Structures of Soluble Amyloid Oligomers
From Computer Simulations

Adrien Melquiond,1 Normand Mousseau,2 and Philippe Derreumaux1* 1Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Institut de Biologie Physico-Chimique et Université Paris 7, 13 rue Pierre et Marie Curie, 75005 Paris, France 2Département de Physique and Regroupement Québécois sur les Matériaux de Pointe, Université de Montréal, C.P. 6128, succursale centre-ville, Montreal (Québec), Canada

ABSTRACT Alzheimer’s, Parkinson’s, and Creutzfeldt-Jakob’s neurodegenerative diseases are all linked with the assembly of normally soluble proteins into amyloid fibrils. Because of experimental limitations, structural characterization of the soluble oligomers, which form early in the process of fibrillogenesis and are cytotoxic, remains to be determined. In this article, we study the aggregation paths of seven chains of the shortest amyloid-forming peptide, using an activated method and a reduced atomic representation. Our simulations show that disordered KFFE monomers ultimately form three distinct topologies of similar energy: amorphous oligomers, incomplete rings with β-barrel character, and cross-β-sheet structures with the meridional but not the equatorial X-ray fiber reflections. The simulations also shed light on the pathways from misfolded aggregates to fibrillar-like structures. They also underline the multiplicity of building blocks that can lead to the formation of the critical nucleus from which rapid growth of the fibril occurs. Proteins 2006;65:180–191. © 2006 Wiley-Liss, Inc.

Key words: protein simulations; amyloid fibril formation; soluble oligomers; aggregation; molecular dynamics; activation-relaxation technique; coarse-grained force field

INTRODUCTION

Several age-related degenerative disorders are linked to normally soluble proteins that assemble into amyloid fibrils.3 Alzheimer’s disease, which affects several million people worldwide and is characterized by memory loss and impairment in other forms of behavior, results from the accumulation of the amyloid β-protein (Aβ) with 40 or 42 amino acids.2 Similarly, Parkinson’s disease is associated with the misfolding of the α-synuclein protein of 140 amino acids, while human transmissible spongiform encephalopathies involve misfolding of the prion protein (PrP) of 240 amino acids.3 All these proteins differ in amino acid composition and length, but share a common insoluble cross-β sheet-structure. Very few structures of amyloid fibrils at atomic level of detail are available.4 That of fibrils made of Aβ and PrP, in particular, is unknown.5

Recent experimental studies support the hypothesis that the soluble oligomers, which form early in the process of fibrillogenesis, may be the key cytotoxic species. Remarkably, this toxicity is also detected for proteins not connected to clinical diseases.6 Moreover, soluble oligomers of Aβ, α-synuclein, and fragment 106–126 of the prion protein may have a common structure.7 In an independent study, Cleary et al. went one step further and showed that soluble trimers and dimers of Aβ were both necessary and sufficient to disrupt cognitive function.8 Thus, understanding the first steps of the assembly process and determining the structure of the pre-fibrillar species could be of clinical value.

This is, however, a very difficult experimental task because protein aggregation involves a variety of distinct oligomeric intermediates in dynamic equilibrium during the lag phase, and therefore, these early metastable assemblies cannot be easily identified.9 In addition, the lag phase is rather slow, spanning for instance 60 h for Aβ.

Theoretical studies can be helpful in providing insights into the aggregation processes, as the investigation by experiments is very complicated and sensitive to a wide variety of parameters such as pH, concentration, and metal ions. For instance, on-lattice replica exchange Monte Carlo simulations on four protein chains of 64 beads each examined the effects of concentration on the free energy landscape.10 Following all the detailed interactions between proteins and solvent atoms with time is costly, however, current molecular dynamics (MD) simulations have explored the stability of preformed monomers11 and multimers,12–16 as well as the possible initial steps in aggregation within time scales of 10–20 ns.17–19 All-atom MD simulations with constraints between the chains20–22 or implicit solvent representations23,24 accel-

Grant sponsors: NSERC, Canada Research Chair program, French Ministry of Research, CNRS, University of Paris 7.

*Correspondence to: Philippe Derreumaux, Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Institut de Biologie Physico-Chimique et Université Paris 7, 13 rue Pierre et Marie Curie, 75005 Paris, France. E-mail: philippe.derreumaux@ibpc.fr

Received 27 January 2006; Revised 16 May 2006; Accepted 25 May 2006

Published online 7 August 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21100
erate aggregation, but the time scale sampled is still several orders of magnitude smaller than the lag phase.

To overcome this sampling problem, discontinuous MD (DMD) simulations, which use hard-sphere and square well potentials, were applied to study polyalanine\(^{25}\) and Aβ peptides,\(^{26}\) but their aggregation pathways remain to be confirmed using more elaborate force fields. The approach we use, here, is to define conformational moves that take into account the structure of the energy landscape, bringing the system from one local minimum to another unknown minimum through first-order saddle points. This is carried out using the activation–relaxation technique (ART) and an optimized potential for efficient peptide-structure prediction (OPEP) with solvent effects included. In this study, we specifically investigate the assembly process of the shortest peptide known to form amyloid fibrils in vitro. Besides its simplicity, the KFFE tetrapeptide is also an excellent model system to study amyloid fibril formation, because it contains all the driving forces of large proteins: it has two charged side-chains [lysine (K) and glutamic acid (E)] and two hydrophobic side-chains [phenylalanine (F)]. On the basis of biophysical data, Tjernberg et al. proposed an antiparallel β-strand orientation structure within the sheets.\(^{37}\) In recent ART-OPEP reports, we found that monomeric KFFE is disordered in solution and four chains can populate a four-stranded antiparallel β-sheet in agreement with Tjernberg’s hypothesis.\(^{28}\) Using the same simulation approach, six chains can associate to adopt a bilayer cross-β-sheet structure starting from random positions of the peptides and a barrel-like monolayer starting from a preformed tetramer-dimer. We also proposed a bidirectional growth mode of amyloid fibril, involving alternate lateral and longitudinal growths.\(^{29}\) This growth can be facilitated by the introduction of preformed structures as was shown in octameric simulations, starting from an antiparallel dimer or a tetramer-dimer.\(^{30}\)

Here, our goal is to determine whether the low-energy states and aggregation pathways vary with the number of monomers available and whether the barrel-like monolayer can be explored starting from randomly chosen states. This is achieved by a detailed analysis of the ART-generated aggregation mechanisms of seven KFFE chains (blocked by acetyl and amide groups) starting from disordered peptides. Furthermore, we present ART simulations using the coarse-grained OPEP energy model and all-atom explicit solvent MD simulations starting from end products of in vitro fibrillogenesis.

**METHODS**

**ART-OPEP Simulations**

We use the same off-lattice coarse-grained protein representation, energy model, and sampling approach as in our previous work on Aβ\(^ {16-22}\) and KFFE peptides\(^ {28-30}\) in settings varying from dimer to octamer. A detailed description of the ART algorithm can be found in Refs. 33–36. In essence, ART searches activated mechanisms directly in the conformational space and allows moves of any complexity and size. ART thus differs from real-space Monte Carlo where each move is encoded in a predefined list of perturbations.\(^ {37}\) Each ART event consists of an activation phase, where the system is displaced from its local minimum and is brought to a first-order transition state by following the eigenvector associated with a negative eigenvalue in the Hessian, followed by a relaxation phase to a new local minimum, which is accepted or rejected according to the Metropolis criterion at temperature T. Because the structures are minimized, the Metropolis temperature used is not a real temperature and the trajectories do not obey a well-defined Boltzmann statistical mechanical ensemble. These limitations do not preclude, however, ART-pathways to resemble those generated by MD simulations on simple protein models in solution.\(^ {35,36}\)

OPEP, which includes solvent effects implicitly, uses a simplified protein chain representation with all backbone atoms included (i.e., N, H, C\(_ \alpha\), C, and O) and each sidechain modeled by one bead\(^ {39,40}\) with an appropriate van der Waals radius and geometry with respect to the main chain in agreement with the analysis of side chains in high-resolution structures. OPEP was trained on the structures of small proteins in solution and was used to predict the structure of a β-hairpin,\(^ {41}\) a three helix-bundle,\(^ {40}\) and the domain B1 of protein G.\(^ {41}\) OPEP treats four types of interactions. (i) All bond lengths, bond angles, improper dihedral angles of the peptide bonds and of the side-chains are maintained near their experimental values by harmonic terms. (ii) Interactions between all nonbonded particles except side-chain–side-chain are modeled by an excluded volume potential. (iii) Side-chain–side-chain interactions (a 20 amino acid alphabet is used) are modeled by a 12-6 potential if the interaction is hydrophobic in character or results from two oppositely charged residues, and by a 6 potential, otherwise. (iv) Hydrogen bonding (H-bond) interactions between amide nitrogens and hydrogen atoms are modeled by a combination of two-body (dependent on the O···H distance and the N···H···O angle) and four-body (cooperative) terms.

In this work, a total of 13 ART simulations were conducted using different initial conformations and random number seeds at 600 K to better sample the conformations. Eleven runs (R1–R11) of 9000 events started from two randomly chosen distributions of seven misfolded chains within a sphere of diameter 50 Å. Two control runs of 12,000 events were also performed to guarantee that the unbiased simulations did not locate metastable states: R12 started from three unfolded chains and two parallel dimers (each dimer consisting of two antiparallel β-strands) spaced by 10 Å, while R13 started from a perfect cross-β-sheet structure with two parallel layers (one tetramer and one trimer with antiparallel β-strands).

**Analysis of ART Trajectories**

To follow the assembly process, we analyzed the total energy \(E\), the orientation of the chains using the scalar
product between the end-to-end unit vectors of chains $i$ and $j^{20,21}$ the number of native and non-native H-bonds as determined using the DSSP program,42 the end-to-end $C_{\alpha}$ distance of each chain, the percentage of native side-chain–side-chain contacts within and between the sheets and the number of phenylalanine–phenylalanine side chain contacts. Two side chains $k$ and $l$, of van der Waals radii $R_k$ and $R_l$, are in native contact if they deviate by less than $R_k + R_l + 1$ Å from their positions in the native (lowest-energy) structure. In this work, the list of native interactions varies from one run to another, because distinct arrangements are nearly isoenergetic.

MD Simulations

All-atom MD simulations covering a total time of 50 ns were performed in explicit solvent conditions, using the GROMOS force field and GROMACS package.43 Two starting models corresponding to the end products of in vitro fibrillogenesis were examined: the perfect monolayer β-sheet and perfect bilayer β-sheet. Each model was solvated in a dodecahedral box of $x$ Å side with $y$ SPC water molecules and simulated using periodic boundary conditions at the desired temperature and constant pressure (1 atm), with coupling constants of 0.1 and 0.5 ps, respectively. For instance, $(x, y)$ were set to $(55, 3650)$ for the bilayer β-sheet and $(75, 5633)$ for the monolayer β-sheet. Bond lengths were constrained with the LINCS algorithm and the time step for dynamics was 2 fs. The Particle Mesh Ewald method was used with a cutoff distance of 12 Å for the electrostatics interactions, the nonbonded interactions were updated every 10 fs, and the density equilibrated in all runs to a value ~1.0 g/mL. Each model was equilibrated for 100 ps prior to MD analysis.

Both MD and ART-OPEP have drawbacks and advantages. ART-OPEP simulations allow the system to move through the conformation space rapidly without having to wait for the rare thermal fluctuations to occur as in MD simulations and generate physically-possible aggregation pathways. As with any Monte Carlo method, however, ART does not provide a time scale. On the other hand, explicit solvent MD simulations capture the atomic detailed interactions between the solute and the solvent, include thermal vibrations, but are unable to explore large conformational changes within the nanosecond timescale.

RESULTS

Classification of the Lowest-Energy Structures Predicted by ART

Figure 1 shows the lowest-energy structures predicted by all ART simulations along with their Boltzmann probabilities: $p(C_j) = \exp(-U_j/kT)/Z$, where $k$ is Boltzmann’s factor and $Z$ is the partition function calculated using the lowest-energy states of the 13 simulations. These probabilities are only qualitative, since entropic contributions are neglected. Based on energy similarity between the low-energy states obtained starting from two dimers (run R12), a three-stranded β-sheet parallel to a four-stranded β-sheet (run R13) and various orientations and conformations of the seven peptides (runs R1–R11), we find that three runs (R1, R2, and R10) have not reached equilibrium within 9000 events. The energy of these metastable states varies between $-63.7$ and $-67.2$ kcal/mol. The other 10 runs, with energy between $-78.5$ (run R6) and $-82$ (runs R7, R12, and R13) kcal/mol, can be clustered into three types of molecular arrangements, although there is considerable fluctuation in the atomic details.

The first arrangement (probability 75%), attained independently of the starting structure (run R3 of energy $-80.5$ kcal/mol, R4 of energy $-78$ kcal/mol, R5 of energy $-80$ kcal/mol, R6 of energy $-78.5$ kcal/mol, runs R7, R12, and R13 of energy $-82$ kcal/mol), consists of a cross-β-sheet-like structure with two layers of disordered β-sheets with an average intrasheet $C_{\alpha}–C_{\alpha}$ distance of 4.5 Å. This structure differs, however, from the standard X-ray fiber diffraction and solid-state NMR pattern in that the sheets are not parallel or antiparallel to each other, but rather perpendicular. As a result, the average intersheet $C_{\alpha} – C_{\alpha}$ distance is $\sim 12$ Å, larger than the standard 10 Å, and the predicted network of H-bonds is not parallel to the fibril axis. Further analysis shows that antiparallel β-strands are preferred (R12 and R13), but mixed antiparallel–parallel strands are also possible (R3, R4, and R7 have two parallel strands). This is consistent with all-atom MD simulations of four KFFE peptides in solution, where antiparallel chains were found in dynamics equilibrium with mixed antiparallel–parallel chains.28

The second organization of energy $-81.5$ kcal/mol (probability 11%), observed in run R5, is a monolayer with β-barrel character, which is found in dynamic equilibrium with the cross-β-sheet structure. The third arrangement of energy around $-81.0$ kcal/mol (probability 13%), observed in runs R8, R9, and R11, is associated with amorphous oligomers.

ART-Predicted Aggregation Pathways

Transient amorphous oligomers form in the early steps of all simulations, starting from a random organization and conformation of the peptides. Among 11 unbiased simulations, three runs, R8, R9, and R11, locate low-energy amorphous oligomers with distinctive morphologies: a partially formed four-stranded β-sheet interacting with a dimer and one monomer in R8, two dimers packed against a trimer in R9, and a disordered pentamer interacting with a dimer in R11 (see Fig. 1). We find that these aggregates can be the result of the addition of monomers, dimers, trimers or even tetramers to existing structures. We know that these low-energy amorphous oligomers would evolve toward fibrillar-like structures if the simulations were continued. This is observed in five simulations. Figure 2 shows a detailed analysis of run R7, and Figure 3 shows representative snapshots from
the initial state [Fig. 3(A)] to the cross-β-sheet-like state [Fig. 3(F)].

Folding starts by the formation of amorphous aggregates consisting of a disordered pentamer and a dimeric β-sheet at event 311 [Fig. 3(B)] which diffuse and assemble into disordered heptamers of various shapes [event 335, Fig. 3(C)]. As seen in Figure 2(F), the early steps are dominated by phenylalanine–phenylalanine side-chain contacts. In a few events, the system moves from amorphous oligomers to a bilayer conformation [Fig. 3(D), event 738] composed of one mixed parallel–antiparallel tetramer (chains 2-3-5-7) and one antiparallel dimer (chains 6-4); Chain 1 (blue) remaining essentially mobile. From there, the two layers are dynamic and
Fig. 2. Detailed analysis of the ART aggregation trajectory R7 as a function of ART events. (A) Energy in kcal/mol; (B) number of native H-bonds within the tetramer (blue) and the trimer (red); (C) chain orientations within the tetramer (2-3-5-7); (D) number of nonnative H-bonds within the tetramer (blue), the trimer (red) and between the two layers (black dashed); (E) chain orientations within the trimer (6-4-1) and between the two layers (5-4); (F) percentage of native side-chain–side-chain contacts within the tetramer (blue), the trimer (red) and between the two layers (black dashed); (G) end-to-end distances of the chains in Å; (H) number of phenylalanine–phenylalanine (F-F) side-chain contacts (black). The corresponding number in a perfect monolayer β-sheet is in blue, and that in a perfect bilayer β-sheet is in red. For clarity, the results are given until the lowest-energy state is located.
explore various orientations, and Chain 1 goes from one layer to another. At event 2010, Chain 1 is perpendicular to the tetramer (Fig. 3(E)), while at event 6815, Chain 1 is recruited by the dimer (see the decrease in energy by 8 kcal/mol in [Fig. 2(A)], and the lowest-energy conformation is encountered at event 7238 [Fig. 3(F)].

Further analysis shows that there is considerable reorientation of all the peptides in the early steps of aggregation, as seen by the fluctuation of the orientation vectors in Figure 2(C,E) during the 1000 first events. The plots of native H-bond [Fig. 2(B)] and nonnative H-bond [Fig. 2(D)] interactions show that the four-stranded β-sheet is clearly in nonnative register within the first 2100 events and shifts to its native register by reptation moves of the chains. This can be seen by the abrupt increase of the number of native H-bonds from 5 at event 2160 to 10 H-bonds at event 2310 [Fig. 2(B)], the decrease of nonnative H-bonds from 5 to 0 [Fig. 2(D)], and the cooperative increase of native contacts from 70 to 100% [Fig. 2(F)]. For the three-stranded β-sheet, reptation moves between nonnative and native H-bond patterns occur at event 6775.

A comparison of all trajectories folding to a bilayer β-sheet shows that there does not exist a unique pathway. In run R7, all end-to-end distances [Fig. 2(G)] vary between 6 and 8 Å, except Chain 7 (10 Å), while in runs R12 and R3, they all fluctuate between 7 and 10 Å. Disordered pentamer-dimer aggregates form early in runs R7 and R6, while a disordered heptamer forms early in run R3. Taken together, these results suggest a large number of distinct amorphous oligomers in the early steps of fibrillogenesis. Complexity in the aggregation process is also illustrated by the finding in run R5 of a β-barrel-like structure in dynamic equilibrium with a cross-β-sheet-like structure. We emphasize that higher-enthalpy monolayers with β-barrel structure are also detected in run R6. The assembly process in run R5 is shown in Figure 4.

Starting from a random orientation of the chains shown in Figure 3(A), folding proceeds through aggregates of various complexity: trimer (Chains 2, 3, 4) →
dimer (Chains 5, 7) → two monomers followed by tetramer (Chains 2, 3, 4, 6) → dimer (Chains 5, 7) → one monomer (Chain 1) within 308 events [Fig. 4(A)]. Note that the peptides are fully aligned in the dimer, while they are disordered in the tetramer. Then, a new trimer is created by dissociation of Chains 4 and 6 from the previous tetramer and association of the free Chain 1, and the system forms transient amorphous oligomers consisting of a disordered three-stranded β-sheet (Chains 1, 2, 3), and two disordered two-stranded β-sheets (Chains 4, 6; Chains 5, 7) stabilized by side-chain interactions [event 351, Fig. 4(B)]. Once this amorphous oligomer is found, ART finds successive transition states favoring the exploration of a five-stranded β-sheet (Chains 1, 2, 3, 4, 6) with a nonoptimal network of H-bonds and a two-stranded β-sheet (Chains 5, 7) with a perpendicular inter-sheet arrangement [event 398, Fig. 4(C)]. Chain 2 (red) is, however, very dynamic and moves from one sheet to another. The resulting trimer (Chains 2, 5, 7) starts to rotate with respect to the tetramer [Fig. 4(D), event 2805], and facilitates H-bond interactions between the edges of the β-sheets (Chains 2 and 3) leading to the β-barrel conformation [Fig. 4(E), event 5790]. This structure of energy ~81.5 kcal/mol, referred to R5-B (B for barrel), converts, however, within 2200 events to an isoenergetic cross-β-sheet-like structure (~80.7 kcal/mol, referred to as R5-C, C for cross) with two parallel β-sheets, consisting of a perfect antiparallel trimer and a less regular mixed parallel–antiparallel tetramer separated by average intersheet Cα–Cα distances of 10 Å [Fig. 4(F), event 7932].

Validation of ART-Predicted Soluble Oligomers

We find a cross-β-sheet structure with the sheets perpendicular rather than parallel or antiparallel to each other. In addition, we do not locate a perfect monolayer β-sheet, although there is MD-based evidence that the human calcitonin hormone fragment DFNKF is stable in a parallel-stranded, single β-sheet consisting of nine chains.13 Is this because the ART-OPEP runs were not long enough, the present size system is too small, or because of an artefact of the force field and the method used? Indeed, due to the coarse-grained nature of the side-chains and the implicit solvent representation, OPEP...
cannot capture the full complexity of hydrophobic and electrostatic interactions between all-atom side-chains and does not take into account H-bonds between the peptides and the solvent. In addition, ART in its current form does not include thermal entropic contribution.

Because there is no experimental evidence that the sheets are parallel or antiparallel within the full fibrils, we examine the stability of a perfect bilayer with either antiparallel or parallel β-sheets and antiparallel β-strands for 20 ns at 310 K. Figure 5(A,B) show the Cα root-mean-square deviations (RMSDs) of both models with respect to their minimized structures. We find that parallel sheets stabilize each layer with Cα RMSDs of the trimer and tetramer on the order of 1 and 1.5 Å within the 20 ns time scale [Fig. 5(A)], but the perfect cross-β-sheet structure remains stable only for 5.9 ns (RMSD <4 Å). After 6 ns, the Cα RMSD of the heptamer oscillates around 7.2 Å and is associated with a rotation move of one sheet with respect to another. Figure 5(C) shows the starting cross-β structure and the center of the most populated cluster representing 37% of the trajectory between 10 and 20 ns. Interestingly, the center of this cluster deviates by 2.9 Å RMSD from R12 (same arrangement as in R7 except that the layers are
extended) and by 2.2 Å RMSD excluding one chain. Overall, these results are supported by the second MD simulation, although the trimer and tetramer are clearly more flexible within an antiparallel registry of the sheets, deviating by 2.5 and 4 Å RMSD on average from their minimized structures [Fig. 5(B)]. Figure 5(D) shows a highly populated structure characterized again by a rotation move of one sheet with respect to another. Taken together, these simulations validate the ART-predicted cross-β-sheet-like structure with perpendicular intersheet arrangement and exclude the possibility that the ART runs were not long enough and the OPEP force field is inadequate.

Figure 5(E) shows the Cα RMSD starting from a perfect monolayer β-sheet at 310 K. We see that, after 11.5 ns, the seven-stranded β-sheet monolayer is already partially misfolded with three chains unpacked [Fig. 5(F)]. This result explains why ART-OPEP did not locate a perfect monolayer β-sheet.

Driving Forces and Mechanisms Leading to Low-Energy Oligomers

To obtain a better understanding of the aggregation process, we followed the formation of nonnative and native hydrogen bonding and side-chain–side-chain interactions as a function of ART events. We recall that the list of native interactions varies from one run to another run because of the quasi-degeneracy in the energy of these structures.

Analysis of run R7 in Figure 2 shows that the early steps of aggregation result from the concomitant formation of both (nonnative and native) H-bonds and essentially phenylalanine–phenylalanine (F-F) side-chain interactions. At event 600, the early aggregates in run R7 have 25 F-F side-chain contacts, and 78% are native (see Fig. 6). All other runs show the same behavior; at event 600, aggregates in runs R5, R8, and R9 have between 10 and 14 F-F side-chain contacts (see Fig. 6), and 43–63% are native.

Although the number of F-F side-chain interactions remains relatively constant during the simulation, we can see a considerable rearrangement of the chains by looking at the percentage of all native side-chain–side-chain interactions as a function of ART events (see Fig. 7). In some runs (R5 and R7, for example), the native interactions form rapidly, with about 60% of the final interactions already in place at event 500, the rest of the simulations consisting in a series of optimization moves, in order to satisfy the other competing forces, such as hydrogen bonds. In other runs (i.e., R8 and R9), the initial collapse leads to the formation of a large number of side-chain–side-chain contacts (see Fig. 6) which are mostly nonnative (see Fig. 7). Folding per se then proceeds through association–dissociation of the chains and repagination of the chains [as described in Figure 2(B,D)]. This process can be followed in Figure 7, which
shows jumps in Qc as the chains fall into place. This is fully consistent with recent isotope edited IR spectroscopy and ART-OPEP simulations on the Aβ16-22 peptide.31,32

DISCUSSION AND CONCLUSIONS

In this article, we have revisited the aggregation paths of the shortest amyloid-forming peptide KFFE, using the activation–relaxation technique and an implicit solvent energy model. We have also studied the stability of two end products of in vitro fibrillogenesis by all-atom MD simulations with explicit solvent. The ART scheme, selected to overcome the long experimental aggregation time and go beyond the initial aggregation steps accessible by MD simulations,17,20,22 allows a structural characterization of the early assemblies that cannot be easily identified by biophysical methods. In addition, in contrast to many aggregation MD-based studies and our previous ART study of an octamer, the peptides, here, are not restrained to adopt β conformations21,22,46 and are free of any interchain restraints.20,22,23,30

The simulations presented here show, in agreement with other studies of nucleation phenomena,47 that the structure of the crystallites near or below the critical nucleus is varied and complex. Starting from randomly chosen conformations and orientations of the peptides, KFFE self-assembles into multiple alternative structures. Although there is considerable fluctuation in the atomic details, they fall into three major topologies of similar energy, which are closely related to experimentally identified assemblies of long amyloid-forming proteins.

Topology 1 is a cross-β-sheet-like structure, which varies from that formed in vitro and in our previous study,29 in that, there is a perpendicular rather than parallel or antiparallel intersheet arrangement. This change in orientation is confirmed by our all-atom MD simulations in solvent, starting from parallel and antiparallel β-sheets. We note that orthogonal β-sheets were also found in ART-OPEP simulations of an octamer starting from preformed structures,30 and in MD simulations of hexapeptides starting from a fully formed six-stranded β-sheet17 and random orientations of preformed two-stranded β-sheets.22 In this fibrillar-like arrangement, antiparallel strands are highly populated, but mixed antiparallel/parallel β-strands are also possible. This is in line with previous aggregation and dissociation MD simulations on related peptides.21,23,32

Topology 2 consists of amorphous aggregates with high H-bond contents (~50%) and diameters varying between 3.1 and 3.8 nm. Although these aggregates were also observed in ART simulations of eight peptides,30 their populations were almost negligible using six chains.29 These assemblies correspond to the micelle-like Aβ intermediates observed by Yong et al., using angle neutron scattering.48

Topology 3 is an ordered monolayer with β-barrel character. This structure, discussed in an ART-OPEP study of six chains starting from preformed assemblies29
but never reported by other simulation types, is reminis-
cent of the incomplete α-synuclein protein rings ob-
served by Lashuel et al., using negative stain micro-
cyte.49 Although this topology was not seen either in our
previous octameric simulations, we believe that this was
due to insufficiently long simulations for this system.50

As a result, the ART-generated pathways for a hept-
tamer, starting from a random distribution of the chains,
is more complex than those for a hexamer started from
an equivalent distribution. The early steps of aggrega-
tion are dominated by side-chain–side-chain interac-
tions, which allow partial assembly of distinct disordered
β-sheets in nonnative registries. These aggregates can
consist of two two-peptide β-sheets, disordered tetramers
with a two-peptide β-sheet or a three-peptide β-sheet,
and a disordered five-peptide β-sheet, among others.

Forcing proceeds by concomitant optimization of H-bond-
and hydrophobic interactions through amorphous
aggregates. During this phase, the sheets can associ-
ate–dissociate and reorient with respect to each other,
and the strands can rearrange by reptation moves. This
underlines the multiplicity of the early aggregates that
can form and act as building blocks for the nucleus from
which rapid growth of fibril can occur. Finally, these
amorphous aggregates convert either directly (higher
probability) or indirectly (lower probability through To-
pology 2) to cross-β-sheet-like structures. The ART-gen-
erated aggregation process is in agreement with bio-
physical studies of the prion protein Sup3549 and with
DMD simulations of polyalanine peptides,50 although the
incomplete rings with β-barrel character were not
described.

In summary, the present ART simulations, together
with the simulations on hexamers29 and octamers start-
ing from preformed structures,30 demonstrate the exis-
tence of thresholds for the stabilization of amorphous
aggregates, and the kinetic accessibility of the mono-
layer with β-barrel character from amorphous aggre-
gates. Both the coarse-grained ART and all-atom MD
simulations indicate that the present size system is still
too small to stabilize KFEF peptides in end products
structurally consistent with in vitro fibrillogenesis.
Rather, ART results suggest that orthogonal β-sheets
could be a dominant organized structure in early steps
of aggregation. The population of such “fibril competent”
structures remains, however, to be determined under
physiological solution conditions. In addition, which
soluble assemblies of proteins are most cytotoxic remains
unclear. Nevertheless, simulations, including ART-OPEP,
are providing numerous insights that will help charac-
terize the oligomers at atomic level and design antiamy-
loid agents.

ACKNOWLEDGMENTS

The authors thank the IDRIS and CINES centers in
France for computational support.

REFERENCES

1. Koo EH, Lansbury PT, Jr, Kelly JW. Amyloid diseases: abnor-
mal protein aggregation in neurodegeneration. Proc Natl Acad
2. Selkoe DJ. Alzheimer’s disease is a synaptic failure. Science
3. Prusiner SB. Prion diseases and the BSE crisis. Science 1997;278:
245–251.
Eisenberg D. Structure of the cross-β spine of amy-
5. Petkova AT, Leapman RD, Guo Z, Yau WM, Mattson MP, Tycio
R. Self-propagating, molecular-level polymorphism in Alzhei-
J, Teddei N, Ramponi G, Dobson CM, Stefani M. Inherent-to-
xicity of aggregates implies a common mechanism for protein mis-
7. Kayed R, Head E, Thompson JL, McIntire TM, Milton SC,
Cotman CW, Glabe CG. Common structure of soluble amyloid
oligomers implies common mechanism of pathogenesis. Science
2003;300:486–489.
MA, Selkoe DJ, Ashe KH. Natural oligomers of the amyloid-β pro-
tein specifically disrupt cognitive function. Nat Neurosci 2005;8:79–
84.
34889.
10. Cellmer T, Bratko D, Praunitz JM, Blanch H. Protein-folding
landscapes in multichain systems. Proc Natl Acad Sci USA 2005;
102:11692–11697.
11. Armen RS, DeMarco ML, Alonso DO, Daggett V, Pauling and
Corey’s α-pleated sheet structure may define the prefibrillar
amyloidogenic intermediate in amyloid disease. Proc Natl Acad
12. Ma B, Nussinov R. Stabilities and conformations of Alzheimer’s
β-amyloid peptide oligomers (Aβ16–22, Aβ16–25, and Aβ18–35):
reminiscent of the incomplete a-synuclein protein rings ob-
13. Haspel N, Zanuy D, Ma B, Wolfson H, Nussinov R. A compara-
tive study of amyloid fibril formation by residues 15–19 of the
human calcitonin hormone: a single β-sheet model with a small
indicate a possible role of parallel β-helices in seeded aggre-
15. Tarus B, Straub JF, Thirumalai D. Probing the initial stage of
aggregation of the Aβ16–22 peptide: assessing the propensity for
16. Lopez de la Paz M, de Mori GM, Serrano L, Colombo G.
Sequence dependence of amyloid fibril formation: insights from
molecular dynamics simulations. J Mol Biol 2005;349:
585–596.
17. Colombo G, Daidone I, Gazit E, Amadei A, Di Nola A. Molecular
dynamics simulation of the core of the aggregation motif of the
fusion-peptide single-layer protifibrils as monitored by molecular-
1067.
19. Tsaì HH, Reches M, Tsai CJ, Gunasekaran K, Gazit E, Nussinov
R. Energy landscape of amyloidogenic peptide oligomerization by
parallel-tempering molecular dynamics simulation: signifi-
8179.
20. Klimov DK, Thirumalai D. Dissecting the assembly of Aβ16–22
amyloid peptides into antiparallel β sheets. Structure 2005;13:
21. Paci E, Gsponer J, Salvatella X, Vendruscolo M. Molecular dy-
namics studies of the process of amyloid aggregation of peptide
22. Wu C, Lei H, Duan Y. The role of Phe in the formation of well-or-
dered oligomers of amyloidogenic hexapeptide (NFAGAIL) observed
in molecular dynamics simulations with explicit solvent. Biophys J


47. Moroni D, ten Wolde PR, Bolhuis PG. Interplay between structure and size in a critical crystal nucleus. Phys Rev Lett 2005;94:235703.

