The Journal of Chemical Physics

Protein aggregation rate depends on mechanical stability of fibrillar structure

Cite as: J. Chem. Phys. **157**, 055101 (2022); https://doi.org/10.1063/5.0088689
Submitted: 19 February 2022 • Accepted: 06 July 2022 • Accepted Manuscript Online: 08 July 2022 • Published Online: 03 August 2022

Tran Thi Minh Thu and Mai Suan Li

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Protein aggregation rate depends on mechanical stability of fibrillar structure

Cite as: J. Chem. Phys. 157, 055101 (2022); doi: 10.1063/5.0088689
Submitted: 19 February 2022 • Accepted: 6 July 2022 •
Published Online: 3 August 2022

Tran Thi Minh Thu[1,2,3,4] and Mai Suan Li[4,5,a)]

AFFILIATIONS

[1] Faculty of Materials Science and Technology, University of Science-VNU HCM, 227 Nguyen Van Cu Street, District 5, Ho Chi Minh City 700000, Vietnam
[2] Vietnam National University, Ho Chi Minh City 700000, Vietnam
[3] International Center for Research on Innovative Biobased Materials (ICRI-BioM)-International Research Agenda, Lodz University of Technology, Zeromskiego 116, 90-924 Lodz, Poland
[4] Institute for Computational Science and Technology, SBI Building, Quang Trung Software City, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Viet Nam
[5] Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland

[a)] **Author to whom correspondence should be addressed:** masli@ifpan.edu.pl

ABSTRACT

The formation of the fibrillar structure of amyloid proteins/peptides is believed to be associated with neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis. Since the rate of aggregation can influence neurotoxicity, finding the key factors that control this rate is of paramount importance. It was recently found that the rate of protein aggregation is related to the mechanical stability of the fibrillar structure such that the higher the mechanical stability, the faster the fibril is formed. However, this conclusion was supported by a limited dataset. In this work, we expand the previous study to a larger dataset, including the wild type of Aβ42 peptide and its 20 mutants, the aggregation rate of which was measured experimentally. By using all-atom steered molecular dynamics (SMD) simulations, we can assess the mechanical stability of the fibril structure, which is characterized by the rupture force, pulling work, and unbinding free energy barrier. Our result confirms that mechanical stability is indeed related to the aggregation rate. Since the estimation of the aggregation rate using all-atom simulations is almost forbidden by the current computational capabilities, our result is useful for predicting it based on information obtained from fast SMD simulations for fibrils.

I. INTRODUCTION

Protein aggregation is associated with neurodegenerative diseases, such as Alzheimer's disease (AD), Parkinson's disease (PD), type II diabetes (T2D), and amyotrophic lateral sclerosis (ALS).[1,2] According to the amyloid cascade hypothesis,[3] the accumulation of amyloid beta (Aβ) peptide in the brain parenchyma is a crucial step in AD. Deposition of the intrinsically disordered protein α-synuclein (αS) was hypothesized to induce PD, while the islet amyloid polypeptide (IAPP) or 37 amino acid amylin is related to T2D. The 32 kDa superoxide dismutase 1 (SOD1), TAR DNA binding protein 43 (TPD-43), and 526 amino acids fused in sarcoma protein (FUS) play a key role in ALS.

Understanding the key factors that control the protein aggregation rate is important as it appears to be relevant to neurotoxicity.[4] For example, due to the last hydrophobic residues, Aβ42 (42 amino acids) self-assembles faster than Aβ40 (40 amino acids), which can lead to Aβ42 being more toxic than Aβ40. Thus, revealing the key factors that govern the aggregation pathways and rate is of paramount importance not only from the point of view of basic research but also from the point of view of medical treatment.

The factors influencing the formation of protein fibrils can be divided into two groups: environmental factors and the intrinsic properties of the polypeptide chain. Temperature, pH, crowders, chaperones, metal ions, osmolytes, etc., belong to the first group.

Proteins are stable over a certain pH range, but at extreme pH values, non-covalent interactions between amino acids become weaker, resulting in protein destabilization and aggregation.[5–7] A lipid membrane can modulate both the self-assembly rate and the morphology of fibrils. The dependence of the fibril formation rate on the concentration of crowders was experimentally measured by Linse *et al.*[8] who reported that copolymeric nanoparticles accelerate β2-microglobulin aggregation. The opposite effect was observed for Aβ peptides, the fibril growth of which is retarded by nanoparticles.[9] These experimental facts can be explained by taking into account the competition between energy and entropy.[10] From this perspective, the propensity of proteins to self-assembly does not depend on the details of the systems under study. Using a quartz crystal microbalance assay with high accuracy in measuring the fibril growth rate, White *et al.* showed that cosolutes accelerate fibril elongation.[11] This finding is supported by simulations[12,13] and is in line with the depletion theory.[14,15] Additional environmental factors, including those associated with *in vivo* conditions, were discussed in recent reviews.[16,17]

Intrinsic factors, such as the order of amino acids in the sequence, hydrophobicity, and net charge, also control the self-aggregation rate of proteins as demonstrated by Chiti *et al.*[18] using the mutagenesis technique. This important experimental finding was later confirmed by simulations.[19,20]

Changes in the amino-acid sequence can vary the aggregation rate as well as the toxicity of protein, and some mutations alter the morphology of the fibril structure.[21] The high correlation between the hydration free energy and the aggregation rate of proteins Aβ42, HypF-N, and AcP[20] supports the hypothesis that rich hydrophobic regions level up aggregation.[22–25]

Protein conformation is controlled by specific salt bridges that are formed by charged residues, and the role of charge in prevention of self-assembly is also known through repulsion between chains.[26,27] A recent study by Thu *et al.* found a high correlation between the β content in the monomeric state and the rate of aggregation such that the higher the aggregation propensity, the higher the beta content in the monomeric state.[28] Using lattice models, Colizzi *et al.* showed that the fibril formation time exponentially grows with the population of the so called fibril-prone structure in the monomeric state.[29] This result was confirmed by all-atom[30] and coarse-grained[31] simulations.

Recently, Kouza *et al.*[32] hypothesized that the fibril formation time is related to the mechanical stability of the fibril state in such a way that the faster the fibril formation, the more stable the fibril state. Using steered molecular dynamics (SMD) simulation[33,34] to assess the mechanical stability of the fibril structure, it has been shown that the Aβ42 fibril is more stable than Aβ40,[32] which is consistent with the hypothesis of Kouza *et al.*, as Aβ42 is known to aggregate

TABLE I. Rupture force, pulling work, and unbinding free energy barrier obtained by SMD simulation at 300 K and pulling speed v = 1 nm/ns. The results were averaged over 10 independent trajectories. Theoretical estimate of $\log(\kappa_{mut}/\kappa_{wt})$ using Eq. (7) and the relative experimental aggregation rate of all mutations are shown. References to experimental works are given in the last column.

		Simulation				Experimental aggregation rate	
No	Mutations	Fmax (pN)	Work (kcal/mol)	$\Delta G^{\ddagger}_{unbind}$ (kcal/mol)	$\log(\kappa_{mut}/\kappa_{wt})$ given by Eq. (7)	$\log(\kappa_{mut}/\kappa_{wt})$	References
1	2NAO	1922.9 ± 137.9	473.2 ± 48.2	124.6 ± 18.5	5.3	0	
2	I41D-A42Q	1550.4 ± 178.6	392.5 ± 33.8	72.1 ± 19.2	4.3	−0.964	
3	I41D-A42S	1424.5 ± 89.7	337.3 ± 67.5	64.6 ± 8.3	3.7	−0.913	
4	I41H-A42D	2153.6 ± 228.8	485.4 ± 43.4	164.0 ± 44.6	5.8	−0.708	20, 38
5	I41E-A42L	2117.8 ± 203.4	433.5 ± 66.4	163.6 ± 41.4	5.4	−0.445	
6	I41H-A42N	1632.1 ± 256.8	437.0 ± 55.0	90.9 ± 34.5	4.6	−0.837	
7	A21G	1648.1 ± 114.6	362.2 ± 17.9	83.8 ± 17.4	4.2	−0.671	20, 37
8	I41T-A42N	1923.6 ± 155.1	447.2 ± 32.1	132.2 ± 31.1	5.2	−0.605	
9	I41T-A42Q	1757.1 ± 141.4	462.9 ± 36.8	99.5 ± 35.5	4.9	−0.59	
10	I41Q-A42Y	1876.7 ± 294.6	434.9 ± 39.3	137.0 ± 51.7	5.1	−0.382	
11	I41L-A42N	1666.9 ± 171.9	399.0 ± 35.7	118.0 ± 33.0	4.6	−0.561	
12	I41Q- A42L	1842.7 ± 198.1	410.7 ± 32.8	117.8 ± 35.7	4.8	−0.295	
13	I41T-A42M	1635.9 ± 316.3	419.2 ± 64.1	142.5 ± 49.9	4.8	−0.292	20, 38
14	I41T- A42I	1792.5 ± 201.3	454.5 ± 44.3	169.6 ± 28.6	5.1	−0.075	
15	I41K	1821.7 ± 247.9	461.4 ± 35.0	118.8 ± 40.4	5.1	−0.518	
16	I41K-A42L	2126.6 ± 237.4	443.0 ± 59.1	149.5 ± 46.9	5.4	−0.379	
17	I41R-A42R	1543.8 ± 204.0	437.6 ± 54.1	105.4 ± 27.9	4.6	−0.324	
18	A42R	1972.3 ± 337.0	470.1 ± 43.4	111.1 ± 30.6	5.3	−0.034	
19	E22G	2240.1 ± 323.3	519.0 ± 47.0	198.0 ± 55.3	6.3	0.209	20, 37
20	D23N	2376.2 ± 208.9	534.9 ± 38.0	212.1 ± 46.0	6.9	0.238	36
21	E22K	2641.4 ± 312.7	574.0 ± 51.2	279.8 ± 63.8	7.5	0.545	20, 37

faster than Aβ40.[35] Conventional MD and SMD simulations of formation of trimers of the two short peptides KLVFF and FVFLM and their stability also confirmed this hypothesis.[32] Thus, the relationship between the mechanical stability of fibrils and fibril formation time was supported by only two examples, which prompted us to conduct this study on a larger dataset.

Since the estimation of the fibril formation time of proteins, τ_{fib}, using all-atom models is computationally prohibited, proving that a correlation between τ_{fib} and the mechanical stability of the fibril state is very useful. This is due to the fact that such a relation allows one to determine τ_{fib} from mechanical stability, which can be obtained with low computational resources using non-equilibrium SMD simulation. Therefore, the main goal of this paper is to test the hypothesis of Kouza *et al.*[32] by performing molecular simulations on a large set of sequences, the aggregation rate of which has been measured experimentally. Chong and Ham[20] used experimental aggregation rates obtained for various Aβ42 mutations to find a correlation between hydrophobicity and aggregation rate.[20] Following this idea, we perform SMD simulations on a similar Aβ42 dataset, but to confirm our hypothesis of a correlation between the mechanical stability of the fibrillar state and fibril formation time. Thus, the set contains wild-type (WT) Aβ42 and 20 mutants, and their experimental aggregation rates were collected from various sources and shown in our previous work[28] (Table I). Mechanical stability, which is characterized by rupture force, non-equilibrium work, or unbinding free energy, was probed by pulling the outermost chain from the rest of the fibrillar structure using all-atom SMD simulations with a CHARMM 36m force field and TIP3P water model. We showed that the experimentally determined aggregation rate is correlated with the mechanical stability of the fibril state in such a way that the higher the rupture force, pulling work, or unbinding free energy, the faster the formation of fibrils.

II. MATERIALS AND METHOD

A. Choice of Aβ42 fibril structure and its mutations

The sequences of Aβ42 and its 20 mutations are shown in Fig. 1. Their aggregation rate has been measured experimentally[36–38] (Table I). Due to polymorphism,[39] the fibril structure of Aβ42 adopts different morphologies, including the U-shape [PDB (protein database) ID: 2BEG[40]], LS-shape (PDB ID: 5OQV[41]), and S-shape (PDB ID: 2NAO[42]). 2BEG was not considered as it is the fibril structure of the truncated sequence $A\beta_{17-42}$. 5OQV was also excluded because it was determined under low pH conditions. Therefore, in this study, we use 2NAO of the full-length Aβ42 and its structure is shown in Fig. 2(a).

Mutations: I41D-A42Q, I41D-A42S, I41H-A42D, I41E-A42L, I41H-A42N, A21G, I41T-A42N, I41T-A42Q, I41Q-A42Y, I41L-A42N, I41Q- A42L, I41T-A42M, I41T- A42I, I41K, I41K-A42L, I41R-A42R, A42R, E22G, D23N, E22K.

FIG. 1. Sequences of Aβ42-WT and 20 mutations (see also Table I).

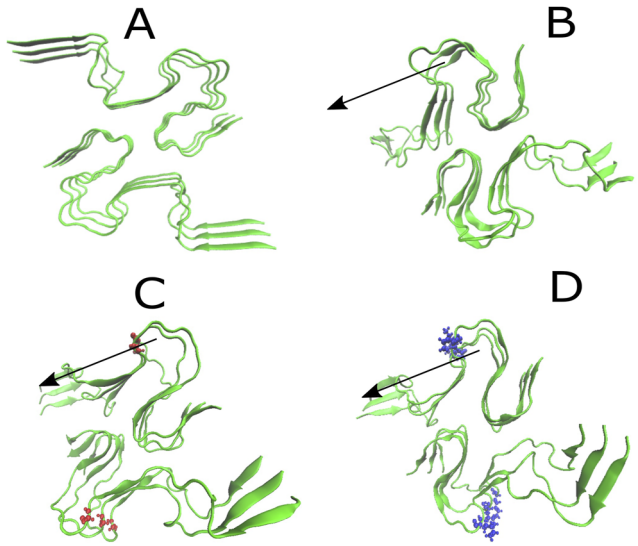

FIG. 2. 2NAO structure retrieved from PDB (a), 2NAO structure after equilibration (b), structure of 2NAO with A21G mutation after equilibration (c), and structure of 2NAO with E22K mutation after equilibration (d). Red and blue denote mutations A21G and E22K, respectively.

B. MD simulation

The Aβ42 protofibril structure was solvated in a 12 × 10.6 × 22.6 nm^3 rectangular box filled with water molecules. In the wild type case, the system contains 281 662 atoms, including 3762 Aβ42 atoms, 92 512 water molecules, and 191 Na^+ and 173 Cl^- ions. The box size and the number of atoms changed slightly with the mutations. PBC boundary conditions were applied for all systems.

CHARMM 36m[43] proved to be one of the best force fields for describing intrinsically disordered proteins.[44,45] In combination with the TIP3P water model,[46] this force field can produce a conformational ensemble of Aβ peptide.[47] Therefore, we chose CHARMM36m + TIP3P water. Since our goal is to compare the mechanical stability of 21 sequences with the same force field and water model, their choice should not qualitatively affect our result.

The leapfrog algorithm[48] was used to integrate the equations of motion with a time step of 2 fs. The length of all bonds was constrained by the LINCS algorithm.[49] The velocity of atoms was changed periodically by a v-rescale temperature coupling, but the system's temperature was kept stable. The relaxation time is 0.1 ps. We calculated the van der Waals (vdW) and electrostatic forces with a short range cutoff radius of 1.4 nm. The long-range interaction was calculated using the particle mesh Ewald (PME) method. The temperature was maintained at 300 K using the V-rescale thermostat algorithm,[50] and the pressure was kept at 1 bar using the Parrinello–Rahman algorithm.[51] MD simulations were conducted using the GROMACS package, version 5.1.2.[52]

C. Preparing initial structures for SMD simulation

The fibril structure of 20 variants was created by mutation in all chains of 2NAO[42] using the Pymol package.[53] Wild type (WT) and

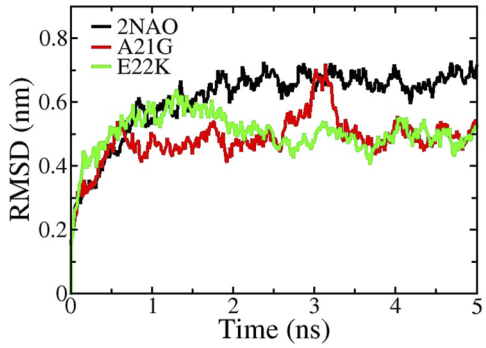

FIG. 3. Root mean square deviation of the fibril structure of WT, A21G, and E22K as a function of time. The last snapshots selected for SMD simulations are shown in Fig. 2.

mutants was equilibrated in several steps. First, the energy of the system was minimized by the steepest descent method and equilibrated by performing a constant volume (NVT) MD simulation for 1 ns followed by a constant pressure (NPT) simulation for 1 ns, maintaining the temperature at 300 K and the pressure at 1.0 bar. Finally, the production NPT MD run was carried out for 5 ns at the same temperature and pressure. Figure 3 shows the time dependence of the root mean square deviation (RMSD) of WT and A21G and E22K variants. Since the RMSD saturates after about 2 ns, these systems were equilibrated and the last snapshot (Fig. 2) was chosen as the initial structure for the SMD simulation.

D. SMD simulation

Proteins were placed in a rectangular box that has the size large enough to have space for pulling simulations and satisfy the minimum image convention condition. An external force was applied to the dummy atom, which is connected with the atom closest to the center of mass (COM) of the outermost chain, through a spring with a stiffness k (Fig. 4). The pulling direction was chosen parallel to the line connecting COMs of the outermost chain and its neighbor [Fig. 4; see also Figs. 2(b), 2(c), and 2(d)]. To prevent the non-pulled part from moving during the pull, all Cα atoms of the chain next to the chain being pulled were restrained.

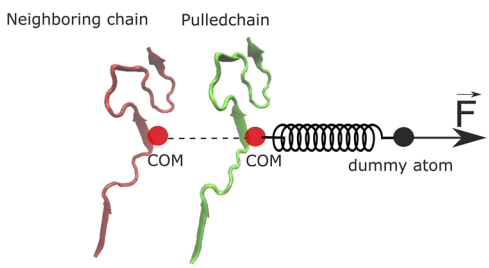

FIG. 4. Protocol of SMD simulation. An external force F is applied to the dummy atom (black) connected with the atom closest to the center of mass (COM) (red) of the pulled chain through a spring with a stiffness k. The pulling direction is parallel to the line connecting the two COMs.

The force experienced by the pulled atom is $F = k(vt - x)$, where x is its displacement from the initial position and v is the pulling speed. The spring constant k was set to 1000 kJ/mol/nm^2, which is the typical value used in the AFM experiment.[54] Like preceding studies,[55–57] in all SMD simulations, the pulling speed was chosen to be $v = 1$ nm/ns, which is much higher than the experimental speed. However, the pulling speed will affect the absolute value of the rupture force and work, but the relative mechanical stability should not depend on v.[58,59] This choice of k and v led to a qualitative agreement with experiments on the mechanical stability of protein–ligand,[58,59] protein–protein,[60,61] and fibrillar[62] systems.

As shown previously,[62] the rupture force F_{max}, which appears when the pulled monomer detaches from the core, can be used to characterize the mechanical stability of the entire fibril. In addition, the pulling work W, which is also a useful scoring function, is defined by the following equation:

$$W = \int f(x)dx = \frac{1}{2}\sum_{i=1}^{n-1}(f_{i+1} + f_i)(x_{i+1} - x_i), \tag{1}$$

where n is the number of simulation steps and f_i and x_i are the force experienced by the pulled chain and position at step i. W is more robust than F_{max} because the pulling work is a function of the entire process, while the rupture force is determined only in a single state.[58]

The mechanical stability of fibrils can be assessed either though the unbinding or binding free energy barriers,[62] which can be obtained using Jarzynski's equality.[63] Extending this equality to the case when an external force is applied to the system at a constant loading rate, Hummer and Szabo showed that the free energy ΔG is related to W by the following equation:[64]

$$\exp\left(\frac{-\Delta G}{k_B T}\right) = \left\langle \exp\left(\frac{W(t) - \frac{1}{2}k(z_t - vt)^2}{k_B T}\right)\right\rangle_N, \tag{2}$$

where $\langle\cdots\rangle N$ stands for averaging over N trajectories, z_t is the displacement of the pulled atom along the puling direction, and W is given by Eq. (1). We performed 10 independent SMD simulations ($N = 10$). $W(t)$ can easily be calculated, and its typical time dependence is shown in Fig. 5(c). Since z_t is also known from simulation, we can calculate $\exp(-[W - k(z_t - vt)^2/2]/k_B T)$ for each trajectory and then average it over all trajectories to get ΔG vs time (or displacement). A typical profile of ΔG is shown in Fig. 5(d), where the maximum corresponds to the transition state (TS).[56] The free energy of the bound state can be defined as $\Delta G_{bound} = \Delta G\ (t = 0)$, while the free energy of the unbound state is $\Delta G_{bound} = \Delta G\ (t = t_{end})$, where t_{end} is the time when the simulation terminates and the dissociation process has been completed. Then, the unbinding $\Delta G^{\ddagger}_{unbind}$ and binding $\Delta G^{\ddagger}_{bind}$ barriers are defined as follows:[56]

$$\begin{aligned}\Delta G^{\ddagger}_{unbind} &= \Delta G_{TS} - \Delta G_{bound} = \Delta G_{TS} - \Delta G(t = 0),\\ \Delta G^{\ddagger}_{bind} &= \Delta G_{TS} - \Delta G_{unbound} = \Delta G_{TS} - \Delta G(t = t_{end}).\end{aligned} \tag{3}$$

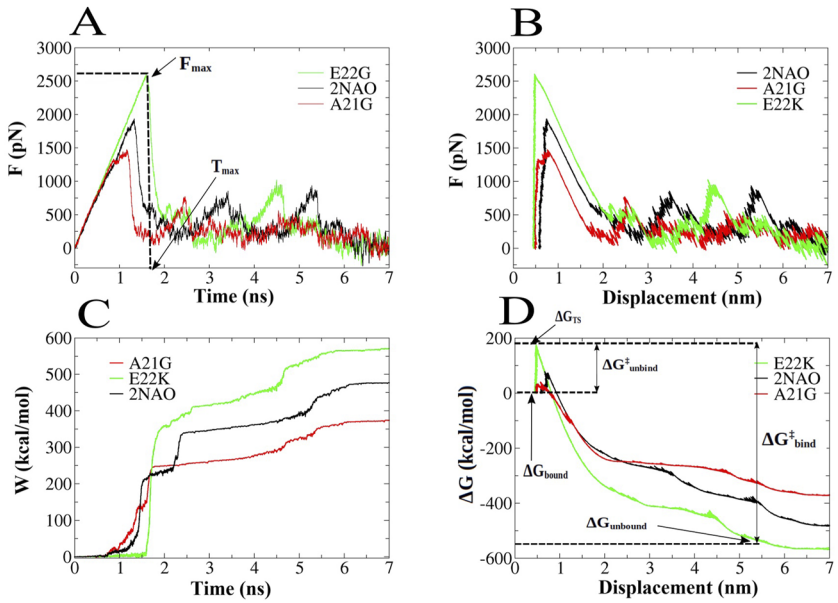

FIG. 5. Force of WT (2NAO), E22K, and A21G as a function of time (a) and displacement (b). The rupture force F_{max} occurs at time t_{max}. (c) Pulling work as a function of time. (d) Free energy profiles with a maximum corresponding to the transition state (TS). The unbinding and binding barriers $\Delta G^{\ddagger}_{unbind}$ and $\Delta G^{\ddagger}_{bind}$ are shown by double-headed arrows.

III. RESULT AND DISCUSSION

A. Evaluation of rupture force, pulling work, and free energy barriers

The dependence of the force experienced by the tensioned chain of three representative sequences on time and displacement is shown in Figs. 5(a) and 5(b). At the breaking point, the appearance of the maximum force F_{max} at time t_{max} indicated a rupture event with critical interactions disrupted. Before this event, the force exerted on the pulled chain increased rapidly and almost linearly depends on the displacement. The mutations do not change the overall shape of the force profiles that are less complex compared to the oligomer case, where the entanglement of chains may occur due to a compact structure.[62] The structure at t_{max} does not differ significantly from the initial structure (Fig. 6). Furthermore, stretching results in complete separation of the pulled chain from the preformed template. The mean value of F_{max} was calculated by averaging over 10 independent SMD trajectories for each system (Table I).

At the beginning, the work is almost zero because the pulled chain cannot leave the fibril. As time increases, the work increases and reaches a stable value [Fig. 5(c)]. Therefore, for each run, W was defined as the work obtained at the end of the simulation, and its average value is shown in Table I. Using the free energy profiles [Fig. 5(d)] and Eq. (3), we can evaluate the unbinding $\Delta G^{\ddagger}_{unbind}$ and binding $\Delta G^{\ddagger}_{bind}$ barriers (Table I).

Aggregation rates were collected from previous experimental works with references shown in the last column of Table I (see also Thu *et al.*[28]). Experiments have shown that E22K speeds up the aggregation (by 1.7 times) and A21G slows it down (by 0.5 times) compared to WT.[20,37] A21G decreases the hydropathy index from 1.8 to −0.4 (https://en.wikipedia.org/wiki/Amino_acid), probably resulting in the reduced aggregation rate, which is in line with the fact that the lower the hydrophobicity, the slower the aggregation.[18,19] Understanding why E22K promotes aggregation is more challenging. Since the hydropathy index of E (−3.5) is compatible with that of K (−3.9), a small change in hydrophobicity cannot answer this question. Moreover, because this mutation increases the net charge of Aβ42 from +3e to +5e, this should slow fibril formation but not speed it up. Recently, Yang *et al.*[65] demonstrated that the size of the side chain at position E22 plays a more crucial role than its charge in the increase in aggregation caused by E22K.

In our simulation, the rupture force of A21G (1648 pN) is lower than that of WT (1923 pN) and E22K (2641 pN). This result is consistent with the trend observed in our previous work that the faster the formation of fibrils, the higher their mechanical stability.[32] Data on the pulling work, binding, and unbinding barriers of WT, E22K, and A21G also support this conclusion.

1. Rupture force increases with the aggregation rate: Evidence from a dataset of 21 sequences

Figure 7 shows $\ln(\kappa_{mut}/\kappa_{wt})$ as a function of F_{max}, where experimental aggregation rates κ_{mut} and κ_{wt} are shown in Table I. The linear fit works with the correlation level of 0.74, which is reasonable taking into account the complexity of the fibril formation process and the fact that data were collected from different experimental groups. The relationship between κ and F_{max} can be described by an exponential function,

$$\kappa = \kappa_0 \exp(cF_{max}), \quad c = 0.000\,965\,14, \tag{4}$$

where κ_0 is a fitting constant and F_{max} is measured in pN.

The main conclusion followed from Eq. (4) is that the aggregation rate increases with the rupture force, implying that the faster the aggregation, the higher the mechanical stability characterized by F_{max}. This hypothesis was formulated in our previous work,[32] but its validity is further supported by a larger dataset of 21 Aβ42 variants.

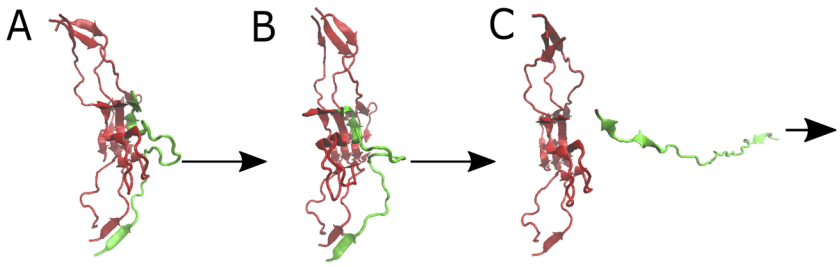

FIG. 6. Unbinding events of WT. (a) Initial structure obtained in the 5 ns conventional MD simulation as described in Sec. II [see Fig. 2(b)]. The distance between COMs of the pulled chain (green) and its neighbor is $x = 0.575$ nm and $t = 0$. (b) Structure at $t = t_{max} = 1207$ ps and $x = 0.731$ nm. (c) Structure in the unbound state at $x = 6.479$ nm and $t = 6000$ ps. At this distance, the pulled chain is entirely detached from the core.

B. Correlation between pulling work and aggregation rate

The pulling work for all systems ranges from 337.3 to 574 kcal/mol (Table I), where the smallest and largest values correspond to I41D-A42S and E22K, respectively. These variants also have a minimum and maximum rupture force. The correlation level between W and the experimental aggregation rate is $R = 0.79$ (Fig. 8).

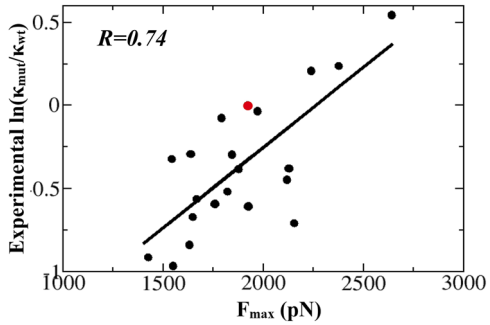

FIG. 7. Dependence of the logarithm of relative aggregation rate (upper panel) and the relative aggregation rate (lower panel) on rupture force. The red circle refers to WT. Linear fit is y = −2.185 + 0.000 965 14 *x. Correlation level $R = 0.74$.

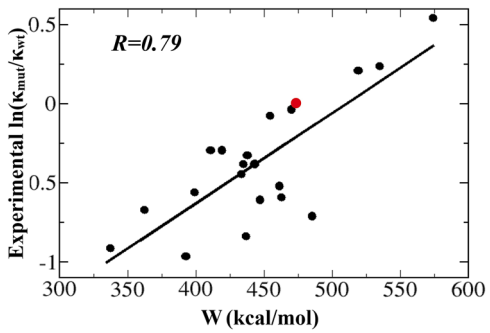

FIG. 8. Dependence of the logarithm of the relative aggregation rate (upper panel) and the relative aggregation rate on the pulling rate (lower panel). The red circle refers to WT. Linear fit is y = −2.9188 + 0.005 7183x. Correlation level $R = 0.79$.

The dependence of the logarithm of the relative aggregation rate on the pulling work can be expressed as

$$\kappa = \kappa_0 \exp(cW), \quad c = 0.005\,718\,3, \tag{5}$$

where κ_0 is a prefactor and W is measured in kcal/mol.

As in previous studiess,[56,58] W has a better correlation with the experiment than the rupture force, and our result shows the same trend. Equation (5) confirms the relationship between the aggregation rate and the mechanical stability of the fibril state.

1. Free energy unbinding barrier has a better correlation with experiment than W and F_{max}

It can be shown that the unbinding and binding barriers for the studied systems have almost the same correlation with the experiment. However, we present data only for $\Delta G^{\ddagger}_{\text{unbind}}$ (Table I), since it is directly related to the unbinding process. $\Delta G^{\ddagger}_{\text{unbind}}$ greatly varies from 64.6 (I41D-A42S) to 279.8 kcal/mol (E22K), and its large value is associated with the fact that it was obtained out of equilibrium due to fast pulling.

The correlation level between the unbinding free energy barrier and the experimental aggregation rate is $R = 0.81$ (Fig. 9), which is higher than that for W and F_{max}. Thus, $\Delta G^{\ddagger}_{\text{unbind}}$ is the best scoring function not only for protein–ligand unbinding[56] but also for mechanical stability of fibrils.

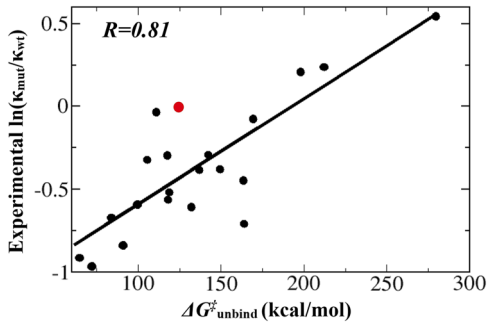

FIG. 9. Dependence of the logarithm of the relative aggregation rate (upper panel) and the relative aggregation rate on the unbinding free energy barrier (lower panel). The red circle refers to WT. Linear fit is y = −1.2274 + 0.006 3659 x. Correlation level $R = 0.81$.

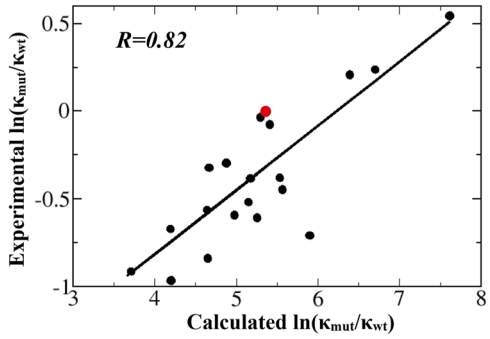

FIG. 10. Aggregation rate obtained by Eq. (7) as a function of the experimental rate. The linear fit gives the correlation level $R = 0.82$.

The exponential fit gives the following dependence of the aggregation rate on $\Delta G^{\ddagger}_{\text{unbind}}$:

$$\kappa = \kappa_0 \exp(c\Delta G^{\ddagger}_{\text{unbind}}), \quad c = 0.009\,795\,4, \tag{6}$$

where $\Delta G^{\ddagger}_{\text{unbind}}$ has a unit of kcal/mol. Again, the increase in the rate of aggregation with the unbinding free energy barrier [Eq. (6)] confirms our hypothesis[32] that the mechanical stability of the fibrillar state is directly related to the rate of the fibril formation process.

2. Dependence of the aggregation rate simultaneously on F_{max}, W, and $\Delta G^{\ddagger}_{unbind}$

Following Dobson *et al.*,[18] we combine the effect of rupture force, work, and unbinding free energy, representing the aggregation in the following form:

$$\ln(\kappa_{\text{mut}}/\kappa_{\text{wt}}) = A \cdot F_{\max} + B \cdot W + C \cdot \Delta G^{\ddagger}_{\text{unbind}}, \tag{7}$$

where A, B, and C are the slopes resulting from a linear fit between $\ln(\kappa_{\text{mut}}/\kappa_{\text{wt}})$ and $F_{\max}$ and W and $\Delta G^{\ddagger}_{\text{unbind}}$, respectively. From Figs. 7–9, we obtain $A = 0.001$, $B = 0.006$, and $C = 0.006$. The predicted aggregation rate obtained using Eq. (7) is shown in column 6 of Table I and correlates with experimental aggregation rates, with $R = 0.82$ (Fig. 10). Thus, the correlation between theoretical [Eq. (7)] and experimental rates is sufficiently high.

IV. CONCLUSION

In this study, we have performed the SMD simulation for studying the mechanical unbinding of a single chain from the fibril structure of 21 variants of Aβ42 including WT. The correlation between the rupture force, pulling work, unbinding free energy barrier and the experimental aggregation rate has been thoroughly analyzed. Since κ increases with $F_{\max}$, W, and ΔG_{unbind} [Eqs. (4)–(6)], which can be used to characterize the mechanical stability, our results support the hypothesis that the higher the mechanical stability of the fibrillary state, the faster the fibril formation.[32] The rationale for this hypothesis is that the kinetic stability can be assessed through the mechanical stability of the fibril state.

It is known that the formation of protein fibrils takes several hours to months, which makes the calculation of the fibril formation time using all-atom models impossible within the framework of present computing facilities. From this point of view, Eqs. (4)–(6) are very useful as they allow for predicting the aggregation relative rate through $F_{\max}$, W, and $\Delta G^{\ddagger}_{\text{unbind}}$, which can easily be obtained using SMD simulations. Among them, Eq. (7) is best suited for this purpose not only because the correlation between theory and experiment is high ($R = 0.82$, Fig. 10) but also because it incorporates the dependence on all three quantities, including the rupture force, work, and unbinding barrier. Together with previous work,[32] this study enriches our knowledge of the key factor that controls the kinetics of protein fibril formation. In particular, faster protein aggregation leads to a fibrillar state with higher mechanical stability.

ACKNOWLEDGMENTS

This research was funded by Vietnam National University, Ho Chi Minh City (VNU-HCM) (Grant No. C2020-18-19); Department of Science and Technology, Ho Chi Minh city, Vietnam (Grant No. 44/2020/HĐ-QPTKHCN); and Narodowe Centrum Nauki in Poland (Grant No. 2019/35/B/ST4/02086). Support from the supercomputer center TASK in Gdansk and PLGrid infrastructure, Poland, is highly appreciated.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Tran Thi Minh Thu: Data curation (lead); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Writing – original draft (equal); Writing – review & editing (supporting). **Mai Suan Li:** Conceptualization (lead); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (lead); Project administration (lead); Supervision (lead); Writing – original draft (equal); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] F. Chiti and C. M. Dobson, Annu. Rev. Biochem. **86**, 27 (2017).
[2] J. Nasica-Labouze *et al.*, Chem. Rev. **115**, 3518 (2015).
[3] J. A. Hardy and G. A. Higgins, Science **256**, 184 (1992).
[4] D. J. Selkoe, Physiol. Rev. **81**, 741 (2001).
[5] M. Fändrich, M. A. Fletcher, and C. M. Dobson, Nature **410**, 165 (2001).
[6] L. Nielsen *et al.*, Biochemistry **40**, 6036 (2001).
[7] L. Adler-Abramovich and E. Gazit, Chem. Soc. Rev. **43**, 6881 (2014).
[8] S. Linse *et al.*, Proc. Natl. Acad. Sci. U. S. A. **104**, 8691 (2007).
[9] C. Cabaleiro-Lago *et al.*, J. Am. Chem. Soc. **130**, 15437 (2008).
[10] N. T. Co, C.-K. Hu, and M. S. Li, J. Chem. Phys. **138**, 185101 (2013).
[11] D. A. White *et al.*, J. Am .Chem. Soc. **132**, 5170 (2010).
[12] A. Magno, A. Caflisch, and R. Pellarin, J. Phys. Chem. **1**, 3027 (2010).
[13] E. P. O'Brien *et al.*, J. Phys. Chem. **2**, 1171 (2011).
[14] H.-X. Zhou, G. Rivas, and A. P. Minton, Annu. Rev. Biophys. **37**, 375 (2008).
[15] S. Asakura and F. Oosawa, J. Chem. Phys. **22**, 1255 (1954).

[16]S. K. Chaturvedi *et al.*, Process Biochem. **51**, 1183 (2016).
[17]W. Wang and C. J. Roberts, Int. J. Pharm. **550**, 251 (2018).
[18]F. Chiti *et al.*, Nature **424**, 805 (2003).
[19]M. S. Li *et al.*, Phys. Rev. Lett. **105**, 218101 (2010).
[20]S.-H. Chong and S. Ham, Angew. Chem. **53**, 3961 (2014).
[21]T. T. Huang *et al.*, Neurochem. Int. **129**, 104512 (2019).
[22]C. Tanford, Science **200**, 1012 (1978).
[23]K. A. Dill, Biochemistry **29**, 7133 (1990).
[24]A. Y. Ben-Naim, *Hydrophobic Interactions* (Springer Science & Business Media, 2012).
[25]W. Kim and M. H. Hecht, Proc. Natl. Acad. Sci. U. S. A. **103**, 15824 (2006).
[26]F. Chiti *et al.*, Proc. Natl. Acad. Sci. U. S. A. **99**, 16419 (2002).
[27]A. T. Petkova *et al.*, Proc. Natl. Acad. Sci. U. S. A. **99**, 16742 (2002).
[28]T. T. M. Thu *et al.*, J. Chem. Phys. **150**, 225101 (2019).
[29]F. Colizzi *et al.*, J. Am. Chem. Soc. **132**, 7361 (2010).
[30]H. B. Nam *et al.*, J. Chem. Phys. **132**, 165104 (2010).
[31]D. Chakraborty *et al.*, Proc. Natl. Acad. Sci. U. S. A. **117**, 19926 (2020).
[32]M. Kouza *et al.*, J. Chem. Phys. **148**, 215106 (2018).
[33]H. Grubmüller, B. Heymann, and P. Tavan, Science **271**, 997 (1996).
[34]S. Izrailev *et al.*, *Computational Molecular Dynamics: Challenges, Methods, Ideas* (Springer, 1999), p. 39.
[35]G. Meisl *et al.*, Proc. Natl. Acad. Sci. U. S. A. **111**, 9384 (2014).
[36]K. Murakami *et al.*, J. Biol. Chem. **278**, 46179 (2003).
[37]C. Nilsberth *et al.*, Nat. Neurosci. **4**, 887 (2001).
[38]W. Kim and M. H. Hecht, J. Biol. Chem. **280**, 35069 (2005).
[39]R. Tycko, Protein Sci. **23**, 1528 (2014).
[40]T. Luhrs *et al.*, Proc. Natl. Acad. Sci. U. S. A. **102**, 17342 (2005).
[41]L. Gremer *et al.*, Science **358**, 116 (2017).
[42]M. A. Walti *et al.*, Proc. Natl. Acad. Sci. U. S. A. **113**, E4976 (2016).
[43]J. Huang *et al.*, Nat. Methods **14**, 71 (2017).
[44]J. Huang and A. D. MacKerell, Jr., Curr. Opin. Struct. Biol. **48**, 40 (2018).
[45]N. T. Co, M. S. Li, and P. Krupa, in *Computer Simulations of Aggregation of Proteins and Peptides*, edited by M. S. Li *et al.* (Springer, New York, 2022), p. 51.
[46]P. Mark and L. Nilsson, J. Phys. Chem. A **105**, 9954 (2001).
[47]S. Samantray *et al.*, J. Chem. Inf. Model. **60**, 6462 (2020).
[48]D. Fincham, Mol. Simul. **8**, 165 (1992).
[49]B. Hess *et al.*, J. Comput. Chem. **18**, 1463 (1997).
[50]G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).
[51]M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).
[52]M. J. Abraham *et al.*, SoftwareX **1**, 19 (2015).
[53]L. Schrodinger, Version 1, 2010.
[54]G. Binnig, C. F. Quate, and C. Gerber, Phys. Rev. Lett. **56**, 930 (1986).
[55]B. K. Mai, M. H. Viet, and M. S. Li, J. Chem. Inf. Model. **50**, 2236 (2010).
[56]D. T. Truong and M. S. Li, J. Phys. Chem. B **122**, 4693 (2018).
[57]N. Q. Thai *et al.*, Mol. Simul. **44**, 335 (2018).
[58]Q. V. Vuong, T. T. Nguyen, and M. S. Li, J. Chem. Inf. Model. **55**, 2731 (2015).
[59]B. Heymann and H. Grubmüller, Chem. Phys. Lett. **303**, 1 (1999).
[60]H. L. Nguyen *et al.*, J. Phys. Chem. B **124**, 7336 (2020).
[61]H. Nguyen *et al.*, J. Phys. Chem. B **126**, 2812 (2022).
[62]A. B. Poma *et al.*, J. Phys. Chem. B **19**, 28195 (2021).
[63]C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[64]G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U. S. A. **98**, 3658 (2001).
[65]X. Yang *et al.*, Proc. Natl. Acad. Sci. U. S. A. **115**, E5849 (2018).