ACCOUNTS
of chemical research

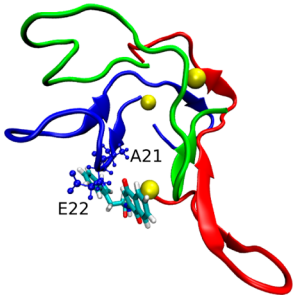
Understanding Amyloid Fibril Nucleation and Aβ Oligomer/Drug Interactions from Computer Simulations

PHUONG NGUYEN[†] AND PHILIPPE DERREUMAUX[*,†,‡]

[†]*Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Université Paris Diderot, Sorbonne Paris Cité, IBPC, 13 rue Pierre et Marie Curie, 75005 Paris, France, and* [‡]*Institut Universitaire de France, IUF, 103 Boulevard Saint-Michel, 75005 Paris, France*

RECEIVED ON AUGUST 21, 2013

CONSPECTUS



Evolution has fine-tuned proteins to accomplish a variety of tasks. Yet, with aging, some proteins assemble into harmful amyloid aggregates associated with neurodegenerative diseases, such as Alzheimer's disease (AD), which presents a complex and costly challenge to our society. Thus, far, drug after drug has failed to slow the progression of AD, characterized by the self-assembly of the 39−43 amino acid β-amyloid (Aβ) protein into extracellular senile plaques that form a cross-β structure. While there is experimental evidence that the Aβ small oligomers are the primary toxic species, standard tools of biology have failed to provide structures of these transient, inhomogeneous assemblies. Despite extensive experimental studies, researchers have not successfully characterized the nucleus ensemble, the starting point for rapid fibril formation. Similarly scientists do not have atomic data to show how the compounds that reduce both fibril formation and toxicity in cells bind to Aβ42 oligomers. In this context, computer simulations are important tools for gaining insights into the self-assembly of amyloid peptides and the molecular mechanism of inhibitors.

This Account reviews what analytical models and simulations at different time and length scales tell us about the dynamics, kinetics, and thermodynamics of amyloid fibril formation and, notably, the nucleation process. Though coarse-grained and mesoscopic protein models approximate atomistic details by averaging out unimportant degrees of freedom, they provide generic features of amyloid formation and insights into mechanistic details of the self-assembly process. The thermodynamics and kinetics vary from linear peptides adopting straight β-strands in fibrils to longer peptides adopting in parallel U shaped conformations in fibrils. In addition, these properties change with the balance between electrostatic and hydrophobic interactions and the intrinsic disorder of the system. However, simulations suggest that the critical nucleus size might be on the order of 20 chains under physiological conditions. The transition state might be characterized by a simultaneous change from mixed antiparallel/parallel β-strands with random side-chain packing to the final antiparallel or parallel states with the steric zipper packing of the side chains.

Second, we review our current computer-based knowledge of the 3D structures of inhibitors with Aβ42 monomer and oligomers, a prerequisite for developing new drugs against AD. Recent extensive all-atom simulations of Aβ42 dimers with known inhibitors such as the green tea compound epigallocatechin-3-gallate and 1,4-naphthoquinon-2-yl-L-tryptophan provide a spectrum of initial Aβ42/inhibitor structures useful for screening and drug design. We conclude by discussing future directions that may offer opportunities to fully understand nucleation and further AD drug development.

1. Introduction

Alzheimer's disease (AD) is the most common form of senile dementia, currently affecting 24 million people with an annual worldwide cost of 1 trillion US dollars. AD is characterized pathologically by abnormally high levels of neurofibrillary tangles resulting from the accumulation of tau protein in dead and dying neurons and by elevated numbers of extracellular amyloid plaques with cross β-structure in the cortex and hippocampus of the human brain. The major component of senile plaques is a small protein of 39−43 amino acids called amyloid-β (Aβ).[1] Any step to target the tau protein or to interfere with the production, self-assembly, and clearance of Aβ is considered a potential treatment for preventing or delaying the onset of AD, but thus far despite the number of inhibitors, vaccines, and compounds tested, all of them have failed to slow the

Published on the Web 12/24/2013 www.pubs.acs.org/accounts
10.1021/ar4002075

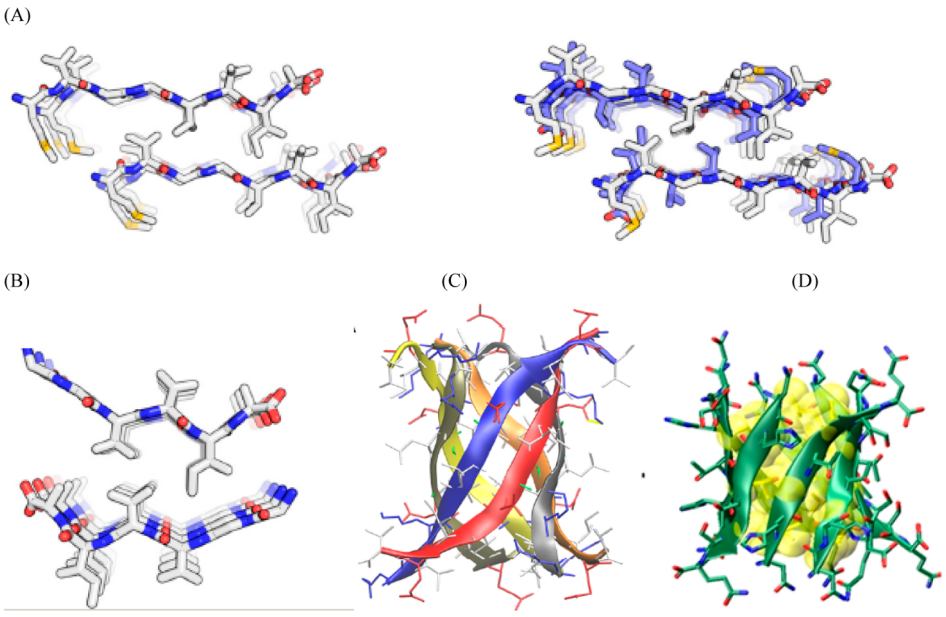


FIGURE 1. Polymorphisms of amyloid fibrils and structures of toxic oligomers: (A) microcrystal structures of Aβ35−42 fibrils composed of β-sheets with either parallel or antiparallel β-strands; (B) the in-register antiparallel β-strands of Aβ37−42 fibrils; (C) the hexamer of the KV11 peptide consisting of six antiparallel β strands forming a barrel;[14] (D) the predicted OPEP β-barrel of the β2m(82−87) peptide.[27,28]

progression of sporadic and familial AD.[2] Many scientists now believe that Aβ is the main culprit in the initiation with the discovery of the mutation A2T protecting against AD.[3] They also think that the drugs are given too late, they lack specificity and optimal binding affinities, and though Aβ40/42 small oligomers are the most toxic species, larger aggregates and fibril fragmentation contribute as well.[4,5]

The mechanism by which Aβ42 peptides of sequence DAEFRHDSGY^{10}EVHHQKLVFF^{20}AEDVGSNKGA^{30}IIGLMVGGVV^{40}IA, produced as soluble monomers, self-assemble reliably into amyloid fibrils is described experimentally and theoretically by a nucleation−condensation polymerization process.[6−8] Aβ contains two hydrophobic patches L17−A21 (central hydrophobic core, CHC) and A30−A42 separated by a hydrophilic patch E22−G29. Aβ42 mature fibrils display U-shaped conformation with β-strands formed by residues 17−20 and 31−40 and the 16 N-terminal residues disordered, whereas Aβ40 mature fibrils have β-strands at 10−23 and 30−38. Despite extensive experimental reports, the atomistic description of the nucleus from which Aβ fibril formation occurs rapidly after the lag phase remains to be determined. The physical basis of amyloid formation is also difficult to describe due to the sensitivity of the process to experimental conditions (pH, agitation, temperature, concentration, ionic strength, fibril growth, sample preparation), and the use of synthetic or AD-brain derived Aβ peptides.[5,9,10]

In addition, it is fully appreciated that the aggregation intermediates preceding the emergence of fibrils, called micellar or amorphous soluble oligomers, annular aggregates with central water-filled channels, and protofibrils, each class representing a group of states and comprising multiple conformations, vary from Aβ40 to Aβ42, and in the presence of familial (H6R, D7N, K16N, A21G, E22G, E22Q, E22K, E22Δ, and D23N) and nonfamilial (K16A, F19P, G33A, and G33I) mutations.[5,11] Both Aβ40 and Aβ42 fibrils possess a long, straight, and highly regular topology visible by transmission electron microscopy (EM) with a cross-β-sheet structure as demonstrated by X-ray diffraction with spacings at 4.7 and 10 Å, and in-register parallel β-sheets as deduced by solid-state NMR spectroscopy. However, the detailed molecular structures depend on the details of experimental conditions.[10] The variations that underline fibril polymorphism include 3-fold and 2-fold symmetries about the long Aβ40 fibril axis, and the transition from antiparallel to parallel β-sheets for the Aβ40-D23N fibril under conditions of repeating seeding.[9,10,12] The models in Figure 1 illustrate another polymorphism type for the Aβ35−42 fibrils and the current structure of Aβ37−42 fibrils, all displaying polar zipper interactions of the side chains.[13] Figure 1 also shows the microcrystal structure of the linear KV11 peptide oligomer with a β-barrel conformation.[14]

Despite all these data providing strong evidence that the morphology and the detailed structures of fibrils may

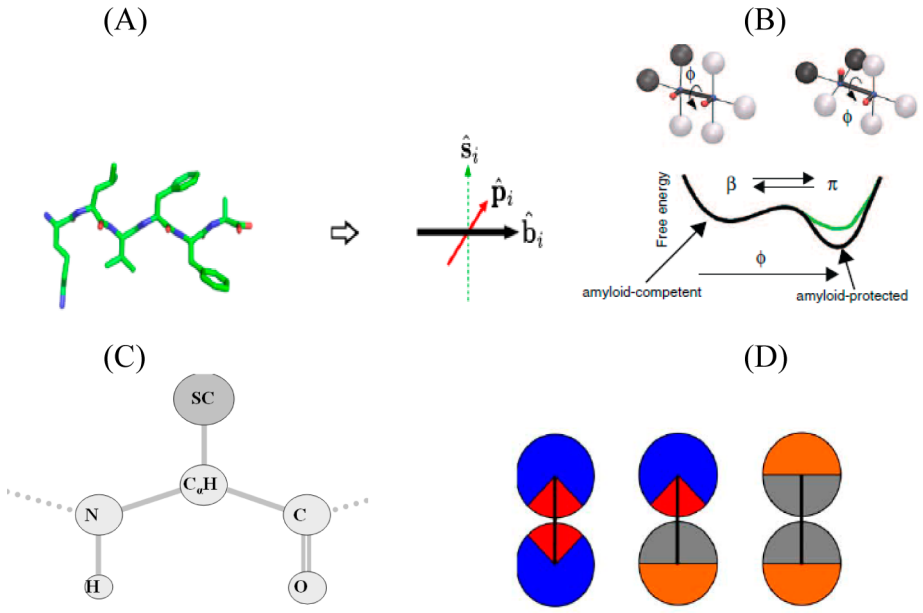


FIGURE 2. Main models. (A) Each peptide is centered at a lattice site, and its orientation is defined by the backbone vector **b** and the H-bond direction **p**, and the side chain direction is defined by the vector product **b** × **p**.[19] (B) Caflisch's model with the β and π minima of each monomer. The large spheres are hydrophobic (black) and hydrophilic (gray), and the two dipoles are shown with small red and blue spheres.[29] (C) OPEP model.[24] (D) Each peptide is represented as cylinder[30] occurring in R-state (random coil, blue circle with a 90° red patch) or β-state (a half orange/half gray circle). Here, we show the R2, Rβ, and β2 dimers.

originate from kinetics rather than thermodynamics and should be encoded in the diverse Aβ oligomers, the exact molecular assembly of a nascent fibril and the atomic structures of the monomers and oligomers of Aβ40/42 in the presence of drugs remain to be determined. Here, we report on recent computational results obtained on these two issues.

2. Coarse-Grained and Mesoscopic Models for Amyloid Proteins

In principle, classical molecular dynamics (MD) can contribute to a better understanding of how proteins fold or misfold. By invention of Anton, which can only do one thing, MD, but 100 faster than the best supercomputer, it has been possible using 1 ms trajectories to elucidate the atomic detail of how 15 proteins with 10−80 amino acids fold into their native states in explicit solvent, providing thereby insights into quantitative thermodynamic and kinetic data.[15]

Despite continuous progress in computer science and algorithmic development to enhance conformational sampling,[16,17] atomistic simulations of amyloid plaque formation are still out of reach because the process includes too many degrees of freedom and spans time scales of days *in vitro*. To overcome this limitation, several simplified models have been developed by averaging out most (on-lattice coarse grained, CG) or many unimportant degrees of freedom (off-lattice CG) or by retaining one single intramolecular degree of freedom (mesoscopic models). The basic idea of CG, which is to replace groups of atoms by a single bead, poses the problem of how to derive effective potentials that maintain the all-atom physical behavior in a physiological environment. Note that CG simulations accelerate the atomistic MD time in explicit solvent by 1−3 orders of magnitude depending on the level of granularities used.

The simplest CG model is on-lattice, where each chain consists of *M* connected beads that are confined to the vertices of a cube using an approximate description of the hydrophobic and electrostatic inter- and intrachain forces.[18] Another on-lattice cubic model uses peptides represented by unit-length sticks interacting by H-bonding interactions favoring slightly parallel over antiparallel β-strands and hydrophobic forces to mimic the zipper packing of the side chains (Figure 2A).[19] While various off-lattice models have been developed such as the flexible tube[20] and the Shea's model,[21] the most two widely used CG's for amyloids are the four-bead "Urbanc" or PRIME20 models[22,23] and the six-bead OPEP[24,25] (Figure 2C) model.

In PRIME20, each residue consists of three spheres for the backbone (NH, CαH, and CO) and one for the side chain. Backbone bond angles and Cα−Cα distances are enforced by pseudobonds, and the potential includes backbone H-bonding and side-chain−side-chain square-well potentials.

Using discontinuous MD (DMD), Urbanc reproduced differences in oligomer size distributions between Aβ40 and Aβ42 consistent with experiments.[22] Using PRIME20/DMD, Hall studied successfully a series of hexapeptides known to form nonamyloid and amyloid fibrils *in vitro*.[23] These two models were not tested, however, on systems with well-defined NMR structures in solution.

On the other hand, OPEP is basically an all-atom backbone with CG side chains. The nonbonded potential consists of two-body and four-body backbone H-bonding terms, attractive−repulsive or repulsive side-chain terms depending on the pairs of uncharged amino acids,[24] and all-atom PMF-derived potentials for the salt-bridges.[25] In contrast to other CG models, the OPEP force field has been extensively tested, by recovering the NMR structures of several nonamyloid peptides[24,25] and the temperature transition of peptides from coiled coils to amyloid aggregates.[25] Using OPEP, we were the first to identify reptation moves between the strands and observe β-barrels[26,27] (Figure 1D) during self-assembly of amyloid peptides that were validated by FTIR,[26] X-ray crystallography,[14] and other simulations.[21,28]

Finally, it is possible to use a mesoscopic description, where each peptide has one degree of freedom with two minima corresponding to amyloid-competent (β) and amyloid-protected (π or R) states and differing in free energy by dE.[29,30] Caflisch's model (Figure 2B), coupled to Langevin dynamics, offers the possibility to generate fibril topologies resembling those observed experimentally (e.g., twist and multifilament composition).[29] Frenkel's model (Figure 2D), coupled to dynamic Monte Carlo (MC), allows computing the free energy landscape of oligomers.[30] While both methods enable the calculation of the nucleus size, they lack sequence and atomistic details.

3. Insights into the Nucleation Process from Simulations

While master equations allow interpretion of the experimental sigmoidal kinetic profiles of amyloid formation by means of primary or secondary (fragmentation or lateral) nucleation processes,[7] they do not provide any information on the topology and size of the primary nucleus. What is clear from all simulations is that self-assembly starts by a two-step mechanism that is dependent on the hydrophobic character of the peptides. The first step involves a hydrophobic collapse and the formation of molten oligomers. In the second step, interpeptide hydrogen bonding interactions drive the system to highly flexible β-rich oligomers.[21,31,32]

In recent studies, the nucleus size, N^*, was determined from the dependence of the free energy on the number of GNNQQY and STVIYE monomers with OPEP[33] and all-atom[34] simulations, respectively. N^* was found to be 4−5 and 5−8 for these two linear peptides (LP). By following the aggregation of a mesoscopic LP by Langevin dynamics, Pellarin et al. showed the nucleus size or critical β-domain size with a probability of 50% to form a fibril to vary with dE, ranging from 4 (dE = −1.5 kcal/mol) to 35 peptides (dE = −2.5 kcal/mol, that is, for a π state 100 times more populated than the β state) and indicating the strong dependence of N^* and the lag phase on the energy landscape of the monomer.[29] Using lattice models, Thirumalai evidenced that the balance between electrostatic and hydrophobic interactions not only modulates the population of the amyloid-competent monomeric state and the lag phase, but also the topology of the fibrils.[35] Based on multiscale simulations, De Simone showed that a comprehensive description of the flexibility of the all states is also a modulator of self-assembly.[36]

Using on-lattice MC simulations of N chains with each chain consisting of eight particles as a model of Aβ42, Li found that N^* is 11, independent of the metrics used. Though the concentration is 2 orders higher than that used in experiments,[18] this result is similar to N^* of 16 found by Fawzi using CG simulations of Aβ40.[37] Notably, the atomistic nucleation theory shows that N* is 15 for Aβ40 at a protein concentration of 120 μM, but variation in the supersaturation of the phase can cause N^* to increase to 50.[38]

Though these simulations quantify the nucleus size, they do not provide any structural insights into the nucleus. One first move toward connecting macroscopic observables and microscopic assembly events with the nucleus size was provided by two simulations using low-resolution models. In the first, the aggregation thermodynamics and kinetics of a linear peptide were studied by using generalized-ensemble techniques and MC for 256 peptides on a cubic lattice with a concentration of 10^{-3} per unit volume. The aggregate mass distribution, $p(m)$, which gives the probability of an aggregate of size m, is bimodal at the midpoint temperature with $p(m<6) = 81.4\%$ and $p(m>62) = 16.7\%$. The free energy surface projected on the length and width of the aggregates and the evolution of the mass of the largest aggregate as a function of MC steps show that the free energy barriers faced by an aggregate to form a fibril are associated with changes in width, and the width of the aggregate prior to nucleation is 3.5.[19] These results, in line with the self-assembly of cuboids, show that nucleation is

directly linked to the formation of multiple sheets and, without intersheet interactions, the kinetics does not follow a sigmoidal shape.[39]

The second simulation, based on Frenkel's mesoscopic model with a dE of 9 kcal/mol, shows that the self-assembly of Aβ42 is described by a nucleation−conformational conversion process with a nucleus size of four chains and a nucleus characterized by a mixed aggregate with Rβ3 character.[30] The lowest energy path to the fibril consists of the following states of increasing oligomer size: R, R2, R3, R2β, R2β2, Rβ3, and β4. Given that N^* is much lower than what has been predicted by other models,[18,37,38] the real topology of N^* for Aβ42 remains to be determined.

A second move toward characterizing the nucleus at an atomistic level was provided by three simulations in explicit solvent. In the first two studies, Laio investigated the aggregation of 18 Val8 and 18 Aβ35−40 peptides employing bias-exchange metadynamics with eight collective variables so as to explore increasing number of fully parallel (P) or antiparallel (AP) β-sheets and steric zipper contacts.[17,40] In both systems, the basin of lowest free energy, B1, includes disordered or amorphous aggregates. In the Val8 system, an effective nucleus size on the order of 14 was proposed and the crossing of the maximum free energy is characterized by a transition from mixed P/AP to parallel β-strand orientations and only when a sufficient number of parallel sheets are formed the free energy starts to decrease to a minimum with fully parallel β-sheets (Figure 3A).[17] Interestingly, this transition to parallel sheets has been found also by CG simulations.[41] In the Aβ35−40 system, a similar but more complex picture emerges. Transitions between the basin B1 and the basin B2 (containing antiparallel β-sheets with up to 10−12 β-strands and random side-chain packing) take a few tens of nanoseconds, and the rate-limiting step in the nucleation pathway involves crossing a barrier of 40 kcal/mol associated with the formation of parallel β-sheets closely packed with a steric zipper of the side chains, slightly different however from the experimental final product.[40]

In the third study, an unbiased all-atom REMD simulation of 16 Aβ37−42 peptides in explicit water was performed using 48 replicas, each for 460 ns.[42] This peptide blocked by opposite charge at both extremities is particularly intriguing because it forms fibrils with antiparallel sheets and parallel β-strands (Figure 1B).[13] Despite frequent β-sheet formation/fragmentation events and a significant reservoir of free monomers, the population of four to five fully P β-strands, consistent with the fibril structure, is 1−2%. The population of four to five fully AP β-strands is higher (3−8%), and overall

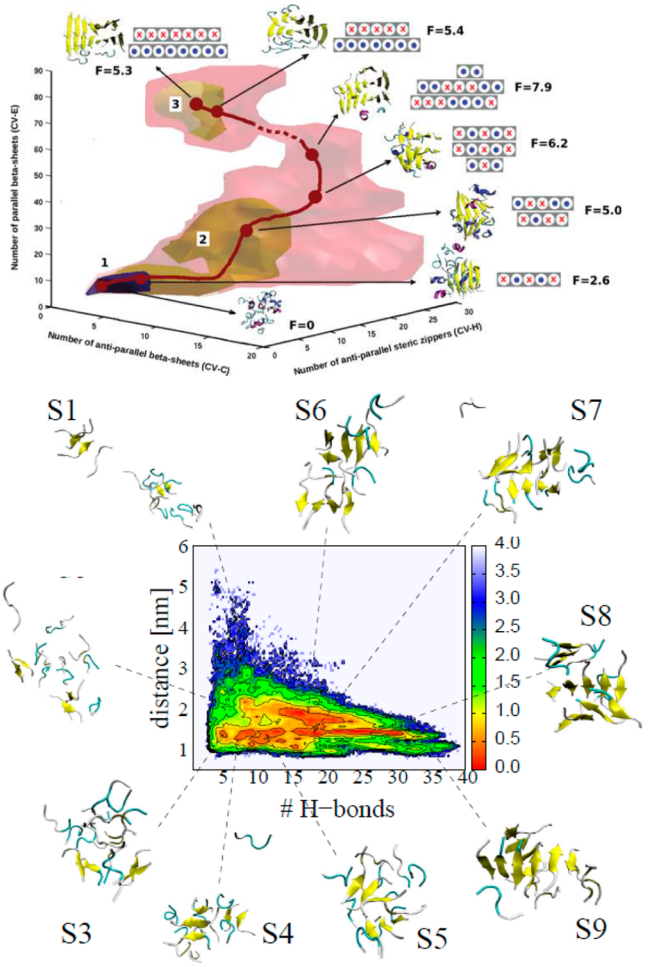


FIGURE 3. Free energy surfaces (FES) of the all-atom systems formed by 18Val8 and 16Aβ37−42 peptides in explicit solvent. (Top) FES for 18Val8 as a function of the number of parallel and antiparallel β-sheets and zipper contacts.[17] The red line shows the lowest free energy path from disordered states (region 1, $F = 0$) to the nucleus ($F = 7.9$), from which fully antiparallel β-sheets with parallel strands form ($F = 5.3$, region 3). Each box depicts a β-strand, and the central blue dot (or circle with a cross) shows a β-strand pointing outside (inside) the plane. (Bottom) FES for 16Aβ37−42 as a function of the mean distance between the centers of mass of β-sheets and the number of H-bonds within β-sheets.[42] The main free energy minima are shown with β-strands in yellow, and the color scale is given in kcal/mol.

the global free energy minimum consists of two and three β-sheets each of two to three β-strands with mixed AP/P strands and a variety of sheet-to-sheet pairing angles surrounded by random coil peptides (Figure 3B). Looking at higher free energy minima, the aggregates also consist of mixed P/AP β-strands, with few P β-strands surrounded by AP β-strands, in agreement with metadynamics simulations of Val8 and Aβ35−40 peptides[17,40] and REMD of GNNQ peptides.[43] Extending the simulation to 800 ns per replica does not change the picture and indicates that N^* is >5−8 for Aβ37−42. Whether N^* is around 15, as reported for other

LPs,[17,38] cannot be determined due to finite-size effects and the fact that polymorphism is under kinetic control.[9,10,29] Indeed, while Aβ35−42 fibril can display either P or AP β-strands within the sheets (Figure 1A),[13] this polymorphism has not been reported yet for Aβ37−42, a surprise given that Aβ34−42 fibril forms AP β-strands.[44]

4. Insights into the 3D Structures of Drug/Aβ Oligomers from Simulations

Many molecules have been screened against Aβ aggregation and toxicity, including Zn chelators, peptide-based derivatives of Aβ, and heterocyclic compounds such as cucurmin and the green tea compound epigallocatechin-3-gallate (EGCG). Recently, Landau reported microcrystal structure of Aβ16−21 fibril with the dye Orange G.[45] Simulations on Aβ16−22 protofibril reveal that small molecules can block both β-sheet elongation and lateral association of layers but also sequesters Aβ16−22 peptides.[46] Based on ^{13}C NMR spectroscopy and all-atom MD simulations, it was possible to obtain direct information on an inhibitor candidate in its direct complex with Aβ42 fibrils.[47]

Unfortunately, because Aβ42 has a high propensity to associate and explore a heterogeneous ensemble of conformations, standard tools of biology have failed to provide atomic structures of the monomer and oligomers of Aβ42 with inhibitors, and we only have low-resolution experimental data on oligomers. Using ion-mobility mass spectrometry, Bernstein reported a collision cross-section (CCS) of 1256 Å^2 for Aβ42 dimers.[5] Using different preparation methods and CD analysis, Teplow reported a β-strand content between 12% and 25% and an α-helix content between 3% and 9% at 295 K, pH 7.5, and day 0, for a mixture of various aggregates.[5,48] Using isothermal titration calorimetry at different EGCG and salt concentrations, Sun showed that the interactions between Aβ42 and EGCG are mainly hydrogen bonding in the region 1−16 and hydrophobic in the region 17−42.[49]

Among all compounds tested, 1,4-naphthoquinon-2-yl-L-tryptophan (NQTrp) reduces the toxicity of Aβ42 oligomers toward a cultured neuronal cell line and transgenic AD *Drosophila* model. The NMR structure of NQTrp bound to Aβ12−28 monomer at a molar ratio 0.5:1 combined with MD simulation shows three dominant binding sites between NQTrp and the Aβ18−21 (VFFA) region,[50] but whether these conformations hold for Aβ42 and oligomers remains to be determined. To address this issue, extensive CG simulations of Aβ17−42 trimer followed by all-atom docking of five molecules on the most populated Aβ structures show that

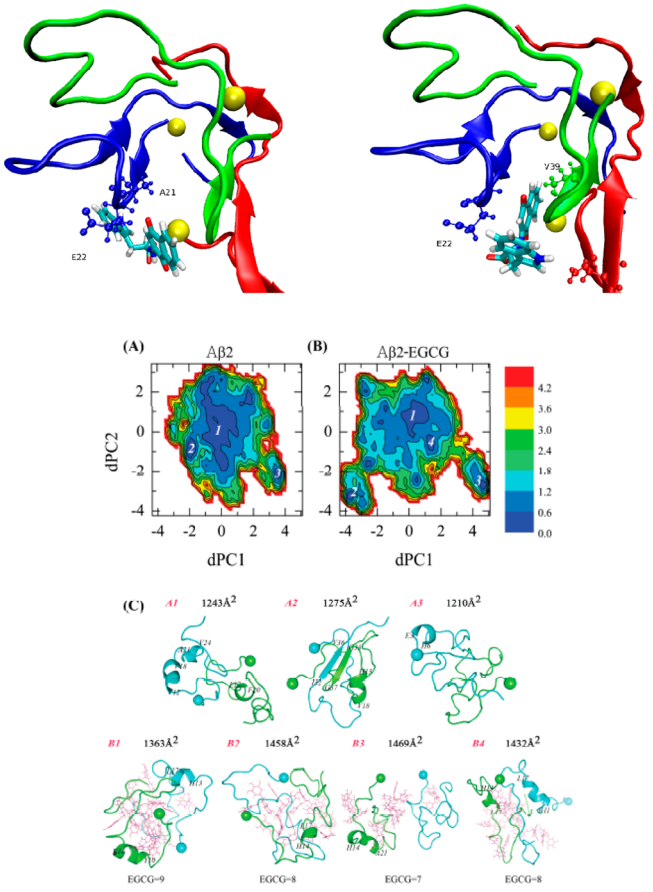


FIGURE 4. Aβ oligomers/drugs. (Top) Two binding modes of NQTrp to the Aβ17−42 trimer.[51] N-termini are located by yellow balls. (Bottom) FES in kcal/mol of Aβ42 dimer and Aβ42 dimer/EGCG with their three and four most populated structures. For each state, we give the CCS in Å^2 and the number of EGCG bound.[54]

the five drugs including cucurmin, EGCG, 2002-H20, and resveratrol have a lower binding affinity for the fibril than the other states, and NQTrp is the most favorable inhibitor.[51] The results also show that NQTrp can bind to four Aβ17−42 poses with similar energies and in all cases with the side chains of F19/F20 and the main chain atoms of F19−E22. The results of this hierarchical procedure, consistent with ref 50, show that NQTrp can have multiple binding modes to Aβ17−42 trimer in a 100 ns all-atom MD even within a given pocket (Figure 4A).

Experimentally, EGCG has been reported to redirect the aggregation of the Aβ42 peptide into unstructured, off pathway oligomers.[52] In contrast to this study, by using solution- and solid-state NMR spectroscopies, EM, and a 10-fold molar excess of EGCG, Reif captured stable spherical EGCG-induced Aβ40 oligomers of 2.1 nm in height that are not fully random coil. While the residues 1−20 are disordered, the salt bridge D23−K28 is formed, and the residues G29−V36 adopt

a β-strand as in the fibril. EGCG interacting with the CHC region prevents the formation of β-strand in the region 17−20 as seen in the fibril.[53]

Recently, an all-atom REMD study on Aβ42 dimer with the OPLS force field and explicit solvent in the absence and presence of EGCG molecules with a molar ratio 2:10 (Aβ/EGCG) as used experimentally was performed using 64 replicas, each for 200 ns.[54] Though different force fields may generate different free energy landscapes,[55] the results on the pure Aβ42 dimer are consistent with available low-resolution experimental data.[5,48] Analysis of secondary structure shows that the free Aβ42 dimer mostly populates coil/turn (79%), and then α-helix (11%) and β-strand (8%), in agreement with CD analysis. The centers of the three most populated clusters have collision cross sections of 1243, 1275, and 1210 Å^2, in agreement with the experimental value of 1256 Å^2 (Figure 4B). Upon EGCG binding, the bend, turn, and coil remain constant and the α-helix content spanning residues 12−18 is increased from 11% to 13%. Though the overall β-content is reduced from 8% to 4%, the reduction is significant in the N-terminal region (residues 1−16, varying from 10% to 1%), the CHC, and residues 39−42 (from 20% to 5%) resulting in a less-fibril state, while the β-strand of the residues 30−36 with a probability of 20% is not impacted. Note this β-strand distribution along the amino acid sequence in the complex is consistent with the NMR results of the EGCG−Aβ40 oligomers.[53]

Interestingly, in the presence of EGCG, the CHC/CHC and C-terminal/C-terminal interactions observed in pure Aβ42 dimer are greatly reduced, resulting in an increase of 8% of CCSs of Aβ dimer and a population of 5% of two separated monomers. EGCG are buried in the interface between the Aβ42 peptides and bind mostly to the hydrophobic residues of the CHC and C-terminal region through van der Waals interactions. EGCG also bind to the hydrophilic N-terminal amino acids D1/E3/R5/D7 and E11 through H-bonds, albeit H-bonds with residues A21/E22/D23 and I32 are present, consistent with isothermal titration calorimetry.[49]

Overall, the results of this REMD simulation on Aβ42 dimer with 10 EGCG are particularly interesting in designing new drugs for three reasons. First, they show that at equilibrium, there is a population of 5% of free Aβ42 monomers that would be able to associate to larger toxic and nontoxic aggregates. This indicates that though inhibiting Aβ42 aggregation at the level of the monomer is an attractive therapeutic approach, the results of simulations of Aβ42 monomer with inhibitor do not preclude the formation of Aβ42 larger aggregates. Using the 45 most populated clusters from REMD simulations of Aβ42 monomer in water, Zhu carried out fragment mapping calculations to identify binding hot spots. A total of 35 clusters display binding pockets made essentially of the CHC and the residues F4/Y10/I31 and M35, though hydrophilic residues participate in 8 binding pockets.[56] Whether the same key interactions would remain in a simulation of Aβ42 oligomers has to be determined.

Second, while simulations on Aβ12−28 or Aβ16−22 peptides with inhibitors may help rank compounds,[50,57,58] the results are certainly not transposable to Aβ40/42, because the N-terminal and C-terminal residues contribute to the binding energies[49,53,54] and a toxic Aβ40 oligomer of high molecular weight was discovered with a β-sheet at the N-terminus.[59]

Third, simulations on Aβ42/EGCG show that there is room for a better inhibitor that would bind more tightly to and sequester Aβ42 dimers, preventing therefore Aβ42 monomers to self-associate into toxic species. In this context, NQTrp is much more efficient since the population of free Aβ42 monomers is 0% using a molar ratio 1:1 and REMD simulation of 64 replicas, each of 250 ns, with the Amber ff99sb*-ildn force field and TIP3P water model.[60]

5. Conclusions

Overall, this report reviews over a decade computational results on amyloid formation obtained with all-atom, coarse-grained, and mesoscopic representations coupled with state-of-the-art sampling techniques. First, we provide a critical analysis of what we can tell on the nucleus size, N^*, and the conformational ensemble from which rapid fibril formation starts. While the thermodynamics and kinetics vary from linear peptides adopting straight β-strands in fibrils to longer peptides adopting in parallel U shaped conformations in fibrils and change with the balance between electrostatic and hydrophobic interactions and the intrinsic disorder of the system, it follows that the critical nucleus size might be on the order of 20 chains under physiological conditions. It also emerges that the transition state might be characterized by a simultaneous change from mixed antiparallel/parallel β-strands to the final antiparallel or parallel states and random packing to the steric zipper packing of the side chains. Second, we discuss the various predictions made by computer simulations and experiments to determine how current inhibitors bind to Aβ42 monomer and small oligomers, a prerequisite for developing more efficient drugs. Using multiple approaches, ranging from REMD simulations to multiscale approaches combining

exhaustive sampling, docking, and dynamics and with increasing computer resources, we should soon have on hand the structures of Aβ42 monomer to tetramer with many known inhibitors. Building on the full spectrum of Aβ42/inhibitor structures, we should be in position to screen and predict new drugs with much higher efficacy by using virtual screening and all-atom steered and metadynamics simulations. Along these lines, a combined experimental and theoretical effort must be made to understand the effects of hydrodynamics, shear, metals, and crowding (cell membrane and protein receptors) on the kinetics and thermodynamics of Aβ42 oligomers interacting with inhibitors.

This work was supported by grants from ANR GRAL SIMI 12-BS07–0017 and Pierre Gilles de Gennes Foundation.

BIOGRAPHICAL INFORMATION

P. Nguyen has been permanent CR1 CNRS researcher for three years. His research focuses on the development of methods to construct free energy surfaces and enhance sampling for protein aggregation.

P. Derreumaux is Distinguished Professor at Sorbonne Paris Cité, Director of UPR9080 and Senior Member of IUF. His research interests are in the areas of coarse-grained, force field, and algorithm developments for the simulation of proteins, RNA, and drug/protein complexes.

FOOTNOTES

*Corresponding author. E-mail: philippe.derreumaux@ibpc.fr.
The authors declare no competing financial interest.

REFERENCES

1 Selkoe, D. J. Folding proteins in fatal ways. *Nature* **2003**, *426*, 900–904.

2 Abbott, A. Neuroscience: The plaque plan. *Nature* **2008**, *458*, 161–164.

3 Jonsson, T.; Atwal, J. K.; Steinberg, S.; Snaedal, J.; Jonsson, P. V.; Bjornsson, S.; Stefansson, H.; Sulem, P.; Gudbjartsson, D.; Maloney, J.; Hoyte, K.; Gustafson, A.; Liu, Y.; Lu, Y.; Bhangale, T.; Graham, R. R.; Huttenlocher, J.; Bjornsdottir, G.; Andreassen, O. A.; Jönsson, E. G.; Palotie, A.; Behrens, T. W.; Magnusson, O. T.; Kong, A.; Thorsteinsdottir, U.; Watts, R. J.; Stefansson, K. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* **2012**, *488*, 96–99.

4 Shankar, G. M.; Li, S.; Mehta, T. H.; Garcia-Munoz, A.; Shepardson, N. E.; Smith, I.; Brett, F. M.; Farrell, M. A.; Rowan, M. J.; Lemere, C. A.; Regan, C. M.; Walsh, D. M.; Sabatini, B. L.; Selkoe, D. J. Amyloid-beta protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nat. Med.* **2008**, *14*, 837–842.

5 Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplow, D. B.; Shea, J. E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nat. Chem.* **2009**, *1*, 326–331.

6 Hellstrand, E.; Boland, B.; Walsh, D. M.; Linse, S. Amyloid β-protein aggregation produces highly reproducible kinetic data and occurs by a two-phase process. *ACS Chem. Neurosci.* **2010**, *1*, 13–18.

7 Knowles, T. P.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. An analytical solution to the kinetics of breakable filament assembly. *Science* **2009**, *326*, 1533–1537.

8 Schmit, J. D.; Gosh, K.; Dill, K. What drives amyloid molecules to assemble into oligomers and fibrils? *Biophys. J.* **2011**, *100*, 450–458.

9 Qiang, W.; Kelley, K.; Tycko, R. Polymorph-specific kinetics and thermodynamics of beta-amyloid fibril growth? *J. Am. Chem. Soc.* **2013**, *135*, 6860–6871.

10 Tycko, R.; Wickner, R. B. Molecular structures of amyloid and prion fibrils: Consensus versus controversy. *Acc. Chem. Res.* **2013**, *46*, 1487–1496.

11 Kaden, D.; Harmeier, A.; Weise, C.; Munter, L. M.; Althoff, V.; Rost, B. R.; Hildebrand, P. W.; Schmitz, D.; Schaefer, M.; Lurz, R.; Skodda, S.; Yamamoto, R.; Arlt, S.; Finckh, U.; Multhaup, G. Novel APP/Aβ mutation K16N produces highly toxic heteromeric Aβ oligomers. *EMBO Mol. Med.* **2012**, *4*, 647–659.

12 Tycko, R.; Sciarretta, K. L.; Orgel, J. P.; Meredith, S. C. Evidence for novel β-sheet structures in Iowa mutant β-amyloid fibrils. *Biochemistry* **2009**, *48*, 6072–6084.

13 Colletier, J. P.; Laganowsky, A.; Landau, M.; Zhao, M.; Soriaga, A. B.; Goldschmidt, L.; Flot, D.; Cascio, D.; Sawaya, M. R.; Eisenberg, D. Molecular basis for amyloid-β polymorphism. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 16938–16943.

14 Laganowsky, A.; Liu, C.; Sawaya, M. R.; Whitelegge, J. P.; Park, J.; Zhao, M.; Pensalfini, A.; Soriaga, A. B.; Landau, M.; Teng, P. K.; Cascio, D.; Glabe, C.; Eisenberg, D. Atomic view of a toxic amyloid small oligomer. *Science* **2012**, *335*, 1228–1231.

15 Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How fast-folding proteins fold. *Science* **2011**, *334*, 517–20.

16 Nguyen, P. H.; Okamoto, Y.; Derreumaux, P. Simulated tempering with fast on-the-fly weight determination. *J. Chem. Phys.* **2013**, *138*, No. 061102.

17 Baftizadeh, F.; Biarnes, X.; Pietrucci, F.; Affinito, F.; Laio, A. Multidimensional view of amyloid fibril nucleation in atomistic detail. *J. Am. Chem. Soc.* **2012**, *134*, 3886–3894.

18 Co, N. T.; Li, M. S. New method for determining size of critical nucleus of fibril formation of polypeptide chains. *J. Chem. Phys.* **2012**, *137*, No. 095101.

19 Irbäck, A.; Jonsson, S.ó.; Linnemann, N.; Linse, B.; Wallin, S. Aggregate geometry in amyloid fibril nucleation. *Phys. Rev. Lett.* **2013**, *110*, No. 058101.

20 Auer, S.; Dobson, C. M.; Vendruscolo, M.; Maritan, A. Self-templated nucleation in peptide and protein aggregation. *Phys. Rev. Lett.* **2008**, *101*, No. 258101.

21 Bellesia, G.; Shea, J. E. Diversity of kinetic pathways in amyloid fibril formation. *J. Chem. Phys.* **2009**, *131*, No. 111102.

22 Urbanc, B.; Betnel, M.; Cruz, L.; Bitan, G.; Teplow, D. B. Elucidation of amyloid β-protein oligomerization mechanisms: Discrete molecular dynamics study. *J. Am. Chem. Soc.* **2010**, *132*, 4266–4282.

23 Wagoner, V. A.; Cheon, M.; Chang, I.; Hall, C. K. Fibrillization propensity for short designed hexapeptides predicted by computer simulation. *J. Mol. Biol.* **2012**, *416*, 598–609.

24 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.

25 Sterpone, F.; Nguyen, P. H.; Kalimeri, M.; Derreumaux, P. Importance of the ion-pair interactions in the OPEP coarse-grained force field: parametrization and validation. *J. Chem. Theory Comput.* **2013**, *9*, 4574–4584.

26 Santini, S.; Mousseau, N.; Derreumaux, P. In silico assembly of Alzheimer's Aβ16−22 peptide into beta-sheets. *J. Am. Chem. Soc.* **2004**, *126*, 11509–11516.

27 Mousseau, N.; Derreumaux, P. Exploring the early steps of amyloid peptide aggregation by computers. *Acc. Chem. Res.* **2005**, *38*, 885–891.

28 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*, No. 165103.

29 Pellarin, R.; Schuetz, P.; Guarnera, E.; Caflisch, A. Amyloid fibril polymorphism is under kinetic control. *J. Am. Chem. Soc.* **2010**, *132*, 14960–14970.

30 Bieler, N. S.; Knowles, T. P. J.; Frenkel, D.; Vacha, R. Connecting macroscopic observables and microscopic assembly events in amyloid formation using coarse grained simulations. *PLoS Comput. Biol.* **2012**, *8*, No. e1002692.

31 Matthes, D.; Gapsys, V.; de Groot, B. L. Driving forces and structural determinants of steric zipper peptide oligomer formation elucidated by atomistic simulations. *J. Mol. Biol.* **2012**, *421* (2−3), 390–416.

32 Lu, Y.; Derreumaux, P.; Guo, Z.; Mousseau, N.; Wei, G. Thermodynamics and dynamics of amyloid peptide oligomerization are sequence dependent. *Proteins* **2009**, *75* (4), 954–963.

33 Nasica-Labouze, J.; Mousseau, N. Kinetics of amyloid aggregation: a study of the GNNQQNY prion sequence. *PLoS Comput. Biol.* **2012**, *8*, No. 1002782.

34 Hills, R. D.; Brooks-III, C. L. Hydrophobic cooperativity as a mechanism for amyloid nucleation. *J. Mol. Biol.* **2007**, *368*, 894–901.

35 Li, M. S.; Co, N. T.; Reddy, G.; Hu, C. K.; Straub, J. E.; Thirumalai, D. Factors governing fibrillogenesis of polypeptide chains revealed by lattice models. *Phys. Rev. Lett.* **2010**, *105* (21), No. 218101.

36 De Simone, A.; Kitchen, C.; Kwan, A. H.; Sunde, M.; Dobson, C. M.; Frenkel, D. Intrinsic disorder modulates protein self-assembly and aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (18), 6951–6956.

37 Fawzi, N.; Okabe, Y.; Yap, E. H.; Head-Gordon, T. Determining the critical nucleus and mechanism of fibril elongation of the Alzheimer's Aβ(1−40) peptide. *J. Mol. Biol.* **2007**, *365*, 535–550.

38 Cabriolu, R.; Auer, S. Amyloid Fibrillation kinetics: Insight from atomistic nucleation theory. *J. Mol. Biol.* **2011**, *411*, 275–285.

39 Zhang, J.; Muthukumar, M. Simulations of nucleation and elongation of amyloid fibrils. *J. Chem. Phys.* **2009**, *130*, No. 035102.

40 Baftizadeh, F.; Pietrucci, F.; Biarnés, X.; Laio, A. Nucleation process of a fibril precursor in the C-terminal segment of amyloid-β. *Phys. Rev. Lett.* **2013**, *110*, No. 168103.

41 Nasica-Labouze, J.; Meli, M.; Derreumaux, P.; Colombo, G.; Mousseau, N. 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*, No. e1002051.

42 Nguyen, P. H.; Derreumaux, P. Conformational ensemble and polymorphism of the all-atom Alzheimer's Aβ(37−42) amyloid peptide oligomers. *J. Phys. Chem. B* **2013**, *117*, 5831–5840.

43 Lu, Y.; Wei, G.; Derreumaux, P. Structural, thermodynamical, and dynamical properties of oligomers formed by the amyloid NNQQ peptide: Insights from coarse-grained simulations. *J. Chem. Phys.* **2012**, *137* (2), No. 025101.

44 Lansbury, P. T.; Costa, P. R.; Griffiths, J. M.; Simon, E. J.; Auger, M.; Halverson, K. J.; Kocisko, D. A.; Hendsch, Z. S.; Ashburn, T. T.; Spencer, R. G.; Tidor, B.; Griffin, R. G. Structural model for the beta-amyloid fibril based on interstrand alignment of an antiparallel-sheet comprising a C-terminal peptide. *Nat. Struct. Biol.* **1995**, *11*, 990–998.

45 Landau, M.; Sawaya, M. R.; Faull, K. F.; Laganowsky, A.; Jiang, L.; Sievers, S. A.; Liu, J.; Barrio, J. R.; Eisenberg, D. Towards a pharmacophore for amyloid. *PLoS Biol.* **2011**, *9* (6), No. e1001080.

46 Chebaro, Y.; Derreumaux, P. Targeting the early steps of Aβ16−22 protofibril disas-sembly by N-methylated inhibitors: A numerical study. *Proteins* **2009**, *75* (2), 442–452.

47 Hochdörffer, K.; März-Berberich, J.; Nagel-Steger, L.; Epple, M.; Meyer-Zaika, W.; Horn, A. H.; Sticht, H.; Sinha, S.; Bitan, G.; Schrader, T. Rational design of β-sheet ligands against Aβ42-induced toxicity. *J. Am. Chem. Soc.* **2011**, *133*, 4348–4358.

48 Kirkitadze, M. D.; Condron, M. M.; Teplow, D. B. Identification and characterization of key kinetic intermediates in amyloid β-protein fibrillogenesis. *J. Mol. Biol.* **2001**, *312*, 1103–1119.

49 Wang, S.-H.; Dong, X.-Y.; Sun, Y. Thermodynamic analysis of the molecular interactions between amyloid β-protein fragments and (−)-epigallocatechin-3-gallate. *J. Phys. Chem. B* **2012**, *116*, 5803–5809.

50 Scherzer-Attali, R.; Pellarin, R.; Convertino, M.; Frydman-Marom, A.; Egoz-Matia, N.; Peled, S.; Levy-Sakin, M.; Shalev, D. E.; Caflisch, A.; Gazit, E.; Segal, D. Complete phenotypic recovery of an Alzheimer's disease model by a quinone-tryptophan hybrid aggregation inhibitor. *PLoS One* **2010**, *5*, No. e11101.

51 Chebaro, Y.; Jiang, P.; Zhang, T.; Mu, Y.; Nguyen, P. H.; Mousseau, N.; Derreumaux, P. Structures of Aβ17−42 trimers in isolation and with five small-molecule drugs using a hierarchical computational procedure. *J. Phys. Chem. B* **2012**, *116*, 8412–8422.

52 Ehrnhoefer, D. E.; Bieschke, J.; Boeddrich, A.; Herbst, M.; Masino, L.; Lurz, R.; Engemann, S.; Pastore, A.; Wanker, E. E. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat. Struct Mol. Biol.* **2008**, *15*, 558–566.

53 Lopez del Amo, J. M.; Fink, U.; Dasari, M.; Grelle, G.; Wanker, E. E.; Bieschke, J.; Reif, B. Structural properties of EGCG-induced, nontoxic Alzheimer's disease Aβ oligomers. *J. Mol. Biol.* **2012**, *421*, 517–524.

54 Zhang, T.; Zhang, J.; Derreumaux, P.; Mu, Y. Molecular mechanism of the inhibition of EGCG on the Alzheimer Aβ(1−42) dimer. *J. Phys. Chem. B* **2013**, *117*, 3993–4002.

55 Nguyen, P. H.; Li, M. S.; Derreumaux, P. Effects of all-atom force fields on amyloid oligomerization: replica exchange molecular dynamics simulations of the Aβ16−22 dimer and trimer. *Phys. Chem. Chem. Phys.* **2011**, *13*, 9778–9788.

56 Zhu, M.; De Simone, A.; Schenk, D.; Toth, G.; Dobson, C. M.; Vendruscolo, M. Identification of small-molecule binding pockets in the soluble monomeric form of the Aβ42 peptide. *J. Chem. Phys.* **2013**, *139*, No. 035101.

57 Scherzer-Attali, R.; Convertino, M.; Pellarin, R.; Gazit, E.; Segal, D.; Caflisch, A. Methylations of tryptophan-modified naphthoquinone affect its inhibitory potential toward Aβ aggregation. *J. Phys. Chem. B* **2013**, *117*, 1780–1789.

58 Viet, M. H.; Ngo, S. T.; Lam, S.; Li, M. S. Inhibition of aggregation of amyloid peptides by beta-sheet peptides and their binding affinity. *J. Phys. Chem. B* **2011**, *115*, 7433–7446.

59 Lopez del Amo, J. M.; Schmidt, M.; Fink, U.; Dasari, M.; Fändrich, M.; Reif, B. An asymmetric dimer as the basic subunit in Alzheimer's disease amyloid β fibrils. *Angew. Chem., Int. Ed.* **2012**, *51*, 6136–6139.

60 Zhang, T.; Xu, W.; Mu, Y.; Derreumaux, P. Atomic and dynamic insights into the beneficial effect of the 1,4-naphthoquinon-2-yl-L-tryptophan inhibitor on Alzheimer's Aβ1−42 dimer in terms of aggregation and toxicity. *ACS Chem. Neurosci.* **2013**,10.1021/cn400197x.