



HAL
open science

Aggregation of disease-related peptides

Phuong H. Nguyen, Fabio Sterpone, Philippe Derreumaux

► **To cite this version:**

Phuong H. Nguyen, Fabio Sterpone, Philippe Derreumaux. Aggregation of disease-related peptides. *Progress in Molecular Biology and Translational Science*, 170, Elsevier, pp.435-460, 2020, Computational Approaches for Understanding Dynamical Systems: Protein Folding and Assembly, 10.1016/bs.pmbts.2019.12.002 . hal-03087681

HAL Id: hal-03087681

<https://hal.science/hal-03087681>

Submitted on 24 Dec 2020

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

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

Aggregation of short disease-related peptides

Phuong H. Nguyen,^{1,2} Fabio Sterpone,^{1,2} Philippe Derreumaux^{3,4*}

¹ CNRS, Université de Paris, UPR 9080, Laboratoire de Biochimie Théorique, 13 rue Pierre et Marie Curie, F-75005, Paris, France

² Institut de Biologie Physico-Chimique-Fondation Edmond de Rothschild, PSL Research University, Paris, France

³ Laboratory of Theoretical Chemistry, Ton Duc Thang University, Ho Chi Minh City, Vietnam

⁴ Faculty of Pharmacy, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Abstract

Protein misfolding and aggregation of amyloid proteins is the fundamental cause of more than twenty diseases. Molecular mechanisms of the self-assembly and the formation of the toxic aggregates are still elusive. Computer simulations have been intensively used to study the aggregation of amyloid peptides of various amino acid lengths related to neurodegenerative diseases. We review atomistic and coarse-grained simulations of short amyloid peptides aimed at determining their transient oligomeric structures and the early and late aggregation steps.

Key words: amyloid, aggregation, structure, dynamics, thermodynamics, hydrodynamics, computer simulations, all-atom, coarse-grained force field

1. Introduction

Numerous diseases affecting either the central nervous system or a variety of peripheral tissues result from the self-assembly of amyloid proteins. Disorders range from Alzheimer's disease (AD, implication of A β proteins of 39-43 amino acids and tau protein of 441 amino acids), Parkinson (α -synuclein protein of 140 amino acids) to amyotrophic lateral sclerosis or ALS (superoxide dismutase, SOD of 154 amino acids) and type II diabetes (IAPP or islet amyloid polypeptide of 37 amino acids). All these proteins differ in amino acid sequence and length, yet they all form amyloid fibrils with cross- β structure.¹ This propensity to self-assembly into amyloid fibrils under given conditions is also observed for heptapeptides (e.g., A β 16-22)², tetrapeptides (e.g., KFFE),³ dipeptides⁴ and even by single amino acids.⁵

Experimentally, aggregation kinetics of amyloid proteins display a sigmoidal curve with a lag phase, during which monomers self-assembly into oligomers and undergo structural rearrangements, until the growth phase where fibril elongation and primary/secondary nucleation occur, followed by the saturation phase where the system is in equilibrium between fibrils and a small concentration of monomers. Note that our understanding of amyloid aggregation kinetics goes much beyond classical nucleation theory where primary nucleation event is sufficient to fit the experimental curves.¹

Structural determination of all species along amyloid fibril formation pathways is challenging by experimental means because of their transient character despite the use of a large variety of biophysical techniques.⁶ Complexity comes from the metastable character of each species (a very large number of conformations for each aggregate), but also from the sensitivity of the kinetics to experimental conditions (solution pH, temperature, salt concentration, agitation, ions) and external conditions (presence of membrane, crowding, etc.).⁷

To add further complexity, amyloid aggregation kinetics is also modulated by the exact composition of the lipid bilayers. For example, although A β 1-40, A β 1-42 and tau oligomers are key players in AD,⁸ dietary PUFA (polyunsaturated fatty acids) supplementation change molecular phospholipids, and there is evidence that increased intake of omega-3 PUFA slows the progression of AD, while omega-6 PUFA is linked to higher risk of AD.⁹ Another example of the modulation of the aggregation upon membrane presence can be taken from experiments on α -synuclein. On the one hand, DPLS lipid bilayers significantly augment its aggregation rate, while DOPE lipid bilayers have no impact on its aggregation rate.¹⁰

Simulations on amyloids at different time and length scales can complement experiment, but require accurate potential energy models ranging from atomistic in explicit aqueous solution/lipid bilayer, coarse-grained with implicit solvent/membrane to mesoscopic or super-mesoscopic representations.¹¹⁻¹⁴ Molecular dynamics (MD), replica exchange molecular dynamics (REMD)

simulations and other sampling methods are often used to generate the conformational ensemble of intrinsically disordered proteins (IDPs). These methods are described in Section 2 and we will see that their results vary with the protein force field used. Section 3 focuses on the application of computer simulations to a better understanding of small oligomers of short amyloid peptides. While an implicit solvent representation reduces the number of degrees of freedom, hydrodynamics, which deals with the motion of fluids and the forces acting on solid bodies immersed in fluids and in motion relative to them, plays a significant role in the early aggregation steps of amyloids using simplified models (Section 4). Section 5 reports on the simulations aimed at understanding the primary nucleation and the surface-catalysed secondary nucleation. The final section reports on recent advances in the determination of oligomer structures of A β 40/42 peptides in aqueous solution and lipid bilayers based on computer simulations. Our review will be mainly centred on A β peptides. For computer simulations of tau, hIAAP and synuclein, see Ref. 14.

2. Computer simulation models for amyloid protein aggregation

Numerous conformational sampling methods are used to study amyloid proteins. We will not describe them in detail but rather summarize their main features. Atomistic molecular dynamics (MD) simulations in explicit environment offer the most detailed dynamic and energetic pictures of protein folding and aggregation. The longest simulation on the fastest computer (Anton) does not exceed, however, 1 millisecond in explicit solvent, sufficient for sampling the monomeric state of amyloid proteins, but far too short for exploring the aggregation during the experimental lag phase.¹⁵ To overcome the multiple minima problem and converge faster to equilibrium properties, replica exchange MD (REMD) simulations, which run multiple MD simulations at different temperatures in parallel and exchange consecutive replicas according to the Metropolis criterion, are commonly used as their implementations are easy. Interestingly, it is possible to determine peptide dimerization-dissociation rates from REMD simulations.¹⁶ Recent modeling methods based on Markovian descriptions of conformational states have also led to frameworks allowing accurate description of dynamics. In conjunction with REMD simulations, these frameworks allow systematic and accurate determination of transition probabilities between the corresponding states, in the case of Markov state models, and of transition rates, in the case of the corresponding coarse master equations.¹⁷

Among other sampling techniques, Monte Carlo, replica exchange MC (REMC), replica exchange with solute tempering, simulated tempering, discrete or discontinuous MD, and metadynamics are also employed.

On-lattice (where the position of the particles are limited to nodes) and off-lattice Monte Carlo simulations of models having a small number of internal degrees of freedom have been used to investigate the nucleation mechanisms.¹⁸⁻²⁰ On-lattice REMC simulations were also used to determine the free energy surfaces

of various oligomers of the A β 16-22 and A β 35-42 peptides^{21,22} and the nucleus size of A β 42.²³

A drawback of REMD is its sensitivity to the overlap between the replicas total energies, and to the system size, the energy overlap scaling with the number of particles as $N^{-1/2}$. Optimal exchange rate between replicas is thus limited by the number of solvent molecules. Because protein-protein and protein-solvent interactions are dominant in protein recognition,²⁴ it is interesting to focus on these interactions and simulation schemes based on potential energy rescaling and Hamiltonian replica exchange exploit this feature. Replica exchange simulation with solute tempering (REST) belongs to this class of schemes.²⁵⁻²⁷ Smith et al. performed REST and standard REMD simulations to study the A β 25-35 peptide in water and a system consisting of two A β 10-40 peptides binding to a DMPC lipid bilayer.²⁷ Their study led to three conclusions. First, the number of replicas in REST simulations can be reduced four to five times. Second, REST produces much fewer conformational states than REMD, but the number of unique states is very similar between the REST and REMD simulations. Third comparing REST and REMD equilibrium conformational ensembles, REST reproduces REMD data extremely well for the system of two A β peptides bound to the DMPC lipid bilayer, and REST demonstrates much better convergence for the A β 1-40 system bound to the lipid bilayer rather than for the small unstructured A β 25-35 peptide in water.²⁷

In simulated tempering, temperature is a dynamical variable with discrete values T_n . Simulated tempering scheme requires the determination of *a priori* unknown weight parameters to ensure a random walk in temperature space, namely the Helmholtz free energies at T_n . We have developed a simulated tempering scheme with on-the-fly weight determination. The weights are self-updated via a trapezoid rule during the simulation, eliminating therefore the need for trial simulations.²⁸ The advantage of this simulated tempering method over REMD was illustrated on the equilibrium states of A β 16-22 trimer using the OPEP CG force field and the folding of the 20-residue Trp-cage using atomistic force field starting from randomly chosen states.²⁹ Of interest is that it is possible to include different atomistic force fields into a single simulated tempering run.³⁰

Unlike classical MD simulations, which integrate the equation of motion with a time step of 1-2 ns, DMD (discontinuous MD) is event-driven and keeps track of particle positions and velocities only at collision times. As a result, DMD is several orders of magnitude faster than the traditional continuous MD if coarse-graining is used. The main disadvantage of DMD is that pairs of particles must interact by means of spherically symmetric potentials that consist of square wells. The DMD scheme has been coupled to all-atom model with implicit solvent,³¹ and various CG models such as PRIME20.³²

Finally in metadynamics, enhanced conformational sampling is realized by an external history-dependent bias potential affecting few selected degrees of

freedom, referred to as collective variables. The bias is adaptively constructed as a sum of Gaussians deposited along the system trajectory to prevent the system from revisiting previous conformational regions. Although metadynamics requires a limited yet effective set of collective variables, this is not a trivial task even for the monomeric state of long amyloid peptides,³³ and of course for amyloid fibril formation.³⁴

There is a constant effort in improving atomistic force fields in explicit water. Many versions of CHARMM and AMBER with modified water force fields have been designed over the years to reproduce experimental data on folded proteins (early versions) and both folded and disordered proteins (CHARMM36m-TIP3P modified,³⁵ and AMBERsb99-disp,¹⁵ where disp is a variant of TIP4P water model. Other current force fields include OPLS-AA and GROMOS force fields.³⁶ We summarize some recent results on short amyloid peptides and the A β 40 and A β 42 peptides.

Following a previous study on A β 16-22 dimers and trimers using three force fields (AMBER99, GROMOS96 and OPLS-AA/TIP3P),³⁷ Nguyen *et al.* studied the effects of 17 widely used atomistic molecular force fields on the structures and kinetics of the A β 16-22 dimer in aqueous solution, for a total MD time of 0.34 millisecond.³⁸ While the AMBER94, AMBER99 and AMBER12SB force fields fail to predict β -sheet, the AMBER96, GROMOS45a3, GROMOS53a5, GROMOS53a6, GROMOS43a1, GROMOS43a2, and GROMOS54a7 force fields form β -sheets very rapidly. In contrast, the AMBER99-ILDN, AMBER14SB, CHARMM22* and CHARMM36m force fields are the best candidates for studying amyloid peptide assembly, as they provide good balances in terms of structures, thermodynamics and kinetics (assembly and dissociation).

In the same line, Strodel *et al.* investigated the effect of five different force fields (GROMOS54a7, OPLS-AA, CHARMM22*, AMBER99SB*-ILDN and AMBER03WS) on the hexamer formation of the wild-type A β 16-22 peptide and three mutants: A β 16-22 (F19V, F20V) which do not form fibrils, and A β 16-22 (F19L) which forms fibrils faster than the wild-type.³⁹ Note that the AMBER03WS force field reduces the over-stabilization of protein-protein interactions generated by older force fields.³⁶ They found that the structural properties are independent of the force field, but oligomer formation kinetics depends strongly on the force field. It is very important to note that the kinetics generated by the various force fields differ more from each other than the kinetics between amyloid and non-amyloid peptides simulated with a single force field.³⁹

Siwy *et al.* performed comparative analysis of the REMD conformational ensembles of A β 10-40 monomer produced by five force fields combining CHARMM36, CHARMM22*, CHARMM22/cmap, and OPLS-AA and two water models (standard and modified TIP3P).⁴⁰ As expected, all force fields predict that A β 10-40 remains unfolded dominated by turn and random coil structure. The

secondary structures generated by CHARMM36 and CHARMM22/cmap simulations are qualitatively similar, but their tertiary interactions show little consistency. Among all force fields, CHARMM22* differs the most from CHARMM36, and based on the analysis of 3JHNH α -coupling and RDC constants, the CHARMM36 force field with standard TIP3P model produces the most accurate representation of A β 10-40 conformational ensemble.⁴⁰

The A β 42 peptide has also been subject of particular attention over the years.⁴¹⁻⁴³ Strodel *et al.* used REMD simulations on the monomeric form with the five OPLS, AMBER99SB, AMBER99SB*ILDN, AMBER99SBILDN-NMR and CHARMM22* force fields and compared the equilibrium ensembles to nuclear magnetic resonance (NMR) experimental data. They found that all force fields except AMBER99SBILDN-NMR reproduce local NMR observables, and CHARMM22* is slightly better than the other force fields.⁴²

Nguyen *et al.* generated the A β 42 dimer ensembles obtained by four force fields (OPLS-AA, CHARMM22*, AMBER99sb-ildn and AMBERsb14 with TIP3P model) coupled to REMD simulations.⁴⁴ The simulation results were compared to experimental circular dichroism (CD) and ion-mobility mass-spectrometry data. On the basis of 144 μ s, the four force fields lead to random coil ensembles with cross-collision sections, hydrodynamics properties, and small-angle X-ray scattering profiles independent of the force field. Overall, there are significant differences in secondary, tertiary and quaternary conformations among the four force fields, for example the intramolecular beta-hairpin content spanning residues 17-21 and 30-36 varying between 1.5% and 13%.

Finally, the stability of tetrameric A β 40 and A β 42 β -barrel structures was studied considering four atomistic force field-water model combinations: Amber99SB-ILDN/TIP3P, OPLS/TIP3P, CHARMM36m/TIP3P-modified, and Amber99sb-disp. For each force field, Nguyen *et al.* found that A β 42 samples barrel structures in aqueous solution, while this is less the case for A β 40. It is to be noted that the authors used the two latest force fields that were developed for IDPs (CHARMM36m and AMBER99SB-disp) and spent a lot of computing time to converge the sampling by using REMD.⁴⁵

There is also a constant effort in designing coarse-grained models with implicit solvent for amyloid aggregation. Using a two-state model for each peptide and Langevin dynamics, Caflisch *et al.* found that the nucleus size varies between 4 and 35 depending on the energy difference between the amyloid-competent and amyloid-protected minima. This simple model is able to generate fibril topologies with twist and multifilament composition as observed experimentally.⁴⁶

Shea and co-workers designed a three-bead model (two for the backbone, and one for the side-chain) that reproduces the polar, hydrophobic and charged nature of each residue. Using a simple energy function that controls the beta-strand propensity of each peptide and Langevin dynamics simulations,⁴⁷ they found that,

when the beta-strand propensity is not too high, the oligomeric states consist of disordered aggregates, multiple β -sheet layers with different orientations including the cross- β structure, and β -barrels consistent other coarse grained simulations.⁴⁸⁻⁵⁰

Frenkel and co-workers describe the monomers as single spherocylinders able to coexist in two states: a soluble state and an amyloid-like state. Using this model and Monte Carlo simulations, they estimated a critical nucleus size of 4 peptides for A β 42. They also established the concentration, temperature and interpeptide interactions conditions upon which aggregation follows the one-step nucleation (where two peptides in the amyloid-competent state binds) and the two-step nucleation (where the amorphous aggregates reorganize into beta-rich nuclei).^{19,20}

The four-bead “Urbanc” and PRIME20 models consist of three spheres for the backbone (N, Ca, and CO) and one bead for the side chain. Using DMD, Urbanc *et al.* reproduced differences in oligomer size distributions between the A β 40 and A β 42 peptides consistent with experimental data.⁵¹ Using PRIME20/DMD, Hall reproduced the propensity of a series of hexapeptides known to form amyloid and non-amyloid fibrils in vitro and studied the mixing of A β 40 and A β 16-22 peptides.³²

Wolynes *et al.* also designed the AWSEM (associative memory, water mediated, structure and energy model) three-bead model with Ca, Cb and O and applied it to the aggregation of A β 40 and A β 42 peptides.^{52,53} Derreumaux and co-workers designed the OPEP (optimized potential for efficient peptide structure prediction) model, where each amino acid is represented by six beads: one bead for each side-chain and the atomic resolution is used for the backbone (N, H, Ca, C and O).^{54,55} This model has been used for on-lattice and off-lattice simulations of many short-amyloid peptides, and the A β 40 and A β 42 peptide.^{56,57} Note it is beyond the scope of this review to describe the potential energy functions used by all models.

3. Structures of small aggregates

Among all amyloid sequences, the A β 16-22 peptide has been the most studied due to its simplicity, its amyloid fibril at neutral pH consists of antiparallel β -strands and parallel β -sheets,² and the amino-acids 17-21 constitute the main driving force for the aggregation of the full-length A β peptides.⁸ Based on MD and REMD simulations with all force fields except AMBER99sb-disp, it is clear that the probability of A β 16-22 fibril formation following the one-step nucleation is very small.

Using the modified AMBER94 force field and REMD, Nussinov *et al.* found six configurations including shifted parallel strand and parallel loop, antiparallel strand, parallel strand, shifted antiparallel strand, cross, and tight cross/lock.⁵⁸

Employing the OPEP coarse-grained force field, in-register and out-of-register parallel and antiparallel strands were generated along with amorphous states.⁵⁹ Nguyen *et al.* demonstrated with 0.34 millisecond MD that the probability of the amyloid-competent state is on the order of at most 10% using the most recent atomistic force fields. Of interest is that, in principle there are 9 out-register antiparallel β -sheets in addition to the native in-register β -sheet, but only five are populated at the dimer level.³⁸ This feature is important as experiments and simulations have shown that large aggregates consist of out-of-register β -strands and transition to the fibril state involves the reptation of the β -strands at high concentration and detachment/binding of the peptides at low concentration.^{60,61} A second important aspect from the A β 16-22 dimer simulations is that the parallel β -sheet conformation is a state of low probability. This feature explains why the E22Q mutation can induce antiparallel to parallel β -sheet transition of A β 16-22 fibrils as observed by solid-state NMR and all-atom simulation studies.⁶²

Simulations of A β 16-22 trimer reveal that the main configurations found at the dimer level remain and we have a low population of mixed antiparallel – parallel beta-sheets.⁶⁰ In terms of dynamics, we can infer a very likely scenario from other peptides known to form fibrils with antiparallel beta-strands. By studying the aggregation of the KFFE peptide up to 4-mers by REMD and Markov models, Sengupta *et al.* found that disordered and β -sheet oligomers do not interconvert, leading to separate pathways for their formations.⁶³

De Groot *et al.* used unbiased explicit solvent MD simulations to investigate the structural and dynamical features of aggregates formed by A β 16-22 at atomistic resolution for 2 microseconds. Simulations of 12 peptides were performed at 2 and 20 mM peptide concentration.⁶⁴ To characterize the structures, they computed collision cross sections of the individual aggregates sampled from the MD trajectories and found heterogeneous ensemble of aggregates lacking a defined quaternary structure, fully consistent with the same simulations of other amyloid peptides ranging from 6- to 12-mers. In particular, they found a very large variety of ordered intermolecular β -structure motifs including distinct fibril-like configurations and β -barrel oligomers of various sizes already prevalent in smaller aggregates. This conformational ensemble is fully consistent with all-atom A β 16-22 DMD simulations of 6 to 12 peptides where β -barrel oligomers of various sizes were found with a probability varying between 2 and 4%,³¹ CG Langevin dynamics simulation results of Shea *et al.*,⁴⁷ and a multitude of all-atom and CG simulations of other sequences.^{26,65,66}

The second peptide we want to discuss is the A β 37-42 peptide because, in contrast to many short-amyloid peptides, its fibril structure displays in parallel β -strands, as formed by the full length A β 40/42 peptides. Nguyen *et al.* performed a REMD simulation of 16 all-atom A β 37-42 peptides in explicit water starting from dispersed orientations.⁶⁷ The total simulation time of 23 microseconds allows them to obtain the conformational distribution of oligomers and β -sheet sizes at 300 K and a peptide concentration of 12 mM. Self-assembly is described by the

condensation-polymerization conversion from micelle to high β -sheet structures. When equilibrium is reached, the oligomer distribution consists of large aggregates (70%) and free monomers (30%). The population of 4-5 fully parallel β -strands consistent with the native fibril is low, the population of 4-5 fully antiparallel β -strands is non-negligible, and the system has a large population of mixed parallel/antiparallel β -strands in equilibrium with amorphous oligomers. This distribution is in agreement with all-atom metadynamics results of Laio *et al.*³⁴ We now report on A β 16-22 simulations of unprecedented sizes that further support the second-step nucleation.

4. Exploring the early aggregates of amyloid peptides at quasi-atomic level with hydrodynamics

Coarse-grained simplified molecular models with implicit solvent have been extensively used to explore the aggregation process of amyloid systems and to inspect the impact of the peptide beta propensity and amino acid sequence, peptide-peptide interactions, concentration, temperature,^{46,47,52,53,68,69} and crowding.⁷⁰ However these simulations based on simplified models neglect the effect of solvent mediated interactions on the kinetic behaviour of the system and the diffusion limited aggregation process. It is therefore important to include these hydrodynamic interactions (HI). This can be done using specialised simulation techniques, e.g. Brownian dynamics with HI, the stochastic rotational method, and the lattice Boltzmann molecular dynamics (LBMD).

Recently we implemented an effective LBMD approach to study biomolecular systems based on the OPEP CG model. This technique, which includes naturally hydrodynamics in implicit solvent coarse-grained molecular dynamics of biomolecules, has been used to explore the mobility of proteins in a crowded solution, the unfolding and dynamics of proteins under shear flow, the multi-scale dynamics of amyloid peptides aggregation.^{22,71,72} Before describing the results of interest, we pause to describe the methodology, see Figure 1A.

In LBMD, the particle-based dynamics is coupled to a kinetic representation of the solvent, simulated via the Lattice Boltzmann (LB) technique.⁷³ The coupling between the particles and the solvent has the form of a Stokes-like drag force acting on a per-particle basis:

$$\vec{F}_i^D = -\gamma(\vec{v}_i - \vec{u}_i) \quad (1)$$

where the first term is the velocity of i -th particle, the second term is the fluid velocity smeared over a finite extension of the i -th particle, and γ is the friction, an adjustable parameter in the methodology taken to be equal for all particles of a given molecule. The drag force adds up to the usual conservative forces derived

from the Hamiltonian of the system and to a random white noise that represents thermal fluctuations. The reader can find more technical details in Ref. 71 and 74.

In our simulations, the LB evolution was solved by using the Bhatnagar-Gross-Krook (BGK) collisional operator, with a lattice grid spacing that was varied depending on the size of the system and the wished resolution (1-5 Å), and the bulk water kinematic viscosity.^{71,74} When using the flexible OPEP force field,⁷⁵ the molecular and LB dynamics were evolved synchronously using a time step of 1.5 fs, a value defined by the properties of the molecular mechanics force field. When focusing on the mesoscale behaviour of aggregate growth, an elastic model was used in tandem with a time step of 10 fs. The friction γ was optimised in order to reproduce the diffusion of an isolated peptide in dilute condition.

We first considered a system composed of 100 A β 16-22 peptides placed in a box corresponding to a concentration of about 60 mM.⁷⁴ The first issue we addressed was to estimate the characteristic time of complete aggregation with HI compared to a Langevin simulation. It turns out, as already anticipated by Skolnick and co-workers in the study of lipid assembly,⁷⁶ that the presence of HI speeds up the aggregation process. In the LBMD simulation, after 100 ns we observe the formation of a large aggregate containing about 80% of the peptides. At the same time scale and in the absence of HI, the aggregation proceeds more slowly, and only 50% of peptides form the largest cluster, see Figure 1B. The result can be understood considering that by including HI the diffusivity of peptide aggregates scales more favourably with the aggregate size than in their absence. Also, when focusing on the fluctuations of the size of the formed aggregates, it results that the presence of HI increases the fast exchange of peptides.

The aggregation process of A β 16-22 was further investigated by considering a system of much larger size, 1000 peptides placed in a box of $L=300$ Å, corresponding to a concentration of 60 mM.⁷⁷ Using atomistic and explicit solvent, this system would count for 2.7 millions of particles, becoming prohibitive for classical MD simulations.

The aggregation process was followed at the microsecond time scale, and for the first time three different regimes were clearly individuated. At the short timescale, $<10^2$ ns, the growth of peptides aggregates proceeds continuously in time, with progressive absorption of single molecules. When in the solution the presence of finite size oligomers (10-20 peptides in each) becomes relevant, the aggregation proceeds via the fusion of these entities, and the growth mechanism is discontinuous manifesting sudden size jumps (see the arrows in the Figure 1C). Several fusion mechanisms were individuated, from head-tail elongation to branching, involving coherently from two up to four aggregates. Interestingly, the time evolution of the aggregate size distribution shows the emergence of a shoulder around the pentamer and decamer states, a signature that can be related to the presence of a critical nucleus size. Finally, the system reaches a steady

regime for time > 350 ns, with the formation of a large amorphous protofibril and three minor ones, no residual monomers are present in the system at this timescale. The progress of the amorphous protofibril elongation is limited by the diffusion of these very large (massive) clusters. To be noted that very large size fluctuations of the largest amorphous protofibril are visible and are caused by fragmentation events.

Having individuated these relevant timescales it is worth to stress that the simulated system is at a very high concentration. If one considers a lower concentration regime, e.g. reproducing the upper limit of experiments, μM , the early phase of the aggregation would be on the microsecond timescale and would approach the second timescale at in vivo concentrations (nM).

Looking at the early phase of aggregation, following the formation of the small oligomers, the peptides acquire in part beta-sheet organisation, going up to 15% of the whole system. The further conversion is then rate-limited by the size of the formed aggregates that constraints the monomer conformational sampling. The possibility to explore the growth of very large amorphous protofibril at the quasi-atomistic resolution allowed to individuate critical structural features like the formation of a large asymmetric pore (dimensions of 3-5 nm) at the edge of the structure, see Figure 1D. It is interesting that amyloid plaques with annular pores have been detected in the brain of Alzheimer's disease patients by antibodies and electron microscopy experiments.⁷⁸

Evolution of the amorphous protofibril elongation was further explored by deploying an ad-hoc multi-scale strategy. The four final clusters individuated in the LBMD simulation based on the fully flexible OPEP force field were converted into an elastic-network and placed in much larger simulation box (60 nm) to avoid finite-size artefact. The elastic network allowed the use of a larger integration time-step for the particle dynamics, and scaling arguments were applied to ensure a kinetic coherence when changing fluid and particle resolutions. With this tool in hands, it was possible to follow the hydrodynamic sustained formation of a larger branched amorphous protofibril, and the critical role of the amorphous protofibril surface acting as an entropically favourable seed for the formation of branched structures via aggregate fusion, see Figure 1E.

5. Primary and secondary nucleation from simulations

The foundation of the nucleation for amyloid fibrils dates from the polymerisation studies of actin by Oosawa⁷⁹ and deoxyhemoglobin by Weaton.⁸⁰ They theoretically explained the nucleation and subsequent polymerisation processes by employing the homogeneous nucleation theory developed for vapour condensation. Their theories have served as a background for the development of kinetic and thermodynamic analyses for amyloid nucleation.⁸¹

Basically, in the first step, the fluctuations in the system induce spontaneous association of monomers, leading to the formation of small clusters of monomers. Due to small sizes, these clusters are thermodynamically unstable with high positive interfacial free energy. With increasing size, the intra-cluster interactions become more significant, and the clusters become more stable, the bulk free energy is negatively decreased, resulting in the decrease in the total free energy. A smallest cluster, whose total free energy is highest along the nucleation pathway, is usually defined as a critical nucleus. The probability of forming a larger cluster started from a nucleus is higher than the probability of dissociation of the nucleus. The initial formation of nuclei in a system where there are still no formed aggregates present is called primary nucleation. The primary nucleation may occur in homogeneous solution, called homogeneous nucleation, or at a surface, called heterogeneous nucleation. A surface can be the wall of the vessel, an air-water interface or the wall of a lipid vesicle etc. The secondary nucleation is induced by the surface of existing aggregates, referred to as surface-catalysed nucleation.⁸²

From the experimental side, the chemical kinetic and thermodynamic analyses play an important role in revealing the mechanism of nucleation of protein aggregation.⁸³ These analyses were applied to the aggregation of IAPP, alpha-synuclein and, notably the A β peptides with wild-type sequences and many variants.^{7,84-86} These studies revealed the role of mutations and truncations at the N-terminus on the concentration-dependent time course of A β fibril formation and lag phase from a solution of monomers.

From the theoretical side, master equations have been used to interpret experimental data. Basically, rate equations with increasing complexity are fitted to the experimental sigmoidal curve. It has been shown that the best fit can only be obtained if all processes, including primary nucleation, fibril growth and secondary nucleation are taken into account. The use of only the first two processes can fail to reproduce the experimental curve.⁸³ The master equations do not provide, however, information on the nucleus. As a remedy, simulations have been carried out, employing various protein models so as to provide insights into the structure and size of the primary nucleus.^{6,87}

At the coarse-grained level, Li *et al.* developed an on-lattice protein model, where each residue is entered on a site of a cubic lattice.^{23,88} MC simulations of multiple A β 42 chains, each chain consisting of eight residues, show that the nucleus size N^* is 11 chains. Irback *et al.* developed another lattice model, where a residue is represented by an unit vector, located on a site of a cubic lattice, and interacts with other peptides by hydrogen bonding and hydrophobic forces. The simulation of 256 peptides shows that fibril formation occurs with a sigmoidal kinetics shape, and the width of the aggregate prior to nucleation is 3.5.¹⁸

Abeln and coworkers developed a more sophisticated lattice model by considering the formation of hydrogen bonds, the directions of side-chains and

pairwise interaction energies between the 20 amino acids. REMC simulations captured the amyloid formation with $N^* = 10$ for a designed 7-residue TFFFTFT peptide at low concentration.⁸⁹ We have developed further this model by incorporating the OPEP force field and notably the four-body hydrogen bonding interaction so as to explore the critical nucleus size of fibril formation for real amyloid peptides. REMC simulations of oligomers formed by 10 - 20 chains of A β 16-22 or A β 37-42 peptides allowed us to characterise the structure, thermodynamics and nucleus sizes.^{21,22} We show that A β 16-22 decamer forms a stable fibril with two anti-parallel 5-stranded beta-sheets, consistent with the NMR fibril structure, indicating that the nucleus size of A β 16-22 is about 10 chains. For a system of 20 A β 37-42 peptides, Figure 2 shows the population of one, two and three layers of β -sheets composed of n -strands. Overall, the system is mainly disordered with a low population of two and three layers of long n -strand beta-sheets. It is to be noted that the free energy surface changes for 30 A β 37-42 peptides with the presence of well-defined native fibrils (with parallel beta-strands) indicating that $20 < N^* < 30$.²²

In addition to the Caffish's and Frenkel's models described in Section 2, Linse et al. developed an algorithm for MC simulation of amyloid formation from flexible peptides. The algorithm is based on kinetic discrimination among fibrillar and non-fibrillar contacts. The simulations of up to 500 hexapeptides show that the fibrillation follows a sigmoidal curve, but there is no requirement for any complicated structural pathway to explain the lag phase and sigmoidal shape. This shape arises from cooperative among multiple interactions within each pair of peptides in a fibril.⁹⁰ Another kinetic model developed by Hsieh *et al.* also shows a two-step nucleation for A β 16-22.⁹¹ Using PRIME20 model with DMD method, Hall *et al.* show that at low temperatures, the aggregation of 48 A β 16-22 chains follows a primary nucleation mechanism with $N^* = 8$.⁹²

Finally, Laio *et al.* investigated the all-atom aggregation of 18 Val8 and 18 A β 35-40 peptides employing the bias-exchange metadynamics method in explicit solvent with the Amber99 force field. The simulations show that the aggregation follows a primary nucleated conformational conversion mechanism and the nucleus size is on the order of 14. The most populated aggregate is amorphous. The free energy landscape projected on appropriate variables reveals that the maximum free energy is characterized by a transition from mixed parallel/antiparallel to parallel β -strand orientations and only when a sufficient number of parallel sheets are formed that the free energy decreases to a free energy minimum with fully parallel β -sheets.^{34,93}

To study the secondary nucleation mechanisms, large system sizes are necessary. However, due to the limitation of current computer power, most atomistic simulations start from preformed fibrils and free monomers. Strodel *et al.* carried out 1 microsecond all-atom simulations of six A β 42 monomers placed near the surface of a preformed fibril fragment composed of 20 A β 17-42 peptides. During the simulations, the monomers slide on the fibril surface, leading to

oligomer formation. The results show that the hydrophobic surface at the C-terminal of the fibril causes the unfolding of the monomers which catalyses the formation of beta-sheet rich oligomers on the fibril surface.⁹⁴ This suggests one possible molecular basis of the secondary nucleation pathway.

Zacharias *et al.* carried out atomistic simulations using the CHARMM27 force field to study the association of monomer, dimer and larger oligomers to the surface of a A β 9-40 fibril model. The results show that the association during the second nucleation follows a dock-lock mechanism consisting of a fast initial docking phase and a slow structural phase, and the hydrophobic effect is a major driving force for the association. Also, the surface induced association process is faster than the elongation at the fibril ends, indicating that secondary nucleation process can become the dominant process of amyloid formation in agreement with experiment.⁹⁵

Very interestingly, Wilson and colleagues studied the co-aggregation of A β 40 and A β 16-22 to obtain molecular insights into the surface-catalysed secondary nucleation of A β 40 by A β 16-22.³² Employing DMD with the CG PRIME20 force field, simulations were performed on different systems: six A β 40 monomers, six A β 16-22 monomers and six A β 40 monomers, and six A β 40 peptides in the presence of preformed two, three and four beta-sheet A β 16-22 protofilaments. The simulations show that while all primary/secondary nucleation and growth processes occur at the same time, secondary nucleation is dominant in A β 40 fibril formation kinetics during co-assembly with A β 16-22. Interestingly, although the fibril surface of A β 16-22 catalyses the assembly of A β 40, the assembly mechanism resembles that of A β 40 alone in solution.³²

6. Recent advances in structures of A β 40/42 oligomers from simulations

In this last section, we would like to review some recent and significant contributions in the field of A β 40/42 peptide simulations. We already discussed the simulation results of A β 42 dimer in aqueous solution (see Section 2).

In the case of A β 40 dimer in aqueous solution, Nguyen *et al.* performed atomistic REMD simulations on the wild-type (WT) sequence, the A2V/A2V mutant and the mixed WT/A2V mutant.⁹⁶⁻⁹⁸ Experimentally, the A2V mutation protects from Alzheimer's disease in its heterozygous form and leads to early AD in its homozygous form. Biophysical experimental showed that the aggregation rate follows the order A2V > WT > A2V-WT. It is found by REMD that the secondary structure content (18% β -strand and 10% α -helix) is invariant, but the patterns of intramolecular and intermolecular conformations drastically change upon single and double A2V mutation. In all sequences, the antiparallel and perpendicular peptide orientations are preferred over the parallel organization. However, upon single A2V mutation, the intermolecular potential energies are reduced, and the population of intramolecular three-stranded β -sheets and of multiple beta-

hairpins spanning the residues 17-21 and 30-36 is increased, providing an explanation for its slower aggregation kinetics. Similar computational results were obtained for the A β 40 WT/A2T dimer and the A β 42 A2T/A2T dimer, A2T mutation known to reduce A β aggregation kinetics and the risk of AD.^{98,99} Of particular interest is the application of a Markov state model to the transition pathways and associated kinetics resulting from a hybrid-resolution model and adaptive sampling technique that explores over 2.7 milliseconds of simulations of formation of A β 40 dimers. Indeed beta-hairpin-structures are formed by direct binding of soluble A β in β -hairpin conformations; while formation of parallel beta-structures resembling the U-fibril state results from A β monomers in arbitrary conformations and occurs 100-fold more slowly.¹⁰⁰

For larger aggregates, two computational studies in aqueous solution are worth discussing. First, using the predictive coarse-grained protein AWSEM force field, Wolynes *et al.* calculated and compared the free energy landscapes and relative stabilities of A β 42 and A β 40 in their monomeric up to octameric forms. At the same concentration, the aggregation free energy profile of A β 42 is more downhill and A β 42 has a computed solubility 10 times smaller than that of A β 40. While, at a concentration of 40 μ M, there is a clear free energy barrier between the pre-fibril tetramer form and the fibril pentamer in the A β 40 aggregation landscape, this is not the case for A β 42. It is to be noted that the fibril state depicts the U-shape. Using oligomerization maps that capture the paths of conversion between similar states of oligomers, several key differences are revealed. Moreover, the two C-terminal residues stabilize the oligomeric structures of A β 42, and facilitate the conversion from pre-fibril trimers to fibril tetramers.^{52,53} Second, atomistic MD simulations of 20 A β disordered chains in implicit solvent followed by transition networks analysis showed that pathways of A β 40 and A β 42 aggregation depend on oligomer shape, with compact and extended configurations and different solvent-exposures of hydrophobic residues.¹⁰¹

Numerous experiments indicate that toxicity could involve pore-forming A β oligomers in membranes, but aggregate heterogeneity has prevented experimental high-resolution structure determination. Nussinov and coworkers have modeled annular pores of 1.7-2.5 nm inner diameters consistent with atomic force microscopy experiments.¹⁰² Their channels were built from the assembly of the U-shape fibril conformation of A β peptide. Recent biophysical experiments in membrane-mimicking environments also suggest A β 42 pores with different oligomer sizes, and notably tetramers¹⁰³ and hexamers.¹⁰⁴ Using atomistic REMD simulations, low-resolution data obtained in lipid bilayers, and other theoretical factors, Derreumaux and coworkers designed 3D structures of A β 40 and A β 42 β -barrels in a bilayer mimicking a neuronal membrane.¹⁰⁵ The tetrameric model with two distinct β -hairpin motifs, eight antiparallel β -strands and an inner pore diameter of 0.7 nm is highly stable in all-atom MD and REMD simulations for A β 42 and is much less likely for A β 40.¹⁰⁵ Overall, this result sheds light on the amyloid pore hypothesis and explains the higher toxicity of A β 42. The authors also reported

novel findings as the stability of this A β 40/42 beta-barrel in aqueous solution has not been tested before at the atomistic level, and is very different from previous computational studies.⁴⁵

7. Conclusions

Understanding how amyloid aggregates actually become toxic is truly a real challenge in developing a treatment for neurodegenerative diseases, as only monomers are nontoxic. Thus far, all molecules (antibodies and drugs) targeting amyloid-beta oligomers have failed to pass clinical trials. Many reasons have been put forward to explain this repetitive failure.^{106,107}

Atomistic and coarse-grained simulations with increased computer efficiency, improved force field accuracy,^{15,108} coupling to machine learning¹⁰⁹ and integration of the main cellular partners¹¹⁰⁻¹¹² are likely to lead to a better molecular understanding of all events involved in amyloid fibril formation and structural characterization of the A β and tau proteins aggregates in Alzheimer's disease. These molecular insights coupled to experiments such as microbubbles in combination with focused ultrasound to deliver drugs through blood-brain-barrier¹¹³ may help design more efficient inhibitors of AD toxicity.

AUTHOR INFORMATION

Corresponding Author

E-mail: philippe.derreumaux@tdtu.edu.vn

ORCID

Phuong H. Nguyen: 0000-0003-1284-967X

Fabio Sterpone: 0000-0003-0894-8069

Philippe Derreumaux: 0000-0001-9110-5585

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge support from DYNAMO ANR-11-LABX-0011. PhD thanks Université de Paris, CNRS and PSL.

REFERENCES

- (1) Chiti F, Dobson CM. Protein misfolding, amyloid formation, and human disease: A summary of progress over the last decade. *Annu Rev Biochem.* 2017; 86:27-68.
- (2) Petkova AT, Buntkowsky G, Dyda F et al. Solid state NMR reveals a pH-dependent antiparallel beta-sheet registry in fibrils formed by a beta-amyloid peptide. *J Mol Biol.* 2004; 335:247-60.
- (3) Melquiond A, Boucher G, Mousseau N et al. Following the aggregation of amyloid-forming peptides by computer simulations. *J Chem Phys.* 2005; 122:174904.
- (4) Gazit E. Reductionist approach in peptide-based nanotechnology. *Annu Rev Biochem.* 2018; 87:533-553.
- (5) Gour N, Kanth P C, Koshti B et al. Amyloid-like structures formed by single amino acid self-assemblies of cysteine and methionine. *ACS Chem Neurosci.* 2019 ; 10:1230-1239.
- (6) Nasica-Labouze J, Nguyen PH, Sterpone F et al. Amyloid beta protein and Alzheimer's disease: when computer simulations complement experimental studies. *Chem Rev.* 2015; 115:3518-3563.
- (7) Meisl G, Yang X, Frohm B et al. Quantitative analysis of intrinsic and extrinsic factors in the aggregation mechanism of Alzheimer-associated A β -peptide. *Sci. Rep.* 2016; 6 : 18728.
- (8) Selkoe DG, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* 2016; 8: 595-608.
- (9) Lu Y, Shi XF, Nguyen PH et al. Amyloid- β (29-42) dimeric conformations in membranes rich in omega-3 and omega-6 polyunsaturated fatty acids. *J Phys Chem B.* 2019; 123:2687-2696.
- (10) Galvagnion C, Brown JW, Ouberai MM et al. Chemical properties of lipids strongly affect the kinetics of the membrane-induced aggregation of α -synuclein. *Proc Natl Acad Sci U S A.* 2016 ; 113:7065-7070.
- (11) Nguyen PH, Tarus B, Derreumaux P. Familial Alzheimer A2 V mutation reduces the intrinsic disorder and completely changes the free energy landscape of the A β 1-28 monomer. *J Phys Chem B.* 2014; 118:501-510.
- (12) Levine ZA, Shea JE. Simulations of disordered proteins and systems with conformational heterogeneity. *Curr Opin Struct Biol.* 2017 ; 43:95-103.
- (13) Zhang M, Ren B, Chen H et al. Molecular simulations of amyloid structures, toxicity and inhibition. *Isr. J. Chem.* 2017 ; 57: 586-601.

- (14) Ilie IM, Caflisch A. Simulation studies of amyloidogenic polypeptides and their aggregates. *Chem Rev.* 2019, Apr 11. doi: 10.1021/acs.chemrev.8b00731.
- (15) Robustelli P, Piana S, Shaw DE. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A.* 2018; 115:E4758-E4766.
- (16) Leahy CT, Kells A, Hummer G et al. Peptide dimerization-dissociation rates from replica exchange molecular dynamics. *J Chem Phys.* 2017; 147:152725.
- (17) Narayan B, Herbert C, Yuan Y et al. Conformational analysis of replica exchange MD: Temperature-dependent Markov networks for FF amyloid peptides. *J Chem Phys.* 2018; 149:072323.
- (18) Irback A, Jonsson S, Linnemann N et al. Aggregate geometry in amyloid fibril nucleation. *Phys. Rev. Lett.* 2013; 110: 058101.
- (19) Šarić A, Michaels TCT, Zaccone A et al. Kinetics of spontaneous filament nucleation via oligomers: Insights from theory and simulation. *J Chem Phys.* 2016; 145:211926
- (20) Bieler NS, Knowles TP, Frenkel D et al. Connecting macroscopic observables and microscopic assembly events in amyloid formation using coarse grained simulations. *PLoS Comput Biol.* 2012; 8:e1002692
- (21) Tran TT, Nguyen PH, Derreumaux P. Lattice model for amyloid peptides: OPEP force field parametrization and applications to the nucleus size of Alzheimer's peptides. *J Chem Phys.* 2016; 144:205103.
- (22) Sterpone F, Doutréline S, Tran TT et al. Multi-scale simulations of biological systems using the OPEP coarse-grained model. *Biochem Biophys Res Commun.* 2018; 498:296-304.
- (23) Li MS, Klimov DK, Straub JE et al. Probing the mechanisms of fibril formation using lattice models. *J. Chem. Phys.* 2008 ; 129 : 175101.
- (24) Tuffery P, Derreumaux P. Flexibility and binding affinity in protein-ligand, protein-protein and multi-component protein interactions: limitations of current computational approaches. *J R Soc Interface.* 2012; 9:20-33.
- (25) Stirnemann G, Sterpone F. Recovering protein thermal stability using all-atom hamiltonian replica-exchange simulations in explicit solvent. *J Chem Theory Comput.* 2015; 11:5573-5577.

- (26) Nasica-Labouze J, Meli M, Derreumaux P, et al. A multiscale approach to characterize the early aggregation steps of the amyloid-forming peptide GNNQQNY from the yeast prion sup-35. *PLoS Comput Biol.* 2011 ; 7(5):e1002051.
- (27) Smith AK, Lockhart C, Klimov DK. Does replica exchange with solute tempering efficiently sample A β peptide conformational ensembles? *J Chem Theory Comput.* 2016; 12 : 5201-5214.
- (28) Nguyen PH, Okamoto Y, Derreumaux P. Communication: Simulated tempering with fast on-the-fly weight determination. *J Chem Phys.* 2013; 138:061102.
- (29) Zhang T, Nguyen PH, Nasica-Labouze J et al. Folding atomistic proteins in explicit solvent using simulated tempering. *J Phys Chem B.* 2015; 119:6941-6951.
- (30) Hoang Viet M, Derreumaux P, Nguyen PH. Communication: Multiple atomistic force fields in a single enhanced sampling simulation. *J Chem Phys.* 2015; 143:021101.
- (31) Ge X, Sun Y, Ding F. Structures and dynamics of β -barrel oligomer intermediates of amyloid-beta₁₆₋₂₂ aggregation. *Biochim Biophys Acta Biomembr.* 2018; 1860:1687-1697.
- (32) Bunce SJ, Wang Y, Stewart KL et al. Molecular insights into the surface-catalyzed secondary nucleation of amyloid- β_{40} (A β_{40}) by the peptide fragment A β_{16-22} . *Sci Adv.* 2019; 5(6):eaav8216.
- (33) Granata D, Baftizadeh F, Habchi J et al. The inverted free energy landscape of an intrinsically disordered peptide by simulations and experiments. *Sci Rep.* 2015; 5:15449.
- (34) Baftizadeh F, Biarnes X, Pietrucci F et al. Multidimensional view of amyloid fibril nucleation in atomistic detail. *J Am Chem Soc.* 2012; 134:3886-3894.
- (35) Huang J, Rauscher S, Nawrocki G, Ran T et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods.* 2017; 14:71-73.
- (36) Best RB. Atomistic force fields for proteins. *Methods Mol Biol.* 2019; 2022:3-19.
- (37) Nguyen PH, Li MS, Derreumaux P. Effects of all-atom force fields on amyloid oligomerization: replica exchange molecular dynamics simulations of the A β (16-22) dimer and trimer. *Phys Chem Chem Phys.* 2011; 13:9778-9988.

- (38) Man VH, He X, Derreumaux P et al. Effects of all-atom molecular mechanics force fields on amyloid peptide assembly: The case of A β ₁₆₋₂₂ dimer. *J Chem Theory Comput.* 2019 ; 15:1440-1452.
- (39) Carballo-Pacheco M, Ismail AE, Strodel B. On the applicability of force fields to study the aggregation of amyloidogenic peptides using molecular dynamics simulations. *J Chem Theory Comput.* 2018; 14:6063-6075.
- (40) Siwy CM, Lockhart C, Klimov DK. Is the conformational ensemble of Alzheimer's A β ₁₀₋₄₀ peptide force field dependent? *PLoS Comput Biol.* 2017; 13:e1005314.
- (41) Rosenman DJ, Wang C, García AE. Characterization of A β monomers through the convergence of ensemble properties among simulations with multiple force fields. *J Phys Chem B.* 2016 ; 120 : 259-277.
- (42) Carballo-Pacheco, M.; Strodel, B. Comparison of force fields for Alzheimer's A β ₄₂: a case study for intrinsically disordered proteins. *Protein Sci.* 2017; 26: 174-185.
- (43) Weber OC, Uversky VN. How accurate are your simulations? Effects of confined aqueous volume and AMBER FF99SB and CHARMM22/CMAP force field parameters on structural ensembles of intrinsically disordered proteins: Amyloid- β ₄₂ in water. *Intrinsically Disord Proteins.* 2017; 5:e1377813.
- (44) Man VH, Nguyen PH, Derreumaux P. High-resolution structures of the amyloid- β 1-42 dimers from the comparison of four atomistic force fields. *J Phys Chem B.* 2017; 121:5977-5987.
- (45) Nguyen PH, Campanera JM, Ngo ST et al. Tetrameric A β ₄₀ and A β ₄₂ β -barrel structures by extensive atomistic simulations. II. In aqueous solution. *J Phys Chem B.* 2019 ; 123:6750-6756.
- (46) Pellarin R, Caflisch, A. Interpreting the aggregation kinetics of amyloid peptides. *J. Mol. Biol.* 2006; 360 : 882 – 892.
- (47) Bellesia G, Shea, J-E. Effect of β -sheet propensity on peptide aggregation. *J. Chem.Phys.* 2009 ; 130 : 145103.
- (48) Melquiond A, Mousseau N, Derreumaux P. Structures of soluble amyloid oligomers from computer simulations. *Proteins.* 2006; 65:180-191.
- (49) Song W, Wei G, Mousseau N, Derreumaux P. Self-assembly of the beta2-microglobulin NHVTLSQ peptide using a coarse-grained protein model reveals a beta-barrel species. *J Phys Chem B.* 2008 ; 112:4410-4418.

- (50) Lu Y, Derreumaux P, Guo Z et al. Thermodynamics and dynamics of amyloid peptide oligomerization are sequence dependent. *Proteins*. 2009; 75:954-963.
- (51) Urbanc B, Betnel M, Cruz, L et al. Elucidation of amyloid β -protein oligomerization mechanisms: discrete molecular dynamics study. *J. Am. Chem. Soc.* 2010, 132: 4266-4280.
- (52) Weihua Zheng, M. C., Min-Yeh Tsai; Wolynes, P. G. Exploring the aggregation free energy landscape of the amyloid-b protein (1-40). *Proc. Natl. Acad. Sci. USA* 2016 ; 113 : 11835-11840.
- (53) Zheng, W.; Tsai, M.-Y.; Wolynes, P. G. Comparing the aggregation free energy landscapes of amyloid beta(1-42) and amyloid beta(1-40). *J. Am. Chem. Soc.* 2017 ; 139 : 16666-16676.
- (54) Derreumaux P. Generating ensemble averages for small proteins from extended conformations by Monte Carlo simulations. *Phys Rev Lett* 2001; 85:206-209.
- (55) Wei GH, Derreumaux P, Normand M. Sampling the complex energy landscape of a simple beta-hairpin. *J. Chem Phys.* 2003; 119: 6403-6406.
- (56) Sterpone F, Melchionna S, Tuffery P. et al. The OPEP protein model: from single molecules, amyloid formation, crowding and hydrodynamics to DNA/RNA Systems. *Chem. Soc. Rev.* 2014; 43: 4871-4893.
- (57) Côté S, Laghaei R, Derreumaux P et al. Distinct dimerization for various alloforms of the amyloid-beta protein: $A\beta(1-40)$, $A\beta(1-42)$, and $A\beta(1-40)(D23N)$. *J Phys Chem B.* 2012; 116: 4043-4055.
- (58) Gnanakaran S, Nussinov R, García AE. Atomic-level description of amyloid beta-dimer formation. *J Am Chem Soc.* 2006; 128:2158-2159.
- (59) Santini S, Wei G, Mousseau N, Derreumaux P. Pathway complexity of Alzheimer's beta-amyloid Abeta16-22 peptide assembly. *Structure.* 2004 ; 12:1245-1255.
- (60) Santini S, Mousseau N, Derreumaux P. In silico assembly of Alzheimer's Abeta16-22 peptide into beta-sheets. *J Am Chem Soc.* 2004; 126:11509-11516.
- (61) Petty SA, Decatur SM. Experimental evidence for the reorganization of beta-strands within aggregates of the Abeta(16-22) peptide. *Proc Natl Acad Sci U S A.* 2005; 102:14272-14277.

- (62) Li X, Lei J, Qi R, Xie L et al. Mechanistic insight into E22Q-mutation-induced antiparallel-to-parallel β -sheet transition of A β ₁₆₋₂₂ fibrils: an all-atom simulation study. *Phys Chem Chem Phys*. 2019 ; 21 :15686-15694.
- (63) Sengupta U, Carballo-Pacheco M, Strodel B. Automated Markov state models for molecular dynamics simulations of aggregation and self-assembly. *J Chem Phys*. 2019 ; 150 :115101.
- (64) Matthes D, Gapsys V, Brennecke JT, de Groot BL. An atomistic view of amyloidogenic self-assembly: structure and dynamics of heterogeneous conformational states in the pre-nucleation phase. *Sci Rep*. 2016 ; 6:33156.
- (65) Sun Y, Ge X, Xing Y, Wang B, Ding F. β -barrel oligomers as common intermediates of peptides self-assembling into cross- β aggregates *Sci Rep*. 2018 ; 8:10353.
- (66) De Simone A, Derreumaux P. Low molecular weight oligomers of amyloid peptides display beta-barrel conformations: a replica exchange molecular dynamics study in explicit solvent. *J Chem Phys*. 2010 ; 132:165103.
- (67) Nguyen PH, Derreumaux P. Conformational ensemble and polymorphism of the all-atom Alzheimer's A β (37-42) amyloid peptide oligomers. *J Phys Chem B*. 2013; 117:5831-5840.
- (68) Barz B, Olubiyi O, Strodel, B. Early amyloid b-protein aggregation precedes conformational change. *Chem. Commun*. 2014 ; 50:5373–5375.
- (69) Saric A, Chebaro YC, Knowles TPJ et al. Crucial role of non specific interactions in amyloid nucleation. *Proc. Natl. Acad. Sci. USA* 2014 ; 111 : 17869–17874.
- (70) Latshaw DC, Cheon M, Hall CK. Effects of macromolecular crowding on amyloid beta (16-22) aggregation using coarse-grained simulations. *J Phys Chem B*. 2014; 118:13513-13526.
- (71) Sterpone F, Derreumaux P, Melchionna S. Protein simulations in fluids: coupling the OPEP coarse-grained force field with hydrodynamics. *J. Chem. Theory Comput*. 2015 ; 11 : 1843–1853.
- (72) Languin-Cattoën O, Melchionna S, Derreumaux P et al. Three weaknesses for three perturbations: comparing protein unfolding under shear, force, and thermal stresses. *J Phys Chem B*. 2018 ; 122:11922-11930.
- (73) Succi S. The lattice Boltzmann equation for fluid dynamics and beyond; ClarendonPress, 2001.

- (74) Chiricotto M, Melchionna S, Derreumaux P et al. Hydrodynamic effects on β -amyloid (16-22) peptide aggregation. *J. Chem. Phys.* 2016; 145 : 035102.
- (75) Chebaro Y, Pasquali S, Derreumaux P. The coarse-grained OPEP force field for non-amyloid and amyloid proteins. *J Phys Chem B.* 2012; 116:8741-8752.
- (76) Ando T, Skolnick J. On the importance of hydrodynamic interactions in lipid membrane formation. *Biophys. J.* 2013 ; 104 : 96 – 105.
- (77) Chiricotto M, Melchionna S, Derreumaux P et al. Multiscale aggregation of the amyloid Ab16–22 peptide: From disordered coagulation and lateral branching to Amorphous prefibrils. *J. Phys. Chem. Lett.* 2019; 10:1594–1599.
- (78) Lasagna-Reeves CA, Glabe CG, Kaye R. Amyloid- β annular protofibrils evade fibrillar fate in Alzheimer disease brain *J Biol Chem.* 2011 ; 286:22122-22130.
- (79) Oosawa F, Kasai M. A theory of linear and helical aggregations of macromolecules. *J. Mol. Biol.* 1962 ; 4 : 10–21.
- (80) Hofrichter J, Ross PD, Eaton W. A. Kinetics and mechanisms of deoxyhemoglobin S gelation: a new approach to understanding sickle cell disease. *Proc. Natl. Acad. Sci. USA* 1974 ; 71 : 4864–4868.
- (81) Ferrone, F. Analysis of protein aggregation kinetics. *Methods Enzymol* 1999 ; 309 : 256– 274.
- (82) Tornquist M, Michaels TCT, Sanagavarapu K. et al.; Secondary nucleation in amyloid formation. *Chem. Commun* 2018 ; 54 : 8667–8684.
- (83) Linse, S. Mechanism of amyloid protein aggregation and the role of inhibitors. *Pure Apply. Chem.* 2019 ; 91 : 211–229.
- (84) Meisl G, Yang X, Hellstrand E et al. Differences in nucleation behavior underlie the contrasting aggregation kinetics of the A β 40 and A β 42 peptides. *Proc. Natl. Acad. Sci. USA* 2014 ; 111 : 9384–9389.
- (85) Dammers C, Schwarten M, Buell AK et al. Pyroglutamate-modified A β (3-42) affects aggregation kinetics of A β (1-42) by accelerating primary and secondary pathways. *Chem. Sci.* 2017 ; 8 : 4996.
- (86) Yang X, Meisl G, Frohm B et al. On the role of sidechain size and charge in the aggregation of A β 42 with familial mutations. *Proc. Natl. Acad. Sci. USA* 2018 ; 115 :

E5849–E5858.

(87) Nguyen PH, Derreumaux, P. Understanding amyloid fibril nucleation and A β oligomer/drug interactions from computer simulations. *Acc Chem Res.* 2014 ; 47 : 603-611.

(88) Co NT, Li MS. New method for determining size of critical nucleus of fibril formation of polypeptide chains. *J. Chem. Phys.* 2012 ; 137 : 095101.

(89) Abeln S, Vendruscolo M, Dobson C et al. Frenkel, A simple lattice model that captures protein folding, aggregation and amyloid formation. *PloS ONE* 2014 ; 9 : e85185.

(90) Linse B, Linse S. Monte Carlo simulations of protein amyloid formation reveal origin of sigmoidal aggregation kinetics. *Mol. BioSyst.* 2011 ; 7 : 2296–2303.

(91) Hsieh MC, Lynn DG, Grover MA. Kinetic model for two-step nucleation of peptide assembly. *J. Phys. Chem. B* 2017 ; 121 : 7401–7411.

(92) Cheon M, Chang I, Hall CK. Spontaneous formation of twisted A β _{16–22} fibrils in large-scale molecular dynamics simulations. *Biophys J.* 2011 ; 101 : 2493–2501.

(93) Baftizadeh F, Pietrucci F, Biarnes X et al. Laio, Nucleation process of a fibril precursor in the C-terminal segment of amyloid- β . *Phys. Rev. Lett.* 2013 ; 110, 168103.

(94) Barz B, Strodel B. Understanding amyloid- β oligomerization at the molecular level: the role of the fibril surface. *Chem. Eur. J.* 2016 ; 22 : 8768–8772.

(95) Schwierz N, Frost CV, Geissler PL et al. From A β filament to fibril: molecular mechanism of surface activated secondary nucleation from all-atom MD simulations. *J. Phys. Chem. B* 2017 ; 121 : 671–682.

(96) Tarus B, Tran TT, Nasica-Labouze J et al. Structures of the Alzheimer's wild-type A β ₁₋₄₀ dimer from atomistic simulations. *J Phys Chem B.* 2015 ; 119 : 10478-10487.

(97) Nguyen PH, Sterpone F, Campanera JM et al. Impact of the A2V mutation on the heterozygous and homozygous A β ₁₋₄₀ dimer structures from atomistic simulations. *ACS Chem Neurosci.* 2016; 7: 823-832.

(98) Nguyen PH, Sterpone F, Pouplana R et al. Dimerization mechanism of Alzheimer A β ₄₀ peptides: the high content of intrapeptide-stabilized

conformations in A2V and A2T heterozygous dimers retards amyloid fibril formation. *J Phys Chem B*. 2016 ; 120 : 12111-12126.

(99) Das P, Chacko AR, Belfort G. Alzheimer's protective cross-interaction between wild-type and A2T variants alters A β ₄₂ dimer structure. *ACS Chem Neurosci*. 2017 ; 8 :606-618.

(100) Cao Y, Jiang X, Han W. Self-assembly pathways of β -sheet-rich amyloid- β (1-40) dimers: markov state model analysis on millisecond hybrid-resolution simulations. *J Chem Theory Comput*. 2017; 13:5731-5744.

(101) Barz B, Liao Q, Strodel B. Pathways of amyloid- β aggregation depend on oligomer shape. *J Am Chem Soc*. 2018; 140:319-327.

(102) Jang H, Arce FT, Ramachandran S et al. Disordered amyloidogenic peptides may insert into the membrane and assemble into common cyclic structural motifs. *Chem Soc Rev*. 2014 ; 43: 6750-6764.

(103) Serra-Batiste M, Ninot-Pedrosa M, Bayoumi M et al. A β ₄₂ assembles into specific β -barrel pore-forming oligomers in membrane-mimicking environments. *Proc. Natl. Acad. Sci. U. S. A*. 2016 ; 113 : 10866–10871.

(104) Österlund N, Moons R, Ilag LL et al. Native ion mobility-mass spectrometry reveals the formation of β -Barrel shaped amyloid- β hexamers in a membrane-mimicking environment. *J Am Chem Soc*. 2019 ; 141:10440-10450.

(105) Nguyen PH, Campanera JM, Ngo ST et al. Tetrameric A β ₄₀ and A β ₄₂ β -barrel structures by extensive atomistic simulations. I. In a bilayer mimicking a neuronal Membrane. *J Phys Chem B*. 2019 ; 123:3643-3648.

(106) Doig AJ, Derreumaux P. Inhibition of protein aggregation and amyloid formation by small molecules. *Curr Opin Struct Biol*. 2015 ; 30 : 50-56.

(107) Doig AJ, Del Castillo-Frias MP, Berthoumieu O et al. Why is research on amyloid- β failing to give new drugs for alzheimer's disease? *ACS Chem Neurosci*. 2017; 8:1435-1437

(108) Sterpone F, Nguyen PH, Kalimeri M et al. Importance of the ion-pair interactions in the OPEP coarse-grained force field: parametrization and validation. *J Chem Theory Comput*. 2013, 9 : 4574-4584.

(109) Noe F, Olsson S, Kohler J et al. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science* 2019, 365, eaaw1147.

(110) Menon S, Sengupta N. Influence of crowding and surfaces on protein amyloidogenesis: A thermo-kinetic perspective. *Biochim Biophys Acta Proteins*

Proteom. 2019; 1867:941-953.

(111) Owen MC, Gnutt D, Gao M et al. Effects of in vivo conditions on amyloid aggregation. *Chem Soc Rev.* 2019 ; 48:3946-3996.

(112) Chiricotto M, Sterpone F, Derreumaux P et al. 10. Multiscale simulation of molecular processes in cellular environments. *Philos Trans A Math Phys Eng Sci.* 2016; 374(2080).

(113) Man VH, Truong PM, Li MS et al. Molecular mechanism of the cell membrane pore formation induced by bubble stable cavitation. *J Phys Chem B.* 2019, 123 : 71-78.

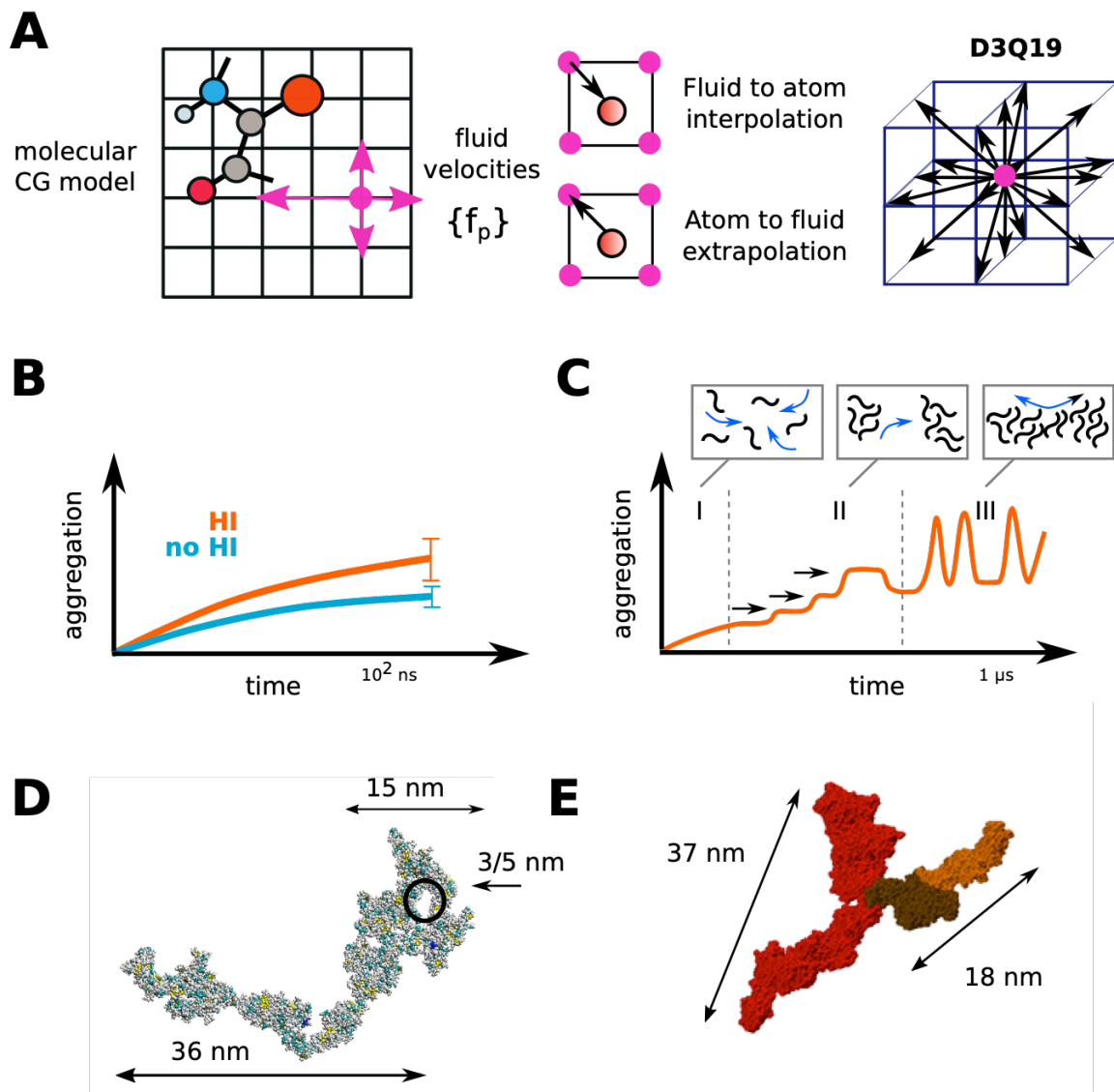


Figure 1: LBMD study of amyloid aggregation. (Panel A): Scheme describing the LBMD technique with particles embedded in a solvent. The set of velocities of the fluid (f_p) is represented by displacement in the lattice space (pink arrows). The drag force acting on the particle and its feedback on the evolution of the fluid velocity field require an interpolation/extrapolation communication strategy. In lattice Boltzmann, the typical 3D set of displacements is represented by 18 directions and the zero displacement (D3Q19). (Panel B): simulations of 100 A β 16–22 peptides. The curves represent the growth of the largest cluster during the simulation time including or not HI. The vertical bars schematically indicate the fluctuations of the cluster size. (Panel C): simulations of 1000 A β 16–22 peptides. The growth of the largest cluster manifests three regimes: individual peptide inclusion, cluster fusion, and fragmentation. (Panel D): amyloid amorphous protofibril sampled in LBMD, capped by an annular pore. (Panel E): Aggregation of a large amorphous protofibril can occur via lateral branching.

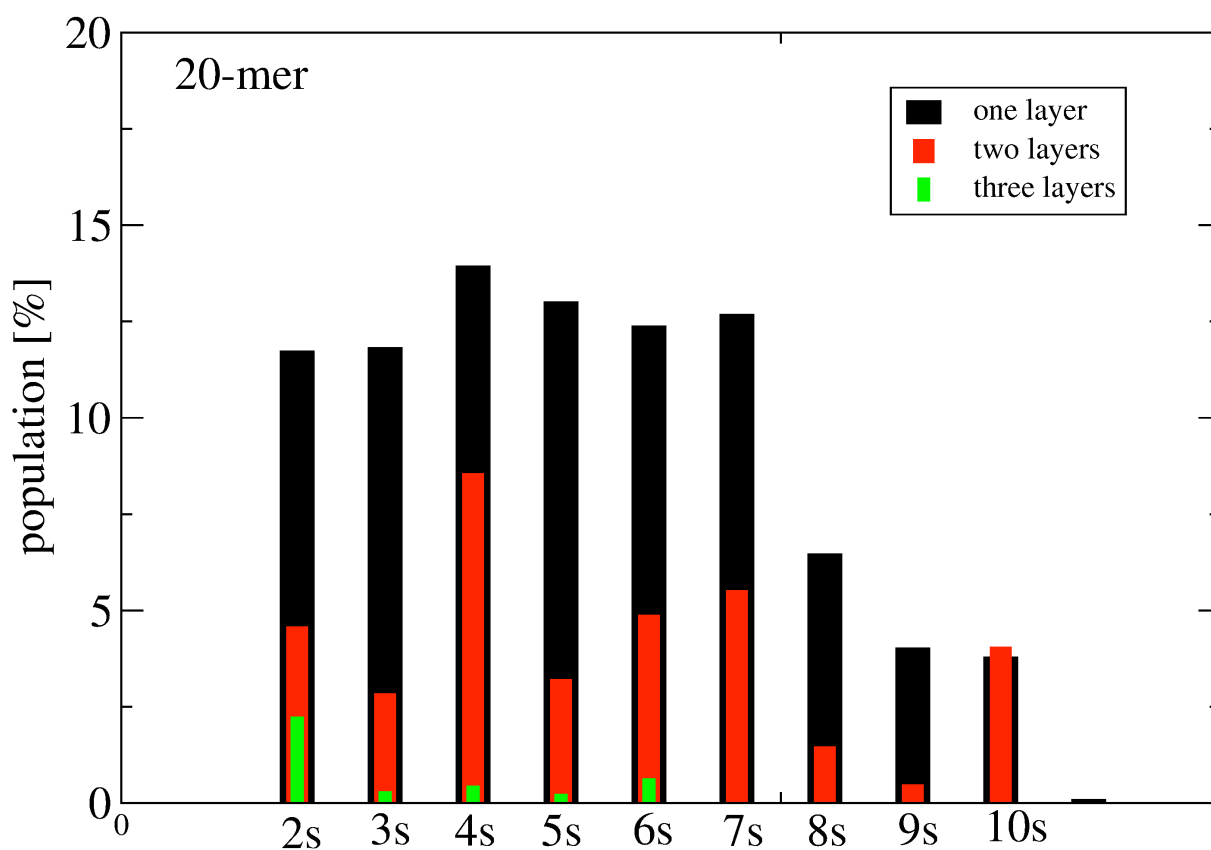


Figure 2. The population of one (black), two (red) and three (green) beta-sheet layers as a function of the number of beta-strands. Shown is the REMC result of 20-mer A β 37-42 below the melting temperature of the aggregates.