THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Structures of Aβ17−42 Trimers in Isolation and with Five Small-Molecule Drugs Using a Hierarchical Computational Procedure

Yassmine Chebaro[†]

Laboratoire de Biochimie Théorique, UPR9080 CNRS, Université Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

Ping Jiang, Tong Zang, and Yuguang Mu

School of Biological Sciences, Nanyang Technological University, Singapore

Phuong H. Nguyen

Laboratoire de Biochimie Théorique, UPR9080 CNRS, Université Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

Normand Mousseau

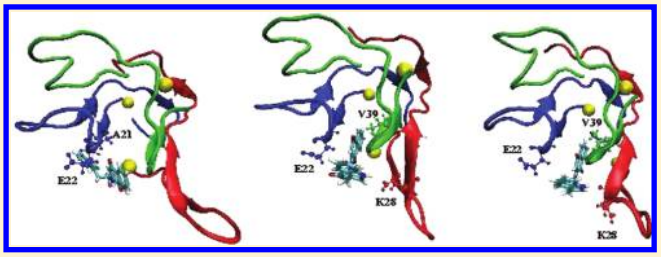
Département de Physique and GEPROM, Université de Montréal, Montréal, Québec, Canada

Philippe Derreumaux*

Laboratoire de Biochimie Théorique, UPR 9080 CNRS, Université Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris; Institut Universitaire de France, 103 bd Saint-Michel, 75005 Paris, France

Supporting Information

ABSTRACT: The amyloid-β protein (Aβ) oligomers are believed to be the main culprits in the cytoxicity of Alzheimer's disease (AD) and p3 peptides (Aβ17−42 fragments) are present in AD amyloid plaques. Many small-molecule or peptide-based inhibitors are known to slow down Aβ aggregation and reduce the toxicity in vitro, but their exact modes of action remain to be determined since there has been no atomic level of Aβ(p3)−drug oligomers. In this study, we have determined the structure of Aβ17−42 trimers both in aqueous solution and in the presence of five small-molecule inhibitors using a multiscale computational study. These inhibitors include 2002-H20, curcumin, EGCG, Nqtrp, and resveratrol. First, we used replica exchange molecular dynamics simulations coupled to the coarse-grained (CG) OPEP force field. These CG simulations reveal that the conformational ensemble of Aβ17−42 trimer can be described by 14 clusters with each peptide essentially adopting turn/random coil configurations, although the most populated cluster is characterized by one peptide with a β-hairpin at Phe19−Leu31. Second, these 14 dominant clusters and the less-frequent fibril-like state with parallel register of the peptides were subjected to atomistic Autodock simulations. Our analysis reveals that the drugs have multiple binding modes with different binding affinities for trimeric Aβ17−42 although they interact preferentially with the CHC region (residues 17−21). The compounds 2002-H20 and Nqtrp are found to be the worst and best binders, respectively, suggesting that the drugs may interfere at different stages of Aβ oligomerization. Finally, explicit solvent molecular dynamics of two predicted Nqtrp−Aβ17−42 conformations describe at atomic level some possible modes of action for Nqtrp.



INTRODUCTION

An increasing effort is made nowadays to determine the main culprits in Alzheimer's disease (AD). Although the pathological aggregation of the Tau protein in neurofibrillary tangles within the neurons is one hallmark of AD, there is increasing in vivo evidence that the amyloid-β (Aβ) protein of 39−43 residues,

Special Issue: Macromolecular Systems Understood through Multiscale and Enhanced Sampling Techniques

Received: December 9, 2011
Revised: January 24, 2012
Published: January 27, 2012

dx.doi.org/10.1021/jp2118778 | J. Phys. Chem. B 2012, 116, 8412−8422

principal constituent of the amyloid or senile plaques, is the main culprit, and Aβ oligomers are the proximate neurotoxic agents.[1] The Aβ1−42 aggregates are found more toxic than Aβ1−40 aggregates and the ratio Aβ1−40/Aβ1−42 differs by 10-fold between brains from nondemented controls and those with sporadic AD. Neurotoxicity of Aβ peptides is also induced by small changes in the Aβ42 to Aβ40 ratio.[2,3] These Aβ oligomers span a wide molecular weigth distribution ranging from <10 kDa (dimer and trimer) to >100 kDa (high order oligomers).[4] In addition to the Aβ1−40/1−42 peptides, the Aβ17−42 fragment, also called p3 peptide, is a constituent of diffuse plaques in AD and cerebellar preamyloid in Down's syndrome, and it has been shown to constitute an additional source of toxicity contributing to the neuronal cell loss characteristic of AD.[5]

Although AD toxicity may originate from various factors, including pore formation in membranes, oxidative stress, fibril fragmentation, and Aβ binding to and influencing protein receptors,[4,6−10] if effective drug design strategies targeting early formed oligomers are to be developed, detailed structure knowledge of these Aβ oligomers in aqueous solution and interacting with small-molecule drugs must be obtained. This is, however, a very complex task by experimental and theoretical means.

The overall picture of Aβ oligomerization is that the monomer exists as an ensemble of predominantly random coil structures with some ordered α-helical or β-sheet conformations.[11−15] Subsequently, these monomers form transient oligomers of increasing order with various Boltzmann populations depending on the sequence and the experimental preparation conditions, preventing therefore their isolation and structural analysis by NMR, until the formation of a nucleus from which rapid amyloid fibril growth occurs.[16] At a secondary structure level, oligomerization is characterized by an increase in β-strand, but its content varies from one experimental study to another. Using filtration though 10 000 molecular weight cutoff, circular dichroism (CD) of low molecular weight Aβ1−40 aggregates gives 88% of random coil/turn and 12% of β-strand at 295 K, pH 7.5, and day 0, emphasizing the heterogeneity of the energy landscape.[16] A different preparation gives, however, a higher β-strand content by CD spectroscopy, varying from 25% in the monomer, 38.6% and 40.8% in the dimer and trimer, to 45% in the tetramer.[17] On the theoretical front, structure characterization of small Aβ oligomers is not simpler and remains to be validated due to the approximations used (levels of granularity of protein model and/or treatment of solvent effects)[11,18−20] to go beyond limited sampling of all-atom MD simulations in explicit solvent.[21−26] For instance, simulations of Aβ1−42 based on the six-bead coarse-grained (CG) OPEP protein model with REMD[11] and a four-bead CG protein model with DMD[18] converge in the heterogeneous picture of the dimeric ensemble with high turn/random coil percentages and the negligible probability of the aggregation-prone N* state with β-strand−loop−β-strand formation at positions 17−42. This small population of N* has also been discussed by simulations of Aβ10−35[27] and Aβ16−35[28] dimers. In addition, all-atom extensive simulations based on two implicit solvent models found that the dimeric configuration consists essentially of intramolecular antiparallel β-sheets for Aβ1−42[19] or that it is described by three clusters with an averaged α-helix and β-strand content of 21% and 37%, respectively, for Aβ10−40.[20] A very different configurational picture also emerges between the tetramers of Aβ1−42[18] and Aβ10−40,[20] but whether this reflects reality, the experimental oligomer distribution varying from Aβ1−40 to Aβ1−42 peptides,[29] or depends on the accuracy of the force field, remains to be determined.

Many small-molecule or peptide-based inhibitors are known to slow down Aβ aggregation and reduce toxicity in vitro. The diversity in their structures suggest that these compounds act by binding to multiple sites in Aβ (although mostly targeting the amino acids 16−20 and 32−37) or by various modes of action at different oligomerization levels.[30] These experimental assays do not report, however, directly on the population and structure of the Aβ-drug oligomers. For instance, the small compound Nqtrp reduces Aβ1−42 fibril formation as well as the cytotoxic effect of Aβ42 oligomers toward cultured neuronal cell line and has been tested successfully on a transgenic AD Drosophila model. The NMR structure of Nqtrp bound to Aβ12−28 monomer showed the interaction with the region 18−21,[31] but whether the conformation holds for Aβ42 and oligomers remains to be determined. The same question applies to the results of computational studies examining the interaction of various drugs on Aβ16−22,[32−34] Aβ14−20,[35] and Aβ1−28[36] peptides. All-atom MD simulations starting from a preformed protofilament consisting of five Aβ15−42 peptides, along with transmission electron microscopy, CD spectroscopy, and cell viability allowed to test β-sheet ligands against Aβ1−42-induced toxicity.[37] Again, whether the designed ligands are the most effective in terms of binding affinities remains to be tested.

To increase our knowledge on the exact action mechanism of known Aβ drugs, it is important to go beyond MD simulations and/or short Aβ peptides. Recently, all-atom REMD studies in implicit solvent found that nanoproxen has stronger antiaggregation potential against Aβ10−40 fibrils rather than against dimers, suggesting potential limitations of this drug for the treatment of AD.[20] CG-DMD simulations examined the formation of Aβ1−42 oligomer in the presence of C-terminal Aβ inhibitors and found that Aβ31−42 and Aβ39−42 can be considered as therapeutic leads for the treatment of AD.[38]

In this study, we have determined the structure of Aβ17−42 trimers in aqueous solution and interacting with five small molecules identified previously as inhibitors. Drugs are labeled as D1 (2002-H20),[39] D2 (curcumin),[40] D3 (EGCG),[40,41] D4 (Nqtrp),[31] and D5 (resveratrol).[39] The choice of this Aβ fragment and trimer is motivated by three bodies of data in addition to the convergence problem to equilibrium ensemble. First, the p3 peptides are found in amyloid plaques and their contribution to AD toxicity might be critical.[42] Second, the region 17−42 displays a β-strand−loop-−β-strand conformation in Aβ1−40/42 fibrils allowing us to determine the probability of the N* state within the trimer. Third, Ono et al. showed that toxicity ranks as follows: tetramer > trimer > dimer > monomer, and trimers are almost as active as preformed fibrils in the nucleation of fibril formation.[17]

Ideally, we would have liked to study the formation of Aβ17−42 trimer in the presence of each inhibitor starting from spatially separated Aβ monomers. Even using coarse-grained potential, this is very time-consuming computationally. Moreover, the reduced representation is not appropriate for all of these small molecules.[43] To overcome these limitations, we used here a multiscale computational approach, consisting of CG REMD-OPEP simulations of Aβ trimers followed by all-atom Autodock simulations of five drugs using all dominant predicted Aβ17−42 conformations and a fibrillar state.

dx.doi.org/10.1021/jp2118778 | J. Phys. Chem. B 2012, 116, 8412−8422

Two Aβ−Nqtrp complexes predicted by Autodock were also subjected to all-atom MD simulations and their binding modes of Nqtrp identified.

MATERIAL AND METHODS

Coarse-Grained OPEP Model. Briefly, OPEP uses a reduced representation of the amino acids consisting of all atoms of the main chain (N, C_α, C, O, and H) and one bead for each side chain (Sc) with appropriate van der Waals radius, geometry with respect to the main chain, and hydrophobicity/hydrophilicity character.[44,45] The force field, with implicit solvent representation, is expressed as a sum of bonded terms (bond lengths, bond angles, improper torsions of the side chains, amide bonds, and backbone torsions) and nonbonded terms (van der Waals interactions and two-body and four-body hydrogen-bonding interactions).[46]

Though coarse-graining of the side chains cannot offer the high frequency vibrational mode accuracy of atomistic models[47] and the use of an implicit solvent cannot reproduce the protein dynamics in explicit solvent, OPEP offers the advantage to have a good trade-off between chemical details, accuracy of the energy landscape, and computer speed. For instance, activation−relaxation technique (ART) and greedy simulations using OPEP have proven accurate in the structure prediction of various nonamyloid peptides with chain lengths varying between 9 and 50 amino acids.[48−51] In addition, MD-OPEP,[52] REMD-OPEP,[53] and ART-OPEP simulations also helped identify the free energy landscape of amyloid peptides by revealing aggregation[54] or inhibition mechanisms[33] that were confirmed experimentally, and transient CG β-barrel structures[55] that were also seen in atomistic simulations.[56] In this work, we use OPEP version 4, which varies from version 3.2 by the use of a new analytical formulation for the nonbonded interactions, and is found to be more accurate than version 3.2[46] in preserving the native structure of some proteins by MD simulations (manuscript in preparation).

CG REMD Simulations and Analysis. The primary structure of Aβ17−42, LVFFAEDVGSNKGAIIGLMVGGVVIA, contains two hydrophobic patches Leu17−Ala21 (or central hydrophobic core, CHC) and Ala30−Ala42 separated by a hydrophilic patch Glu22−Gly30. The initial configuration of the unblocked trimer taken from Luhr's amyloid fibrillar model[57] was first minimized and equilibrated for 0.5 ns at 300 K. REMD simulations were carried out using 22 replicas and a logarithmic-like temperature distribution spanning 250−700 K. Exchanges between consecutive replicas were attempted every 7.5 ps on the basis of the Metropolis criterion, leading to an acceptance ratio of 30−40%. Each replica was simulated for 1.2 microsecond, leading to a total time of 25.4 microseconds. Simulations were carried out using the RATTLE[58] algorithm with a time step of 1.5 fs, and the temperature was fixed by the Berendsen's thermostat with a coupling constant of 100 fs.[52,59] OPEP-REMD simulations showed little variations in the equilibrium structures and heat capacity curves of two model peptides (monomer and trimer) using both the Berendsen and Langevin thermostats.[60]

Several properties were monitored for analysis of the CG structures. Secondary structure was determined using the STRIDE program.[61] Error bars for the secondary structure propensities show the interval of confidence on the mean value given by the bootstrap statistical analysis method. Intramolecular and intermolecular contacts between the side chains (Sc) were considered formed if their centroid distances were less than the sum of their van der Waals radii + 1 Å. Hydrogen bonds were considered formed when the angle between the donor and acceptor is less than 60° and when the distance between donor and acceptor was less than 3 Å. We used the PTWHAM version (weighted-histogram method for parallel tempering) to compute the density of states as a function of temperature.[62] Cluster analysis was done using the $g_{cluster}$ tool in the Gromacs 4.5.3[63] package with a C_α-rmsd cutoff of 3 Å. As discussed in the Results and Discussion section, the Aβ conformational ensemble at 300 K can be described by 14 clusters essentially random coil/turn in character and a very unlikely fibrillar state.

Docking of Five Drugs onto Aβ17−42 Trimers Using Autodock. Prior to docking, the center of each of the 15 OPEP-derived trimeric structures was subjected to MD simulations after mapping to atomic models using Gromacs 4.5.3[63] in a cubic box (5 × 5 × 5 nm) filled with TIP3P explicit water molecules.[64] Relaxation of all Aβ structures was performed at 300 K using 1 ns MD simulations with position restraints of all heavy atoms, followed by 1 ns MD simulations free of any restraints. The modified AMBER ff99SB force field, namely, ff99SBildn, was used.[65] At the end of the simulations, the backbone RMSDs of 9 protein models were <0.5 and of 6 protein models were within 0.5−0.8 nm from the OPEP structures. These final 15 structures, target proteins for docking, are labeled from P1 to P15.

The initial topologies of the five drugs were generated in Spartan and energy-minimized.[66] The conformations were further optimized at the HF/6-31G* level using Gaussian09,[67] and the atomic partial charges were derived using the RED III package.[68]

We used AutoDockTools 1.5.4 to prepare the pdbqt files for ligands/targets and the control-parameter files for griding and docking.[69] Autodock 4.2 was used to implement docking and calculate binding energies.[70] Partial charges assigned by R.E.D III were used for the atoms on ligands, and the atomic partial charges of protein models were derived from the force field ff99SBildn. Nonpolar hydrogen atoms were merged into the connected heavy atoms. The dimension of the grid box was that used for the MD simulation box (5 × 5 × 5 nm), allowing the ligands to dock all around the target surfaces. Targets were centered in the box. The grid maps had a spacing of 1 Å.

During the docking simulations with the genetic algorithm, the Aβ trimer was kept fixed and all torsional angles of ligands were allowed to move. Specifically, the number of degrees of freedom were 5 for D1, 10 for D2, 12 for D3, 6 for D4, and 5 for D5. The number of energy evaluations per docking run was set to 2 500 000. The binding sites and their populations were analyzed by the integrated clustering method in Autodock. The number of docking runs for each drug/protein target was varied from 100 to 200, and we found that they gave converged results for the dominant binding sites and their populations. Cluster analysis used a rmstol = 2.0 Å for only the atoms in the ligands, and rmsnosym was turned off (molecule symmetry was considered). In this study, a significant binding site is identified only when more than 5% of the total poses are located in the same binding site.

All-Atom MD Simulations of Two Aβ−Nqtrp Complexes. Two Aβ−Nqtrp complexes predicted by Autodock were also studied using MD simulations with the ff99SBildn force field. Starting from their Autodock structures, the models were placed in an octahedral box containing 5866 TIP3P waters[71] and 5 Na ions. Each solvated system was minimized using the steepest descent method and equilibrated for 1 ns at

dx.doi.org/10.1021/jp2118778 | J. Phys. Chem. B 2012, 116, 8412−8422

constant pressure (1 atm) and temperature (T = 300 K), using the Berendsen coupling[59] and velocity rescaling methods.[72] Subsequently, each system was simulated at constant temperature (T = 300 K) and constant volume for 80 ns using GROMACS. The equations of motion were integrated by using a leapfrog algorithm with a time step of 2 fs. Covalent bond lengths were constrained via the SHAKE[73] procedure. We used the particle-mesh Ewald method to treat the long-range electrostatics interactions.[74] The nonbonded interaction pairlists were updated every 5 fs, using a cutoff of 1.2 nm. Data were collected every 2 ps.

RESULTS AND DISCUSSION

Convergence of the CG Aβ Trimer Simulation. The first 400 ns of each replica were eliminated for analysis to avoid any bias due to the starting conformation. Convergence of the REMD simulation was assessed by block analysis, i.e., using the three time intervals 400−800, 800−1200, and 400−1200 ns. Figure 1A shows the evolution of the total energy and Figure 1B,C

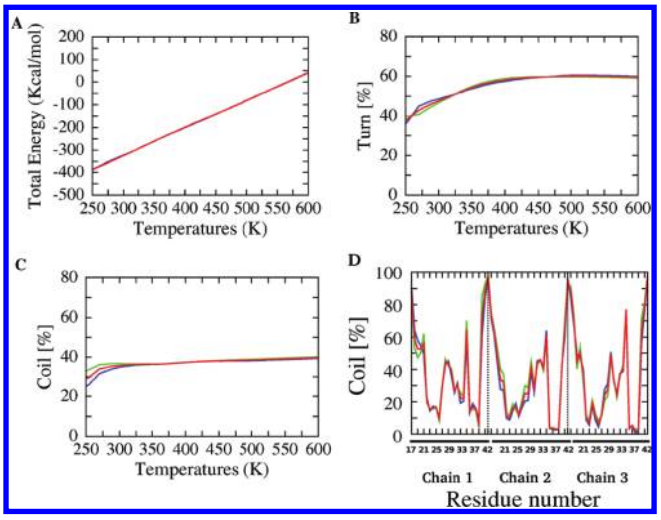


Figure 1. Evolution of some trimeric Aβ(17−42) properties as a function of the time interval. We used three intervals: 400−800 ns (in blue), 800−1200 ns (in green), and 400−1200 ns (in red). Total energy (A), turn percentage (B), and coil percentage (C) as a function of temperature. (D) Coil propensity of each amino acid in each chain at 300 K with dashed lines showing the positions of Ala42.

shows the evolution of the turn and coil components from 250 to 600 K. We see that the three quantities are time independent when averaged over 400 ns time windows. Figure 1D shows the coil percentage along the amino acid sequence of the three chains at 300 K. Again, we see that the coil profiles remain unchanged using the three time intervals. These results confirm that, with our simulation protocol, the sampling of the trimer has fully converged. In what follows, we focus on the results at 300 K and the averaged properties of all conformations during 400−1200 ns.

Secondary Structure of the CG Aβ Trimer. Averaged over all conformations and amino acids, the secondary structure at 300 K is mostly turn (46.1 ± 0.5%) and coil (35.4 ± 0.5%) and has little β-strand (7 ± 0.5%). Figure 2A,B shows the turn and β-strand propensities as a function of the amino acid index for the three chains at 300 K. Combined with Figure 1D, we see that the turn probability varies from 40% to 90% for all amino acids of the three chains except at positions Leu17−Phe20, Ala30, Met35, and Val39−Ala42 where the coil probability is

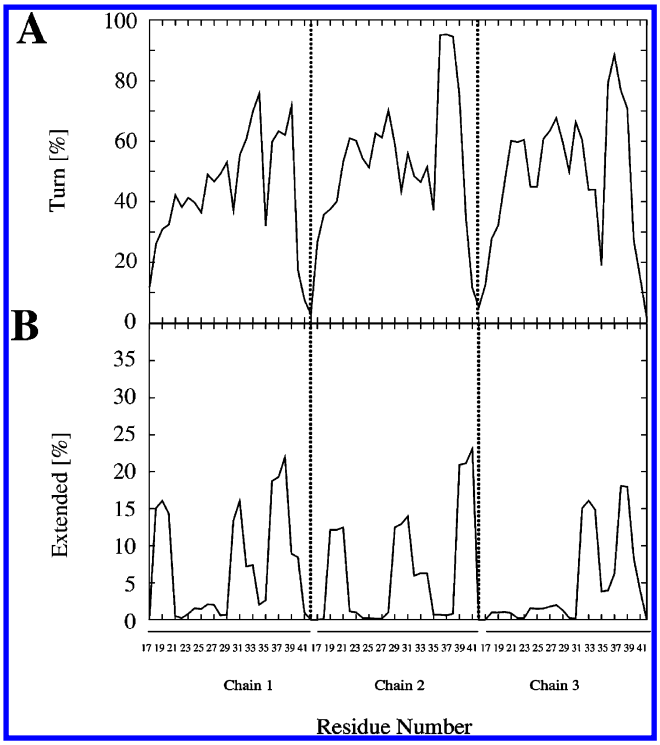


Figure 2. Turn (A) and β-strand (B) propensities as a function of the amino acid index using all conformations between 400 and 1200 ns at 300 K. The dashed lines show the positions of Ala42.

high. There is also a β-strand probability ranging from 7.5% to 22% at positions 18−20, 30−32, and 36−40 in chain 1, at positions 19−21, 29−34, and 39−41 in chain 2, and at positions 32−34 and 38−40 in chain 3 (Figure 3B). We also find an α-helix with a probability of 8.7 ± 0.5% spanning essentially the region Glu22−Lys27 (data not shown).

Quaternary CG Aβ Structures. Using a C_α-rmsd cutoff of 3 Å, the quaternary structures can be described by a total of 15 clusters, representing 97.8% of all conformations. Figure 3 shows the center of the five most populated clusters (panels A−E), representing 64.6% of all conformations, and the center of the 15th cluster with a three-stranded parallel β-sheet (panel F) and a population of 0.3%. This amyloid fibrillar state deviates by 5.2 Å rmsd from the Luhr's PDB model.

The first cluster containing 19% of all generated conformations displays a β-hairpin spanning residues Phe19−Leu31 (strands at Phe19−Ala21 and Gly29−Leu31) in chain 2, and a β-α-β-turn-β motif in the other two chains. In this motif, the α-helix spans residues Glu22−Lys27, the turn spans Gly37−Glys38, and the β-strand signal is rather weak, however, with one H-bond between the two β-strands (Figure 3A).

The second cluster, in Figure 3B, with a population of 15%, is very disordered though there is an interpeptide antiparallel β-sheet spanning Leu17−Phe20 of CHC (chain 1) and Ile31−Leu34 of the C-terminal (chain 3), an α-helix spanning Ala21−Asn26 (chain 1), and turns at positions Gly37−Gly38 (chains 2 and 3). In this cluster, chain 2 is free of any secondary structure.

The centers of the third and fourth clusters are almost fully random coil/turn in character. The fourth cluster in Figure 3D displays, however, a short α-helix at positions Glu22−Ser26 in one chain, while the third cluster in Figure 3C displays one intermolecular antiparallel β-sheet between Val36−Gly38 (chain 1) and Val39−Ile41 (chain 2), and one intermolecular parallel β-sheet between Ala30−Ile31 (chain 1) and Gly38−Val39

dx.doi.org/10.1021/jp2118778 | J. Phys. Chem. B 2012, 116, 8412−8422

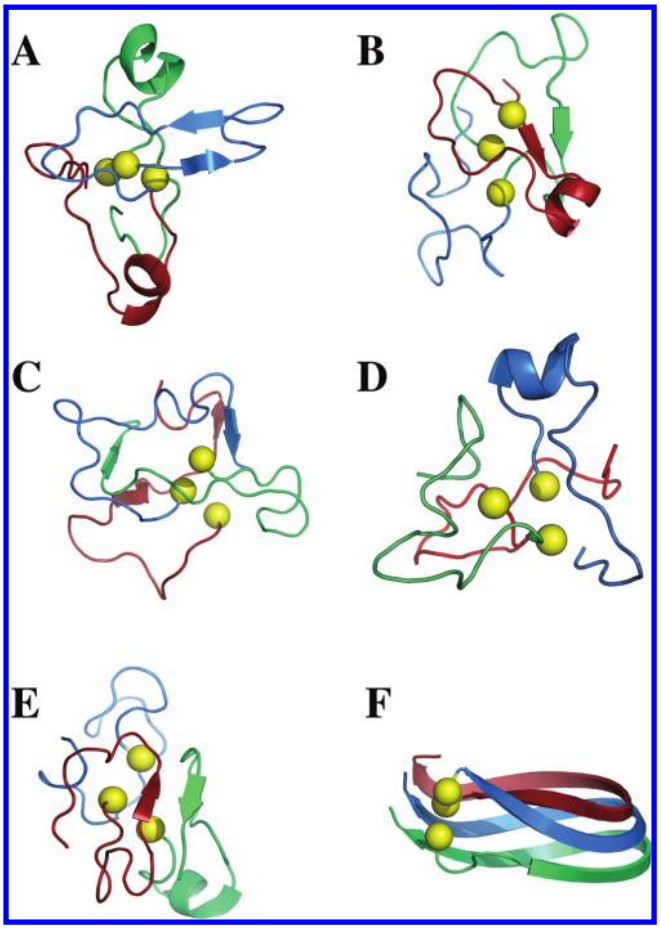


Figure 3. Cluster Analysis of CG Aβ structures. The centers of the first five dominant clusters are shown: 19% (A), 15.4% (B), 13.3% (C), 8.4% (D), and 8.2% (E). The fibrillar-like state with β-strand−loop−β-strand conformations is shown: 0.35% (F). The first chain is red, the second blue, and the third green. The population of clusters 6, 7, 8, and 9 are 7.4%, 7%, 6.2%, and 3.8%, respectively. The population of clusters 10, 11, 12, 13, and 14 represent 8.5% of all conformations. Yellow balls show the positions of the N-terminal residues.

(chain 3). Finally, the center of the fifth cluster (Figure 3E) displays an antiparallel β-sheet between Ile31−Gly33 of chain 1 and Gly38−Val40 of chain 3 and an α-helix of chain 3 at positions Glu22−Asn27.

The intermolecular Sc−Sc contact maps are shown in Figure 4, and the intra- and intermolecular hydrogen bond occupancy maps are shown in Figure 5. All maps use all conformations generated within 400−1200 ns. As seen in Figure 4, the binding interfaces are rather complex, even though there is a non-negligible probability for the three CHCs to interact with each other. Looking at the intermolecular hydrophobic Sc interactions, we find there is high probability of contacts between the CHC (residues 17−19) of chain 2 and the C-terminal (residues 33−36 and 40−42) of chain 3 and a lower probability between the CHC (residues 19−20) of chain 3 and the C-terminal (residues 35−36) of chain 2 and between the two CHC (Figure 4c). The Sc−Sc contact map between chains 1 and 2 in Figure 4a also shows a high formation probability between the two CHC regions and between the CHC of chain 1 and the residues 31−34 of chain 2. In contrast, chain 3 interacts with chain 1, as seen in Figure 4b, through the amino acids 30−31 (chain 3) and 39−42 (chain 1), and the CHC of chain 3 with the C-terminal of chain 2. Figure 5 clearly shows that the intramolecular H-bonds contribute much more to the equilibrium quaternary structure than the interpeptide H-bonds. While there are many intramolecular H-bonds with probabilities varying between 25% and 10%, the highest probability reaching 25% between residues 21−25 and 22−26 of chain 1, the maximal interpeptide H-bond occupancy is only 8% between residue 39 of chain 2 and residue 38 of chain 1.

As a final step toward characterizing the quaternary conformational ensemble, the first 14 CG centers (thus excluding the fibrillar state) were transformed into all-atom structures using the Scrwl4 program,[75] and the solvent accessible area surface (SASA) of each residue was calculated using the dssp algorithm[76] by considering the Boltzmann probability of each cluster. It is striking that the SASA profiles of the three chains superpose so well (Figure 6). We find four peaks in the SASA profiles with values varying between 90 and 160 Å^2 for the amino acids Phe19 and Phe20, Glu22, Ser26−Asn27−Lys28, and Met35. We also note that while Leu17 and Val18 are well buried, the C-terminal amino acids 40−42 and notably Ala42 are more exposed to solvent.

Binding Modes of the Five Drugs to Trimeric Aβ Using Autodock. Having identified and characterized the 14 dominant clusters and the less-frequent fibril state for trimeric Aβ17−42, we can predit the binding sites for our five small-molecule drugs. Each all-atom drug, labeled from D1 to D5, was docked onto the 15 all-atom Aβ relaxed structures, labeled from P1 to P15. For each drug−protein pair, 200 Autodock runs were performed. Figure 7 reports the energies of the dominant binding modes for each drug docked onto the 15 Aβ targets. The binding energies of the drug−Aβ complexes vary between −4.2 and −1.5 kcal/mol. Table S1 in Supporting Information provides the same information with the population of each binding site.

As seen in Figure 7, the drug Nqtrp is the most favorable in binding to most of the protein targets; while 2002-H20 is the worst scored indicating a lower probability to bind to Aβ trimers. The drugs D1 (2002-H20) and D5 (resveratrol) bind to all 15 Aβ structures, while the drugs D2−D4 (curcumin, EGCG, and Nqtrp) do not have good binding modes for 5, 8, and 3 Aβ structures, respectively (see Table S1, Supporting Information, for details). We also find that the drugs have a lower binding affinity for the fibril state P15 than most of the other states. From Table S1 (Supporting Information), it is also clear that there is not a unique binding mode for all drugs on each protein target. For instance, Nqtrp (D4) finds three binding modes, while curcumin (D2) finds five modes on the Aβ P1 structure within 0.6 kcal/mol.

In each drug−Aβ complex, a good representative binding mode was defined as the lowest-energy complex of the cluster with the largest population. The five drugs, if with a nonrandom binding mode, are simultaneously shown bound to each Aβ target (Figure S1, Supporting Information). The side chains of Phe19 and Phe20 in the three Aβ peptides are also explicitly shown. Despite structure variability of the drugs, it is interesting to note that for a given Aβ conformation, the same binding mode (see the complexes P3-D and P7-D) or the same binding region can be identified (see the complexes P1-D and P2-D). Looking at the complexes P15-D, we find that the four drugs 2002-H20, EGCG, Nqtrp, and resveratrol prefer the same side of the fibrillar assembly, but they have distinct binding sites.

To get structural insights into the binding sites, Figure 8 zooms on the drug environment for the lowest-energy pose of each drug among all 15 protein targets. In all cases, aromatic

dx.doi.org/10.1021/jp2118778 | J. Phys. Chem. B 2012, 116, 8412−8422

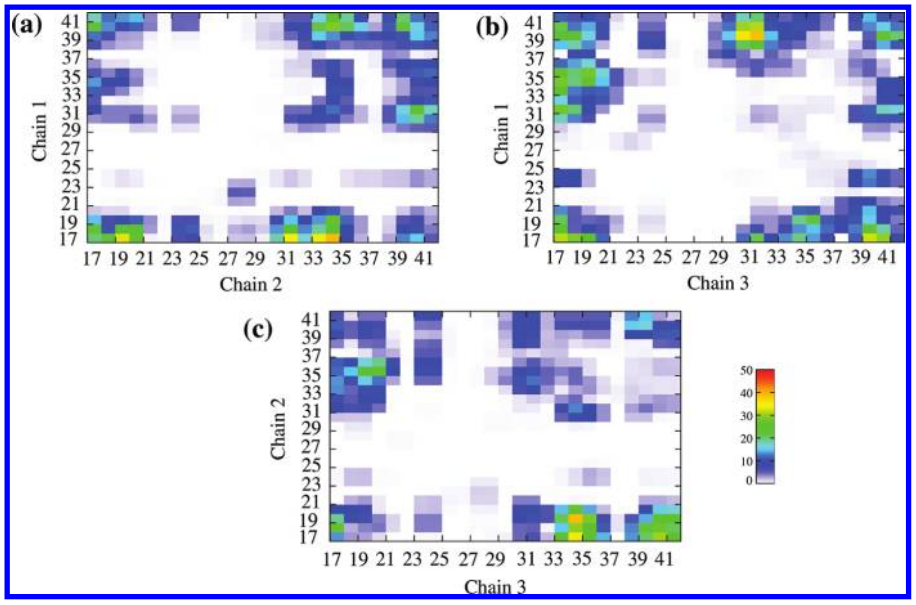


Figure 4. Intermolecular side−chain side−chain contact maps. We show the contact probabilities between chain 1 and chain 2 (a), chain 1 and chain 3 (b), and chain 2 and chain 3 (c). The strenght of the probability is color-coded varying from red (50%) to white (0%).

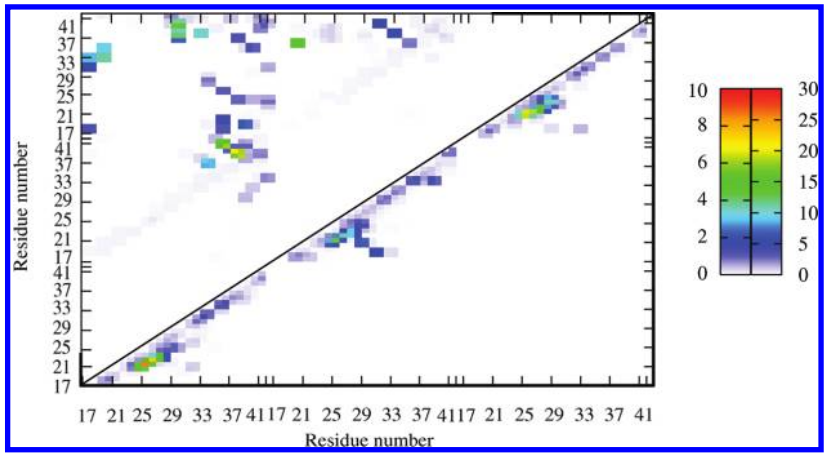


Figure 5. Intramolecular and intermolecular hydrogen bond occupancies. The upper triangle shows the intermolecular H-bonds and the lower triangle shows the intramolecular H-bonds. The strength of the probability is color-coded: 0−10% and 0−30% for the inter- and intrapeptide H-bonds, respectively.

interactions between the ring motifs of the drugs and the Phe residues of Aβ clearly contribute to the stabilization of the complexes. To further quantify this observation, we calculated the averaged contact numbers between the Aβ amino acids and the five drugs in their most populated binding modes. As seen in Figure 9, the amino acids Phe19, Phe20, and then Leu17 and Val18 show the largest number of contacts with drugs, even though the side chains of the C-terminal residues, and backbone hydrogen bonding interactions contribute as well as described in the next section.

Finally, Figure 8 shows that curcumin (D2) and resveratrol (D5) bind most preferentially to P1 (the most populated trimeric state predicted by OPEP), while D1 (2002-H20) and D3 (EGCG) bind preferentially to marginally populated OPEP models, namely, P10 and P14. In contrast, Nqtrp (D4) binds to P5, P1, P7, and P14 with similar energies (−3.9 kcal/mol) and to the other Aβ structures with lower energies.

Insights into Two Aβ−Nqtrp Complexes Using All-Atom MD Simulations. To validate the stability of the structures and the binding modes predicted by Autodock, we performed MD simulations on the first (P1) and eighth (P8) Aβ−Nqtrp models. These two models were selected because of their high and low Autodock binding energies of −3.9 and −3.3 kcal/mol, respectively.

We first looked at the conformational changes of the Aβ17−42 conformation by calculating the C_α-rmsd of the individual peptides with respect to the Autodock structure. Results shown in Figure S2 (Supporting Information) indicate that in both models, two peptides are very stable and that the third one is more flexible. We then calculated the probability contact maps between the atoms of Nqtrp and the centers of mass of side chains for each peptide (see methods in Supporting Information), and results are displayed in Figures S3 and S4 (Supporting Information) for models 1 and 8, respectively. For simplicity, in what follows, we denote $R_i(r_j)$ a set of residues r_j of the peptide i-th, which form side chain contacts with Nqtrp, $H_{CO}(r_j)$ a hydrogen bond between a carbonyl group of Nqtrp and an amide group of residue r_j, and $H_{NH}(r_j)$ a hydrogen bond between an amide group of Nqtrp and a carbonyl group of residue r_j.

dx.doi.org/10.1021/jp2118778 | J. Phys. Chem. B 2012, 116, 8412−8422

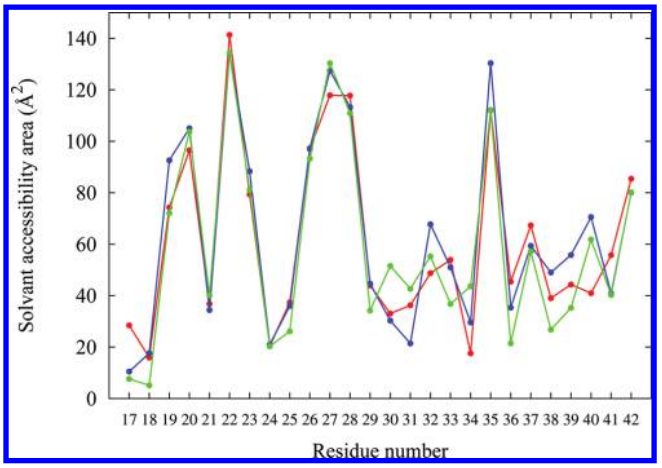


Figure 6. Solvent accessibility surface area of each residue in the trimer at 300 K. Results are obtained using the 14 most populated clusters. The first chain is red, the second blue, and the third green. Note that the SASA values of Ala42 in chains 2 and 3 superpose.

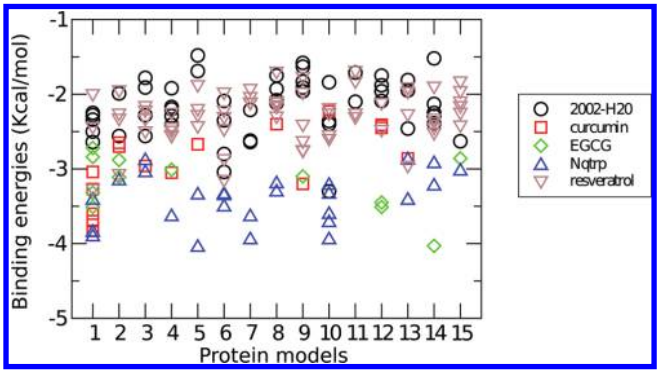


Figure 7. Binding energies of all significant drug−Aβ protein target poses. We only show the clusters with a population greater than 5%.

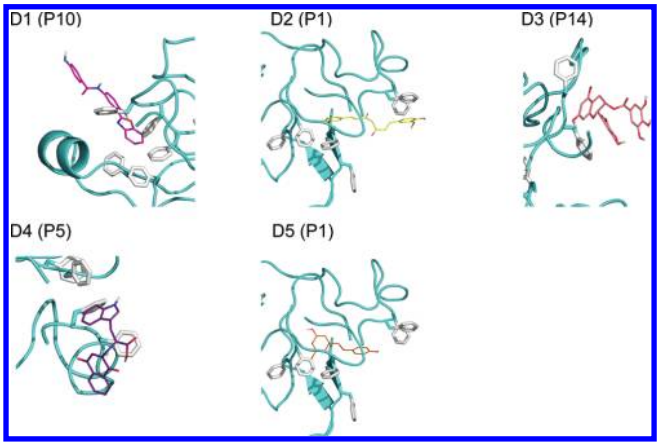


Figure 8. Snapshots of the best binding mode of the five drugs to Aβ trimer. For each drug (Di), we show the structure of the complex and the Aβ model (Pj) with the lowest binding energy among all fifteen protein models. Note that drugs 2 and 5 bind most preferentially to P1, the lowest energy conformation predicted by OPEP.

As seen from Figure S2 (Supporting Information), the binding site of the Autodock structure of model 1 is stabilized by the dominant contacts Nqtrp−R_1(Leu17−Phe19), Nqtrp−R_2(Phe19, Phe20), and Nqtrp−R_3(Gly37−Val40). Averaged over 80 ns, the maps show that the binding site is mainly formed by contacts Nqtrp−R_1(Leu17−Ala21), Nqtrp−R_1(Ser26−Lys28), and

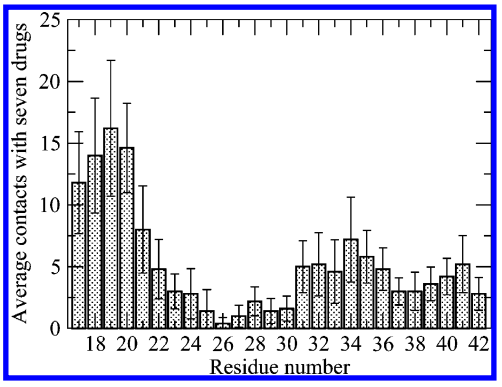


Figure 9. Averaged contact numbers between the Aβ amino acids and the drugs in the most populated binding modes. Error bars represent the standard deviation among the five drugs. The aromatic region (Phe19 and Phe20) shows a higher binding preference than the other region.

Nqtrp−R_2(Val18−Glu22), and Nqtrp−R_3(Gly37−Val40). Compared to the Autodock structure, the binding site is rearranged with more contacts formed and shifted to the residues Ala21 and Glu22. However, the overall structure of the binding site remains.

For model 8, the contact maps displayed in Figure S3 (Supporting Information) show that the binding site of the Autodock structure is formed by the contacts Nqtrp−R_1(Ile31−Leu34), Nqtrp−R_3(Leu17−Phe19), and there is no contact with a second peptide. Averaged over 80 ns, the binding site is characterized by contacts Nqtrp−R_1(Leu17−Glu22, Asn27− Gly33), Nqtrp−R_2(Leu17) and Nqtrp−R_3(Leu17−Phe20, Gly33−Gly38).

Finally, to fully characterize the binding in the two models, we constructed the free energy landscape of the binding sites by performing PCA[77] using the distances between the center of mass of Nqtrp with residues that make contacts with Nqtrp (see methods in Supporting Information).

Figure 10 displays the free energy landcape of the binding site for model 1. We can identify three main minima. The three states are very similar in the structures of the trimer, but are different in the structure and orientation of Nqtrp. To characterize these states in more detail, we analyzed the contact maps and hydrogen bonds of conformations belonging to those states. In the state S1 (42% of population), Nqtrp is in planar configuration (Figure10) and forms contacts Nqtrp−R_1(Leu17−Ala21), Nqtrp−R_2(Val18−Glu22,Lys28), and Nqtrp−R_3(Val36−Val40). On average, only a few hydrogen bonds are formed; H_{CO}(Glu22) (27%), H_{NH}(Ala21) (3%), and H_{NH}(Glu22) (6%). In the state S2 (20%), Nqtrp is in perpendicular configuration (Figure 10) and forms contacts Nqtrp−R_1(Leu17−Ala21, Ser26−Lys28), Nqtrp−R_2(Val18−Glu22), and Nqtrp−R_3(Gly37−Val40). Hydrogen bonds are formed; H_{CO}(Glu22) (18%), H_{CO}(Val39) (2%), and H_{NH}(Lys28) (8%). Structures of state S3 (8%) are quite similar to those in state S2 but differ in the percentage of formed hydrogen bonds.

For model 8, there are also three dominant binding states on the free energy landscape shown in Figure 11. Again, the trimeric Aβ structures in the three states are conserved. The binding site of state S1 (51%) is stabilized by contacts Nqtrp−R_1(Ala21, Asn27−Ile31), Nqtrp−R_3(Leu17, Leu34−Val36), and hydrogen bond H_{CO}(Gly29) (4%). In state S2 (15%), Nqtrp makes more contacts with the trimer Nqtrp−R_1(Leu17−Glu22, Asn27−Ile31), Nqtrp−R_3(Leu17, Leu34−Val36), and

dx.doi.org/10.1021/jp2118778 | J. Phys. Chem. B 2012, 116, 8412−8422

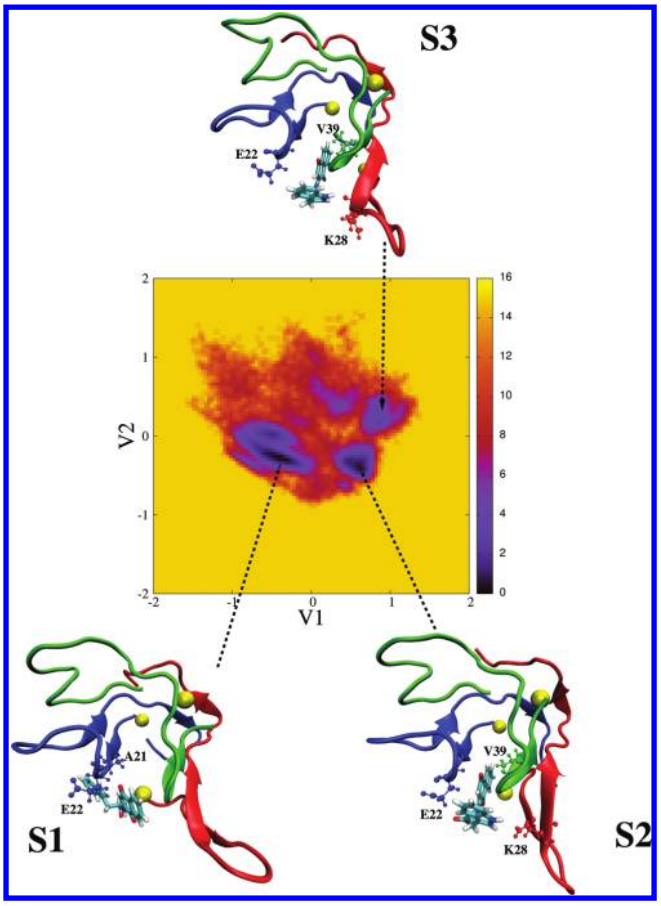


Figure 10. Free energy landscapes (in kcal/mol) of the binding site in the P1−Nqtrp model as obtained from 80 ns MD simulation. Shown are results along the first two principal components obtained from the PCA analysis. The color gradient from black to yellow is indicative of an increase in free energy. The centers of each cluster corresponding to the main minima are shown. For clarity, the yellow balls refer to the N-terminal residues. In each state, we show the Aβ residues that form H-bonds with Nqtrp.

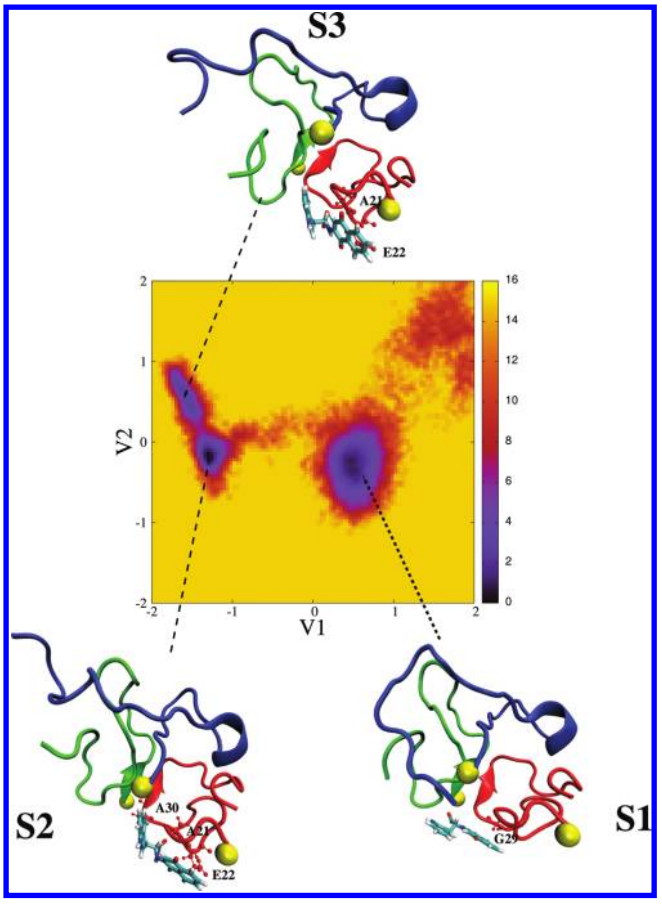


Figure 11. Free energy landscapes (in kcal/mol) of the binding site in the P8−Nqtrp model as obtained from 80 ns MD simulation. Shown are results along the first two principal components obtained from the PCA analysis. The color gradient from black to yellow is indicative of an increase in free energy. The centers of each cluster corresponding to the main minima are shown. For clarity, the yellow balls refer to the N-terminal residues. In each state, we show the Aβ residues that form H-bonds with Nqtrp.

also more hydrogen bonds are formed; H_{CO}(Ala21) (20%), H_{CO}(Glu22) (22%), and H_{CO}(Ala30) (20%). State S3 (11%) is similar to state S2 but only two hydrogen bonds are formed; H_{CO}(Ala21) (23%) and H_{CO}(Glu22) (10%).

Recently, Scherzer-Attali et al. have studied the interactions between Nqtrp with Aβ14−20, Aβ16−22, and Aβ18−24 trimers. By performing very long (25 microseconds) MD simulations using the CHARMM force field and a solvent-accessible surface-based implicit model, they showed that Nqtrp interacts with the peptides mainly through the amide group of residues Phe20, Ala21, and Glu22.[31] To compare with this work, we averaged the total number of hydrogen bonds between the carbonyl (amide) groups of Nqtrp and amide (carbonyl) groups of peptides over 80 ns MD trajectory, and results are shown in Figure S5, Supporting Information. For model 1, the interaction is more localized on the hydrogen bond H_{CO}(Glu22) (20%), and less on H_{NH}(Ala21) (2%), H_{NH}(Glu22) (3%), and H_{NH}(Lys28) (5%). For model 8, the interaction is delocalized on 4 hydrogen bonds: H_{NH}(Ala21) (8%), H_{NH}(Glu22) (5%), H_{NH}(Gly29) (4%), and H_{NH}(Ala30) (4%). Overall, our findings that the main H-bonds involve residues Phe19−Glu22 are consistent with Scherzer-Attali et al.[31] They are also consistent with 300 ns MD simulations of the Aβ42 momomer starting from an essentially α-helical conformation with 5 and 10 EGCG inhibitors and showing major binding sites with Phe19, Phe20, Glu22, Lys28, and Gly29.[78]

CONCLUSIONS

One key aspect supported by a large body of clinical and experimental evidence is that early Aβ oligomers are the primary toxic species, and in this respect, Aβ42 is more toxic than Aβ40. Many small molecule compounds are known to reduce both fibril formation and oligomers toxicity in cells, but their success in late stage clinical trials remains to be determined. Structural characterization of these Aβ-inhibitor oligomers using standard tools of biology is hampered by the complexity of the conformational dynamics of Aβ. One solution to this problem is the application of multiscale computational methods combining coarse-grained and all-atom models. In this study, we have determined the structures of Aβ17−42 trimers in isolation and with five known drugs: 2002-H20,[39] curcumin,[40] EGCG,[40] Nqtrp,[31] and resveratrol.[39] In addition to their toxicity, trimers are particularly intriguing since they are almost as efficient as preformed fibrils in nucleating assembly.[17] Our findings can be summarized as follows.

First, our extensive REMD-OPEP simulations of the Aβ17−42 trimer show that the preference for parallel β-sheet geometries of

dx.doi.org/10.1021/jp2118778 | J. Phys. Chem. B 2012, 116, 8412−8422

the peptides is not populated in aqueous solution at 300 K. The population of the fibril-prone conformation, N*, is 0.3%, and thus within statistical errors. Rather, each peptide behaves essentially as a compact coil/turn chain with little regular secondary structures, consistent with DMD simulations,[18] although there is a non-negligible configuration, referred to as P1, with a population of 19%, characterized by one chain with a β-hairpin spanning residue Phe19−Leu31, and the other two chains with a weak β-α-β-turn-β motif. The role of the present β-hairpin in oligomerization has been recently hypothesized based on the solution structure of Aβ40 in complex with a binding protein.[79]

Overall, our predicted coil/turn percentage of 84% is consistent with a previous CD study on small oligomers.[16] All generated conformations at 300 K can be described by a total of 14 clusters, and though the three peptides are asymmetric as seen by comparing their secondary structure profiles, they display very similar SASA profiles and a much higher preference for antiparallel than parallel orientation. The peptide's propensity for antiparallel geometry in the low order oligomers has been recently discussed by using infrared spectroscopy of Aβ40.[80] This orientational feature is also observed in the crystal structure of Aβ18−41 complexed/fused within the CDR3 loop of the IgNAR protein.[81] Note that it is not surprising that the construction of this Aβ−IgNAR chimera does not reveal a β-hairpin consisting of residues 19−31 but rather at another position, namely, residues 32−41.

Second, our extensive Autodock simulations reveal that the five drugs, 2002-H20, curcumin, EGCG, Nqtrp, and resveratrol, have different binding affinities for the conformational ensemble of trimeric Aβ17−42 and multiple binding modes although they interact preferentially with the CHC motif and notably the side chains of Phe19 and Phe20. This preference for the CHC region, confirmed by many experiments as a recognition element even in the absence of atomic level models for Aβ−drug oligomers,[30,82] is encouraging because the side chains of the predicted protein models were not free to move, and it is possible to go beyond the Autodock free-energy scoring function by using MD-based methods, though more time-consuming.[34,36,83]

Comparing the binding affinity of the five drugs on the possible conformational ensemble of Aβ17−42, 2002-H20 has a poor binding affinity, while Nqtrp is the best binder. Our result for 2002-H20 does not contradict the fact that it was identified among the top five compounds reducing Aβ40 toxicity from a collection of 17 905 compounds by using small molecule microarrays.[39] Rather, it suggests that 2002-H20 may be more efficient at the monomer level or at other oligomerization levels. Our result for Nqtrp is fully consistent with experiments showing that it is one of the most potent inhibitors of Aβ aggregation[31] and MD simulations showing that Nqtrp has the highest affinity for monomeric Aβ12−28 among a total of ten inhibitors.[36] Our Autodock analysis indicates that Nqtrp can bind strongly to 4 different trimeric Aβ17−42 conformations with similar binding affinities. In all cases, the side chains of Phe19 and Phe20 constitute the pharmacophore (see Figure S1, Supporting Information), but Nqtrp has multiple binding modes to trimeric Aβ17−42 conformation as revealed by our all MD atom simulations. This suggests that Nqtrp and the four other drugs are not optimized for Aβ and may bind other exposed β-sheet conformations in the extracellular space. In addition, we already know that these drugs are not specific to AD. For instance, resveratrol may be useful in the prevention of a wide range of pathologies, such as cardiovascular diseases and cancers and Parkinson's and Huntington's diseases.[84]

Of potential importance is the binding of Nqtrp to the P1 Aβ17−42 conformation. First, our picture of the binding sites is similar to that observed by NMR spectrocopy and MD simulations of Aβ12−28 monomer with Nqtrp.[31] Second, our dominant binding sites also suggest that Nqtrp may stabilize the β-hairpin conformation of one peptide preventing, therefore, higher-order β-sheet interactions. We are currently using all-atom REMD simulations of Aβ1−42 trimer with various concentrations of Nqtrp and with different force fields to evaluate the robustness of our results.

■ ASSOCIATED CONTENT

Supporting Information

Data analysis of the Nqtrp-Aβ MD simulations; structures of the fifteen protein models binding to drugs in the most populated poses; binding energies of five drugs upon protein models scored by Autodock; time evolution and distribution of the RMSD of three chains in the P1 and P8 A-Nqtrp models; contact maps between atoms of the Nqtrp and those of the peptides obtained from 80 ns trajectory and from the Autodock structures of P1 and P8 A-Nqtrp; averaged number of hydrogen bonds between the carbonyl (amide) groups of Nqtrp and amide (carbonyl) groups of peptides. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: philippe.derreumaux@ibpc.fr.

Notes

The authors declare no competing financial interest.

Present Address

[†]Département de Biologie et de Génomique Structurale, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de Santé et de Recherche Médicale (INSERM), U964/Centre National de Recherche Scientifique (CNRS), UMR 1704/Université de Strasbourg, Illkirch, France.

■ ACKNOWLEDGMENTS

Y.S. and P.D. were supported by the Centre National de la Recherche Scientifique and the Université Paris 7 Denis Diderot. P.D. also acknowledges the support of the "Institut Universitaire de France". N.M. is supported in part by the Canada Research Chair Foundation and the Natural Science and Engineering Research Council of Canada. P.D. and Y.M. acknowledge the support of the French-Singapore Merlion program 2009 (project 05.09.09) and the 2010 Ph.D. Merlion program.

■ REFERENCES

(1) Shankar, G. M.; Bloodgood, B. L.; Townsend, M.; Walsh, D. M.; Selkoe, D. J.; Sabatini, B. L. *J. Neurosci.* **2007**, *27*, 2866−2875.
(2) Paravastu, A. K.; Qahwash, I.; Leapman, R. D.; Meredith, S. C.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 7443−7448.
(3) Kuperstein, I.; Broersen, K.; Benilova, I.; Rozenski, J.; Jonckheere, W.; Debulpaep, M.; Vandersteen, A.; Segers-Nolten, I.; van der Werf, K.; Subramaniam, V.; et al. *EMBO J.* **2010**, *29*, 3408−3420.
(4) Sakono, M.; Zako, T. *FEBS J.* **2010**, *277*, 1348−1358.
(5) Wei, W.; Norton, D. D.; Wang, X.; Kusiak, J. W. *Brain* **2002**, *125*, 2036−2043.

(6) Renner, M.; Lacor, P. N.; Velasco, P. T.; Xu, J.; Contractor, A.; Klein, W. L.; Triller, A. *Neuron* **2010**, *66*, 739−754.
(7) Laurén, J.; Gimbel, D. A.; Nygaard, H. B.; Gilbert, J. W.; Strittmatter, S. M. *Nature* **2009**, *457*, 1128−1132.
(8) Calella, A. M.; Farinelli, M.; Nuvolone, M.; Mirante, O.; Moos, R.; Falsig, J.; Mansuy, I. M.; Aguzzi, A. *EMBO Mol. Med.* **2010**, *2*, 306−314.
(9) Xue, W.-F.; Hellewell, A. L.; Gosal, W. S.; Homans, S. W.; Hewitt, E. W.; Radford, S. E. *J. Biol. Chem.* **2009**, *284*, 34272−34282.
(10) Hureau, C.; Coppel, Y.; Dorlet, P.; Solari, P. L.; Sayen, S.; Guillon, E.; Sabater, L.; Faller, P. *Angew. Chem., Int. Ed.* **2009**, *48*, 9522−9525.
(11) Melquiond, A.; Dong, X.; Mousseau, N.; Derreumaux, P. *Curr. Alzheimer Res.* **2008**, *5*, 244−250.
(12) Yang, M.; Teplow, D. B. *J. Mol. Biol.* **2008**, *384*, 450−464.
(13) Sgourakis, N. G.; Merced-Serrano, M.; Boutsidis, C.; Drineas, P.; Du, Z.; Wang, C.; Garcia, A. E. *J. Mol. Biol.* **2011**, *405*, 570−583.
(14) Zhuang, W.; Sgourakis, N. G.; Li, Z.; Garcia, A. E.; Mukamel, S. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15687−15692.
(15) Cote, S.; Derreumaux, P.; Mousseau, N. *J. Chem. Theory Comput.* **2011**, *7*, 2584−2592.
(16) Kirkitadze, M. D.; Condron, M. M.; Teplow, D. B. *J. Mol. Biol.* **2001**, *312*, 1103−1119.
(17) Ono, K.; Condron, M. M.; Teplow, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 14745−14750.
(18) Urbanc, B.; Betnel, M.; Cruz, L.; Bitan, G.; Teplow, D. B. *J. Am. Chem. Soc.* **2010**, *132*, 4266−4280.
(19) Mitternacht, S.; Staneva, I.; Härd, T.; IrbÃäck, A. *J. Mol. Biol.* **2011**, *410*, 357−367.
(20) Kim, S.; Chang, W. E.; Kumar, R.; Klimov, D. K. *Biophys. J.* **2011**, *100*, 2024−2032.
(21) Huet, A.; Derreumaux, P. *Biophys. J.* **2006**, *91*, 3829−3840.
(22) Zheng, J.; Jang, H.; Ma, B.; Nussinov, R. *J. Phys. Chem. B* **2008**, *112*, 6856−6865.
(23) Masman, M. F.; Eisel, U. L. M.; Csizmadia, I. G.; Penke, B.; Enriz, R. D.; Marrink, S. J.; Luiten, P. G. M. *J. Phys. Chem. B* **2009**, *113*, 11710−11719.
(24) Horn, A. H. C.; Sticht, H. *J. Phys. Chem. B* **2010**, *114*, 2219−2226.
(25) Miller, Y.; Ma, B.; Nussinov, R. *J. Am. Chem. Soc.* **2011**, *133*, 2742−2748.
(26) Blinov, N.; Dorosh, L.; Wishart, D.; Kovalenko, A. *Biophys. J.* **2010**, *98*, 282−296.
(27) Tarus, B.; Straub, J. E.; Thirumalai, D. *J. Mol. Biol.* **2005**, *345*, 1141−1156.
(28) Chebaro, Y.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2009**, *113*, 7668−7675.
(29) Bitan, G.; Kirkitadze, M. D.; Lomakin, A.; Vollers, S. S.; Benedek, G. B.; Teplow, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 330−335.
(30) Amijee, H.; Madine, J.; Middleton, D. A.; Doig, A. J. *Biochem. Soc. Trans.* **2009**, *37*, 692−696.
(31) 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.; et al. *PLoS One* **2010**, *5*, e11101.
(32) Soto, P.; Griffin, M. A.; Shea, J.-E. *Biophys. J.* **2007**, *93*, 3015−3025.
(33) Chebaro, Y.; Derreumaux, P. *Proteins* **2009**, *75*, 442−452.
(34) Viet, M. H.; Ngo, S. T.; Lam, N. S.; Li, M. S. *J. Phys. Chem. B* **2011**, *115*, 7433−7446.
(35) Convertino, M.; Pellarin, R.; Catto, M.; Carotti, A.; Caflisch, A. *Protein Sci.* **2009**, *18*, 792−800.
(36) Convertino, M.; Vitalis, A.; Caflisch, A. *J. Biol. Chem.* **2011**, *286*, 41578−41588.
(37) Hochdörffer, K.; März-Berberich, J.; Nagel-Steger, L.; Epple, M.; Meyer-Zaika, W.; Horn, A. H. C.; Sticht, H.; Sinha, S.; Bitan, G.; Schrader, T. *J. Am. Chem. Soc.* **2011**, *133*, 4348−4358.
(38) Fradinger, E. A.; Monien, B. H.; Urbanc, B.; Lomakin, A.; Tan, M.; Li, H.; Spring, S. M.; Condron, M. M.; Cruz, L.; Xie, C.-W.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 14175−14180.
(39) Chen, J.; Armstrong, A. H.; Koehler, A. N.; Hecht, M. H. *J. Am. Chem. Soc.* **2010**, *132*, 17015−17022.
(40) Frydman-Marom, A.; Levin, A.; Farfara, D.; Benromano, T.; Scherzer-Attali, R.; Peled, S.; Vassar, R.; Segal, D.; Gazit, E.; Frenkel, D.; et al. *PLoS One* **2011**, *6*, e16564.
(41) Ehrnhoefer, D. E.; Bieschke, J.; Boeddrich, A.; Herbst, M.; Masino, L.; Lurz, R.; Engemann, S.; Pastore, A.; Wanker, E. E. *Nat. Struct. Mol. Biol.* **2008**, *15*, 558−566.
(42) Jang, H.; Arce, F. T.; Ramachandran, S.; Capone, R.; Azimova, R.; Kagan, B. L.; Nussinov, R.; Lal, R. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 6538−6543.
(43) Derreumaux, P.; Wilson, K.; Vergoten, G.; Peticolas, W. *J. Phys. Chem.* **1989**, *93*, 1338−1350.
(44) Derreumaux, P. *J. Chem. Phys.* **1997**, *107*, 1941.
(45) Forcellino, F.; Derreumaux, P. *Proteins* **2001**, *45*, 159−166.
(46) Maupetit, J.; Tufféry, P.; Derreumaux, P. *Proteins* **2007**, *69*, 394−408.
(47) Derreumaux, P.; Vergoten, G. *J. Chem. Phys.* **1995**, *102*, 8586−8605.
(48) Wei, G.; Mousseau, N.; Derreumaux, P. *Biophys. J.* **2004**, *87*, 3648−3656.
(49) Chen, W.; Mousseau, N.; Derreumaux, P. *J. Chem. Phys.* **2006**, *125*, 084911.
(50) Maupetit, J.; Derreumaux, P.; Tufféry, P. *J. Comput. Chem.* **2010**, *31*, 726−738.
(51) Maupetit, J.; Derreumaux, P.; Tufféry, P. *Nucleic Acids Res.* **2009**, *37*, W498−W503.
(52) Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2007**, *126*, 025101.
(53) Chebaro, Y.; Dong, X.; Laghaei, R.; Derreumaux, P.; Mousseau, N. *J. Phys. Chem. B* **2009**, *113*, 267−274.
(54) Santini, S.; Mousseau, N.; Derreumaux, P. *J. Am. Chem. Soc.* **2004**, *126*, 11509−11516.
(55) Song, W.; Wei, G.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2008**, *112*, 4410−4418.
(56) Simone, A. D.; Derreumaux, P. *J. Chem. Phys.* **2010**, *132*, 165103.
(57) Luhrs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dobeli, H.; Schubert, D.; Riek, R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17342−17347.
(58) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24.
(59) Berendsen, H.; Postma, J.; van Gunsteren, W.; Nola, A. D.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(60) Spill, Y.; Pasquali, S.; Derreumaux, P. *J. Chem. Theory Comput.* **2011**, *7*, 1502−1510.
(61) Frishman, D.; Argos, P. *Proteins* **1995**, *23*, 566−579.
(62) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. *J. Chem. Theory Comput.* **2007**, *3*, 26−41.
(63) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(64) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(65) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78*, 1950−1958.
(66) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172−3191.
(67) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox,

J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.01; Gaussian, Inc.: Wallingford, CT, 2009.
(68) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821−7839.
(69) Morris, G. M.; Huey, R.; Olson, A. J. *Current Protocols in Bioinformatics*; Wiley: New York, 2008; Chapter 8, Unit 8.14.
(70) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. *J. Comput. Chem.* **2007**, *28*, 1145−1152.
(71) Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1991**, *30*, 3864−3871.
(72) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
(73) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327−341.
(74) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(75) Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L. *Proteins* **2009**, *77*, 778−795.
(76) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.
(77) Jolliffe, I. T. *Principal Component Analysis*; Springer: Berlin, 2002.
(78) Liu, F. F.; Dong, X. Y.; He, L.; Middelberg, A. P.; Sun, Y. *J. Phys. Chem. B.* **2011**, *115*, 11879−87.
(79) Hoyer, W.; Gronwall, C.; Jonsson, A.; Stahl, S.; Hard, T. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5099−104.
(80) Sarroukh, R.; Cerf, E.; Derclaye, S.; Dufrene, Y. F.; Goormaghtigh, E.; Ruysschaert, J. M.; Raussens, V. *Cell. Mol. Life Sci.* **2011**, *68*, 1429−38.
(81) Streltsov, V. A.; Varghese, J. N.; Masters, C. L.; Nuttall, S. D. *J. Neurosci.* **2011**, *31*, 1419−1426.
(82) Gordon, D. J.; Sciarretta, K. L.; Meredith, S. C. *Biochemistry* **2001**, *40*, 8237−8245.
(83) Tuffery, P.; Derreumaux, P. *J. R. Soc. Interface* **2012**, *9*, 20−33.
(84) Richard, R.; Pawlus, A. D.; Iglesias, M. L.; Pedrot, E.; Waffo-Teguo, P.; Merillon, J.; Monti, J. P. *Ann. N.Y. Acad. Sci.* **2011**, *1215*, 103−108.