LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-837698

Self-assembly of the full-length amyloid A beta 42 protein in dimers

Y. Zhang, M. Hashemi, Z. Lv, Y. L Lyubchenko

July 19, 2022

Nanoscale

Disclaimer

Nanoscale

ARTICLE

Self-Assembly of the Full-length Amyloid Aβ42 Protein in Dimers

Yuliang Zhang,[a,b] Mohtadin Hashemi,[a] Zhengjian Lv,[a] and Yuri L. Lyubchenko*[a]

Received 00th August 2016,
Accepted 00th August 2016

DOI: 10.1039/x0xx00000x

www.rsc.org/nanoscale

The self-assembly of amyloid (Aβ) proteins into nano-aggregates is a hallmark of Alzheimer's disease (AD) development, yet the mechanism of how disordered monomers assemble into aggregates remains elusive. Here, we applied long-time molecular dynamics simulations to fully characterize the assembly of Aβ42 monomers into dimers. Monomers undergo conformational changes during their interaction, but the resulting dimer structures do not resemble those found in fibril structures. To identify natural conformations of dimers among a set of simulated ones, validation approaches were developed and applied, and a subset of dimer conformations were characterized. These dimers do not contain long β-strands that are usually found in fibrils. The dimers are stabilized primarily by interactions within the central hydrophobic regions and the C-terminal regions, with a contribution from local hydrogen bonding. The dimers are dynamic, as evidenced by the existence of a set of conformations and by the quantitative analyses of the dimer dissociation process.

Introduction

The primary model for the development of Alzheimer's disease (AD) is based on the self-assembly of amyloid β (Aβ) proteins, specifically the formation of Aβ aggregates.[1,2] Similar protein self-assembly mechanisms contribute to the development of other diseases, such as Parkinson's and Huntington's.[2-4] The final products of the amyloid self-assembly process are fibril structures that are typically characterized by traditional techniques such as solid state NMR.[5, 6] These studies showed that Aβ proteins in fibrils form long β-strands that help maintain the fibrillar structure. However, Aβ monomers are largely unstructured[7-9], which leads to the question of how the conformational transition occurs. Self-assembly is a kinetic process during which oligomers of various sizes are formed; therefore, understanding the self-assembly process requires experimental methods capable of characterizing transient species of the aggregation reaction. The finding that amyloid oligomers rather than fibrils are neurotoxic[10-14] suggests that the secondary structure of oligomers is different from that of fibrils. This assumption is supported by the characterization of oligomers using spectroscopic techniques that showed conformational changes occurring at early aggregation stages and no extensive formation of β-structures were reported.[12,15,16] Additionally, it is unclear how amyloid structures differ in oligomers of different types. The photo cross-linking methodology developed in ref.[17] allowed the authors to isolate individual oligomers to characterize their structure with circular dichroism spectroscopy; however, the potential contribution of the photo cross-linking procedure to the oligomer structure is a concern.

Single molecule method has been utilized to mechanical unfold the cysteine-engineered polyproteins.[18] We have recently developed an approach to probe amyloid dimers based on measuring the interactions of amyloid monomers with AFM force spectroscopy.[19-23] In this approach, monomers are end-immobilized to the AFM tip and the surface, and the interaction between them is measured by approaching the tip to the surface to allow dimer formation. The subsequent retraction step allows the rupture force required to dissociate the dimer to be measured. Using AFM dynamic force spectroscopy, we demonstrated that the dimer lifetimes were in the range of seconds. These values are several orders of magnitude larger than typical structural dynamics of Aβ monomers, suggesting that dimerization requires structural transitions within monomers that enable them to form stable complexes.[20-24] Similar results were obtained for other proteins and short amyloid peptides, leading us to conclude that the formation of a stable arrangement within dimers is a general phenomenon of the amyloid self-assembly process. However, these experiments did not allow us to analyze the structural features of the dimers. We recently performed a long-time scale molecular dynamics (MD) simulation for Aβ(14-23) peptide that enabled us to identify structural changes in the peptide upon dimer formation.[24] These simulations revealed many different structures, suggesting that validation procedures are needed to identify structures within experiments. In our recent publication,[25] we developed a validation approach termed Monte Carlo Pulling (MCP), in

[a] *Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, NE, 69198, USA. E-mail: ylyubchenko@unmc.edu*
[b] *Biology and Biotechnology Division, Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, 94550, USA.*

† Electronic Supplementary Information (ESI) available: Figures S1-S5: The Aβ42 monomer simulation; the time-dependent conformational changes from aMD simulations; the free energy landscape constructed after REMD simulations; the rupture patterns of additional major structures for Aβ42 dimer from aMD; the fluctuation of the Radius of Gyration (Rg) during the course of asymmetric and symmetric unraveling; movie S1-S3: the symmetric and asymmetric unraveling processes of dimers. See DOI: 10.1039/x0xx00000x

which we compared the rupture forces for MD models of dimers with the experimental data obtained in AFM force spectroscopy studies of the same Aβ(14-23) peptides. The rupture forces are strongly dependent on the pulling rates, and the developed approach allowed us to perform MCP simulations at rates identical to those in the experiment. As a result, the native dimer structure of Aβ(14-23) was identified.

Here, we applied MD simulation to analyze the dimer formation of biologically relevant full-size Aβ42 protein using the specialized Anton supercomputer.[26,27] A variety of dimer structures were identified and characterized. To validate the dimer structures, dimers were pulled using the MCP approach, and the simulated MCP rupture data were compared with the AFM experimental data. To further improve the stringency of the force spectroscopy based validation process, we developed a procedure utilizing experimental rupture pattern data to allow the most suitable dimers to be selected. An important characteristic of these dimers was the absence of long β-strands that are typically found in fibrils. Multiple interactions in the Central Hydrophobic Core (CHC) segments (Leu17 through Ala21) and the C-terminal regions (Ile31 and Ala42) stabilize the dimers. As an additional control, MCP simulations were performed with dimer structures adopted from the fibrils. Comparing the results with experimental data ruled out the possibility of dimer formation with high β-content in the self-assembled Aβ42 protein. The role of dimer structures and their dynamics in additional aggregation processes and possible dimer neurotoxicity are also discussed.

Simulation Methods

Monomer simulation procedure

To generate the initial structure of the monomers used for the dimer simulation, we conducted conventional molecular dynamics (cMD) simulation using GROMACS ver. 4.5.5[28] employing Amber ff99SB-ILDN force field[29] and the TIP3P water model.[30] The initial monomer structure (Fig. S1a) was adopted from NMR data[7] (PDB ID: 1IYT), obtained in hexafluoroisopropanol (HFIP):water (80:20 ratio), see details in Supplementary document. Five hundred nano seconds long NPT (constant Number, constant Pressure and constant Temperature) cMD simulation, at 1 bar and 300 K, were carried out using the Holland Computing Center (HCC). Following which, cluster analysis was performed through g_cluster command in the GROMACS package, with the GROMOS method of clustering and the root-mean square deviation (RMSD) for the protein backbone with a 3 Å cut-off value, as previously described.[24] Due to large structural fluctuations of residues 1–9 and 36–42, only data for residues 10–35 were selected for cluster analysis.

We addressed secondary structure dynamics according to the method developed by Thirumalai's group.[31] Briefly, if the dihedral angles from two consecutive residues satisfy the definition of an α-helix (-80° ≤ ϕ ≤ -48° and -59°≤ ψ ≤ -27°) and β-strand (-150° ≤ ϕ ≤ -90° and 90 ≤ ψ ≤ 150°), the structures are considered to be α and β conformations, respectively. The changes of secondary structure over time are monitored by, $\alpha(t) = \frac{1}{\Delta}\int_t^{t+\Delta} \alpha_s(s)ds$ and $\beta(t) = \frac{1}{\Delta}\int_t^{t+\Delta} \beta_s(s)ds$, where $\alpha(s) = \frac{1}{41}\sum_{i=1}^{41} \delta_{i,\alpha}$ and $\beta(s) = \frac{1}{41}\sum_{i=1}^{41} \delta_{i,\beta}$ at t =s and Δ=1 ns. When the residues adopt α or β conformations, the $\delta_{i,\alpha} = 1$ or $\delta_{i,\beta} = 1$.

Dimer simulation on the specialized supercomputer Anton

For simulations on Anton, we used the Maestro-Desmond software package Version 4.0 (Schrödinger, New York, NY, 2014) to build the initial dimer, using the same force field and water model as for the monomer cMD simulation. The dimer was created from copies of monomer with different orientations from most representative cluster in Fig. S1b, with the angle between the long axes of each monomer at 90°. More details can be found in the supplementary information. After running 4 μs cMD simulation on Anton, the last 150 ns trajectory of the simulation was used to determine the intermolecular contacts between the two monomers with the g_mdmat command from the GROMACS package.[28]

The calculation of the time-dependent secondary structure changes follows the same methods as described in the monomer simulation section. Here, $\alpha(s) = \frac{1}{82}\sum_{i=1}^{82} \delta_{i,\alpha}$ and $\beta(s) = \frac{1}{82}\sum_{i=1}^{82} \delta_{i,\beta}$ at t = s and Δ=1.2 ns.

Accelerated molecular dynamics simulation

To further extend conformational sampling, the resulting structures from the cMD simulations on Anton were subjected to the accelerated molecular dynamics (aMD) simulation method using STAMPEDE at the Texas Advanced Computing Center.[32] The simulation procedures were adapted from the description by Pierce et. al.[33] and the website (URL: http://ambermd.org/tutorials/advanced/tutorial22/). The procedure details are shown in the Supplementary document. The dimer systems were then submitted to STAMPEDE for a 500 ns NVT aMD simulation.

The Principal Component analysis of backbone dihedrals (dPCA),[34] in which the artifacts from combining internal and overall motion are minimized, was used to acquire the representative structures after the aMD simulation. In total, 250,000 structures were used for the dPCA analysis. The dihedral angles of the terminal residues were ignored. The following equation for the free energy calculations was used:

$$\Delta G(V1,V2) = -k_B T \ln\left(\frac{P(V1,V2)}{P_{max}}\right) \quad (1)$$

where V1 and V2 are the 1st and 2nd largest Principal Components; P(V1,V2) represents the distribution obtained from the MD trajectories, P_{max} is the maximum value of the distribution; and k_B and T are the Boltzmann constant and the absolute temperature, respectively. The Fortran program written by Dr. Yuguang Mu was used to perform this analysis.

Replica exchange MD simulations

Replica exchange MD (REMD)[35] simulation with a hybrid-resolution model, PACE,[36] were performed with the NAMD v2.9 package.[37] The procedure was adapted from Han and Schulten's method described in ref.,[36] in which amyloid fibril elongation was modeled. The initial structures were the same as in the aMD simulations, but were converted into a united atom description. Then, the systems were solvated with 2632 Martini[38] Coarse-Grained (CG) water and 150 mM NaCl. Thirty-two replicas were simulated in the temperature range 300-650 K, with exchange attempted every 4 ps; the exchange probability was 20-25%. After a 600 ns NVT (constant Number, constant Volume and constant Temperature) simulation for each replica, 12,500 structures at 300 K from the last 500 ns trajectory were used for the dPCA analysis.

Monte Carlo pulling (MCP) simulations

The MCP method that allows us to simulate AFM force spectroscopy experiments has been described in our recent publication.[25] MCP method is a modification of the PROFASI approach described in refs.[39,40] with FF08 force field and implicit water. The major feature of MCP method is that the pulling rate in the simulation is the same as in the experiment. Therefore comparison with the experiment is the validation test for the simulated dimer structure. Considering that the results from aMD and REMD simulation are more reliable than cMD; only the dimer structures obtained from these simulations were chosen for MCP. The two Cα of the N-terminal Cys residues of each monomer were defined as the pulling groups. A virtual spring was attached onto each pulling group and used to stretch them along a vector during the pulling process. The energy dynamics of the spring were calculated by the A2A spring function and the total energy in the course of pulling was described by the following equation:

$$E_{tot} = E(x) + \frac{k}{2}[L_0 + vt - L(x)]^2 \quad (2)$$

where E(x) indicates the energy without an external force, k and t are the spring constant of the virtual spring. L_0 is the initial distance between two Cα atoms of the N-terminal Cys residues of each monomer. L(x) represents the real-time distance between the Cα atoms of Cys residues during pulling, and x denotes a protein conformation. When v = 0.1 fm per MC step, the value is equivalent to 600 nm/s. Here, v = 0.083, equivalent to 500 nm/s, was used for all MCP simulations.

Results

Equilibrated structures of Aβ42 monomers

Prior to the simulation of Aβ42 dimer structures, we performed MD simulations of Aβ42 monomers to identify equilibrated structures of monomers. We applied this approach in our recent simulations of the dimer structures for Aβ(14-23) peptide[24]. We have shown that the equilibrated structure of the monomer does not depend on the initial conformation of Aβ(14-23) peptide. The Aβ42 monomer candidates were obtained by running 500 ns all-atom MD simulations using the explicit TIP3P water model. The initial structure was adopted from the NMR coordinates of the Aβ42 monomer (Fig. S1a, PDB ID: 1IYT) in an organic solution.[7] After completion of the simulation, we performed cluster analysis to identify the most representative structures for the Aβ42 monomer. The classification of clusters in the trajectory was performed with the approach described in ref.[41] by calculating the root-mean-square deviation (RSMD) of backbone atoms between all pairs of structures with a cut-off at 0.3 nm. Twelve clusters were identified, with the largest cluster comprising 53.73% of the entire population, the representative structure for this cluster is shown in Fig. S1b.

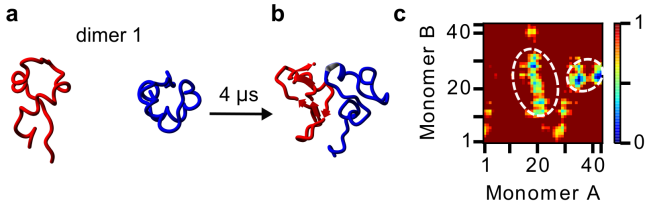

Fig. 1. Results of the MD simulation of Aβ42. (a) Schematic for the assembly of a dimer in which one monomer is rotated by 90° relative to another. (b) The snapshots of the dimer corresponding to the last frames of the 4 μs simulation runs. In all snapshots (a, b), the red and blue colors indicate monomer A and B, respectively. (c) The contact map for the dimer. The colors in the contact map represent the distance in nm between the pairwise residues. The regions of interest are encircled with dashed lines.

To structurally characterize monomer dynamics, we monitored the overall secondary structural changes according to the method described in ref.[31] (details in Methods). The analysis of the time-dependent contents of α-helix and β-structures (α(t) and β(t)) reveals that the α-to-β transition occurs after 200 ns (Fig. S1c), indicating that the ability to detect such a conversion requires long simulation times. The monomer structure remains quite dynamic, so in the time-interval between 200 ns and 500 ns, the α-helix and β-strand contents fluctuate rather broadly, 0.08 ± 0.05 and 0.03 ± 0.03, respectively. The conformational variability is primarily defined by the conversion of two stable helical conformations of the initial structure encompassing residues Ser8-Gly25 and Lys28-Met35 to conformers with low α-helix and low β-strand contents. These findings are consistent with simulation results for Aβ42 monomers by other groups that showed that the representative monomer structure is primarily a random coil with a small helical segment around the Arg5-Gln15 region.[8, 42] Based on the analysis of the conformation and dynamics of Aβ42 monomers, we selected monomers from the 1st cluster to analyze Aβ42 dimer conformational dynamics with the Anton supercomputer.

Structure of Aβ42 dimers

Two Aβ42 monomers were placed at a center of mass distance (COM) of 4 nm to generate the initial dimer configurations (Fig. 1a). The snapshot for dimers after 4 μs is shown in Fig. 1b. In the dimer, interactions between monomers are limited to short stretches of the protein, as depicted in the dimer interaction map in Fig. 1c. The primary interactions between the monomers in dimers are within the CHC region spanning from Leu17 to Ala21, and the C-terminal region between Ile31 and Ala42. To determine whether the structure reached equilibrium, the time-dependent change in the dimers secondary structure was calculated. The variability of α-helix (α(t)) and β-structures (β(t)) is shown in Fig. 2a. The graph shows that both parameters in the dimers initially (200 ns – 500 ns) fluctuate in the range of 50% or above, but remain constant after the 2 μs simulation time point, which suggests that the structure reached the local equilibrium state. These findings are further confirmed by the time-resolved change in the angle between the two monomers, depicted in Fig. 2b. However, we cannot exclude the possibility that the dimers are trapped in local energy minima, meaning that the conformational space is not sufficiently sampled during the simulations.

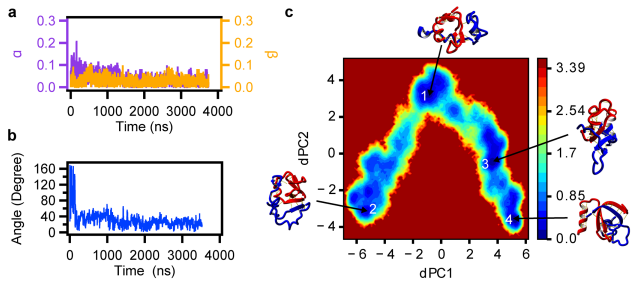

Fig. 2. (a) Time-dependent dynamics of the α-helix and β-strand contents in the dimer obtained from the cMD simulations (Anton computer). The α-helix and β-strand content variations over time are represented by the purple and orange colors, respectively. (b) Time-dependent orientation of one monomer with respect to the other during 4 μs simulation on Anton. In the calculation, the orientation was obtained according to the relative angle between two vectors, which were selected to be between Cα atoms of Y10 and M35 on each monomer. (c) The results from the aMD simulation of Aβ42 dimer. The free energy landscapes were constructed after a 500 ns aMD. The local energy minima are indicated with arrows. The corresponding snapshots are shown, with monomers A and B represented by red and blue, respectively. The dashed lines indicate the hydrogen bonds.

To address this issue, we extended the dimer simulation using accelerated MD (aMD) simulation (see specifics in Methods), in which two boost energies are introduced to the system to enhance conformational sampling.[43-45] According to ref.,[33] sampling during several hundred nanoseconds of aMD simulation is equivalent to sampling in the millisecond time scale for conventional MD (cMD) simulations, suggesting that we should be able to extend the sampling efficiency by several orders of magnitude with aMD. Of note, the aMD approach was recently used to analyze the larger Aβ-αSyn co-assembly system,[46] thereby justifying the suitability of this approach for Aβ42 dimer simulation.

The results of the 500 ns aMD simulation for the dimer, depicted as an energy landscape, are shown in Fig. 2c, in which the dihedral Principle Component Analysis (dPCA) was applied to generate the energy landscape.[34] A series of well-defined energy minima, shown in blue, are identified. Snapshots of representative structures corresponding to these local minima are indicated in the figure. As seen from the structures, even an extended MD simulation did not lead to formation of dimer structures with long β-strands. A common feature of all structures is the dimer stability provided by the interactions of the CHC regions (Leu17-Ala21) and the C-terminal hydrophobic regions of the monomers, which is consistent with the results of the cMD simulations. The dimer is stabilized by intermolecular β-sheet structures. A conformational analysis of the α-helix and β-strand content of the dimers as a function of time, following aMD simulations, revealed a slightly higher β-strand content (~0.05 vs. ~ 0.02), as shown in Fig. S2.

We compared the results of the aMD simulations for the Aβ42 dimer structures with those obtained with REMD,[35] which is a more common approach used to improve sampling.[47-49] We utilized a hybrid REMD simulation approach, in which proteins are defined using united atom parameters, and Martini water is used as the solvent. The energy landscape plot is shown in Fig. S3a. A set of minima was also identified and representative snapshots of the dimers are shown in the figure. Similar to the results obtained with aMD simulations, the dimers do not contain β-sheet stretches; rather, a few residues are involved in intermolecular β-sheet formation. Similar to aMD simulation, the CHC segments (Leu17 through Ala21) and the C-terminal regions (Ile31 and Ala42) stabilize the dimer structure (Fig. S3b). Interestingly, in REMD simulation, the C-terminal residues tend to form β-hairpin structures that increase the interaction with the CHC region (Fig. S3a, minimum 2).

Validation of MD simulations

To validate the MD simulation data, we used our recently developed MCP approach,[25] in which the simulated structures are pulled and the calculated rupture forces are compared with experimental results. Such a comparison is rigorous, as the rupture process is simulated with MCP at pulling rates similar to those used in experimental conditions. This is a critical issue because the rupture force depends on the pulling rate.[24] Four initial configurations were selected, from the representative dimers obtained in aMD simulations (Fig. 2c), and MCP simulations were carried out for each dimer configuration. In order to obtain a statistically significant dataset, MCP simulations were repeated 500 times for each dimer conformation.

The results for MCP simulations for dimer conformation 1 (Fig. 2b) are shown in Fig. 3, along with experimental results from our previous publication.[23] The data in Fig. 3 demonstrate

that force distributions for simulated dimers are very similar to the rupture force values obtained in the experiments. The geometric mean peak values are 58.1 ± 1 pN (Fig. 3a) and 57.1 ± 1 pN (Fig. 3b) for the simulations and experiments respectively. Statistical analysis performed with the Kolmogorov-Smirnov nonparametric test produced p = 0.346. AFM force spectroscopy experimental data provide an additional parameter–the position of the rupture events that characterizes the location of interacting regions.[20]

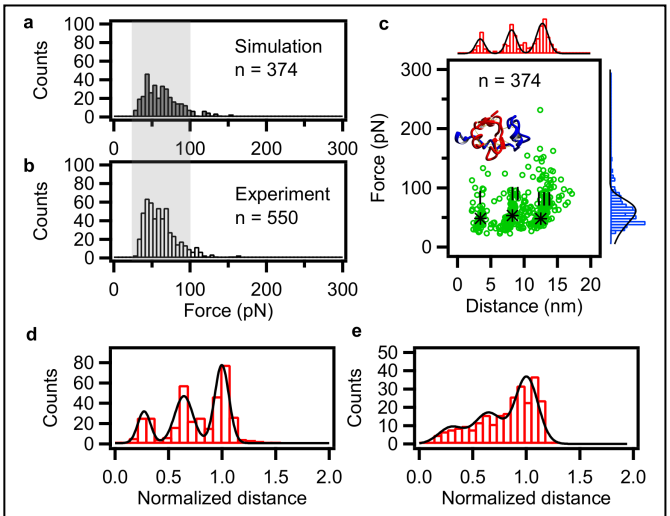

Fig. 3. The rupture patterns. (a, b) The rupture force distributions for the simulation and the experiment, respectively. The number of force curves analyzed (n) is indicated. (c) The rupture patterns of Aβ42 dimers obtained from the MCP simulations corresponding to the first energy minima in Fig. 2b. The scatter plots (green circles) show the rupture forces corresponding to a specific rupture distance. The rupture force distributions (blue) are shown along the y-axis to the right of the plots, and the histograms of the rupture distances (red) are on the topside of the scatter plots. The black * symbols are the centroids from the highest density areas of the scatter points, indicating the different patterned regions, labeled I through III. The black curves are the fits. (d) The rupture patterns of MCP simulations of dimer in c. (e) The experimental data from the previous work.[23] The black curves are obtained from Gaussian fitting.

We applied this approach to characterize the effect of single amino acid substitutions in C-terminus of Aβ42 on the rupture pattern of the dimers.[23] These studies demonstrated that the rupture pattern is very sensitive to the mutations in Aβ42 and suggest that the C-terminal residues provide an important contribution to Aβ42 dimer stability. Given such a high sensitivity of the Aβ42 rupture pattern to the protein sequence, we computed a similar dataset to get additional validation to simulations (see Methods for specifics). The results are summarized as a scatter plot in Fig. 3c in which each calculated data point is shown with a green circle. The data can be divided into three groups numbered in the graph. The grouping is clearly seen in the distributions of the rupture distances obtained from this dataset shown as a red histogram on the top that has three distinct peaks. This dataset suggests that there are three distinct interacting regions, corresponding to the rupture distances of 3.4 ± 0.7 nm, 8.2 ± 1.0 nm, and 12.7 ± 1.3 nm for the dimers (Fig. 3c and Fig. S4). We compared these data with the experiment and the comparison is shown in Fig. 3d and e. Both the experiment and the simulation demonstrate that rupture events for peak 1 are less representative and the highest yield of rupture events corresponds to peak 3. Experimental values for the peaks positions (4.6 ± 1.5 nm, 9.6 ± 1.4 nm and 14.8 ± 1.5 nm) were calculated by subtracting the length of tethers from the contour length; however, the length of the PEG tether provides more than half of the value, thereby resulting in a systematic error in the peak position values.[23] Therefore, it is reasonable to use normalized values of peak positions by dividing the peak positions by the position of the last peak. The area ratio for the simulations (Fig. 3d) is 1:2.1: 2.7, from the shortest to the longest rupture distance peak. The experimental results (Fig. 3e) show a very similar three-peak distribution, with an area ratio of 1:2.1:3.7. The peak positions for the experimental data, 0.31 ± 0.11, 0.65 ± 0.09 and 1 ± 0.1 coincide with the peak positions for the simulations, 0.27 ± 0.06, 0.65 ± 0.08 and 1 ± 0.1.

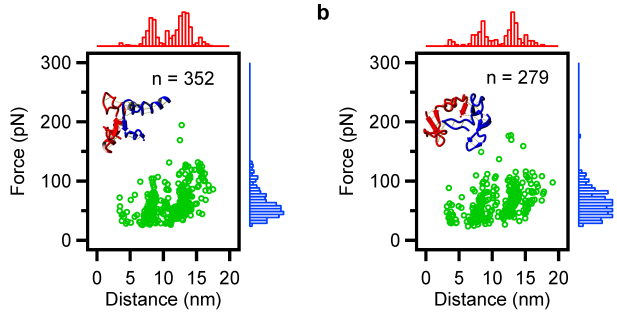

Fig. 4. The rupture patterns (a and b) of Aβ42 dimers obtained in REMD simulations (Fig. S3a; structures 1 and 2) generated by MCP simulations. Similar to Figures 3 and S4, each rupture force and the corresponding rupture distance are indicated with green circles. The force distributions are shown as blue histograms and placed on the right side of the scatter plot. The rupture lengths distributions (red histogram) are placed at the top of the plot. The snapshots of dimers in each figure are the initial structures.

A similar analysis was performed for the dimers obtained in the REMD simulations. These structures were validated using the MCP approach and the results for structures 1 and 2 are shown in Fig. 4a and b, respectively. The force distributions for both structures, shown in blue, are placed to the right of the graphs. The rupture force values are 59.4 ± 1.5 pN and 58.0 ± 1.5 pN for structures 1 and 2, respectively. Both values are close to the experimental values (Kolmogorov-Smirnov nonparametric test, p = 0.4 and p = 0.02, respectively). Next, we calculated the rupture profiles (the data points are shown in green), and the rupture patterns are shown as red histograms at the top of each graph. The rupture length histograms show that structure 1 has two peaks, whereas structure 2 has three peaks, with positions close to those

found in the experiment. Therefore, according to the rupture profile criterion, structure 2 is preferred.

Rupture process for dimer dissociation

To understand the role of various interacting segments in the dissociation process, we characterized the rupture events corresponding to the three peak positions identified in Fig. 3c. The results for all three types of rupture events are shown in Fig. 5. The black lines show force curves, and snapshots of the dimer structures immediately prior to the rupture event are presented above the force curves. The dimer dissociation process corresponding to the shortest rupture distance events (type I events) is illustrated in Fig. 5a.

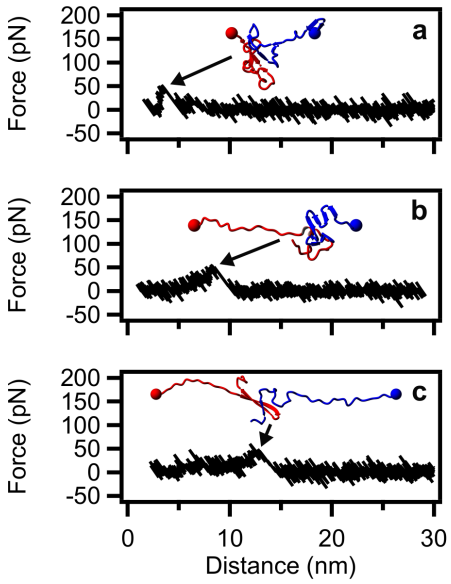

Fig. 5. The typical rupture events for simulations of dimer dissociations. The force curves (black lines in a–c) are taken from the different patterned regions I–III (Fig. 3c), respectively. The snapshots corresponding to the structures immediately prior to dissociation are indicated with arrows. The monomer color-coding is the same as described above. Blue and red balls represent the N-terminal residues for the monomers.

The full dynamic process can be viewed as a movie (movie S1). The analysis shows that the rupture process occurs after the partial unravelling of both N-terminal regions (Fig. 5a).

The dissociation processes, identified as type II events in Fig. 3c, show a pronounced asymmetry during the dimer unraveling. A typical force curve is shown in Fig. 5b. During the dissociation, the N-terminal region of one monomer gradually unfolds, extending the total length of the N-terminal region (Asp1–Lys16), which becomes as long as ~5 nm. The rest of the protein, as well as the other monomer, remain a compact globule with a diameter of 2–3 nm. The full process can be seen in movie S2. During dissociation, the CHC region of the dimer and the C-terminal region on one monomer remain compact and stabilized by hydrophobic interactions. The eventual rupture distance is ~8 nm, calculated from the combined length of the extended N-terminal region (Asp1–Lys16; ~5 nm) and the diameter (~3 nm) of the compact cluster (gray regions in Movie S2). Typically, when asymmetric unraveling occurs, one monomer maintains the compact conformation while the other monomer gradually unfolds, accompanied by an increase in the Radius of gyration (Rg) (Fig. S5a).

The dimers dissociation process for the longest rupture distance events (type III) is illustrated in Fig. 5c. The unraveling process can be viewed as a movie (Movie S3). Type III events are characterized by the extension of almost the entire Aβ42 chain starting from the N-termini. The dimer at the end of the rupture process is stabilized by interacting hydrophobic segments of the C-termini. During the dissociation process, monomer unfolding is accompanied by a gradual increase in Rg values for both monomers (Fig. S5b). The total length of the dimer prior to separation is ~12 nm, composed of the length of two extended N-termini (10 nm) and the diameter of the hydrophobic cluster (~2 nm; gray regions in Movie S3).

Rupture process of dimers within fibrils

The major finding of the MD simulations is that Aβ42 dimers do not contain long β-strand structures as those found in fibrils. To characterize the rupture process of Aβ42 dimers with structures similar to those found in fibrils, an MCP analysis was performed for Aβ42 protein dimers with initial structures adopted from fibrils to mimic high β-content structures. We used the dimer structures from U-shaped fibrils,[5] as well as the dimer structures adopted from the recently published S-shaped fibrils,[6] schematically shown as U-shaped and S-shaped dimers in Fig. 6a–d.

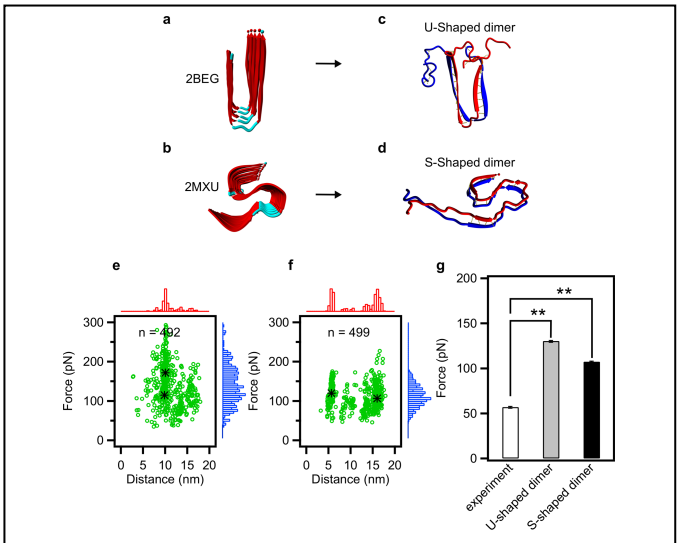

Fig. 6. The pulling (MCP) simulation of Aβ42 dimers from two fibrillar structures – U-shaped (a, PDB ID: 2BEG) and S-shaped (b, PDB ID: 2MXU). The missing residues at N-terminus have been added to mimic the full-length of Aβ42 peptide. The corresponding dimer structures are shown in (c) and (d). The monomers in (c) and (d) are shown in different colors. The rupture force and rupture distance patterns are shown in (e) and (f) for U-shaped and S-shaped dimers, respectively. (g) The Kolmogorov-Smirnov nonparametric statistical analysis for the correlation between the experiment and the simulations. Symbols ** indicate $p < 0.01$. Error bars are S.E. values.

The results for the rupture force distributions are shown in Figs. 6e and f. The mean values of the rupture forces were 130.4 ± 1 pN and 107.4 ± 1 pN for U-shaped and S-shaped dimers, respectively. Both values are significantly different from the experimental data (Fig. 6g), suggesting that the probability of forming these types of dimers is very low.

Discussion

Our thorough computational analysis was able to produce the atomic structure of Aβ42 dimers and reveal their dynamic properties. An important consideration in this analysis is the use of monomers with equilibrated structures. Note that the monomer structures obtained in our simulations were very close to those obtained in refs.,[8,42] in which the REMD approach was used. Dimer formation is accompanied by conformational changes of monomers with the formation of short α-helices and β-structures (Fig. S2). As a result, the dimers are stabilized primarily by interactions of the CHC segments (Leu17 through Ala21). The dimer is stabilized by the interaction of the C-terminal regions, but similar to the CHC region, the C-terminus of the dimer also does not have a high β-content (Fig. 2). The hydrophobic interaction, rather than backbone hydrogen bonding, is crucial for the Aβ42 interaction, consistent with recent evidence.[50] Qualitatively, these findings are in agreement with the characterization of early-stage Aβ42 aggregates by CD and ThT fluorescence that show that the oligomers have a low β-content.[17,51]

The simulated dimer structures were validated by comparison with AFM pulling experimental results for Aβ42 dimers using the MCP approach.[25] Importantly, the simulations were performed at conditions identical to experimental conditions, and the comparison was made over large data sets, enabling the comparison of both the mean rupture force values and the force distributions, as shown in Fig. 3. Both parameters for the simulations are in a concurrence with experimental results, providing a strong validation for the structures of the Aβ42 dimers. We took advantage of the MCP approach to simulate the rupture process for any structure and performed pulling simulations for hypothetical dimers with long β-strands that appear in fibrils.[5,6] The simulated structures (Fig. 6) produce rupture force values that are considerably larger than those obtained in the experiments. Therefore, we can rule out the possibility of Aβ42 monomers self-assembling into dimers with high β-content structures, as found in fibrils. This finding led us to speculate that the formation of β-sheets is a characteristic structural feature of later stages of Aβ42 aggregates associated with the formation of fibrils.

The energy landscape of dimers is rather rough and is characterized by a set of local minima, as shown in Fig. 2. The roughness of the energy landscape suggests that Aβ42 dimers are very dynamic and that various stable states are probed by monomers. These states cannot reliably be distinguished by the rupture forces alone, although the depths of the minima vary. Several types of dimers were obtained, however they have major structural similarities and also produce rather similar rupture forces during MCP simulations. To be able to distinguish between these structures, as demonstrated in Figs. 3, 4, and S3, the rupture profiles were compared. The analysis of the rupture process (Fig. 5) identified fine features of the dynamics of Aβ42 dimers. Upon pulling, the N-terminus of one monomer extends while the other monomer remains unchanged. This asymmetry of the dimer structure may play a role in the next stages of the self-assembly process and create the directionality of aggregate growth, as observed in experiments.[52] Indeed, the substitutions at C termini alter dramatically the Aβ42 monomer-monomer interaction.[51] For instance, (Val36/D-Pro)-(Gly37/L-Pro) replacement destabilize the destabilized the C terminal-turn structure on Aβ42, leading to "Aβ40-like" interaction.[53] These finding are perfectly in line with our AFM force spectroscopy studies that demonstrate that these mutations resulted in dramatic changes in the profiles of rupture forces.[23]

The existence of multiple energy minima on the energy landscape of the dimers has a number of biological implications. Aggregation of Aβ42 and other amyloids is the process in which aggregates with different morphologies are formed. One model suggests the existence of different aggregation pathways for aggregates with different morphologies (ref.[13,54-59] and references therein), with the structure specific dimers serving as origins for these pathways. It is widely accepted that oligomers are the most neurotoxic species of amyloids. This is also supported by the evidence that Aβ42 dimers are neurotoxic.[60] The neurotoxic effect of Aβ42 dimers assumes that dimers interact with a multitude of other proteins and cellular membranes, which can be facilitated by the structural plasticity of Aβ42 dimers.

Aβ42 dimerization has been modeled in the past,[48, 61-65] revealing various conformations including α-helixes and anti-parallel β-sheets. These data are in contrast with our findings. There are two major factors explaining these differences. First, we assembled dimers by using equilibrated monomer structures, which was not performed in any of the prior publications. Second, we performed simulations of the dimer formation on the long-time scale: ~4 μs for the initial cMD simulation with the Anton supercomputer, followed by aMD simulations that extend sampling to the millisecond time scale, which is several orders of magnitude higher than the previously published data.

Through these MD simulations we used Amber ff99SB-ILDN force field that was applied in the all-atom MD simulation of short fragments of Aβ system.[66,67] Our simulations show a good correlation with experimentally data providing additional support for the use if this field amyloid type proteins. Recently, the ff99IDPs has been developed for IDP monomer simulation and demonstrate the good consistency with NMR data.[68] It would be interesting to compare these two force fields and these are long-range plans. For REMD simulations we used

implicit MARTINI water approximation. Water play critical roles in Aβ protein aggregation.[69,70] However, rather good correlations of Aβ42 dimers conformations obtained by REMD and aMD simulations justifies the use MARTINI approximation in this work, although further use of this model needs additional justifications.

Conclusions

MD simulations are widely applied to model various systems, including amyloid aggregates. A major concern in these modeling approaches is the validation of the obtained structure. It is usually necessary to compare simulations with experiments, but selecting an experimentally testable parameter is problematic for transient systems such as amyloid oligomers. The developed approaches fill this gap and provide appropriate tools to test simulated structures. Importantly, the simulation can be performed at conditions identical to those in AFM force spectroscopy experiments11. In the initial validation test, the simulated rupture forces are compared with the experimentally determined values and the selection is made based on the best fit to the experimental rupture forces. The comparison is done for multiple simulations to make the results statistically rigorous. Moreover, we developed and tested another validation method in which a different experimental parameter from the AFM force experiment, the rupture pattern, is compared with the simulated results. The combination of the two validation criteria allowed us to increase the stringency in the selection of the computational models. The proposed approach can be extended to other complexes that can be probed by AFM force spectroscopy experiments.

Acknowledgements

We thank Dr. Dittrich and his colleagues at the Pittsburg Supercomputing Center (PSC) for their training on the Anton supercomputer and their advice; Dr. Schulten for the VMD and NAMD software package; Dr. Swanson at the Holland Computing Center (HCC) for advice regarding the use of the HCC facility; and the Texas Advanced Computing Center at the University of Texas at Austin and the Extreme Science and Engineering Discovery Environment (XSEDE) for the use of their facility for aMD simulations, supported by NSF grant ACI-1053575 for XSEDE. This work was supported by NIH grants GM096039 and GM100156, NSF grant 1004094, and the PSCA14025P award for computer time on Anton at PSC - all to YLL. MH was partially supported by Bukey Memorial Fellowship. Manuscript preparation (YZ) was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Notes

The authors declare no competing financial interest.

References

1 J. Hardy and D. J. Selkoe, *Science*, 2002, **297**, 353-356.
2 C. M. Dobson, *Nature*, 2003, **426**, 884-890.
3 F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2006, **75**, 333-366.
4 A. T. Petkova, R. D. Leapman, Z. Guo, W.-M. Yau, M. P. Mattson and R. Tycko, *Science*, 2005, **307**, 262-265.
5 T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert and R. Riek, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 17342-17347.
6 Y. Xiao, B. Ma, D. McElheny, S. Parthasarathy, F. Long, M. Hoshi, R. Nussinov and Y. Ishii, *Nat. Struct. Mol. Biol.*, 2015, **22**, 499-505.
7 O. Crescenzi, S. Tomaselli, R. Guerrini, S. Salvadori, A. M. D'Ursi, P. A. Temussi and D. Picone, *Euro. J. Biochem.*, 2002, **269**, 5642-5648.
8 N. G. Sgourakis, M. Merced-Serrano, C. Boutsidis, P. Drineas, Z. Du, C. Wang and A. E. Garcia, *J. Mol. Biol.*, 2011, **405**, 570-583.
9 S. Vivekanandan, J. R. Brender, S. Y. Lee and A. Ramamoorthy, *Biochem. Biophys. Res. Commun.*, 2011, **411**, 312-316.
10 C. G. Glabe, *Neurobiol. Aging*, 2006, **27**, 570-575.
11 C. G. Glabe, *J. Biol. Chem.*, 2008, **283**, 29639-29643.
12 L. Yu, R. Edalji, J. E. Harlan, T. F. Holzman, A. P. Lopez, B. Labkovsky, H. Hillen, S. Barghorn, U. Ebert, P. L. Richardson, L. Miesbauer, L. Solomon, D. Bartley, K. Walter, R. W. Johnson, P. J. Hajduk and E. T. Olejniczak, *Biochemistry*, 2009, **48**, 1870-1877.
13 A. Laganowsky, C. Liu, M. R. Sawaya, J. P. Whitelegge, J. Park, M. Zhao, A. Pensalfini, A. B. Soriaga, M. Landau, P. K. Teng, D. Cascio, C. Glabe and D. Eisenberg, *Science*, 2012, **335**, 1228-1231.
14 P. Liu, M. N. Reed, L. A. Kotilinek, M. K. Grant, C. L. Forster, W. Qiang, S. L. Shapiro, J. H. Reichl, A. C. Chiang, J. L. Jankowsky, C. M. Wilmot, J. P. Cleary, K. R. Zahs and K. H. Ashe, *Cell Rep.*, 2015, **11**, 1760-1771.
15 M. Ahmed, J. Davis, D. Aucoin, T. Sato, S. Ahuja, S. Aimoto, J. I. Elliott, W. E. Van Nostrand and S. O. Smith, *Nat. Struct. Mol. Biol.*, 2010, **17**, 561-567.
16 A. Sandberg, L. M. Luheshi, S. Sollvander, T. Pereira de Barros, B. Macao, T. P. Knowles, H. Biverstal, C. Lendel, F. Ekholm-Petterson, A. Dubnovitsky, L. Lannfelt, C. M. Dobson and T. Hard, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 15595-15600.
17 K. Ono, M. M. Condron and D. B. Teplow, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 14745-14750.
18 H. Dietz and M. Rief, *Proc. Natl. Acad. Sci. USA*, 2006, **103**, 1244-1247.
19 Y. L. Lyubchenko, B. H. Kim, A. V. Krasnoslobodtsev and J. Yu, *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.*, 2010, **2**, 526-543.
20 A. V. Krasnoslobodtsev, J. Peng, J. M. Asiago, J. Hindupur, J. C. Rochet and Y. L. Lyubchenko, *PloS one*, 2012, **7**, e38099.
21 A. M. Portillo, A. V. Krasnoslobodtsev and Y. L. Lyubchenko, *J. Phys. Condens. Matter.*, 2012, **24**, 164205.
22 B. H. Kim and Y. L. Lyubchenko, *Nanomedicine: nanotechnology, biology, and medicine*, 2014, **10**, 871-878.
23 Z. Lv, R. Roychaudhuri, M. M. Condron, D. B. Teplow and Y. L. Lyubchenko, *Sci. Rep.*, 2013, **3**.
24 S. Lovas, Y. Zhang, J. Yu and Y. L. Lyubchenko, *J. Phys. Chem. B*, 2013, **117**, 6175-6186.
25 Y. Zhang and Yuri L. Lyubchenko, *Biophys. J.*, 2014, **107**, 2903-2910.
26 D. E. Shaw, R. O. Dror, J. K. Salmon, J. P. Grossman, K. M. Mackenzie, J. A. Bank, C. Young, M. M. Deneroff, B. Batson, K. J. Bowers, E. Chow, M. P. Eastwood, D. J. Ierardi, J. L.

Klepeis, J. S. Kuskin, R. H. Larson, K. Lindorff-Larsen, P. Maragakis, M. A. Moraes, S. Piana, Y. Shan and B. Towles, presented in part at the Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, Portland, Oregon, 2009.
27 D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan and W. Wriggers, *Science*, 2010, **330**, 341-346.
28 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435-447.
29 K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins*, 2010, **78**, 1950-1958.
30 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
31 G. Reddy, J. E. Straub and D. Thirumalai, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 11948-11953.
32 T. A. C. C. (TACC). Texas Advanced Computing Center (TACC). The University of Texas at Austin).
33 L. C. Pierce, R. Salomon-Ferrer, F. d. O. C. Augusto, J. A. McCammon and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 2997-3002.
34 Y. Mu, P. H. Nguyen and G. Stock, *Proteins*, 2005, **58**, 45-52.
35 Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.*, 1999, **314**, 141-151.
36 W. Han and K. Schulten, *J. Am. Chem. Soc.*, 2014, **136**, 12450-12460.
37 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781-1802.
38 S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *J. Phys. Chem. B*, 2007, **111**, 7812-7824.
39 A. Irback and S. Mohanty, *J. Comput. Chem.*, 2006, **27**, 1548-1555.
40 S. A. E. Jonsson, S. Mitternacht and A. Irback, *Biophys. J.*, 2013, **104**, 2725-2732.
41 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chem. Int. Ed. Engl.*, 1999, **38**, 236-240.
42 O. Wise-Scira, L. Xu, T. Kitahara, G. Perry and O. Coskuner, *J. Chem. Phys.*, 2011, **135**, 205101.
43 D. Hamelberg, J. Mongan and J. A. McCammon, *J. Chem. Phys.*, 2004, **120**, 11919-11929.
44 J. Wereszczynski and J. A. McCammon, *J. Chem. Theory Comput.*, 2010, **6**, 3285-3292.
45 Y. Wang, C. B. Harrison, K. Schulten and J. A. McCammon, *Comput. Sci. Discov.*, 2011, **4**.
46 J. C. Jose, P. Chatterjee and N. Sengupta, *PloS one*, 2014, **9**, e106883.
47 N. G. Sgourakis, Y. Yan, S. A. McCallum, C. Wang and A. E. Garcia, *J. Mol. Biol.*, 2007, **368**, 1448-1457.
48 T. Zhang, J. Zhang, P. Derreumaux and Y. Mu, *J. Phys. Chem. B*, 2013, **117**, 3993-4002.
49 T. Zhang, W. Xu, Y. Mu and P. Derreumaux, *ACS. Chem. Neurosci.*, 2014, **5**, 148-159.
50 X. Wang, J. K. Weber, L. Liu, M. Dong, R. Zhou and J. Li, *Nanoscale*, 2015, **7**, 15341-15348.
51 G. Meisl, X. Yang, E. Hellstrand, B. Frohm, J. B. Kirkegaard, S. I. Cohen, C. M. Dobson, S. Linse and T. P. Knowles, *Proc. Natl. Acad. Sci. USA*, 2014, **111**, 9384-9389.
52 T. Watanabe-Nakayama, K. Ono, M. Itami, R. Takahashi, D. B. Teplow and M. Yamada, *Proc. Natl. Acad. Sci. USA*, 2016, **113**, 5835-5840.
53 R. Roychaudhuri, M. Yang, A. Deshpande, G. M. Cole, S. Frautschy, A. Lomakin, G. B. Benedek and D. B. Teplow, *J. Mol. Biol.*, 2013, **425**, 292-308.
54 Y. Miller, B. Ma, C.-J. Tsai and R. Nussinov, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 14128-14133.
55 Y. L. Lyubchenko, *J. Mol. Pharm. Org. Process Res.*, 2013, **1**, e107.
56 D. Huang, M. I. Zimmerman, P. K. Martin, A. J. Nix, T. L. Rosenberry and A. K. Paravastu, *J. Mol. Biol.*, 2015, **427**, 2319-2328.
57 J. C. Stroud, C. Liu, P. K. Teng and D. Eisenberg, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 7717-7722.
58 E. Cerf, R. Sarroukh, S. Tamamizu-Kato, L. Breydo, S. Derclaye, Y. F. Dufrene, V. Narayanaswami, E. Goormaghtigh, J. M. Ruysschaert and V. Raussens, *Biochem. J.*, 2009, **421**, 415-423.
59 C. Liu, M. Zhao, L. Jiang, P.-N. Cheng, J. Park, M. R. Sawaya, A. Pensalfini, D. Gou, A. J. Berk, C. G. Glabe, J. Nowick and D. Eisenberg, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 20913-20918.
60 G. M. Shankar, S. Li, T. H. Mehta, A. Garcia-Munoz, N. E. Shepardson, I. Smith, F. M. Brett, M. A. Farrell, M. J. Rowan, C. A. Lemere, C. M. Regan, D. M. Walsh, B. L. Sabatini and D. J. Selkoe, *Nat. Med.*, 2008, **14**, 837-842.
61 B. Urbanc, L. Cruz, F. Ding, D. Sammond, S. Khare, S. V. Buldyrev, H. E. Stanley and N. V. Dokholyan, *Biophys. J.*, 2004, **87**, 2310-2321.
62 S. Mitternacht, I. Staneva, T. Hard and A. Irback, *J. Mol. Biol.*, 2011, **410**, 357-367.
63 B. Barz and B. Urbanc, *PloS one*, 2012, **7**, e34345.
64 X. Zhu, R. P. Bora, A. Barman, R. Singh and R. Prabhakar, *J. Phys. Chem. B*, 2012, **116**, 4405-4416.
65 M. H. Viet, P. H. Nguyen, P. Derreumaux and M. S. Li, *ACS. Chem. Neurosci.*, 2014, **5**, 646-657.
66 G. Di Fede, M. Catania, E. Maderna, M. Morbin, F. Moda, L. Colombo, A. Rossi, A. Cagnotto, T. Virgilio, L. Palamara, M. Ruggerone, G. Giaccone, I. Campagnani, M. Costanza, R. Pedotti, M. Salvalaglio, M. Salmona and F. Tagliavini, *Sci. Rep.*, 2016, **6**, 20949.
67 W. M. Berhanu and U. H. E. Hansmann, *Protein Sci.* , 2012, **21**, 1837-1848.
68 W. Ye, D. Ji, W. Wang, R. Luo and H.-F. Chen, *J. Chem. Inf. Model.*, 2015, **55**, 1021-1029.
69 G. Reddy, J. E. Straub and D. Thirumalai, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 21459-21464.
70 R. B. Best, W. Zheng and J. Mittal, *J. Chem. Theory Comput.*, 2014, **10**, 5113-5124.