des protéines comme NAC mais est pour $A\beta_{1-40}$ dont l'agrégation est plus lente et moins toxique que celle d' $A\beta_{1-42}$ et peut donc être suivie expérimentalement. Il faudrait aussi reproduire nos simulations afin de savoir quelle est la précisions de la méthode numérique utilisée afin d'avoir un paramètre de comparaison fiable face à d'autres études numériques et expérimentales.

En dépit de ces limitations, la méthode de simulation permet d'avoir des informations détaillées avec une bonne résolution sur des systèmes qui peuvent être désordonnés. Même si les études numériques doivent être confirmées par des mesures expérimentales, elles restent une bonne alternative lorsque cela n'est pas possible ou difficile de faire un suivi *in vitro* à cause de la nature hors-équilibre et de la toxicité de ces systèmes. Les méthodes numériques fournissent des informations difficilement observables à de telles résolutions au niveau expérimental. Cependant, la vue plus globale apportée par l'expérimentation permet d'avoir des informations complémentaires et plus générales qui caractérisent les protéine amyloïdes.

Les observations d'un système mettant en contact deux protéines amyloïdes différentes a permis d'explorer les particularités de telles interactions. Nous pouvons nous demander si nous observerions les mêmes conclusions en prenant la protéine α -synucléine au complet. Sachant que d'autres combinaisons de protéines sont retrouvées dans les fibres de maladie amyloïdes, l'etude d'oligomères hétérogenes avec NAC ou $A\beta_{1-40}$ et la protéine à prions PrP impliquée dans la maladie de Creutzfeld-Jacob serait intéres-

santes. Comme mentionné précédemment, des collaborations avec des groupes expérimentaux apporteraient une autre dimension aux résultats obtenus numériquement. Un autre axe de recherche intéressant consiste a changer l'environnement des protéines afin d'observer les modifications structurales qui en découlent. Par exemple, une variation dans le pH agit sur les charges des résidus composant les protéines. Des hypothèses concernant la formation de pores ouvrent aussi la porte à des simulations qui prendraient en compte des interactions avec la membrane et tenteraient de déterminer quel genre d'oligomère va se fixer sur les membranes et par quel mécanisme.

BIBLIOGRAPHIE

- [1] C. M. Dobson. *Nature*, 426:884–890, 2003.
- [2] C. Soto. Nature Reviews Neuroscience, 4:49-60, 2003.
- [3] F. Chiti and C. M. Dobson. Annual Review of Biochemistry, 75:333–366, 2006.
- [4] T. Härd. The Journal of Physical Chemistry Letters, 5:607–614, 2014.
- [5] D. A. Kirschner, C. Abraham, and D. J. Selkoe. *Proceedings of the National Academy of Sciences of the United States of America*, 83:503–507, 1986.
- [6] N. M. Green, N. G. Wrigley, W. C. Russell, S. R. Martin, and A. D. Mclachlant. *The EMBO journal*, 2:1357–1365, 1983.
- [7] A. W. P. Fitzpatrick, G. T. Debelouchina, M. J. Bayro, D. K. Clare, and M. A. Caporini. *Proceedings of the National Academy of Sciences*, 110:5468–73, 2013.
- [8] E. A Waxman and B. I. Giasson. *Biochimica et Biophysica Acta*, 1792:616–624, 2009.
- [9] A. M. Bodles, D. J. S. Guthrie, B. Greer, and G. B. Irvine. *Journal of Neurochemistry*, 78:384–395, 2001.
- [10] J. P. Taylor, J. Hardy, and K. H. Fischbeck. *Science (New York, N.Y.)*, 296:1991–1995, 2002.
- [11] F. Muñoz Lobato, M. J. Rodríguez-Palero, F. J. Naranjo-Galindo, F. Shephard, C. J. Gaffney, N. J. Szewczyk, S. Hamamichi, K. A Caldwell, G. A Caldwell, C. D Link, and A. Miranda-Vizuete. *Antioxidants & redox signaling*, 20:217–35, 2014.
- [12] S. Baglioni, F. Casamenti, M. Bucciantini, L. M Luheshi, N. Taddei, F. Chiti, C. M. Dobson, and M. Stefani. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26:8160–8167, 2006.

- [13] C. Haass and D. J. Selkoe. *Nature Reviews Molecular Cell Biology*, 8:101–112, 2007.
- [14] G. M. Spillantini, M. L. Schmidt, V. M.-Y. Lee, J. Q. Trojanowski, R. Jakes, and M. Goedert. *Nature*, 388:839–840, 1997.
- [15] K. Wakabayashi, K. Matsumoto, K. Takayama, M. Yoshimoto, and H. Takahashi. Neuroscience Letters, 239:45–48, 1997.
- [16] G. Forloni, I. Bertani, A. M. Calella, F. Thaler, and R. Invernizzi. Annals of Neurology, 47:632–640, 2000.
- [17] K. Uéda, H. Fukushima, E. Masliah, Y. Xia, A. Iwai, M. Yoshimoto, D. A. Otero, J. Kondo, Y. Ihara, and T. Saitoh. *Proceedings of the National Academy of Sciences*, 90:11282–11286, 1993.
- [18] A. Iwai, M. Yoshimoto, E. Masliah, and T. Saitoh. *Biochemistry*, pages 10139–10145, 1995.
- [19] B. Urbanc, L. Cruz, F. Ding, D. Sammond, S. Khare, S. V. Buldyrev, H. E. Stanley, and N. V. Dokholyan. *Biophysical Journal*, 87:2310–2321, 2004.
- [20] X. Zhu, R. P. Bora, A. Barman, R. Singh, and R. Prabhakar. *The Journal of Physical Chemistry*. B, 116:4405–4416, 2012.
- [21] N. G Sgourakis, M. Merced-Serrano, C. Boutsidis, P. Drineas, Z. Du, C. Wang, and A. E. Garcia. *Journal of Molecular Biology*, 405:570–83, 2011.
- [22] D. J. Rosenman, C. R. Connors, W. Chen, C. Wang, and A. E. García. *Journal of Molecular Biology*, 425:3338–3359, 2013.
- [23] J. A. Lemkul and D. R. Bevan. *Biochemistry*, 52:4971–4980, 2013.
- [24] R. Horton, A.L. Moran, K.G. Scrimgeour, M.D. Perry, and J.D. Rawn. *Principles of biochemistry*. 2006.

- [25] R. Kayed, E. Head, J. L. Thompson, T. M. McIntire, S. C. Milton, C. W. Cotman, and C. G. Glabe. *Science (New York, N.Y.)*, 300:486–489, 2003.
- [26] D. J. Selkoe. *Nature cell biology*, 6:1054–1061, 2004.
- [27] M. Sunde, L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys, and C. C. Blake. Journal of Molecular Biology, 273:729–739, 1997.
- [28] J. C. Stroud, C. Liu, P. K. Teng, and D. Eisenberg. *Proceedings of the National Academy of Sciences of the United States of America*, 109:7717–7722, 2012.
- [29] S. K. Maji, L. Wang, J. Greenwald, and R. Riek. Structure-activity relationship of amyloid fibrils. *FEBS letters*, 583:2610–2617, 2009.
- [30] O. S. Makin and L. C. Serpell. *The FEBS journal*, 272:5950–5961, 2005.
- [31] M. Fändrich. Journal of Molecular Biology, 421:427–440, 2012.
- [32] R. Jansen, W. Dzwolak, and R. Winter. *Biophysical Journal*, 88:1344–1353, 2005.
- [33] H. A. Lashuel and P. T. Lansbury. *Quarterly reviews of biophysics*, 39:167–201, 2006.
- [34] S. M. Butterfield and H. A. Lashuel. *Angewandte Chemie (International ed. in English)*, 49:5628–54, 2010.
- [35] M. S. Celej, E. Sarroukh, R.and Goormaghtigh, G. D Fidelio, J.-M. Ruysschaert, and V. Raussens. *The Biochemical Journal*, 443:719–726, 2012.
- [36] R. Brasseur, L. Lins, B. Vanloo, J. M. Ruysschaert, and M. Rosseneu. *Proteins*, 13:246–257, 1992.
- [37] D. F. Clayton and J. M. George. *Trends in neurosciences*, 21:249–2–54, 1998.
- [38] M. Sandal, F. Valle, I. Tessari, S. Mammi, E. Bergantino, F. Musiani, M. Brucale, L. Bubacco, and B. Samorì. *PLoS biology*, 6(1):99–108, 2008.

- [39] S. Kumar, A. Sarkar, and D.i Sundar. *Biochemical and biophysical research communications*, 387(2):305–9, 2009.
- [40] J.-C. Rochet, K. A. Conway, and P. T. Lansbury. *Biochemistry*, 39:10619–10626, 2000.
- [41] K. A. Conway, J. D. Harper, and P. T. Lansbury. *Nature medicine*, 4:1318–1320, 1998.
- [42] V.N. Uversky and A.L. Fink. FEBS letters, 522(1-3):9–13, 2002.
- [43] M. J. Thompson, S. A. Sievers, J. Karanicolas, M. I Ivanova, D. Baker, and D. Eisenberg. Proceedings of the National Academy of Sciences of the United States of America, 103:4074–4078, 2006.
- [44] B. I. Giasson, I. V. Murray, J. Q. Trojanowski, and V. M. Lee. *The Journal of biological chemistry*, 276:2380–2386, 2001.
- [45] N. D. Lazo, M. A. Grant, M. C. Condron, A. C. Rigby, and D. B. Teplow. *Protein Science*, 14:1581–1596, 2005.
- [46] G. M. Shankar, S. Li, T. H. Mehta, A Garcia-Munoz, N. E. Shepardson, I. Smith, F. M. Brett, M. A. Farrell, M. J. Rowan, C. A. Lemere, C. M. Regan, D. M. Walsh, B. L. Sabatini, and D. J Selkoe. *Nat. Med.*, 14(8):837–842, 2008.
- [47] K. Ono, M. M. Condron, and D. B. Teplow. J. Biol. Chem., 285(30):23186–23197, 2010.
- [48] P. Derreumaux. The Journal of chemical physics, 111:2301, 1999.
- [49] P. Derreumaux. Physical Review Letters, 85:206–209, 2000.
- [50] R. Malek and N. Mousseau. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 62:7723–7728, 2000.
- [51] G. Wei, N. Mousseau, and P. Derreumaux. *Biophysical journal*, 87:3648–3656, 2004.

- [52] R. Laghaei, N. Mousseau, and G. Wei. J. Phys. Chem. B, 114:7071–7077, 2010.
- [53] R. Laghaei and N. Mousseau. J. Chem. Phys., 132(16):165102, 2010.
- [54] R. Laghaei, N. Mousseau, and G. Wei. J. Phys. Chem. B, 115:3146–3154, 2011.
- [55] W. Chen, N. Mousseau, and P. Derreumaux. J. Chem. Phys., 125(8):84911, 2006.
- [56] G. Wei, N. Mousseau, and P. Derreumaux. *Prion*, 1(1):3–8, 2007.
- [57] X. Dong, W. Chen, N. Mousseau, and P. Derreumaux. *J. Chem. Phys.*, 128(12):125108, 2008.
- [58] Y. Chebaro, N. Mousseau, and P. Derreumaux. J. Phys. Chem. B, 113(21):7668–7675, 2009.
- [59] A. Melquiond, X. Dong, N. Mousseau, and P. Derreumaux. Curr. Alzheimer Res., 5:244–250, 2008.
- [60] S. Côté, P. Derreumaux, and N. Mousseau. J. Chem. Theory Comput., 7:2584–2592, 2011.
- [61] Y. Lu, G. Wei, and P. Derreumaux. J. Phys. Chem. B, 115:1282–1288, 2011.
- [62] J. Maupetit, P. Tuffery, and P. Derreumaux. Proteins: Struct., Funct., Bioinf., 69:394–408, 2007.
- [63] Philippe Derreumaux and Normand Mousseau. *The Journal of chemical physics*, 126:025101, 2007.
- [64] Norman L. Allinger. Journal of the American Chemical Society, 99:8127–8134, 1977.
- [65] N. L. Allinger, Y. H. Yuh, and J.-H. Lii. CHEMICAL. Journal of the American Chemical Society, 11(23):8551–8556, 1989.
- [66] N. L. Allinger, K. Chen, and J.-H. Lii. *Journal of computational chemistry*, 17(1):642–668, 1996.

- [67] W. R. P. Scott, P. H. Hu, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kru, and W. F. V. Gunsteren. *Journal of physical Chemistry* A, 103:3596–3607, 1999.
- [68] W. L. Jorgensen and J. Tirado-rives. *Journal of the American Chemical Society*, 110:1657–1666, 1988.
- [69] B. R. Brooks, R.E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. *Journal of Computational Chemistry*, 4:187–217, 1983.
- [70] P. K. Weiner and P. A. Kollman. *Journal of Computational Chemistry*, 2:287–303, 1981.
- [71] Thomas A Halgren. Merck Molecular Force Field. *Journal of computational chemistry*, 17:490–519, 1996.
- [72] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. *The Journal of Chemical Physics*, 76:637–649, 1982.
- [73] H.C. Andersen. J. Comput. Phys., 52(1):24–34, 1983.
- [74] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, and J.R. Haak. J. Chem. Phys., 81(8):3684–3690, 1984.
- [75] X. Ji, Y. Zheng, W. Wang, J. Sheng, J. Hao, and M. Sun. *Journal of molecular graphics & modelling*, 46:125–131, 2013.
- [76] N. K. Richards-Henderson, K. M. Callahan, P. Nissenson, N. Nishino, D. J. To-bias, and B. J. Finlayson-Pitts. *Physical chemistry chemical physics : PCCP*, 15:17636–46, 2013.
- [77] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E Mark, and Wilfred F. van G. *European biophysics journal*: *EBJ*, 40:843–56, 2011.
- [78] J.D. Chodera, W.C. Swope, J.W. Pitera, C. Seok, and K.A. Dill. J. Chem. Theory Comput., 3(1):26–41, 2007.

- [79] D. Frishman and P. Argos. *Proteins : Struct., Funct., Genet.*, 23:566–579, 1995.
- [80] S. Kumar, D. Bouzida, H. R. Swendsen, P. A. Kollman, and J. M. Rosenbergl. *Journal of computational chemistry*, 13(8):1011–1021, 1992.
- [81] X. Daura, R. Suter, and van Gunsteren W.F. J. Chem. Phys., 110 :3049–3055, 1999.
- [82] C.A. Ross and M.A. Poirier. *Nature medicine*, 10 Suppl(July):S10–7, 2004.
- [83] C.A. Ross and M.A. Poirier. *Nature reviews. Molecular cell biology*, 6(11):891–8, 2005.
- [84] C.F. Lippa, H. Fujiwara, D.M.A. Mann, B. Giasson, M. Baba, M.L. Schmidt, L.E. Nee, B. O'Connell, D.A. Pollen, P. St. George-Hyslop, B. Ghetti, D. Nochlin, T.D. Bird, N.J. Cairns, V.M.-Y. Lee, T. Iwatsubo, and J.Q. Trojanowski. *The American Journal of Pathology*, 153(5):1365–1370, 1998.
- [85] C.F. Lippa, M.L. Schmidt, V.M. Lee, and J.Q. Trojanowski. Ann. Neurol., 45:353–357, 1999.
- [86] R. Resende, S.C.F. Marques, E. Ferreiro, I. Simões, C.R. Oliveira, and C.M.F. Pereira. *Neurochemical research*, 38(4):797–806, 2013.
- [87] N. Ramakrishna, H.C. Meeker, S. Patel, T.W. Brown, and A. El Idrissi. *Journal of neuroscience research*, 90(8):1589–96, 2012.
- [88] J.E. Duda, V.M. Lee, and J.Q. Trojanowski. *Journal of neuroscience research*, 127(February):121–127, 2000.
- [89] R.P. Seung, J.-H. Lee, D.-H. Kim, C.-S. Chang, and Y.-S. Kim. FEBS 19685, 421:73–76, 1998.
- [90] D.-P. Hong, A.L. Fink, and V.N. Uversky. *Journal of molecular biology*, 383(1):214–223, 2008.

- [91] M.M. Dedmon, K. Lindorff-Larsen, J. Christodoulou, M. Vendruscolo, and C.M. Dobson. *Journal of the American Chemical Society*, 127(2):476–477, 2005.
- [92] K.-P. Wu, D.S. Weinstock, C. Narayanan, R.M. Levy, and J. Baum. *Journal of molecular biology*, 391(4):784–96, 2009.
- [93] O. Ullman, C.K. Fisher, and C.M. Stultz. *Journal of the American Chemical Society*, 133(48):19536–46, 2011.
- [94] J. Yoon, S. Jang, K. Lee, and S. Shin. *Journal of biomolecular structure & dynamics*, 27(3):259–70, 2009.
- [95] I.F. Tsigelny, P. Bar-On, Y. Sharikov, L. Crews, M. Hashimoto, M.A Miller, S.H. Keller, O. Platoshyn, J.X.-J. Yuan, and E. Masliah. *The FEBS journal*, 274(7):1862–77, 2007.
- [96] J.D. Perlmutter, A.R. Braun, and J.N. Sachs. *The Journal of biological chemistry*, 284(11):7177–89, 2009.
- [97] D. Balesh, J. Ramjan, and W.B. Floriano. *Journal of Biophysical Chemistry*, 02(02):124–134, 2011.
- [98] T. Gurry, O. Ullman, C.K. Fisher, I. Perovic, T. Pochapsky, and C.M. Stultz. *Journal of the American Chemical Society*, 135(10):3865–72, 2013.
- [99] A. M. Bodles, O. M. A. El-Agnaf, B. Greer, D. J. S. Guthrie, and G. B. Irvine. *Neurosci Lett*, 359:89–93, 2004.
- [100] O. M. A. El-Agnaf and G. B. Irvine. *Biochem Soc Trans*, 30:559–65, 2002.
- [101] R. Shaltiel-Karyo, M. Frenkel-Pinter, N. Egoz-Matia, A. Frydman-Marom, D. E. Shalev, D. Segal, and E. Gazit. *PLoS One*, 5:e13863, 2010.
- [102] Y. Liu and D. Schubert. J Neurochem, 71:2322–9, 1998.
- [103] O. M.A El-Agnaf, R. Jakes, M. D. Curran, D. Middleton, R. Ingenito, E. Bianchi, A. Pessi, D. Neill, and A. Wallace. FEBS Lett, 440:71–5, 1998.

- [104] A. Kazmierczak, J. B. Strosznajder, and A. Adamczyk. *Neurochem Int*, 53:263–9, 2008.
- [105] A. Kazmierczak, G. A. Czapski, A. Adamczyk, B. Gajkowska, and J. B. Strosznajder. *Neurochem Int*, 58:206–14, 2011.
- [106] O.M.A El-Agnaf and G.B. Irvine. *Journal of structural biology*, 130(2-3):300–309, 2000.
- [107] M. Bisaglia, A. Trolio, M. Bellanda, E. Bergantino, L. Bubacco, and S. Mammi. *Protein Science*, 15:1408–1416, 2006.
- [108] D.P. Karpinar, M.B. Gajula Balija, S. Kügler, F. Opazo, N. Rezaei-Ghaleh, N. Wender, H.-Y. Kim, G. Taschenberger, B.H. Falkenburger, H. Heise, A. Kumar, D. Riedel, L. Fichtner, A. Voigt, G. H. Braus, K. Giller, S. Becker, A. Herzig, M. Baldus, H. Jäckle, S. Eimer, J.B. Schulz, C. Griesinger, and M. Zweckstetter. *The EMBO journal*, 28(20):3256–3268, 2009.
- [109] J. Yoon, S. Jang, K. Lee, and S. Shin. *Journal of biomolecular structure & dynamics*, 27(3):259–269, 2010.
- [110] G. Liang, J. Zhao, X. Yu, and J. Zheng. *Biochemistry*, 52(6):1089–100, 2013.
- [111] S. Côté, R. Laghaei, P. Derreumaux, and N. Mousseau. J. Phys. Chem. B, 116:4043–4055, 2012.
- [112] J.T Jarrett and T.L.Jr. Peter. *Biochemistry*, 31:12345–12352, 1992.
- [113] M. E. J. Newman and G. T. Barkema. *Monte Carlo Methods in Statistical Physics*. Clarendon Press (Oxford), 1999.
- [114] J.-E. Suk, S.B. Lokappa, and T.S. Ulmer. NIH Public Access. *Biochemistry*, 49(7):1533–1540, 2010.

- [115] M. Vilar, H.-T. Chou, T. Lührs, S.K. Maji, D. Riek-Loher, R. Verel, G. Manning, H. Stahlberg, and R. Riek. *Proceedings of the National Academy of Sciences of the United States of America*, 105(25):8637–42, 2008.
- [116] M. Bisaglia, I. Tessari, L. Pinato, M. Bellanda, S. Giraudo, M. Fasano, E. Bergantino, L. Bubacco, and S. Mammi. *Biochemistry*, 44(1):329–339, 2005.
- [117] K.-M. Pan, M Baldwin, J. Nguyen, M. Gasset, A.N.A. Serban, D. Groth, I. Mehlhorn, Z. Huang, R.J. Fletterick, F.E. Cohenu, and S.B. Prusiner. *Proc. Natl. Acad. Sci. USA*, 90(December):10962–10966, 1993.
- [118] D.M. Walsh. Journal of Biological Chemistry, 274(36):25945–25952, 1999.
- [119] M.D. Kirkitadze, M.M. Condron, and D.B. Teplow. J. Mol. Biol., 312:1103– 1119, 2001.
- [120] M.R. Smaoui, F. Poitevin, M. Delarue, P. Koehl, H. Orland, and J. Waldispühl. *Biophysical journal*, 104(3):683–93, February 2013.
- [121] R. Tycko and R.B. Wickner. *Accounts of chemical research*, 46(7):1487–96, 2013.
- [122] R. Kodali and R. Wetzel. *Current opinion in structural biology*, 17(1):48–57, 2007.
- [123] K. Hun Lim. Journal of Physical Chemistry & Biophysics, 03(02):2–3, 2013.
- [124] Y. Zhang, B. Wang, H. Wan, Q. Zhou, and T. Li. *Neuroscience letters*, 541:132–7, 2013.
- [125] J.S. Jeong, A. Ansaloni, R. Mezzenga, H. a Lashuel, and G. Dietler. *Journal of molecular biology*, 425(10):1765–81, 2013.
- [126] B. Urbanc, M. Betnel, L. Cruz, G. Bitan, and D. B. Teplow. J Am Chem Soc, 132:4266–80, 2010.

- [127] B. Barz and B. Urbanc. *PLoS One*, 7:e34345, 2012.
- [128] Y. Atsushi, O. Akisumi, N. Kazuya, H. Shin'ichi, and K. Noriyuki. *Chemical Physics Letters*, 595–596:242 249, 2014.
- [129] H. Han, P. H. Weinreb, and P. T. Lansbury. *Chemistry & biology*, 2:163–169, 1995.